

INFORMALITY EXIT AND EXCLUSION



GUILLERMO E. PERRY • WILLIAM F. MALONEY
OMAR S. ARIAS • PABLO FAJNZYLBER
ANDREW D. MASON • JAIME SAAVEDRA-CHANDUVI



INFORMALITY:
EXIT AND EXCLUSION

THE PRINCIPAL AUTHORS OF THIS BOOK ARE AS FOLLOWS:

- Overview: Guillermo E. Perry and William F. Maloney
- Chapter 1: William F. Maloney and Jaime Saavedra-Chanduvi
- Chapter 2: William F. Maloney and Omar S. Arias
- Chapter 3: Omar S. Arias
- Chapter 4: William F. Maloney and Mariano Bosch
- Chapter 5: Pablo Fajnzylber and William F. Maloney
- Chapter 6: Pablo Fajnzylber
- Chapter 7: Andrew D. Mason
- Chapter 8: Jaime Saavedra-Chanduvi

INFORMALITY:
EXIT AND EXCLUSION

Guillermo E. Perry

William F. Maloney

Omar S. Arias

Pablo Fajnzylber

Andrew D. Mason

Jaime Saavedra-Chanduvi



THE WORLD BANK
Washington, D.C.

©2007 The International Bank for Reconstruction and Development / The World Bank
1818 H Street, NW
Washington, DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org
E-mail: feedback@worldbank.org

All rights reserved

1 2 3 4 10 09 08 07

This volume is a product of the staff of the International Bank for Reconstruction and Development / The World Bank. The findings, interpretations, and conclusions expressed in this volume do not necessarily reflect the views of the Executive Directors of The World Bank or the governments they represent.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgement on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Rights and Permissions

The material in this publication is copyrighted. Copying and/or transmitting portions or all of this work without permission may be a violation of applicable law. The International Bank for Reconstruction and Development / The World Bank encourages dissemination of its work and will normally grant permission to reproduce portions of the work promptly.

For permission to photocopy or reprint any part of this work, please send a request with complete information to the Copyright Clearance Center Inc., 222 Rosewood Drive, Danvers, MA 01923, USA; telephone: 978-750-8400; fax: 978-750-4470; Internet: www.copyright.com.

All other queries on rights and licenses, including subsidiary rights, should be addressed to the Office of the Publisher, The World Bank, 1818 H Street, NW, Washington, DC 20433, USA; fax: 202-522-2422; e-mail: pubrights@worldbank.org.

ISBN-10: 0-8213-7092-8
ISBN-13: 978-0-8213-7092-6
eISBN-10: 0-8213-7093-6
eISBN-13: 978-0-8213-7093-3
DOI: 10.1596/978-0-8213-7092-6

Library of Congress Cataloging-in-Publication Data

Perry, Guillermo.

Informality : exit and exclusion / Guillermo E. Perry and William F. Maloney ; Omar S. Arias ... [et al.].
p. cm.

Includes bibliographical references.

ISBN-13: 978-0-8213-7092-6

ISBN-10: 0-8213-7092-8

ISBN-13: 978-0-8213-7093-3 (electronic (13))

ISBN-10: 0-8213-7093-6 (electronic (10))

1. Informal sector (Economics) I. Maloney, William F. (William Francis), 1959- II. Title.

HD2341.P43 2007
330—dc22

2007013736

Cover photo: "Fenómeno" (1962) by Remedios Varo © 2007 Artists Rights Society (ARS), New York / VEGAP, Madrid.

Cover design: Naylor Design.

Contents

Foreword	xi
Acknowledgments	xiii
Abbreviations	xv
Overview: Informality: Exit and Exclusion	1
The <i>razón de ser</i> of the informal sector: Adding exit to exclusion	1
Workers: A mix of opting out and exclusion	4
Firms: Little gain, high costs, or weak enforcement?	9
Need for more effective and legitimate institutions	12
Summing up: Policy implications of the report	13
Informality and the development agenda	19
Note	19
References	19
Chapter 1: The Informal Sector: What Is It, Why Do We Care, and How Do We Measure It?	21
Introduction: What is informality?	21
Informality and the relationship between the individual and the state	22
Three margins of informality	25
Measuring the informal sector	28
Correlations among measures and trends over time	35
Conclusions	37
Annex	39
Notes	39
References	40
Chapter 2: The <i>Razón de Ser</i> of the Informal Worker	43
Informal work: Adding exit to exclusion	43
The sectors of informal labor: Characteristics and dynamics	48
Motivations for participation in informal work	62
Conclusions	74
Notes	74
References	75
Chapter 3: Informality, Earnings, and Welfare	79
Compensating differentials, comparative advantage, and informal work	80
The question of equal pay for equal work in the informal and formal sectors	85
Informality and self-rated welfare	92
Conclusions and policy implications	98
Notes	99
References	99

Chapter 4: The Informal Labor Market in Motion: Dynamics, Cycles, and Trends	101
Informality through the lens of gross labor force dynamics	102
Drivers of the increase in informality	112
Conclusion	125
Notes	127
References	128
Chapter 5: Microfirm Dynamics and Informality	133
Conceptual framework: Firm dynamics and institutional development	134
Microfirm dynamics in Latin America	140
Informality among microfirms	148
Conclusions	153
Notes	154
References	154
Chapter 6: Informality, Productivity, and the Firm	157
Informality among registered firms	158
Firm-level determinants of informality	160
Impact of informality on firm productivity and economic growth	171
Conclusions	175
Notes	177
References	177
Chapter 7: Informality, Social Protection, and Antipoverty Policies	179
Informality and social protection: Why policy makers should care	179
The state of social protection in Latin America and the Caribbean	184
Private risk management and rationale for public social protection	190
Challenges for social protection in the face of informality	193
Reengineering social protection to protect all citizens	199
Potential costs of social protection reform: Financing essential cover	204
Managing the transition from here to there	206
Conclusion	208
Notes	210
References	211
Chapter 8: The Informal Sector and the State: Institutions, Inequality, and Social Norms	215
Social norms, the state, and informality	217
The tax side of the social contract in Latin America	223
Inequality, taxes, and transfers	235
Informality: A reflection of a broken social contract?	239
Conclusions	244
Notes	245
References	246

Boxes

Overview

- | | | |
|---|--|---|
| 1 | Informal workers in Latin America and the Caribbean: Their profile, motivations, earnings, and welfare | 4 |
| 2 | Government failures in the design of social protection systems in Latin America and the Caribbean | 8 |

Chapter 1

- | | | |
|-----|---|----|
| 1.1 | The ILO definition of informality | 27 |
| 1.2 | Indirect methods of estimating informality | 34 |
| 1.3 | Schneider and Enste in the new world: Checking MIMIC estimates against Mexican data | 35 |

Chapter 2

- | | | |
|-----|---|----|
| 2.1 | Home-based work: Exploitation or flexible work arrangement? | 52 |
| 2.2 | Data from rotating panels in Argentina, Brazil, and Mexico | 56 |
| 2.3 | Informal self-employment: Risky and informal, or risky <i>because</i> informal? | 60 |
| 2.4 | Special informal employment surveys: What can we learn from them? | 64 |

Chapter 4

- 4.1 Conceptual issues in gross worker flows 103
- 4.2 Simulated effects of labor market legislation on the size of the informal sector 106

Chapter 5

- 5.1 Patterns of entry and exit in industrialized countries 137

Chapter 7

- 7.1 Social protection—strengthening people’s abilities to manage risk and promoting long-term productivity, growth, and development 180
- 7.2 The comprehensive insurance framework—providing guidance on social protection policy making 182
- 7.3 Extending health insurance coverage by correctly aligning risk-pooling instruments 201
- 7.4 Old-age protection in the new millennium: Chile’s proposed pension reform 203
- 7.5 De-linking health coverage from employment status: Spain’s shift to general-revenue financing of essential social insurance 207

Chapter 8

- 8.1 The extreme of informality and exclusion: Being undocumented 219
- 8.2 Local taxation and social norms 229
- 8.3 Tax compliance, social norms, and trust in the state: The contrasting cases of Chile and Argentina in the late 1990s 231
- 8.4 Earned income tax credits: Transfers that encourage formal employment 234
- 8.5 Expansion of private security services in Managua 242
- 8.6 Negotiating tax reform and the start of the social contract, Chile, 1990 243

Figures**Overview**

- 1 Labor market informality and income per capita 2
- 2 Trends in informality, by various definitions 3
- 3 Probability of transition between formal salaried and self-employment in Mexico 6
- 4 Rate of urban employment across sectors, by age in Brazil, 2002 7
- 5 Informal workers across firm size 9
- 6 Male entry into and exit from self-employment 10
- 7 Advantages of formalization reported by IFC-surveyed firms 11
- 8 Underreporting of tax and social security contributions, by firm size 11
- 9 Economically active population contributing to the pension system 12
- 10 Informality versus inequality 13
- 11 Tax morale and state capture 14
- 12 Self-employment and quality of institutions (governance) 14

Chapter 1

- 1.1 Margins of informality 25
- 1.2 Methods for measuring the informal sector 28
- 1.3 Selected measures of informality 29
- 1.4 Informality and gender: women versus men (ages 25–64) 31
- 1.5 Informal salaried workers (legal definition) across firm size and time 32
- 1.6 Distribution of manufacturing sector firms according to worker registration in Mexico 33
- 1.7 Distribution of manufacturing sector firms according to tax compliance in Mexico 33
- 1.8 Global correlation of measures of informality with GDP 37
- 1.9 Trends in informality by various definitions 38

Chapter 2

- 2.1 Relative sector sizes and wages under a nominal wage rigidity 44
- 2.2 Relative sector sizes and wages in the presence of a labor tax 48
- 2.3 Rate of urban employment across sectors, by age 53
- 2.4 Probabilities of transition among sectors of employment 54
- 2.5 Absolute mean duration of labor force status 57
- 2.6 Propensity to move to self-employment from different sectors 58
- 2.7 Transition between self-employment and out of labor force, by gender 61
- 2.8 Employment sector allocation by gender, marital status, and parental status in Mexico 61

2.9	Propensity to move to informal salaried status from different sectors	62
2.10	International comparison of desired and actual self-employment rates	65
2.11	Earnings gain from voluntary movement to informality in Mexico	68
Chapter 3		
3.1	Propensity to informal employment by firm size and economic sector, 2005–06	82
3.2	Propensity to informal employment by education and tenure, 2005–06	83
3.3	Propensity to informal employment by age, gender roles, and work preferences, 2005–06	84
3.4	Distribution of hourly earnings for workers in urban areas of Argentina and Bolivia, 2005–06	86
3.5	Earnings cost of informality in urban areas of Argentina, Bolivia, and the Dominican Republic	89
3.6	Fraction of workers who are income-poor and self-rated poor, by education and occupational group	93
3.7	Impact of informality on self-rated poverty	94
3.8	Direct impact of having access to social protection through a family member on self-rated poverty	95
3.9	Differences between actual labor incomes and the level of income needed to avoid self-rating poverty in Argentina	97
Chapter 4		
4.1	Formal share of the labor force and unemployment, Brazil and Mexico	102
4.2a	Probability of transition between formal salaried and self-employment, Brazil and Mexico	104
4.2b	Probability of transition between formal salaried and informal salaried, Brazil and Mexico	104
4.3	Involuntary transition to self-employment in Mexico	105
4.4	Searching while employed, Mexico	105
4.5	Decreased availability of formal sector jobs without segmentation	106
4.6	Probability of transition to unemployment (separation rate), Brazil and Mexico	108
4.7	Probability of transition from unemployment (job-finding rate), Brazil and Mexico	109
4.8	Informal employment reaction to the business cycle	111
4.9	Relative sector shares and earnings, real exchange rate	112
4.10a	Variation of informality rate (social protection definition)	114
4.10b	Variation of informality rate (productive definition)	115
4.11	Decomposition of changes in informality (social protection definition) for urban salaried workers	115
4.12a	Labor force participation, rate of formality, and unemployment, Brazil	116
4.12b	Labor force participation, rate of formality, and unemployment, Colombia	116
4.13	Actual and predicted size of the industrial formal sector, Brazil	117
4.14	Informality rate for salaried workers in Greater Buenos Aires	118
4.15	Informal salaried versus formal salaried, Mexico, relative earnings and sector size	118
4.16	Evolution of informality across age groups	120
4.17	Labor legislation and related variables, Peru	123
Chapter 5		
5.1	Entry and exit into self-employment among men	141
5.2	Entry and exit into self-employment among women	142
5.3	Job creation and destruction in Colombia, 1977–82	144
5.4	Job creation and destruction by entering and exiting versus continuing firms in Colombia, 1977–82	145
5.5	Employment growth of microenterprises in Nicaragua	146
5.6	Size distribution of microenterprises in Mexico and Nicaragua	147
5.7	Informality among Mexican microenterprises by number of paid workers and time in business	150
5.8	Effect of doubling labor productivity on probability of being informal	153
Chapter 6		
6.1	Unreported sales and workers, by firm size	159
6.2	Effect of doubling employment size on underreporting rates	159
6.3	Effect of having started informally on underreporting rates	160
6.4	Effect of doubling labor productivity on underreporting rates	160
6.5	Advantages of formalization reported by IFC-surveyed firms	162
6.6	Disadvantages of formalization reported by IFC-surveyed firms	162
6.7	Impact of changes in costs and benefits of informality	163
6.8	Labor regulations described by firms as biggest obstacle to hiring workers	169
6.9	Estimated impact of informality on growth	172
6.10	Estimated impact of informality on labor productivity	173

6.11	Effects of a 10 percent increase in tax and social security evasion at the industry/region level on individual firm productivity (pooled sample)	174
6.12	Productivity difference between tax-paying Mexican microfirms and informal firms with similar characteristics	175
6.13	Productivity and paid employment effects of exogenous increase in formality driven by Brazil's SIMPLES program	175

Chapter 7

7.1	Pension coverage rates in Latin America and the Caribbean	184
7.2	Social security coverage and GDP per capita	185
7.3	Coverage rates in Latin America and the Caribbean, mid-1990s to early/mid-2000s, by quintile of per capita income	186
7.4	Pension coverage rates in Argentina, by quintile of per capita income	187
7.5	Pension coverage among the elderly in Latin America and the Caribbean, contributory and noncontributory programs	188
7.6	Coverage of two poverty-targeted assistance programs in Mexico: <i>Oportunidades</i> and <i>Seguro Popular</i> (by decile, 2004)	188
7.7	Public social protection spending in Latin America and the Caribbean	189
7.8	Public spending on social insurance and social assistance in Latin America and the Caribbean, early 2000s	190
7.9	Nonpoor population who fell below the national poverty line due to out-of-pocket health expenditures	191
7.10	Distribution of Mexican workers' years insured by IMSS, 1997–2005, for low- and high-wage workers in IMSS in 1997	195
7.11	Density function for years of formal sector work among 60-year-olds in Uruguay (accumulated ages 18–60)	196
7.12	Cumulative distribution of years of formal sector service among 60-year-olds in Uruguay (accumulated ages 18–60)	196

Chapter 8

8.1	Informality and government effectiveness	220
8.2	Informality and corruption	220
8.3	Share of firms confident that the judiciary will enforce contractual and property rights	220
8.4	Share of firms that consider the court system fair, impartial, and uncorrupt	221
8.5	Informality and state competence indicators	221
8.6	Central government tax revenue and GDP per capita	224
8.7	Tax revenue of Latin America and Caribbean countries	226
8.8	Average tax rate of Latin America and Caribbean countries	226
8.9	Average VAT and income tax productivity in Latin American countries	227
8.10	VAT productivity and informality indicators	227
8.11	Tax morale, state capture, and perception that the government spends taxpayers' money wisely	228
8.12	Tax morale and informality indicators	232
8.13	Social spending and taxation by income quintiles	238
8.14	Informality and inequality	239

Tables

Chapter 1

1.1	Correspondence of the “productive” and “legalistic” definitions of informality	31
1.2	Correlations across measures of informality	36
1A.1	Comparisons of ILO and Gasparini-Tornarolli measures of self-employment	39

Chapter 2

2.1	Distribution of formal salaried, informal salaried (social protection/legal definition), and self-employed in urban areas of Latin America (percent)	49
2.2	Distribution of informal salaried, self-employed, and unpaid workers, by firm size (percent)	50
2.3	Characteristics of informal salaried employees in Mexican microfirms	51
2.4	Informal employment and work-life occupational history, Dominican Republic and Argentina (percent of workers)	63
2.5	Preference for independent employment (percent of workers)	64
2.6	Distribution of the motivations/reasons for being in the current job as an independent worker (percent)	66
2.7	Reported reasons to be informal self-employed in Mexico (percent of workers)	67
2.8	Reported reasons to be informal self-employed in Brazil (percent of workers)	67
2.9	Distribution of the motivations/reasons for being in the current salaried job (percent)	69

2.10	Reported reasons to be informal salaried in Brazil (percent of workers)	69
2.11	Principal reason for working without monetary compensation in the Dominican Republic (percent of workers)	70
2.12	Main reasons why the informal do not contribute to social security (percent of workers)	71
2.13	Main reasons why workers do not contribute to health insurance (percent of workers)	73
Chapter 3		
3.1	Unconditional hourly earnings gaps (percent difference in earnings) for formal employees, informal employees, and independent workers in urban areas in Argentina, Bolivia, and the Dominican Republic, 2005	87
3.2	Impact of informality on earnings in Argentina, estimated log-earnings differences and treatment parameters	91
3.3	Correlation between income and self-rated poverty	92
3.4	Impact on the propensity to self-rate poor and be income poor (normalized logit regression coefficients)	95
3.5	Job satisfaction and informal employment in the Dominican Republic	96
3.6	Job satisfaction and informal employment in Colombia	97
Chapter 4		
4.1	Evolution of informality across firm size and time: Distribution of informal and self-employed workforce	119
4.2	Subcontracting in Mexico, 1992–2002 (percent)	119
Chapter 5		
5.1	Entry probabilities into Mexico's self-employment sector (percent)	140
5.2	Exit probabilities from Mexico's self-employment sector (percent)	143
5.3	Growth and contraction probabilities of Mexican microfirms (percent)	146
5.4	Employment size distribution of firms, selected countries (percent)	147
5.5	Employment size distribution of firms with employees, selected countries (percent)	148
5.6	Reasons for starting up, firm prospects, and firm licensing in Brazil	149
5.7	Informality among Mexican and Brazilian microenterprises by age, education, and previous activity of the owner	149
5.8	Informality among Mexican, Brazilian, and Colombian microenterprises by paid employment and time in business	151
5.9	Informality among Mexican, Brazilian, and Colombian microenterprises by age, education, and previous activity of the owner	152
Chapter 6		
6.1	Firm-level correlates of sales and employment underreporting	167
Chapter 8		
8.1	Long-run informality relationships	222
8.2	Indicators of informality and institutional indicators	223
8.3	Comparative perspective of tax burdens and structures (percent of GDP)	224
8.4	Tax structure by region, selected years, 1975–2002 (percentage of total tax revenue)	225
8.5	Tax exemptions in Latin America (percent of GDP)	227
8.6	Comparative efficiency and corruption in tax administration: Survey evidence for 2005–06	233

Foreword

INFORMALITY IS A WAY OF LIFE IN LATIN America and the Caribbean, no matter how we measure it. In some countries, the majority of the workforce is not covered by labor protections; the ubiquitous microenterprises found on every street corner are often not registered with authorities and only partially comply with other regulations; and tax evasion by rich and poor alike is the norm. Although these features of the Latin American and Caribbean landscape are not new, their striking increase in some countries has given new life to the debate concerning what informality tells us about how our economies are functioning.

Informality: Exit and Exclusion, the ninth in our series of regional flagship reports, explores this very heterogeneous sector from a variety of perspectives ranging from a concern with the protection of workers, to the productivity of firms, to the determinants of tax evasion. The report accepts and sheds more light on the exclusionary character of much informality, which leaves citizens outside formal institutions. However, it especially highlights that there is an important “exit” dimension that has been understressed in the literature: workers, firms, and families, dissatisfied with the performance of the state or simply not finding any benefit to interacting with it, opt into informality. Hence, the discussion ranges from the purely economic analysis of microeconomic incentives to reflection on the soundness of the “social contract”—shorthand for how citizens of the region relate to the state and to each other.

It is not surprising, then, that the analysis gives rise to recommendations that span many dimensions of the policy agenda. Because the high level of informality in developing

countries more generally is partly due to the low opportunity cost of opening a microbusiness, the gamut of measures to increase formal sector productivity (improving the business climate, fomenting innovation, and so forth) and to increase workers’ skills is important. Labor market reform is also essential because part of the rise in the share of workers uncovered by labor protections in the 1990s appears to be due to the increased burden of labor costs and restrictions in several countries. But reform also needs to extend to remedying the poor design of many social protections that workers find of poor quality, inefficient, and of little value. Hence, shifting workers’ cost-benefit analysis toward engagement with the institutions of the state—both through improving the benefits of being formal and through better monitoring—is critical to reducing their often voluntary entry into informality. Relatedly, although experiments reducing the costs of registering businesses have led to relatively little formalization of existing businesses, the impact on those few businesses appears substantial and progress in this area appears important. All of these reforms need to take place in a context of improving the perception of fairness and efficiency of the state, which is often seen as serving the needs of particular elites rather than the needs of the population. In the long run, this is the only way to change social norms of compliance and reduce the “culture of informality.”

This is a topic that the Latin America and the Caribbean Region of the World Bank has been working on in a sustained way through various regional studies and reports for more than 10 years. The present volume builds on this work, introduces original data sources collected for the

project and new statistical techniques to analyze them, and brings together research done in our partner countries. Both its findings and its policy recommendations are thus grounded in detailed analysis of the reality of the region.

As a final note, this Flagship is the last undertaken under the direction of Guillermo E. Perry, and I would like to thank him for his leadership, insights, and enthusiasm as

Chief Economist for the Latin America and the Caribbean Region of the World Bank.

A handwritten signature in black ink, reading "Pamela Cox". The signature is fluid and cursive, with the first letters of the first and last names being capitalized and prominent.

Pamela Cox

Vice President for Latin America and the Caribbean
The World Bank

Acknowledgments

INFORMALITY: EXIT AND EXCLUSION IS THE PRODUCT OF A COLLABORATIVE EFFORT BY A NUMBER of professionals from within and outside the Bank. The report was prepared under the direction of Guillermo E. Perry by a core team comprising William F. Maloney, Omar S. Arias, Pablo Fajnzylber, Andrew D. Mason, and Jaime Saavedra-Chanduvi.

Background papers were prepared by Rita Almeida (World Bank), James Alm (Georgia State University), Guillermo Beylis (World Bank), Mariano Bosch (London School of Economics and Political Science [LSE] and University of Alicante), Karla Breceda (World Bank), Monserrat Bustelo (World Bank), Dante Contreras (Universidad de Chile), Barbara Cunha (University of Chicago), Wendy Cunningham (World Bank), Norbert Fiess (University of Edinburgh), Marco Fugazza (United Nations Conference on Trade and Development [UNCTAD]), Sebastian Galiani (Washington University in St. Louis), Rodrigo García-Verdú (World Bank), Leonardo Gasparini (Centro de Estudios Distributivos Laborales y Sociales [CEDLAS] and Universidad Nacional de La Plata), Edwin Goñi (World Bank), Melanie Khamis (LSE), Fernando Landa (Unidad de Análisis de Políticas Sociales y Económicas [UDAPE], Bolivia), Santiago Levy (Brookings Institution), Norman Loayza (World Bank), Leonardo Lucchetti (University of Illinois at Urbana-Champaign), Jorge Martinez (Georgia State University), David McKenzie (World Bank), Truman Packard (World Bank), Esteban Puentes (University of Chicago), Jamele Rigolini (World Bank), Gabriel Montes Rojas (University of Illinois at Urbana-Champaign), Lucas Ronconi (University of California at Berkeley), Claudia Sanhueza (Universidad de Chile), Mariano Tommasi (Universidad de San Andrés), Leopoldo Tornarolli (CEDLAS and Universidad Nacional de La Plata), Federico Weinschelbaum (Universidad de San Andrés), Chris Woodruff (University of California, San Diego), and Patricia Yañez (UDAPE, Bolivia).

The report has also benefited from specific contributions or comments by Juliano Assunção (Pontifícia Universidade Católica [PUC], Brazil), Tito Cordella (World Bank), Jean Jacques Dethier (World Bank), Maria Victoria Fazio (World Bank), Gary Fields (Cornell University), Alvaro Forteza (World Bank), Vicente Fretes Cibils (World Bank), Marcelo Giugale (World Bank), Gustavo Gonzaga (PUC, Brazil), Ricardo Furman (International Finance Corporation [IFC]), Luke Haggarty (IFC), Mary Hallward-Driemeier (World Bank), Marco Hernandez (Oxford University), Tom Kenyon (IFC), Peter Lindert (University of California at Davis), David McKenzie (World Bank), Carolina Mejia (Fedesarrollo), Ezequiel Molina (CEDLAS and Universidad Nacional de La Plata), Carmen Pages (Inter-American Development Bank), Georgina Pizzolito (World Bank), Kristtian Rada (IFC), Lauro Ramos (Instituto de Pesquisa Econômica Aplicada [IPEA], Brazil), Jose Guilherme Reis (World Bank), Helena Ribe (World Bank), Rafael Rofman (World Bank), Jose Scheinkman (Princeton University), Pablo Suarez (Brown University), Judith Tendler (Massachusetts Institute of Technology), Gabriel Ulyssea (IPEA, Brazil), and Evelyn Vezza (World Bank); and by our peer reviewers Jesko Henschel

(World Bank), Pierella Paci (World Bank), Stefano Scarpetta (Organisation for Economic Co-operation and Development [OECD]), and Simeon Djankov (World Bank).

The project benefited from a data collection grant from the World Bank Development Economics Group Research Support Board to undertake surveys in Bolivia, Colombia, and the Dominican Republic, building on the 2005 survey conducted in Greater Buenos Aires, Argentina, by the Ministry of Labor, Employment, and Social Security and the National Institute of Statistics and Census of Argentina, with support from the World Bank. We very warmly thank Clara Else for her guidance and help in shepherding through the proposal and François Bourguignon for his support. Invaluable collaboration to develop these surveys is acknowledged to Carlos Ignacio Becerra (Departamento Administrativo Nacional de Estadística [DANE], Colombia), Dilia Cruz (Central Bank of the Dominican Republic), Olga Diaz (Central Bank of the Dominican Republic), Pedro José Fernández (DANE, Colombia), Eduardo Freire (DANE, Colombia), Rolando Guzman (PARETO Consulting Group), Oscar Lora (Instituto Nacional de Estadística [INE], Bolivia), Gabriel Loza (UDAPE, Bolivia), Juliana Martinez Cuellar (Departamento Nacional de Planeación [DNP], Colombia), José Ramirez (UDAPE, Bolivia), Carmen Julia Reyes (Central Bank of the Dominican Republic), Guillermo Rivas (DNP, Colombia), Norberto Rojas (DNP, Colombia), Luz Magdalena Salas (DNP, Colombia), Mauricio Santamaria (DNP, Colombia), Romilda Santana (Central Bank of the Dominican Republic), Johnny Suño (INE, Bolivia), María Fernanda Tellez (DNP, Colombia), and María Eugenia Villamizar (DANE, Colombia).

We are also grateful to Mauricio Cardenas (Fedesarrollo) and Sebastian Edwards (University of California, Los Angeles); to participants at the National Bureau of Economic Research workshop on Informality in Bogota; to Hartmut Lehmann (University of Bologna and Institute for the Study of Labor [IZA]) and Stefano Scarpetta (OECD); to participants at the World Bank/IZA Conference on Informality in Bertinoro; and to participants at the annual conference of the Latin American and Caribbean Economics Association for helpful feedback on background work for this report.

Finally, Patricia Da Camara, Catherine Russell, and Elena Serrano coordinated the report's publication and dissemination activities, working closely with Dina Towbin, Andrés Meneses, and Shana Waggoner in the World Bank's Office of the Publisher.

Abbreviations

ATE	average treatment effect	LDC	least developed country
AUGE	<i>Acceso Universal para prestaciones integrales y Garantías Explícitas</i> (Plan of Universal Access with Explicit Guarantees)	LSMS	Living Standards Measurement Survey
BONOSOL	<i>Bono solidario</i>	MAROP	<i>Mecanismo de Ahorro para el Retiro Oportunidades</i>
BPS	<i>Banco de Previsión Social</i>	MIMIC	Multiple Indicator–Multiple Cause Model
DANE	<i>Departamento Administrativo Nacional de Estadística</i>	MPG	Minimum Pension Guarantee
ECINF	<i>Pesquisa Economica Informal Urbana</i>	OECD	Organisation for Economic Co-operation and Development
EITC	earned income tax credit	OLF	out of the labor force
ENAH0	<i>Encuesta Nacional de Hogares</i>	PASIS	<i>Pensión Asistencial</i>
ENAMIN	<i>Encuesta Nacional de Micronegocios</i>	PNAD	<i>Pesquisa Nacional por Amostra Domicilios</i>
ENEU	<i>Encuesta Nacional de Empleo Urbano</i>	PPP	purchasing power parity
F-I	formal to informal	RN	<i>Renovación Nacional</i>
FONAHPO	<i>Fondo Nacional de Habitaciones Populares</i>	RTS	simplified tax regime
FONASA	<i>Fondo Nacional de Salud</i>	SARE	<i>Sistema de Apertura Rápida de Empresas</i>
F-SE	formal to self-employed	SE-F	self-employed to formal
GDP	gross domestic product	SENA	Colombian Training Institute
ICBF	Colombian Institute for Family Welfare	SIMPLES	<i>Sistema Integrado de Pagamento de Impostos e Contribucoes as Microempresas e Empresas de Pequeno Porte</i>
ID	identity	SME	small and medium enterprise
I-F	informal to formal	TT	treatment on the treated
ILO	International Labour Organization	TU	treatment on the untreated
IMSS	Mexican Social Security Institute	VAT	value-added tax
INEGI	<i>Instituto Nacional de Estadística, Geografía e Informática</i> (National Statistical Institute [Mexico])	WDI	<i>World Development Indicators</i>
ISSSTE	Social Security Institute for Mexican civil servants at the federal level	WFTC	working families tax credit
		WGI	Worldwide Governance Indicators

OVERVIEW

Informality: Exit and Exclusion

There are two main types of activist reactions to discontent with organizations to which one belongs or with which one does business: either to voice one's complaints, while continuing as a member or customer, in the hope of improving matters; or to exit from the organization, to take one's business elsewhere. —Albert O. Hirschman (1981, p. 246)

INFORMALITY HAS GAINED INCREASING attention as a possible drag on growth and rising social well being, and as a force corrosive to the integrity of our societies. In fact, informality is not especially higher in Latin America and the Caribbean than in other developing countries of similar per capita income, according to two popular measures (figure 1). However, given the long-standing negative connotations of informality—inferior working conditions, low-productivity firms, disrespect for the rule of law, to name a few—it is not surprising that the *rise* in informality over the 1990s across several measures (figure 2) is viewed with concern and as meriting closer investigation.

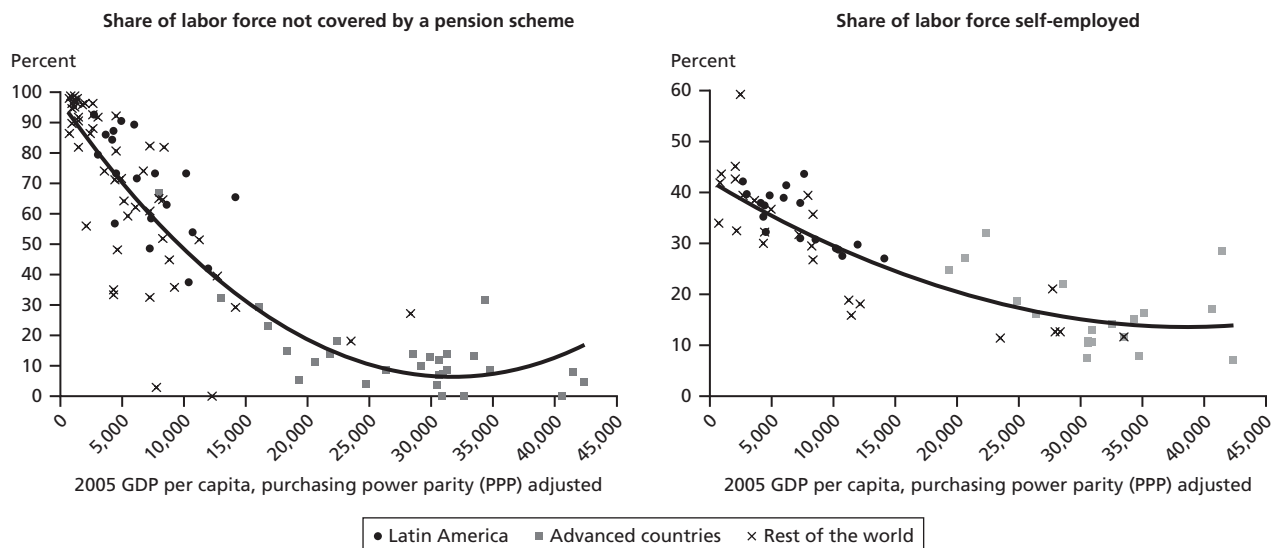
This said, the mere fact that we need to employ multiple measures of informality capturing distinct approaches to the sector suggests that we are not clear on exactly what it is and what we should be studying. In all likelihood, we are dealing with several distinct phenomena under this convenient, but arguably unhelpful, umbrella term. For a report on informality in Latin America and the Caribbean this clearly poses special challenges. Fortunately, recently available data sets and the development of new techniques to analyze them have made possible more solidly grounded analysis of the underlying heterogeneity and reasons for being of the sector, the factors driving its evolution across different countries, and its social and economic ramifications. Such an

analysis, we hope, will serve as a foundation for more effective public policies.

The *razón de ser* of the informal sector: Adding exit to exclusion

The report sees informality as a manifestation of the relationships between economic agents and a state that the economic literature should play an important role in remedying market failures, coordinating the provision of public goods, and maintaining a level and equitable playing field. Of the many lenses through which informal workers have been viewed, the most influential lens has focused on their *exclusion* from critical state benefits or the circuits of the modern economy. This exclusion can be seen as occurring along three margins, or borders, between formality and informality. First, a long tradition in the labor literature sees segmentation in the labor market as preventing workers from leaving their holding pattern in informality and taking jobs in the formal sector that offer state-mandated benefits. Second, de Soto's (1989) seminal work argues that burdensome entry regulations prohibit small firms from crossing the line into formality and thriving. And, third, some larger firms faced with excessive tax and regulatory burdens may remain partially informal as a defensive measure and, as a result, they forgo potential growth and efficiency gains. This report finds that these exclusion factors

FIGURE 1

Labor market informality and income per capita

Sources: Loayza and Rigolini 2006; World Development Indicators 2006.

are indeed important, documents their negative impact on productivity and welfare, and discusses some of the reforms necessary to mitigate them.

However, the report also highlights a second lens through which to view informality. This lens is more akin to Hirschman's (1970) *exit*: many workers, firms, and families choose their optimal level of engagement with the mandates and institutions of the state, depending on their valuation of the net benefits associated with formality and the state's enforcement effort and capability. That is, they make implicit cost-benefit analyses about whether to cross the relevant margin into formality, and frequently decide against it. This view suggests that high informality results from a massive opting out of formal institutions by firms and individuals, and implies a blunt societal indictment of the quality of the state's service provision and its enforcement capability.

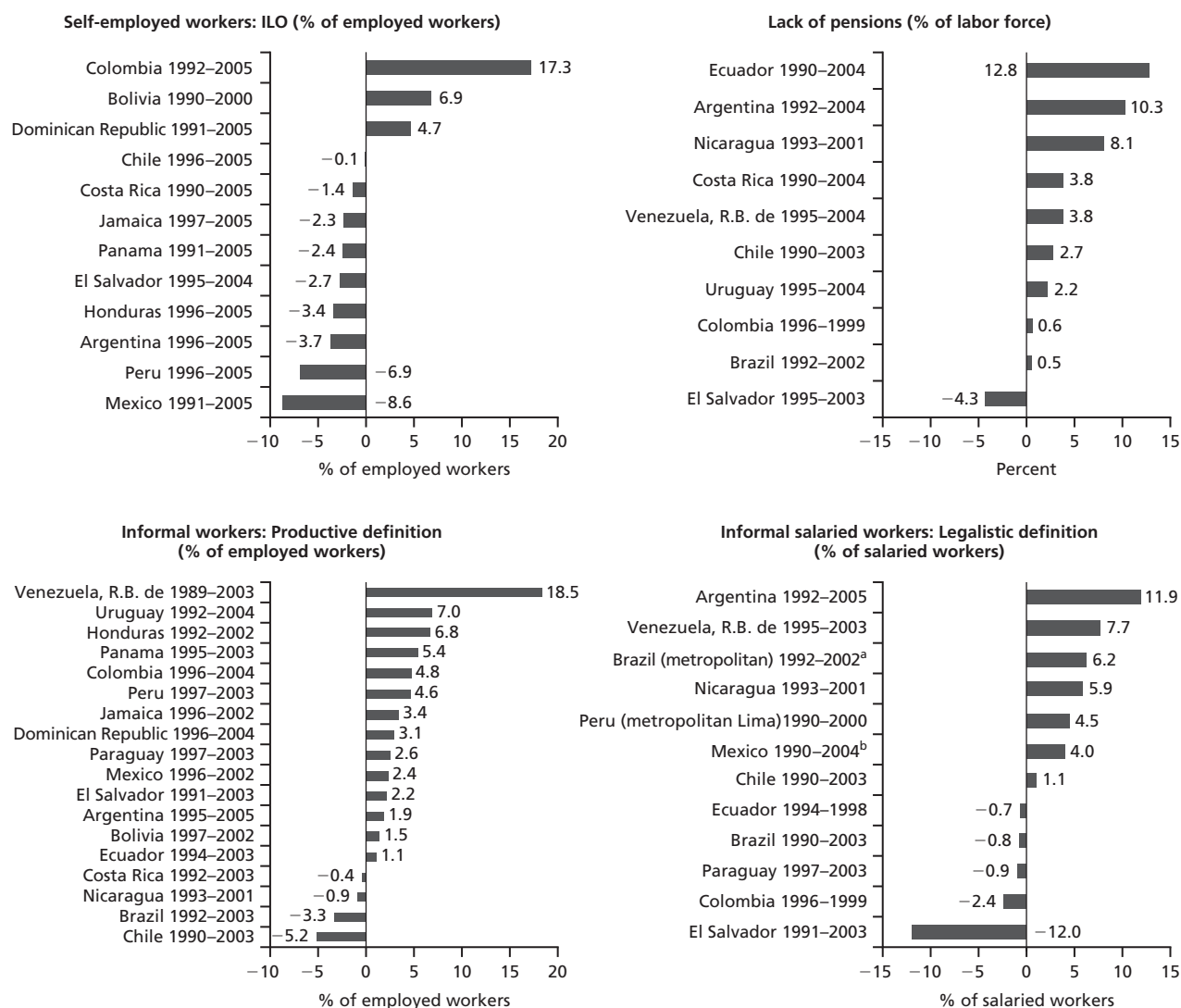
This view leads to important divergences with many popular stylized facts about the informal sector. Microfirm owners with little intention or potential to grow may see insignificant benefits from engaging with unhelpful regulatory and tax institutions. Interpreting Hirschman literally, they take their business elsewhere, knowing that they will fall below the radar screen of enforcing authorities. Unskilled workers, who partially pay for social protection benefits directly and implicitly through lower formal

wages, may find that what they expect to get in the formal jobs for which they are qualified does not outweigh the foregone current consumption or the greater flexibility and earnings they may get as informal workers. This is especially true when they have social protection alternatives through access to universal or noncontributory programs, or through private means. Larger firms or skilled professionals may decide to underreport their operations and incomes, balancing private gains from tax evasion with low detection risks resulting from poor enforcement.

These two lenses, focusing, respectively, on informality driven by exclusion from state benefits and on voluntary exit decisions resulting from private cost-benefit calculations, are complementary rather than competing analytical frameworks. First, individual countries differ greatly in history, institutions, and legal frameworks, so exclusionary mechanisms may be more important in some and exit may be more important in others. Second, the informal sector is tremendously heterogeneous, and arguably, then, there is a continuum in the relative importance of exclusion and exit among individual workers and firms within countries. Third, in some cases the two can be virtually indistinguishable. A microentrepreneur concluding, through a cost-benefit analysis, that formality is not worth the high registration costs may be explicitly excluded or self-excluded—either way, the effect is much the same. A poor worker, excluded

FIGURE 2

Trends in informality, by various definitions



Sources: Gasparini and Tornarolli 2006; International Labour Organization (ILO) Labor Statistics Database 2006; World Development Indicators 2006.

Note: Although global data is available only for the definitions of informality as self-employed workers and workers not covered by a pension scheme, two other measures are calculated here based on regional data sources. In the "productive" definition, a worker is considered informal if he or she is unskilled self-employed, a salaried worker in a small firm, or a zero-income worker. In the "legalistic" definition, a salaried worker is informal if he or she does not have the right to a pension linked to employment upon retirement.

a. Percent of workers without *carteira* (work card).

b. Based on the balanced panel sample (common municipalities) for the period 1990–2004.

from health care services because he or she lives in a remote rural area or a poor urban neighborhood, may see little point in being formal and paying labor taxes for services to which he or she has no access.

Finally, informality is a multidimensional phenomenon: agents interact with the state along some dimensions and not others, creating a large gray area between the extremes of full compliance and noncompliance. Exit and exclusion

can play different roles across different dimensions: a microfirm owner unlicensed because of high registration costs may be de facto excluded from desired formal credit circuits while opting out of contributing to poorly designed state pension funds on behalf of his or her workers. The findings of the report indeed support the use of both lenses to fully understand and address the causes and consequences of informality in the Latin American and Caribbean region.

Workers: A mix of opting out and exclusion

This report examines informal labor using the vast set of regular household surveys conducted in most countries of the region, together with recent special modules on informality collected by statistical agencies in Argentina, Bolivia, Colombia, and the Dominican Republic, in collaboration with the World Bank and line ministries. The analyses shed light on the characteristics of informal workers, their motivations, and their preference for the benefits and nonpecuniary characteristics of jobs (for example, flexibility, autonomy, stability, mobility) by examining their remuneration, self-rated welfare, and job satisfaction. The findings (summarized in box 1) show great heterogeneity within and across countries. They also indicate, however, that the informal sector can generally be thought of as comprising two large groups who differ

significantly in both their motivations and their relative levels of job satisfaction.

First, on average, independent workers—the self-employed or those owning microfirms—in the countries surveyed (with the exception of Colombia) report being as well-off as they would be in jobs they are qualified for in the formal sector. As a consequence, the majority are not looking for formal jobs. Most of these informal workers appear to choose their occupations according to their individual needs (especially their desire for flexibility and autonomy) and abilities (their comparative advantage in terms of entrepreneurship). In fact, when asked, a majority of workers state that they prefer the independence of being self-employed. In Mexico, they were overachievers in salaried work, earning relatively high wages, given their human capital. Many women, in particular, report forgoing

BOX 1

Informal workers in Latin America and the Caribbean: Their profile, motivations, earnings, and welfare

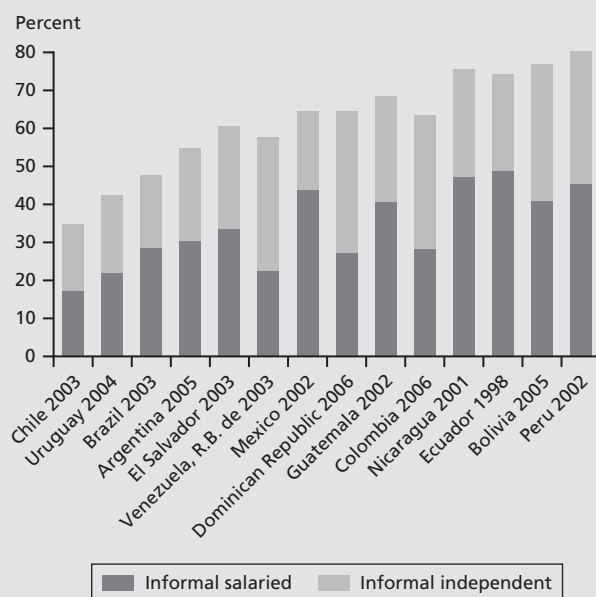
Who are the informal?

Informal employment encompasses a diverse range of people. While the report considers numerous common criteria to define informal employment, it focuses on the social protection definition. Based on whether the job is a salaried relationship and unregistered with social security, informal employment accounts for 54 percent of total regional urban employment and it comprises two groups: (1) informal independent workers, making up 24 percent of total urban employment (ranging from 20 percent or less in Argentina, Brazil, Chile, and Uruguay to more than 35 percent in Bolivia, Colombia, the Dominican Republic, Peru, and the República Bolivariana de Venezuela), and (2) informal salaried workers who account for roughly 30 percent of total regional urban employment and more than half of all informal work (ranging again from 17 percent in Chile to more than 40 percent in Bolivia, Ecuador, Guatemala, Mexico, Nicaragua, Paraguay, and Peru) (figure 1.B.1)

The informal independent work sector includes microfirm owners and self-employed professionals, as well as artisans, handymen, construction laborers, taxi drivers, and street vendors. The informal salaried work sector largely comprises domestic employees, unpaid family

FIGURE 1. B.1

Distribution of informal workers in urban areas contributing to a social security system in Latin America



Source: Author's estimates, based on household surveys.

Note: In Ecuador, Mexico, and República Bolivariana de Venezuela, coverage of independent workers is proxied by having a tertiary education. Argentina denotes Greater Buenos Aires.

workers, microfirm workers, and those who work in larger firms under informal labor arrangements. The characteristics more strongly correlated with informal employment are firm size (10 employees or fewer), education (completion of schooling below the secondary level), industry sector (construction, agriculture, retail, and transport), tenure (less than one year), age (youth predominantly informal salaried, self-employed mostly older workers), and women's household status (married women with children).

What are their motivations, earnings, and welfare?

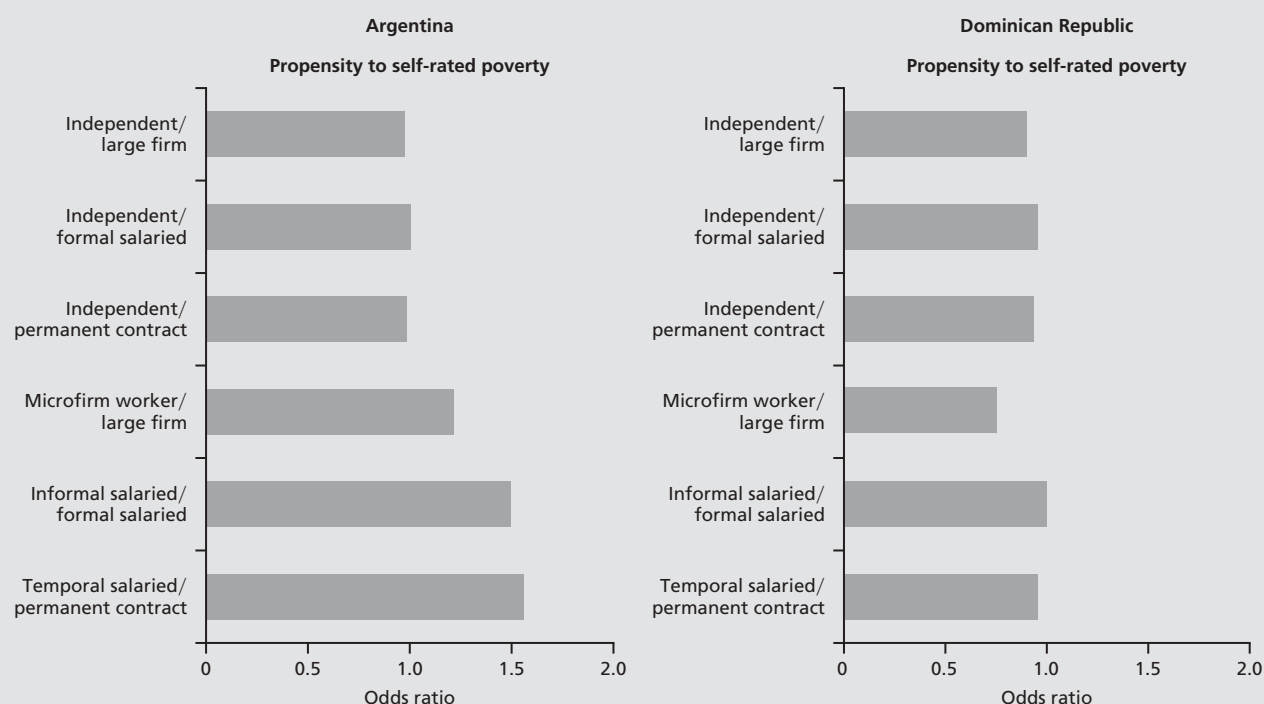
Evidence from workers' patterns of mobility, reported motivations, and self-rated welfare and job satisfaction indicate that (1) the majority of independent workers (approximately two-thirds) entered their jobs voluntarily, attach significant value to the nonpecuniary benefits of autonomous work, and choose to "exit" formal social protection systems; (2) on the contrary, the majority of informal salaried workers appear to be excluded from more desirable jobs, both as formal salaried and as independent workers, although voluntary motives are

still significant for many of the people in this subsector (for example, youth and unpaid family workers). Informal workers state they do not contribute to social security and health insurance plans mainly because of low incomes and also because of their employer's decision not to offer benefits (in the case of most informal salaried workers); because they lack information on the benefits and functioning of social protection programs; and, in the case of health care, because they resort to other means of coverage, including coverage through other family members and universal services.

For both groups of informal workers, there is substantial heterogeneity of motives and demographic characteristics. For instance, in Argentina, the informal self-employed report being as well-off as formal salaried workers, but the informal salaried appear worse-off (figure 1.B.2); in the Dominican Republic, both informal self-employed and informal salaried workers are as well-off as formal employees (figure 1.B.2); and, in Colombia, both informal groups report much lower levels of satisfaction with current jobs.

FIGURE 1. B.2

Impact of informality on self-rated poverty



Source: Arias and Luchetti (2007), based on household surveys, 1997 and 2004.

Note: Self-rated poverty is based on household survey responses.

formal salaried work to better balance home and work responsibilities.

Moreover, even though the desired flexibility associated with self-employment and microfirm ownership is often associated with low income and little security, those jobs are still considered better options than the corresponding jobs that the workers could get in the formal economy, given their minimal human capital, access to other assets, and the low aggregate productivity in the economy. In other words, many self-employed workers opt for informality because their options in the formal sector are at least equally poor. Furthermore, they have the possibility of using informal mechanisms or freely available (noncontributory) social protection programs to substitute, in part, for formal social security benefits that they would otherwise have to pay for, explicitly or implicitly, via (sometimes steep) payroll taxes. In summary, most of the self-employed do not appear to be “excluded” from the formal sector; rather, after implicitly making a cost–benefit analysis, they opt out of formality.

A different picture emerges, however, for the majority of informal salaried workers in the countries studied. Indeed, most of the informal salaried appear to be queuing for more desirable jobs in either the formal salaried sector or as self-employed workers (with Mexico and the Dominican Republic being notable exceptions). For many of these workers, informality largely reflects the decision of the firms for which they work, especially microenterprises, to operate outside the realm of government regulations. On average, these workers are not choosing to opt out of formal contracts and social security institutions, and they would much prefer an equivalent job in the formal sector. In these cases, exclusion rather than exit from formality appears to be the driving force behind their present informal status.

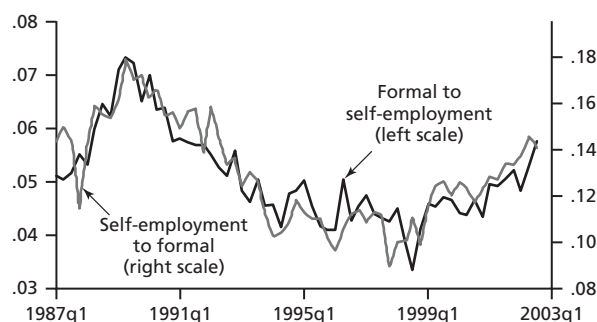
However, considerable heterogeneity exists even within these two subsectors of the informal sector, with both voluntary and involuntary entrants and a great variety of workers’ life trajectories found in each. For example, while, on average, salaried workers in microenterprises appear to want a formal sector job or self-employment, approximately half of the employees of Mexican microenterprises turn out to be unpaid family members whose presence in the sector probably reflects a mix of professional and personal considerations. Similarly, while, on average, independent workers are as well-off as their formal

counterparts, about one-third of the self-employed (over half in Colombia) appear to be so largely involuntarily; they would prefer formal jobs.

There is also considerable variation across countries in the plausible causes of job segmentation for those groups of informal workers who report being involuntary. For example, in Brazil, Colombia, and Peru, the evidence points to an impact of increasing labor costs and rigidities on the expansion of informality in the 1990s. In other countries, such as Argentina, the introduction of temporary contracts and subsequent weakened enforcement of tax and labor regulations seem to have played a role in the observed increases in the share of the labor force without social security contributions. In most countries, trade liberalization appears to have had modest or no effects on informality trends.

The presence of both exit and exclusion factors in informal labor markets—apparent in informal workers’ subjective motivational responses and self-rated welfare measures reported above—is corroborated by the observed patterns of labor market dynamics, an important analytical tool introduced in this report. Indeed, evidence from Mexico (figure 3), suggests that, during good times, the number of workers who leave the formal sector to become self-employed or take an informal salaried job is nearly equal to or sometimes even greater than the number of those who transit from the informal to the formal sector. This is a major piece of evidence supporting the view of integrated labor markets in which workers are freely choosing between formal and informal jobs.

FIGURE 3
Probability of transition between formal salaried and self-employment in Mexico



Source: Bosch and Maloney 2006.

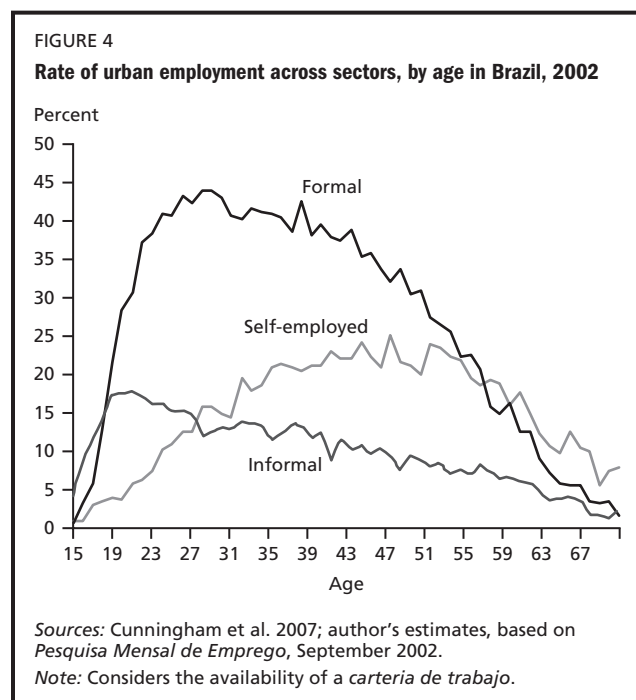
However, although initially following a pattern similar to that of Mexico, after the constitutional reform in Brazil the flows from informal to formal salaried employment became significantly lower than the flows in the reverse direction even in boom times—suggesting a significant degree of market segmentation. Consistent with the literature focused on the U.S. economy, it is also true that formal sector hiring comes almost to a halt during bad times, perhaps reflecting more binding downward wage rigidities. Hence, many entrants have no choice but to join less desirable occupations in the informal sector. As a consequence, the performance of labor markets is asymmetric throughout the business cycles in most countries: labor markets tend to behave more as segmented markets during downturns and recessions, and more as integrated markets during booms.

Finally, both the more exclusionary and exit dimensions of informal salaried work are suggested by workers' life-cycle employment trajectories. In most countries, young workers, especially the less educated, are more likely to be informal employees than formal salaried workers, and very few are informal self-employed. Meanwhile, middle-aged and older workers are more likely to be found in the formal sector or to be self-employed, although many still end up as informal employees (see figure 4). This suggests that

informal salaried work is a point of entry to the labor market for many of the young, and, as they accumulate experience or simply queue, they are more able to find a job in the formal sector or fulfill a surprisingly common desire to be self-employed. Informal salaried employment is also an option for many older workers who lack the skills or capital to become self-employed or get a formal salaried job, or who opt out of informality because, for example, they will never accumulate enough years to secure a meaningful pension. The fact that participation in self-employment rises with age is, again, similar to that of the United States, and similarly may be partly explained by voluntary entry delayed by credit or human capital constraints.

The findings discussed above have critical implications for policy design. Whether because of outright exclusion and market segmentation or massive voluntary opting out of formality, informality may lead to a suboptimal social equilibrium in which many workers go unprotected from health and employment shocks and from poverty in old age. In either case, the imperative for reform is strong. However, our findings caution against one-size-fits-all solutions. Again, whereas labor markets in Mexico and the Dominican Republic show a high degree of integration—for example, most informal workers choose to be informal—except, perhaps during crises—labor markets in Argentina, Colombia, and, to a lesser extent, Brazil show many signs of segmentation. Thus, it is very useful to clarify the relative importance of alternative drivers of informality so that strategists may devise the most suitable policy changes.

For example, if labor informality were essentially driven by inadequate labor legislation leading to labor market segmentation, then the key to reducing informality would be to engage in (politically demanding) labor market reform. If, however, expected low benefits and high costs of social security contributions relative to other forms of protection against shocks are prominent in driving many workers to opt for informality, reform of the design and provision of social protection benefits must be an additional component of the policy package to reduce informality. Indeed, this report finds that there is significant room for improvements in the area of social protection program design (box 2). Further, when considering the case of informal salaried workers, it is also key to understand and address the factors that lead firms to be partially or entirely informal, as we discuss in the next section.



BOX 2

Government failures in the design of social protection systems in Latin America and the Caribbean

There are high human costs associated with the lack of access to appropriate risk management instruments, such as health insurance and old-age security. In addition, missing insurance markets and other market failures—such as negative externalities associated with the absence of health treatment or unsupported poor elders—create a clear rationale for public intervention to ensure basic access to social protection. Therefore, a key challenge facing policy makers in Latin American and Caribbean countries is to ensure that their citizens have suitable access to social risk management instruments, even in the face of significant informality in the region. However, ill-designed interventions actually may serve to make things worse. They not only hinder people's ability to access basic services or to manage risk, but also distort economic incentives that may have adverse effects on productivity and long-term economic growth, for example, by creating strong incentives for many workers to remain informal.

Indeed, in Latin America and the Caribbean there is ample evidence of “government failure” in the design and implementation of social protection, and that failure needs to be addressed as part of any actions to strengthen risk management among the region's citizenry. This report documents these issues in detail. At the level of *specific programs*, design problems raise the costs of participating in social security, relative to the benefits of participation, causing some workers to opt out of the system. Design deficiencies also impede some workers' access to benefits. At the level of social protection *systems*—the constellation of contributory social security and noncontributory social assistance schemes—programs often compete, creating adverse labor market incentives and outcomes.

Among the key weaknesses in the design of the region's social security programs we find the following:

- *High payroll contributions relative to the expected benefits and quality of service:* Evidence shows that high contributions can act as a disincentive to formal employment.
- *Excessive “bundling” of multiple benefit packages:* For example, packages in some countries include not only health care, worker risk, life insurance, and pensions, but also housing, child care, and sports and recreation. Some of these components represent “pure taxes” rather than contributions that give

rise to benefits for many formal sector (or potential formal sector) workers.

- *Rigid, one-size-fits-all approaches to mandated social security programs:* For example, spouses and secondary household workers who opt to enter formal employment generally would have to pay for health coverage to which they already have a right through the household head's affiliation. This double payment for a single benefit creates a powerful incentive to choose informal jobs. Furthermore, social security systems rarely account for different needs and preferences among workers at different stages of the life cycle, and, as a consequence, high initial payroll contributions create a significant disincentive for young workers who have other investment priorities (such as education and housing) to join the formal sector
- *Weak accounting for workers' mobility into and out of the formal sector during their careers and across their life cycles:* In particular, lengthy pension vesting periods make many workers ineligible because of the high mobility between formal and informal sectors, and thus creates, ex ante, a major disincentive to choosing formal jobs. Furthermore, lack of portability often leads to intermittent health insurance coverage gaps for workers who shift from formality to informality and vice versa.

Moreover, recent efforts of a number of governments in the region to use noncontributory or subsidized assistance programs to compensate for low social security coverage inadvertently may have complicated the quest to provide effective social protection to a broader swath of the region's population. Indeed, the evidence suggests that programs offering informal workers services on a noncontributory or highly subsidized basis actually may “compete” with formal social security programs that workers have to pay for through payroll tax contributions, thereby creating additional incentives for informality.

Addressing these design weaknesses will be important not only to ensure broader access to effective risk management instruments by the citizens of Latin America and the Caribbean, but also to make social protection programs consistent with increased productivity and sustained economic growth in the future.

Firms: Little gain, high costs, or weak enforcement?

Labor informality is primarily a small firm phenomenon with the vast majority of workers who are unregistered with social security administrations found in firms of fewer than five workers (figure 5). Hence, understanding the rationale behind small firms' decisions to register their firms and their workers, to pay taxes, and so forth is fundamental to understanding the phenomenon of informality.

The report shows that, in nature and dynamics, these microfirms are closely related to their advanced-country counterparts and should not be treated as a pathology peculiar to developing countries. As figure 6 shows, the patterns of entry, exit, and participation in self-employment are very similar, suggesting that the high levels of self-employment seen in figure 1 may not reflect a fundamentally different phenomenon but simply much more of it. This may arise precisely because the lower labor productivity of the formal sector in developing countries implies a lower opportunity cost of participating in independent work, which worker surveys suggest is as highly regarded in Latin America and the Caribbean as it is in countries of the Organisation for Economic Co-operation and Development (OECD).

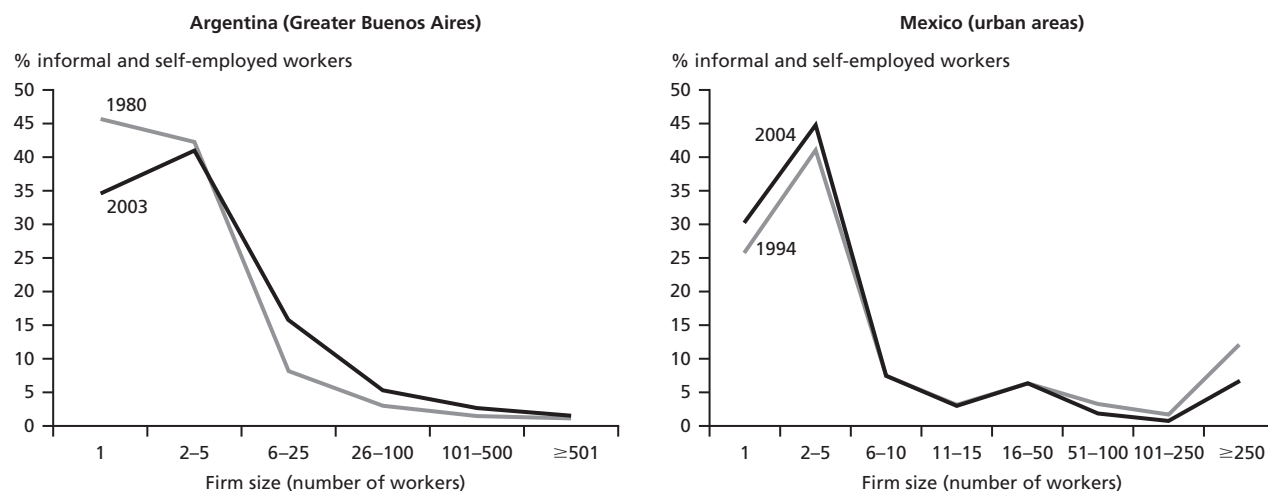
In this light, although a sizable body of literature sees these firms' lack of registration with the authorities through an exclusionary lens, this report argues that

formality can be seen as an input in the production process for which small firms have little need. Formality increases rapidly with firm size and productivity. As an example, whereas 86 percent of Mexican microfirms with only one paid employee do not pay social security contributions, 71 percent of those with five paid workers report paying social security for at least some of their employees. In Brazil, 76 percent of microfirms do not have an operating license and 94 percent do not pay taxes. Those rates fall to 33 and 23 percent, respectively, among firms with five paid workers.

However, the large majority of microfirms remain too small to benefit sufficiently from formality to overcome its various costs. In Brazil, for instance, 87 percent of all microfirms have no paid employees. In Nicaragua, less than 7 percent of microfirms have more than two employees after three years of operation. Most of these firms have no potential to grow (turnover is extremely high in this sub-sector), and thus their credit needs may be marginal. They have a limited number of clients and with most of those clients they have personal relationships, so they would not benefit much from greater access to the organized impersonal markets and courts usually associated with formality. It is not surprising that a survey of informal Mexican microfirms reveals that nearly 75 percent of the microfirms report the main reason for not registering with the authorities is that they are just too small to make it worth their

FIGURE 5

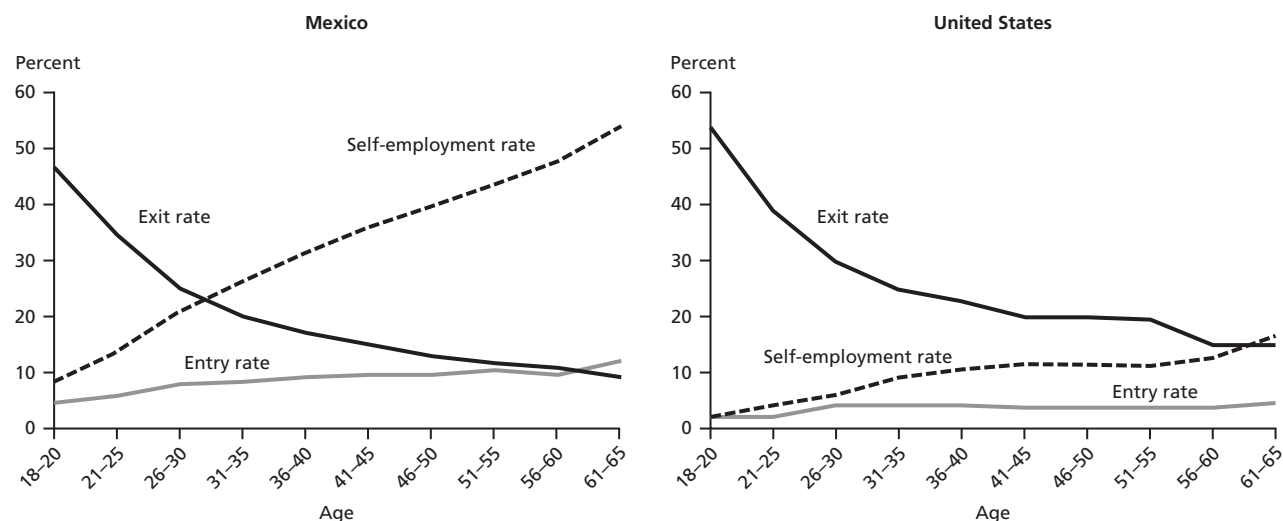
Informal workers across firm size



Source: Author's calculations, based on *Encuesta Nacional de Empleo Urbano* (Mexico) and *Encuesta Permanente de Hogares* (Argentina).

FIGURE 6

Male entry into and exit from self-employment



Sources: Evans and Leighton 1989; Fajnyzber, Maloney, and Rojas 2006; and author's calculations, using *Encuesta Nacional de Empleo Urbano*.

while. In contrast, only 2, 8, and 4 percent of surveyed firms, respectively, answer that the main reason is the high costs and time required to register or the high costs of operating as registered businesses. The same appears true in the Dominican Republic, although prohibitive registration red tape seems to be of greater import in Argentina. In most cases, the degree of formality increases with firm size, suggesting that as firms grow, their demand for the services associated with formal institutions increases, as does the probability of detection by authorities.

Having so much of the workforce in microfirms and having so many of these microfirms unplugged from the formal economy in all likelihood extract some productivity costs. First, there may be scale economies or externalities arising from employment in large firms. To the extent that informality is associated with a preponderance of small firms and that increasing returns to scale are relevant for their size range, informality could lead to considerable efficiency losses. Second, unfair competition from informal firms could slow down the process of creative destruction by which inefficient firms are replaced by their more efficient competitors, and could negatively affect the incentives of formal firms to innovate and adopt new technologies. Third, if reducing the costs and increasing the benefits of formality allow previously informal firms to gain increased access to markets and services, the result could be increased aggregate productivity growth.

The firm-level evidence reported here presents a mixed record on the above hypotheses. The microfirm evidence suggests that firms choosing to register do have better performance or, alternatively, the firms that started operations being registered exhibit, on average, higher levels of labor productivity than their equivalent unregistered peers. In particular, programs in Brazil and Mexico to reduce the costs of becoming formal have had a positive impact on the performance of those firms that decided to cross the frontier from informality to formality. The evidence on the effect of these programs on the aggregate levels of informality and productivity through time remains inconclusive (as will be discussed in chapter 6). Further, evidence from the investment climate reports suggests that firms surrounded by complying firms have higher rates of productivity. However, there is evidence that, in some cases, informality reflects defensive evasion of possibly excessive regulation.

All evidence considered, the report concludes that tilting the cost-benefit analysis toward formality for a substantial percentage of informal firms requires a fair combination of “carrots” and “sticks.” As an example, in Brazil and Mexico, microfirms constrained at the frontier coexist with entrepreneurs who show no demand for the presumed benefits of formality and for whom reducing registration costs would not lead to formalization under present conditions. Similarly, evidence for the Dominican Republic suggests that many informal firms perceived more gains from

staying under the state radar, including not only the tax savings but also the avoidance of interaction with the state bureaucracy altogether. Further, a large fraction of firms that become formal allegedly do so to avoid fines and bribes, according to surveys of many countries in the region (see figure 7). Thus, interventions to address regulatory constraints faced by small firms or to reduce tax rates might not be enough to achieve a major change in their levels of formality. Such a goal would also require improving the positive incentives for joining the formal sector, including improvements in private and public services available to formal firms (for example, credit, contract enforcement, technical assistance, and the like), and enhancing the level of enforcement to increase the opportunity cost of remaining informal.

Nonetheless, the low aggregate productivity in the formal sector would put some limits on the overall effect of even well-designed, integrated programs to induce informal firms to become formal. Hence, a large chunk of the efforts to reduce informality must focus on policies to enhance productivity and growth in the formal sector, which would increase the perceived benefits of formality, along with the opportunity cost of remaining informal.

Finally, in many countries in the region, even large firms show considerable tax evasion and labor informality. They

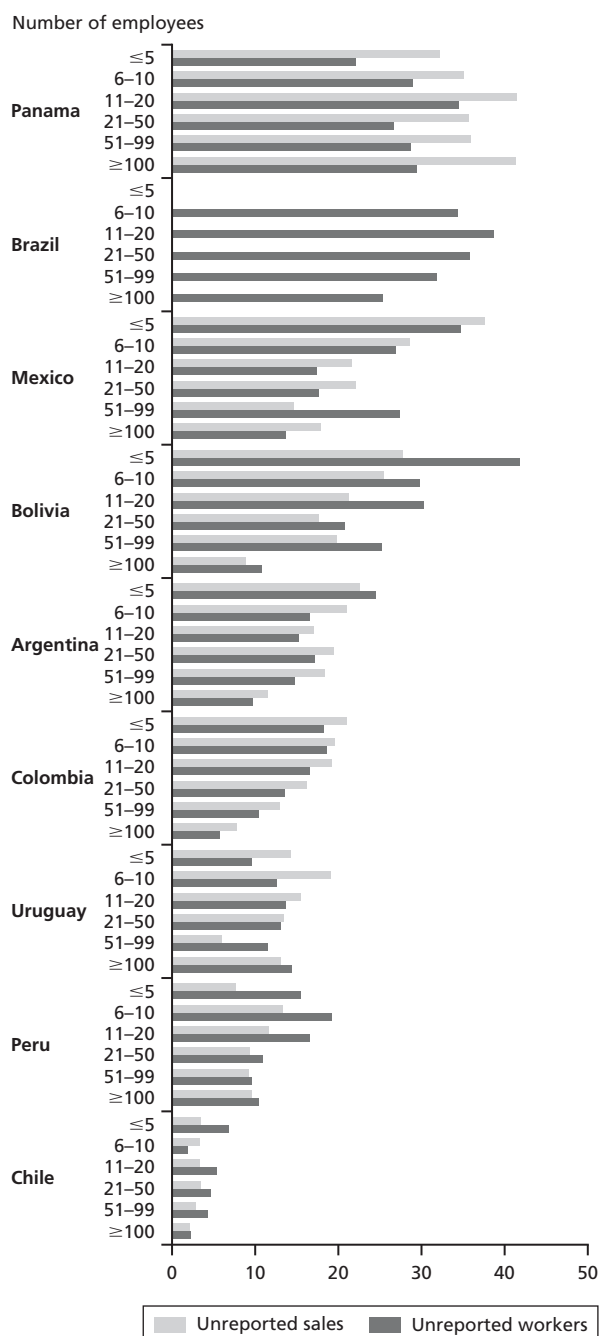
have a significant number of employees without social security contributions, and the fraction of underreported sales for tax purposes is quite high—even as high as in small firms, in some cases (see figure 8). Firms again appear to undertake a careful cost–benefit analysis, weighing the “private” benefits

FIGURE 7

Advantages of formalization reported by IFC-surveyed firms

Source: Investment climate surveys 2006.

FIGURE 8

Underreporting of tax and social security contributions, by firm size

Source: Investment climate surveys 2006.

of informality (evading taxes, avoiding burdensome regulations) against the “private” costs of informality (risk of penalties and bribes, imperfect access to markets and government services) in choosing their “degree” of formality. Administrative and tax simplification programs to reduce red tape and compliance costs, regulatory reviews aimed at eliminating laws and regulations that are either anachronistic or privately motivated, and enhanced enforcement are the key policy responses with respect to the phenomenon of partial informality within large firms.

Need for more effective and legitimate institutions

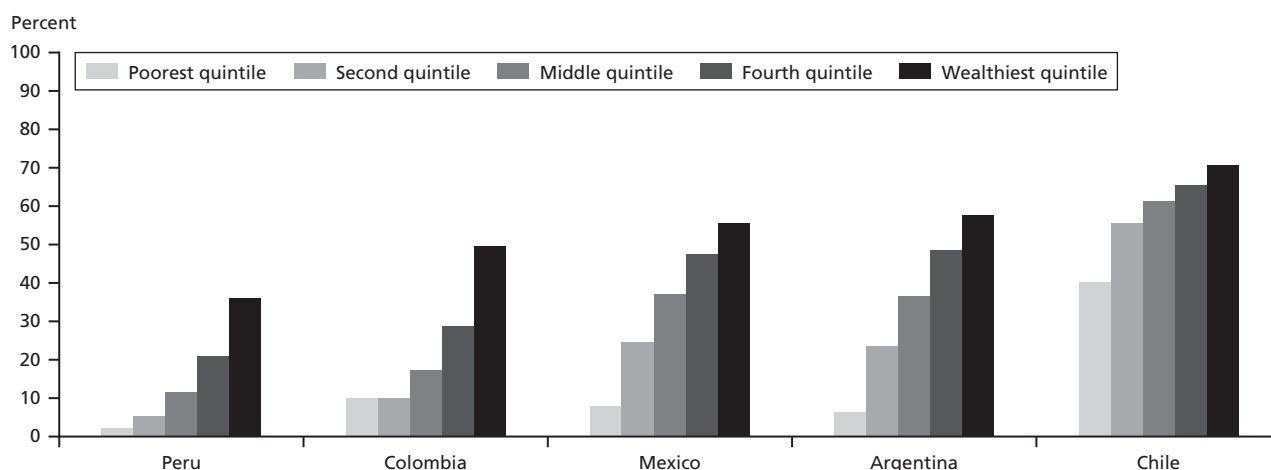
Whether informality is an outcome of exclusionary policies or mechanisms, or a result of cost–benefit decisions by firms and individuals that lead them to opt out of formal institutions, it represents a fundamental critique of the Latin American state at several levels. In the burdensome business and labor market regulations, as with the poorly designed social protection systems, the state is complying inadequately with its designated roles. This failure, in conjunction with an abiding lack of confidence in who the state represents and serves and in its capacity to enforce the law, may intensify the tendency of many Latin Americans—perhaps no more or less than citizens of other emerging regions—to opt out, rendering the fulfillment of the fundamental roles of the state all the more difficult.

As an example, most countries in the region are characterized by “truncated welfare systems,” in which those in the formal sector have access to an often generous multidimensional package of social security, while those in the informal sector have much more limited access to government benefits or formal risk management instruments. Not only has overall progress in increasing social security coverage been slow in the region, but coverage has actually declined in a number of countries throughout the 1990s. And, in nearly all countries in the region, coverage rates are significantly lower among low-income than among high-income workers (see figure 9). Indeed, in many countries the poorest workers and families are practically excluded from the system. At the same time, although independent workers can choose to participate in formal social security systems in many countries, very few do it voluntarily. Actually, there is evidence that many workers opt out from the system once they obtain the right to a minimum pension (see Gill, Packard, and Yermo 2004).

The underperformance of Latin American and Caribbean states may reflect deeper, poorly resolved social tensions that may be thought of as constituting a dysfunctional underlying “social contract”; in Hirschman’s terms, the lack of voice and loyalty. Beyond high informality, this can also be seen in the inability of the state to redress the long-standing high inequality, in the weak rule of law, in the sometimes large share of undocumented citizenry, or in the recurrent

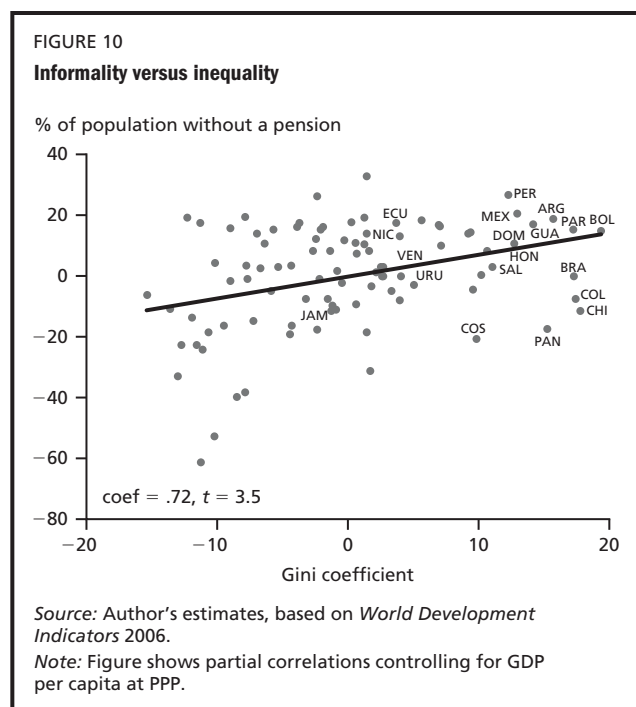
FIGURE 9

Economically active population contributing to the pension system



Sources: Adapted from Rofman and Lucchetti 2006.

Note: Estimates based on national household surveys from the following years: Peru (2003), Colombia (1999), Mexico (2002), Argentina (2004), and Chile (2003).



bouts of macroeconomic instability. For example, high inequality of incomes and power is correlated with informality (see figure 10) and is often associated with weak institutions and state capture by both elites and organized segments of the middle class (see de Ferranti et al. 2004; Perry et al. 2006). State capture leads to the generalized perception that the state is run for the benefit of the few, and thus it reinforces a social norm of noncompliance with taxes and regulations, what might be dubbed a “culture of informality.” Noncompliance is then further compounded by the suspicion that others are not complying either—an absence of what is termed “strong reciprocity—which, in turn, makes enforcement even more difficult.” As an example, in the Dominican Republic, 19 percent of small firm owners report that they do not register or contribute to pension systems because “businesses like them don’t.”

At a more aggregate level, after controlling for per capita income, informality is negatively correlated with “tax morale”—society’s disposition toward tax compliance—which, in turn, depends inversely on perceptions of “state capture” (see figure 11) and positively on perceptions of the quality of public spending. Furthermore, informality measured as the share of the workforce in self-employment is negatively related to the quality of institutions, as indicated in figure 12. This suggests that greater exit to independent work accompanies higher distrust of the state. Inequality and its correlates then lead to inefficient laws and regulations

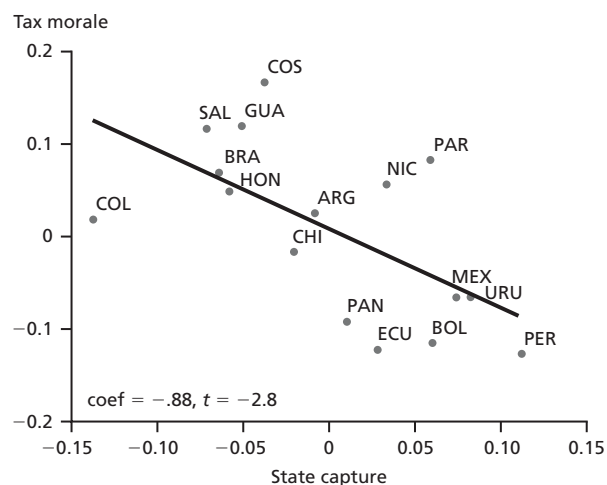
(tax privileges and loopholes, weak competition laws, and cumbersome regulations). Those inefficiencies lead to low tax collections and the abuse of monopolistic power, and they further erode state legitimacy and its capacity to provide public goods and services and to enforce the law. Some recent literature has focused on the possibility that Latin America is locked, in fact condemned, to a bad social equilibrium where entrenched elites have no interest in responding to the voiced petitions of those seeking full participation.

However, this report is more optimistic about the possibilities for incremental reform, despite acknowledging the vital importance of enhancing the legitimacy of the state. Ironically, this conclusion partly emerges from the study of the distressing increases in informality in many countries of the region over the 1990s. These increases can be traced partly to sharp increases in real minimum wages in some countries, inadequate macroeconomic policies (that led to artificial booms in nontradable, highly informal sectors) in some countries in the early 1990s, changes in labor market and social security legislation or weakened enforcement capabilities, and to the increased availability of noncontributory social protection schemes for informal workers. In the meantime, poorly designed social security systems continued to tax heavily workers in the formal sector. Indeed, if informality increased in many countries during a decade that exhibited modest growth due at least partially to a combination of policy missteps, then a shift toward better policies at least can reverse these upsetting trends. Furthermore, although it appears that inequality and informality share important causal determinants, Chile shows that strong and more fair institutions and a fall in informality can be built even in the midst of high income inequality. Quoting Hirschman (1971) again, we find a “bias for hope”—good policies can have important effects.

Summing up: Policy implications of the report

Achieving significant reductions in present informality levels will require, first and foremost, actions to increase the aggregate productivity in the economy. A more enabling investment climate will permit formal firms to expand and pay higher wages. Raising human capital levels, especially for the poor, will permit more workers to find remunerative jobs in a more dynamic formal sector. Without such improvements in aggregate productivity, we will continue to find a very large number of microfirms, characterized by high turnover, weak growth prospects, and low productivity, that would see little benefit in engaging with formal

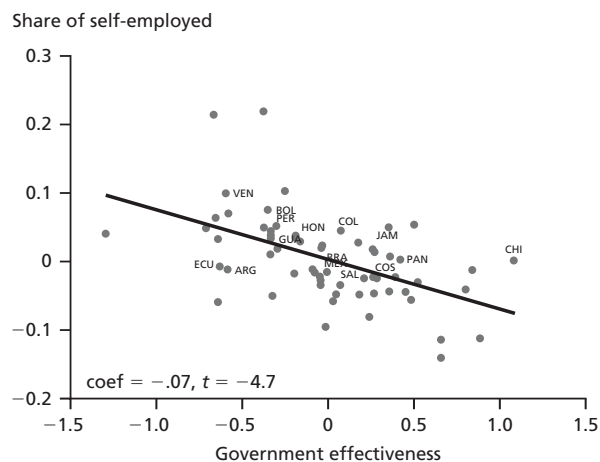
FIGURE 11

Tax morale and state capture

Source: Author's estimates, based on data from *World Development Indicators* 2006 and *Latinobarómetro* 2004.

Note: Figure shows partial correlations controlling for GDP per capita at PPP. State capture is proxy by an indicator of the perception about the economy being run according to the interests of a few. To construct the indicator we ask: In general terms, would you consider that the country is governed according to the interests of a few, or is it governed for the benefit of the country?

FIGURE 12

Self-employment and quality of institutions (governance)

Sources: Author's estimations, based on *World Development Indicators*; *Worldwide Governance Indicators*, World Bank 2005; and investment climate surveys, 2005.

Note: Partial correlations controlling for GDP per capita at PPP. Government effectiveness index measures the quality of public service provision, the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies. Higher values correspond to a more effective government.

institutions. Without generalized improvements in human capital, many unskilled workers would continue to prefer self-employment, even if their earnings are low, because the jobs they could find in the formal sector would also deliver very low earnings. Such improvements in human capital have to be synchronized with those in the investment climate, as stressed in previous studies (particularly, de Ferranti et al. 2003), because otherwise the demand for more skilled workers will not grow at the same speed as the increase in supply—thereby depressing skill premiums (for instance, to secondary education) and further eroding incentives to invest in education.

However, there are many things that can be done to change the balance of incentives for those workers and small firms whose implicit cost–benefit estimates place them close to the margin between formality or informality. The same can be said of larger firms that remain partially informal. To have a significant effect, such actions normally require a good balance of carrots (reforms and actions to facilitate, reduce the costs, and increase the benefits of formalization) and sticks (enhanced, evenhanded enforcement of such improved laws and regulations). Furthermore, those excessive regulations and taxes that create some degree of

labor market segmentation, as evidenced in most countries at least during some periods, need to be reduced. We discuss below some policy changes that may be especially promising in particular circumstances.

Achieving a deeper change in incentives also requires actions to change the pervasive culture of noncompliance that we observe in most countries in the region. Because such social norms are, in part, the result of a lack of trust in the effectiveness of the state and the equity of its actions, overcoming the culture of informality probably requires major improvements in the quality and fairness of state institutions and policies. In short, it requires building an effective and inclusive social contract in which the great majority of individuals feel compelled to participate and comply with state mandates.

Reforming labor market policies to increase productivity of the formal and informal sectors

The findings of this report suggest that labor market policies are important determinants of informality, working through three channels. First, excessive labor costs, whether arising through labor legislation or unrealistic

union demands (such as exaggerated minimum wages, severance costs, or labor taxes and contributions), depress the number of jobs in the formal sector, often creating the classic segmented market. Other recent reports from the World Bank—“Minimum Wages and Social Policies: Lessons from Developing Countries” (Cunningham 2007b) and “Job Creation in Latin America and the Caribbean” (World Bank 2007)—as well as the Inter-American Development Bank (IDB) (2006) publication *Good Jobs Wanted*, have investigated in detail the trade-offs encountered in offering strong protections to some workers at the possible expense of excluding others (see also de Ferranti et al. 2001). Minimum wages in most countries are not extremely binding, but in some—such as Colombia—they seem to be a deeply segmenting force that begs for moderation. Higher non-wage burdens on formal jobs in Brazil, Colombia, and Peru appear responsible for substantial declines in formal employment. Further, the experience in the OECD countries suggests that such regulations have a heavy exclusionary impact on young people who are trying to find jobs and who, in Latin America and the Caribbean, are especially overrepresented among the involuntary informal salaried. Nor does the evidence of a very high degree of integration of the formal and informal sectors in some other countries in the region necessarily imply satisfactory labor codes. Latin America’s globally very high levels of severance costs, for example, may substantially reduce job creation arising from growth without necessarily segmenting the market. Furthermore, as in the case of minimum wages, regulations sometimes can be binding in practice for the informal salaried sector and may inhibit job creation there.

Second, legislation can create incentives for voluntary informality. The design of social safety nets and labor legislation needs to consider a more integrated view of the labor market, taking into account the cost–benefit analysis that workers and firms make in deciding whether to interact with formal institutions. High labor taxes or contributions that do not correspond to benefits that workers value cause workers to opt out of the formal labor market. The difficulties of managing work and children under rigid formal work contracts also lead young mothers to opt for informal independent work offering more flexibility.

Third, the effect of labor market institutions on productivity growth has probably been underemphasized. Theory and anecdotal evidence suggest that excessive restrictions on job reallocation or destruction for just cause, or other

related government- or union-induced inflexibilities, may have a disincentive effect on technology adoption, which accounts for up to half of the differences in levels of economic development. And, as shown in this report, the level of economic development is the most important determinant of observed informality levels.

Overall, the present constellation of often well-intended but heavy-handed labor regulations poorly serves workers and firms, and both could benefit from substantial reform. In particular, stronger enforcement of a redesigned labor code that combines strengthened safety nets, well-designed worker protections, and worker representation with the flexibility that firms need to adapt in a global economy has the potential to expand formal employment and reduce opting out. Simply tightening the enforcement of existing laws, particularly in the largely informal microfirm sector, may just eliminate jobs—many of which the chapters in this volume will show to be of good quality when measured by the worker’s overall welfare. At the other extreme, attempting to reduce the weight of labor legislation by creating special classes of less protective contracts can be problematic. When well designed, such contracts may offer flexibility that helps young people enter the market. However, they can often effectively create a parallel, unregulated “formal” sector that displaces formal contracting. That can lead to higher job turnover and to diminished incentives for training and may contribute to the overall culture of informality. Provisions to accommodate different non-wage costs for smaller firms and flexibility in benefits plans (such as simplified health/pension plans) may offer an extension of the overall rubric of labor protections without prejudicing the viability of these firms. Perhaps more flexible work schedules, or enhanced availability of day care, would keep more women in the formal sector, although care must be taken to avoid measures that might lead firms to discriminate against women in hiring as a result. Finally, institutional strengthening (staffing, training, technical assistance) of the labor ministries and coordinating of relevant public agencies (social security administration, enterprise development agencies, and competitiveness councils) are needed so that those ministries and agencies can assume their increasingly more complex role of facilitating labor productivity growth.

Informality sharply decreases with education, partly because the opportunity cost of being independent rises. However, the substandard education and training systems in the Latin American and Caribbean region both impede

the growth necessary to generate jobs in the more modern sector of the economy and reduce workers' attachment to it. Furthermore, the poor signaling of education quality that results from a lack of uniform certification or accreditation impedes the entry of young workers into formal jobs. Remedying these failures, perhaps along with an expansion of intermediation services, may reduce the information asymmetries that young workers face. Further discussion of necessary reforms in this area is offered in the regional report "Raising Student Learning in Latin America and the Caribbean: The Challenge for the 21st Century" (Vegas and Petrow 2007). Ongoing upgrading of the workforce through training, particularly in rapidly evolving industries, is a central element of the national innovation system, and is critical to developing skills used in the modern sector of the economy and to promoting productivity growth. Training systems in the region, however, need to be more competitive and responsive to market demands.

In sum, the high rate of informal employment in some cases reflects classic segmentation, but in others it simply reflects the high costs and small benefits of employment in the formal sector. Labor regulations need to allow for productivity growth in both sectors, while revisiting the design of regulations and social protection systems that provide incentives for firms and individuals to become informal.

Reengineering social protection to cover all citizens

Central to this agenda is the need to rethink and, in fact, reengineer social protection policy and programs in much of Latin America and the Caribbean. Poor access to basic risk management instruments, coupled with poor design of social security schemes, serves workers poorly and provides an incentive to be informal. Fundamentally, this may require rethinking the traditional model of social protection in which protection depends on the specific form of the labor contract. A broader notion of who has access to basic risk management instruments is needed—one based on ensuring the basic protection and welfare of countries' *citizens* rather than of "workers" as traditionally and narrowly defined.

Drawing insights from the economics of insurance, this report outlines a long-term vision for social protection reform in the region.¹ In the case of health care, because shocks that go "uncovered" can impose significant external

costs on society, there is a case for providing a package of minimum essential direct cover, de-linked from the labor contract and financed through general taxation. In the case of old-age security, there is also a case to provide essential cover in the form of a poverty prevention pension targeted toward the poor as part of a broader multipillar pension system that includes provisions for individual retirement savings. The social costs associated with people falling into poverty in old age create a clear risk management rationale for providing minimum old-age support de-linked from the labor contract; nonetheless, the high probability of income loss in old age suggests that saving should play the main role in earnings replacement during old age.

For a number of reasons, including those related to fiscal and institutional capacity, movement to provide essential cover in health care and old-age security, de-linked from the labor market and financed by general taxes, represents a long-term agenda for many countries in the region. It is important, therefore, for countries to orient their short-to-medium-term policy agendas in ways that are consistent—or at least *not inconsistent*—with their long-term vision. This will be critical if the region's governments are to ensure that the objectives of social policy—and particularly social risk management—are well aligned with those of higher productivity and growth.

To this end, there will be high returns to governments in the region undertaking step-by-step reforms to improve the efficiency of existing programs as well as to establish greater consistency and incentive compatibility across programs. Several sets of actions will contribute to short-term improvements in social protection while moving countries in the direction of achieving essential cover over the long term. These actions would include efforts to improve the cost-benefit ratios of programs in several ways. First, complex, multidimensional social security benefit packages should be unbundled to focus on increasing access and quality of those programs with high public goods content (for example, health care, old-age security). Second, efforts should be made to improve program quality and benefits via microefficiency reforms in health care and country pension systems. Third, program design should be strengthened—for example, by revising overly burdensome pension-vesting periods, fostering benefits portability, eliminating double payments for health insurance, and so on—to account for and enable greater worker mobility. Finally, rules, eligibility requirements, and benefits levels across programs and institutions of social security and

social assistance should be made consistent and compatible with incentives. Such types of measures will provide the foundation for more effective social protection for all citizens and help strengthen the alignment of social and economic policy objectives.

Improving opportunities for workers in the formal sector while reducing barriers to business formalization

As suggested above, the fact that few micro firms grow from small informal entities to large formal firms can be explained in two complementary ways. On one hand, the presence of low opportunity costs for entry into the sector could be to blame because it would lead to a predominance of low-productivity businesses with low growth prospects and high failure rates. In this context, policy makers wishing to reduce informality should focus not only on altering the direct costs and benefits of formality but also on altering the drivers of formal sector productivity, including measures to improve the investment climate and policies aimed at increasing human capital accumulation. On the other hand, however, policy-induced barriers to formalization could impede microfirms' access to technologies, markets, and government services, and that would explain at least some of their low growth and job creation rates. In this second interpretation, reductions in informality and improvements in microfirm performance could be achieved by lowering statutory burdens to formalization and implementing administrative simplification programs aimed at reducing the transaction costs associated with operating legitimate businesses. In particular, programs that facilitate business registration—for instance, through the use of Internet-based technologies and one-stop-shops—should lead to an increase in formality rates and improved microfirm performance.

In practice, the existing evidence from Brazil and Mexico suggests that administrative simplification programs do lead to increases in formal firm registration rates, with large consequent improvements in the revenue- and employment-generating capacity of the corresponding businesses. However, the number of positively affected firms is relatively small in comparison with the overall size of the informal microfirm sector. Moreover, at least in the case of Mexico, it appears that most of the newly registered firms are new entrants—that is former wage workers who decide to open new formal businesses—rather than existing informal firms that become formal as a result of lower

registration costs. Hence, although high entry costs may keep some entrepreneurs from starting new formal businesses, the available evidence suggests that just lowering administrative barriers to business registration may not have a major impact on aggregate levels of informality. This finding does not imply that those interventions are not important. First, lower entry costs do induce at least some entrepreneurs to open new formal businesses, and their improved performance probably would be sufficient justification for the corresponding reforms. Second, the impact of administrative simplification programs may be larger when accompanied by other complementary measures aimed either at increasing the potential benefits of joining the formal sector or at reducing the costs of regulatory compliance—beyond the facilitation of firm registration.

With regard to the first alternative, recent evidence from randomized experiments suggests relatively high returns to capital among very small Mexican microenterprises, and this suggests that considerable increases in income could be obtained through measures to increase small businesses' access to bank credit and other forms of external finance. Similarly, in principle, formality could be made more attractive by improving the provision of business development and training services available to formal firms, and by facilitating access to product markets through public procurement opportunities and supplier development programs aimed at increasing links with larger private firms. Other ways of making formality more attractive include improvements in the quality of legal services available to small businesses so that they may find it less risky to expand beyond local markets, and creating mechanisms to provide information to entrepreneurs wishing to formalize their businesses (from advisory services on taxes and regulations to information on financial and nonfinancial services available to them).

As for measures to reduce other costs of operating legitimate businesses, governments should consider performing comprehensive regulatory reviews aimed at eliminating unnecessary and costly bureaucratic requirements. In this respect, the challenge is that of distinguishing relevant from anachronistic regulations, as well as identifying those regulations that significantly increase the costs of operating private businesses and are not justified by legitimate public interests, such as the protection of public safety or the environment. If successful, such regulatory reforms may help reduce informality both by increasing job creation in the

formal sector and by reducing regulatory noncompliance among formally registered, medium-size and large firms—for example, tax and social security evasion. Indeed, although cross-country differences in the size of the informal sector are largely explained by countries' general levels of development—which, in turn, are driven partially by the quality of governance and institutions—there is also evidence that, for given levels of per capita income, informality tends to be higher where regulatory burdens are heavier. Moreover, not only the quantity but also the quality of regulations appears to matter, as is illustrated by the finding that firms facing higher levels of corruption are more likely to evade taxes and social security. In that context, the objective should be to eliminate unnecessary regulations while enhancing their evenhanded enforcement and streamlining administrative processes to reduce excessive red tape.

Overall, a wider and integrated approach appears to be necessary to switch the incentives of a large fraction of informal firms in the direction of formality. Such an approach would likely have to combine carrots (such as lower costs of formalization, better and more efficient government services, and greater access to government- and market-provided services for formal firms) with sticks (for example, increasing evenhanded enforcement of regulations and, hence, the expected cost of being caught). The correct mix of policies, however, is likely to vary across countries and over time, depending on the relative importance of the various determinants of informality. Moreover, policies aimed at reducing firm informality should be considered in conjunction with the labor market and social protection issues associated with the possibility of large contingents of previously informal workers shifting to other segments of the labor market.

Simplifying tax laws, facilitating compliance, and strengthening enforcement

As mentioned above, underreporting of sales and incomes for tax purposes is common even in large firms, although the degree of compliance and its relation with firm size vary widely among countries. To deal successfully with this dimension of informality, there is also a need for an integrated approach consisting of both sticks and carrots.

First, there is substantial room for measures that fall within the traditional monitoring and “punishment paradigm” of tax compliance. In most countries, improvements are needed in the three main aspects of tax administration: taxpayer registration, audit, and collection. Taxpayer regis-

tration can be enhanced through better use of third-party information (for example, cross-references between tax reporting, social security records, and data from the financial system). Audits can be made more effective by the adoption of modern audit technology, such as has been the case in Chile and Spain. In most countries there is scope for reducing administrative and compliance costs and increasing collections through changes in tax structure, combining reductions in marginal tax rates with the elimination of exemptions and privileged regimes that create loopholes in value-added, income, and property taxes, and simplification of deductions. Collections can also be increased by facilitating payments through the banking system, relying more heavily on source withholding and applying non-harsh penalties for noncompliance often and consistently. Finally, criminalization of certain tax offenses in combination with a modernized tax administration agency has been credited with playing a key role, among other factors, in Spain's success in the late 1970s and 1980s in drastically reducing tax evasion and eventually doubling the ratio of tax revenue to gross domestic product.

Second, reforms should also emphasize the “service paradigm” with policies to enhance the role of the tax administration as a facilitator and a provider of services to taxpayer-citizens. Promoting taxpayer education and developing taxpayer services in filing returns and paying taxes, broadcasting advertisements that link taxes with government services, stimulating voluntary compliance by lowering compliance costs, simplifying taxes and their payment, and promoting a taxpayer—and a tax administrator—“code of ethics” have proved to be useful complementary measures to the punishment paradigm to enhance compliance. The use of semiautonomous revenue authorities has been shown in several countries to improve tax administration with a more service-oriented approach to tax enforcement.

The service paradigm fits squarely with the perspective that emphasizes the role of social norms in tax compliance. Experience from other countries shows that a government's commitment to evenhandedly enforce the tax laws while facilitating taxpayer compliance can have an important effect on the pervasive culture of noncompliance found in many countries in the region. Disclosure of public expenditure information, and the participation of and supervision by citizens—that is, “voice”—in the way taxes will be spent, may also help increase trust in the state and may contribute to social norms of compliance. These factors have been credited with success in increasing tax

compliance (and collections) in Chile and Spain, particularly through widespread consensus among political parties about the need for the reform of the tax system, improved democratic governance, and highly visible enhancements in social and other public services.

Enhancing the effectiveness and legitimacy of the state

Improving the quality of state policy making is one element in a larger agenda of reducing the culture of informality that also requires increasing the perceived efficiency, fairness, and accountability of the state. To spin it positively, a particular policy change (such as greater and more evenhanded enforcement of adequate and equitable taxation laws and other regulations) may reduce informality by more than what would be expected, given individual elasticities, if it induces a change in the social norm of tax and regulatory compliance that permits the state to improve the provision of public goods and services and to enforce the law.

As with informality more specifically, the task of bolstering the legitimacy of the state involves both carrots and sticks. Identifying the correct set of carrots to foster a sense of greater inclusion and responsiveness requires detailed analysis of specific country circumstances to detect the “binding constraints” whose removal would cause a shift in the decision of a large number of firms and workers located close to the three borders between formality and informality. Adequate sticks require modern enforcement techniques (for example, many countries in Latin America and the Caribbean have not yet developed adequate tax-auditing systems like those common in OECD countries), political resolve, and evenhanded enforcement of laws and regulations that are perceived as fair.

Again, to improve the legitimacy of the state through more equitable public policies and programs, it is necessary to consider carefully the formality incentives or disincentives generated by government programs. Greater inclusion or a more participatory social contract may not by themselves reduce informality. Some countries’ laudable attempts to extend free social protection services to informal workers, made without reforming badly designed social security systems that tax many of its formal contributors, have inadvertently reinforced perverse incentives that could actually lead to higher levels of informality. Furthermore, the rise in informality in both metropolitan Brazil and Colombia results partially from well-intentioned

policies emerging from the constitutional exercises that strove to be more inclusive. Hence, however well meaning or inclusive policies may be, they must be well designed. The inclusive state must also be a competent state.

Informality and the development agenda

Informality reflects underdevelopment, and this report seeks to tease out some of the interactions and directions of causality between the phenomena of informality and development. There is evidence that, in part, informality is merely a stage in the development process: the ubiquitous microfirm reflects the unattractive options in the small modern sector and the traditional reliance on family and community. However, other evidence suggests that, in part, informality is a canary in the coal mine—the symptom of poor policies and, more profoundly, a lack of confidence in the state and perhaps in our fellow citizens. To return to Hirschman, confronted with a lack of voice in and relevance of the state, Latin Americans take their business elsewhere; and, in doing this, they further undermine the region’s growth prospects. Hence, redressing the causes of high informality is part and parcel of the broader development agenda.

Note

1. See de Ferranti et al. (2000) for an earlier application of the economics of insurance in the region.

References

- Arias, Omar, and Leonard Lucchetti. 2007. “Informal Employment and Self-Rated Welfare: Measuring Compensating Differentials.” Photocopy. World Bank, Washington, DC.
- Bosch, M., and W. F. Maloney. 2006. “Gross Worker Flows in the Presence of Informal Labor Markets. Evidence from Mexico, 1987–2002.” Policy Research Working Paper 3833, World Bank, Washington, DC.
- Cunningham, W. 2007a. “Life Trajectories of Latin American Workers.” Photocopy. World Bank, Washington, DC.
- . 2007b. “Minimum Wages and Social Policies: Lessons from Developing Countries.” Photocopy. World Bank, Washington, DC.
- Cunningham, W., L. McGinnis, R. G. Berdu, and C. Tesliuc. 2007. “The Promise of Youth: Policy for Youth at Risk in Latin America and the Caribbean.” Photocopy. World Bank, Washington, DC.
- de Ferranti, D., G. E. Perry, F. H. G. Ferreira, and M. Walton. 2004. *Inequality in Latin America and the Caribbean: Breaking with History?* Washington, DC: World Bank.
- de Ferranti, D., G. E. Perry, I. Gill, J. L. Guasch, W. F. Maloney, C. Sanchez-Paramo, and N. Schady. 2003. *Closing the Gap in Education and Technology*. Washington, DC: World Bank.
- de Ferranti, D., G. E. Perry, I. Gill, L. Servén, F. Ferreira, W. F. Maloney, and M. Rama. 2000. *Securing Our Future in the Global Economy*. Washington, DC: World Bank.

- de Ferranti, D., G. Perry, D. Lederman, and W. F. Maloney. 2001. *From Natural Resources to the Knowledge Economy*. Washington, DC: World Bank.
- de Soto, H. 1989. *The Other Path. The Invisibles Revolution in the Third World*. New York: Basic Books.
- Evans, D. S., and L. S. Leighton, 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79 (3): 519–35.
- Fajnzylber, P., W. F. Maloney, and G. V. Montes Rojas. 2006. "Microenterprise Dynamics in Developing Countries: How Similar Are They to Those in the Industrialized World? Evidence from Mexico." *World Bank Economic Review* 20 (3): 389–419.
- Gasparini, L., and L. Tornarolli. 2006. "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata." Photocopy. World Bank, Washington, DC.
- Gill, I., T. Packard, and J. Yermo. 2004. *Keeping the Promise of Social Security in Latin America*. Washington, DC: World Bank.
- Hirschman, A. O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- . 1971. *A Bias for Hope*. New Haven, CT: Yale University Press.
- . 1981. *Essays in Trespassing. Economics to Politics and Beyond*. Cambridge, UK: Cambridge University Press.
- IDB (Inter-American Development Bank). 2006. *Good Jobs Wanted: Labor Markets in Latin America*. Washington, DC.
- Loayza, N., and J. Rigolini. 2006. "Informality Trends and Cycles." Policy Research Working Paper 4078, World Bank, Washington, DC.
- Perry, G., O. Arias, J. H. Lopez, W. F. Maloney, and L. Servén. 2006. *Poverty Reduction and Growth: Virtuous and Vicious Circles*. Washington, DC: World Bank.
- Rofman, R., and L. Lucchetti. 2006. "Pension Systems in Latin America: Concepts and Measurements of Coverage." Social Protection Discussion Paper 0616, World Bank, Washington, DC.
- Vegas, E., and J. Petrow. 2007. "Raising Student Learning in Latin America and the Caribbean: The Challenge for the 21st Century." World Bank, Washington, DC.
- World Bank. 2006. *World Development Indicators*. Washington, DC. Available at <http://www.worldbank.org>.
- . 2007. "Job Creation in Latin America and the Caribbean." Photocopy. Washington, DC.

CHAPTER 1

The Informal Sector: What Is It, Why Do We Care, and How Do We Measure It?

SUMMARY: This chapter seeks to unpack our understanding of the term informality, why we may care about it, and what dynamics may be driving its elements. The number of phenomena it encompasses and the limitations of its measures are manifold, dictating caution in employing the term. Yet two stylized facts remain: First, however measured, informality is high in Latin America, although not obviously so for the region's level of development; and it remains an important phenomenon. Second, in several countries it has experienced striking increases over the last decades. Whatever adverse regulatory, poverty, growth, or social morale implications "informality" may have, they have become more relevant with time.

Introduction: What is informality?

The term *informality* means different things to different people, but almost always bad things: unprotected workers, excessive regulation, low productivity, unfair competition, evasion of the rule of law, underpayment or nonpayment of taxes, and work "underground" or in the shadows.¹ The multiplicity of adjectives from very distinct fields of study suggests that we may have a classic "blind men and the elephant problem"—everybody touches part of the animal, but understands only the part that they touch. More likely still, we are exploring several distinct phenomena as we attempt to describe one ungainly composite "informality." To further complicate things, Keith Hart, the purported coiner of the term *informality* (and one who did not think the sector was necessarily bad), argued that the source of our blindness—the undocumented nature of the sector—left the sector especially vulnerable to being a *tabula rasa* on which analysts projected their particular concerns:

Most enterprises run with some measure of bureaucracy are amenable to enumeration by surveys, and, as such, constitute the "modern sector" of the economy. The remainder—that is, those who escape enumeration—are variously classified as the "low-productivity urban sector," the "reserve army of the *underemployed and unemployed*,"

the "urban *traditional* sector," and so on. These terms beggar analysis by assuming what has to be demonstrated (Hart 1973, p. 68).

With this, Hart cautions us against inquiry with excessively well-informed ideas of what we may find.

Fortunately, the accumulation of rich data sets over the last decades has cast progressively more light on the realm of the informal, permitting us to document the great heterogeneity of actors and their *razón de ser*. Among them we find the following:

Labor:

- workers, particularly the old and young, who would prefer a job with standard labor protections, but are unable to get one;
- workers who have quit formal sector jobs to start a microbusiness to be their own boss, make more money, and avoid paying social protection taxes; and women leaving formal salaried jobs for the flexibility of balancing home and income-raising responsibilities.

Microfirms:

- microentrepreneurs with no intention of or potential for growing, and hence no intention of engaging the institutions of civil society;

- microentrepreneurs stymied in their expansion by excessively high barriers to registering with the government and thereby accessing other inputs offered by the informal sector.

Firms:

- firms and individuals avoiding taxation or other mandated regulations because everybody else does, and because enforcement is weak and uneven;
- firms registering only part of their workers and part of their sales—or declaring only part of the salary of their workers—due to an excessive regulatory burden.

Though far from exhaustive, these three pairs are illustrative of the variety of types of agents captured under the rubric of “informality” and, further, capture three different margins of informality discussed later. They also suggest the reasons we care about informality.

Why do we care about informality?

Each example above has a different underlying logic and reason for being and, hence, a reason that we may care about its existence or size from a policy perspective.

Unprotected families

As the regional flagship report *Securing Our Future* (de Ferranti et al. 2000) noted, while development is often seen as a process of increasing income, in practice we have also seen the emergence of institutions to shelter families from adverse shocks, be they loss of job, illness, or natural calamity. The presence of a large fraction of the workforce in Latin America that does not count on formal mechanisms to hedge or mitigate these shocks is, hence, of intrinsic concern. What complicates policy making is that, as the chapters in this volume will show, workers often choose jobs that lack such benefits or they willingly leave jobs that offered such benefits, valuing more other characteristics of informal jobs. In this case, the worker and his/her family must be at least as well-off as before, but may still be vulnerable to some types of misfortune (in particular, health) for which informal protections are few. There also may be externalities for society in, for instance, the classic case of families undersaving for retirement. Further, if, as is suggested throughout this report, some informality is due to the low valuation of government-provided services compared with their implicit or explicit costs to workers, then a choice to be unprotected may point to a dysfunctional and inefficient social protection system.

Drag on productivity and growth

Rigidities in either the labor or product market that prevent the optimal allocation of workers among sectors generally lead to output and welfare losses. Regulatory failures that lead to higher informality may have a direct impact on productivity. But beyond this, informality itself has been postulated to have adverse impacts on productivity. As noted in *Poverty Reduction and Growth: Virtuous and Vicious Circles* (Perry et al. 2006), workers uninsured against health, old age, and other risks may have lower productivity and fewer incentives to invest in human capital accumulation. Firms unable to access credit, larger sales/product markets, and sources of innovation, and those evading taxes may operate at a suboptimal scale. Competition with noncomplying firms leads to productivity losses at formal firms. At the aggregate level, a large concentration of workers in small firms rather than larger firms may lead to lower productivity growth.

Erosion of the functioning and legitimacy of market- and equity-enhancing institutions

Noncompliance with tax collection and market-supporting regulation erodes the rule of law and the integrity of public institutions, and limits society’s ability to address collective needs that range from infrastructure to the mitigation of inequality. Noncompliance may become a social norm that increases the costs of enforcing the law, undermines the legitimacy of societal institutions, and creates horizontal and vertical inequities (with better-off insiders and worse-off outsiders). This said, compliance with legal norms may be endogenous to the perceptions of the current effectiveness of public institutions and, more profoundly, to the nature of the underlying “social contract.”

Indicators of other problems

A sizable body of literature sees informality as arising from poor regulation or other government failures. To the degree that this is the case, unusually high or increasing informality may be suggestive of poor policy regimes.

Informality and the relationship between the individual and the state

Implicit in each of the examples above is a relationship between the individual or firm and the state. Economic theory posits a legitimate role of the state in a number of areas. The state redresses coordination failures in the provision of public goods (for example, roads, defense, public

security) and in the prevention of social bads (such as pollution). Further, it fills in missing markets—establishing courts, property rights, risk-pooling mechanisms—and sets the rules in the modern economy. Finally, it concerns itself with distributional issues and power asymmetries—redistributing from rich to poor, ensuring that labor–capital relationships are not too one-sided or that no firm or group of firms gains too much economic power, and that all citizens receive equal treatment regarding the provision of key services. To redress the market failures identified above, the state necessarily requires the power to monitor and coerce agents to do things that, privately, they would not do. This view of the state has led to seeing the informal sector through a lens emphasizing lack of compliance with legal norms. Though this is not the only lens (Hart defines informality simply as “undocumented,” an important dimension taken up in chapter 8), it enjoys currency, particularly in the economics field, and will be a central organizing theme of this report.

However, even that fairly narrow definition raises the question of why agents are not in compliance with state norms. Among the many lenses through which this question has been viewed, one of the most influential lenses on the labor and firm side has focused on their exclusion from critical state benefits and, concomitantly, the circuits of the modern economy. However, this report highlights a second lens through which to view informality that is more akin to Hirschman’s (1970) “exit”: many workers, firms, and families choose their optimal levels of engagement with the mandates and institutions of the state on the basis of their valuation of the net benefits associated with formality and the enforcement effort and capability of the state. That is, they make implicit cost–benefit analyses about whether to cross the relevant margin into formality and frequently decide against it. Under this view, high informality results from a massive opting out of formal institutions by firms and individuals, and offers an indictment of the state’s regulations and services and of its enforcement capability.

As a starting point, it is useful to sketch three types of relationships between the individual and the state and, more generally, the institutions of civil society, that capture these two dimensions.

Opportunistic evasion

In the case closest to economic theory’s vision of the state, the informal sector is seen as *evading* legal norms that give rise to additional adjectives: *tax-evading* by those focusing

on lack of compliance with revenue-raising norms, *illegal* when the sector engages in unsanctioned activities, *unfairly competitive* by those focusing on how industrial structure is affected by such evasion (Capp, Elstrodt, and Jones 2005), *unprotected* by those thinking about why workers in least-developed countries (LDCs) are not covered by labor legislation (International Labour Organization [ILO]), and *subcontracted* by those concerned with the potentially exploitative dynamics of globalization (Castells and Portes 1989). Each of those descriptors can be seen as evasion, broadly construed, of the state’s legitimate and efficiently executed brief.² Opportunistic evasion is, of course, the primal form of “opting out,” despite the fact that “voice” through the political system may be perfectly adequate.

Many of the cases above have an exclusionary complement. Those firms avoiding labor legislation, for instance, may be implicitly creating a dual labor market where their employees would prefer to enjoy the full benefits of the social protection system, but find themselves, for at least the present, in inferior jobs.

Defensive evasion and exclusion: Coping with the imperfect, captured, or informal state

However, as a large body of literature has documented, the state often deviates from the economists’ ideal. Simply put, the state does its job badly—ranging from poor regulation to oppressive or exclusionary measures, forcing agents, who perhaps are inclined toward compliance under the ideal state, to cope by defensive evasion. De Soto (1989), Djankov et al. (2002), Friedman et al. (2000), Loayza, Servén, and Oviedo (2005), and Schneider (2005), among many others, have stressed the very high registration costs, the regulatory burden to becoming formal, as well as the high ongoing costs of fully integrating with the state that drive firms to stay off the state’s radar.

The postulated reasons for this state deviation from the ideal range widely both in view of the nature of the state and in implications for policy. A large body of literature stresses that the bureaucracy may be populated by rent seekers and, in principle, defensive evasion in this case could be largely alleviated by regulatory reform. However, in more extreme views (Acemoglu, Johnson, and Robinson 2001; North, Wallis, and Weingast 2005), the state is behaving in a deliberately and coherently *exclusionary* manner, manifesting an underlying stable political-economy equilibrium where incumbent business and labor elites defend their rents and will find ways to offset and nullify any tinkering

with the costs of doing business. In this spirit, the informal firm, as depicted by de Soto (1989), is excluded from the benefits of the state and hobbled in its participation in the market economy; the informal worker is excluded from the benefits enjoyed by a privileged caste of workers. Further, in the absence of a major shock to the political-economy equilibrium, they are permanently so.

Other, more generous views, (for example, Centeno and Portes 2003) see weak Latin American states as assigning themselves an unmanageable—and usually unenforceable—load of regulatory measures. That is, what we see is less a conspiracy to exclude than overwhelmed and poorly coordinated bureaucracies. However, the exclusionary views do touch on a leitmotif in the political science literature that stresses the informality of Latin American *political systems*. In particular, this literature studies the divergence between the formal structures of democracy and the economists' ideal bureaucracy, on one hand, and how governance is really done, on the other. O'Donnell (1996) argues that often behind formal elections and alternation of power lies particularistic access to the state with roots in century-old traditions of patron–client relations. Everyone “understands the model” of particularistic access, distrust of the state and its evenhandedness is high, tax morale³ and the general feeling of social reciprocity are low. Further, as chapter 7 will discuss, the state is perceived as providing little: relatively few citizens are covered by what has been called the *truncated welfare state*—low quality and coverage of public social services, such as health care or education, further erode tax morale and prompt opting out of the system of taxes and transfers. Hence, a Latin American citizen weighing working with a state that diverges substantially from the ideal, or employing other “nonformal” ways of solving social problems and market failures, may not perceive the informal-formal dichotomy as quite so sharp.

At a meta level, it may be argued that the underperformance of Latin American states along these dimensions partly reveals poorly resolved social tensions and manifests what we might use as conceptual shorthand—a dysfunctional underlying “social contract.” Beyond high informality, this can also be seen in the inability of the state to redress the long-standing high inequality, in the weak rule of law, or in the recurrent bouts of macroeconomic instability. Difficult as these phenomena have been to manage in the region, the report is generally optimistic about the possibilities of improving the quality of the state and the

design of its policies and, hence, substantially reducing the distortions that both exclude and encourage exit.

Passive evasion and state irrelevance

This discussion of the limits of the Latin American state brings us back to Hart's (1973) emphasis on how multiple institutional systems coexist within a polity and that the state is only one candidate among many. This may particularly be the case for very rudimentary microenterprises that may not consider themselves part of the modern economy/social order, and whose production requires little in the way of services from the largely irrelevant state. Such firms are described colorfully in Geertz's (1963) seminal *Peddlers and Princes*, which traces the social evolution from the bazaar economy in Indonesia to that of the more rational, modern “firm.” The premodern or bazaar economy encompasses a vast number of proto-firms that are not constrained by access to the benefits that normally are associated with formality but, as Hart stresses, operate within subsystems of institutions that coexist with, substitute for, or compete with the “formal” state institutions. In fact, what is striking in Geertz's description of two Indonesian towns is the significant discussion of institutions for managing credit, risk, and collective issues of irrigation, but very little about the state.

Here we find a conceptual kinship with the literature on social capital that deals with “informal” relations of trust, reciprocity, and the like that exist in the absence of formal institutions (see Alesina and La Ferrara 2000; Glaeser, Laibson, and Sacerdote 2000; Greif 1993, 1998; and Stiglitz 2000, among others). While, generally, such relations are considered positively—that is, as ways of solving problems of contracting and market failure—they may, in addition, be preferable in some dimensions to formal institutions that may eventually displace them. Local institutions are likely to be more closely tailored to the relevant market failure and less subject to moral hazard due to closer monitoring by family or village. As Bentolila and Ichino (2000) argue, the informal safety nets in southern Europe cushion families against employment shocks better than the formal unemployment insurance schemes of northern Europe. Studying financing in Chicago migrant communities, Bond and Townsend (1996) conclude, “We are inclined to view the small role played by the formal sector as stemming, at least in part, from community disinterest as opposed to formal sector negligence” (p. 24) due to the insufficient flexibility of formal institutions. With some caution, they suggest that

“formal sector institutions attempt to create more flexible financial instruments by either using or mimicking existing informal and semi-formal structures” (p. 24). In sum, informal institutions cannot be ruled out *ex ante* as suboptimal, given the type of enterprises operating within them and the level of development of the state.

The demand for formal institutions increases with the sophistication of the firm and, more generally, of society. Geertz’s peddlers become organized firms whose growth will require access to an increasingly sophisticated set of socially provided inputs. In line with de Soto’s anecdote of the Peruvian street vendors who sought to pay taxes so they would be granted *de facto* property rights over their pitches, participation in formal institutions can be seen as a “normal” input, increasing with firm size or sophistication (see Levenson and Maloney 1998). This is entirely consistent with the logic, postulated in the social capital literature, that individuals optimize their investment in informal networks with a view toward long-run returns, except that here the “networks” include formal institutions.

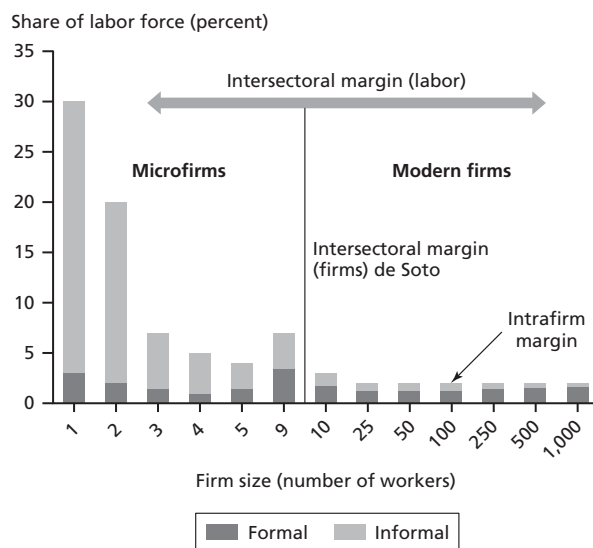
At the economywide level, the reach and density of formal institutions are almost certainly endogenous to the complexity of the society, and, as Stiglitz (2000) suggests, we may expect a greater need and density for formal institutions with development—a natural evolution from informal to more formal institutions.⁴ At each point along the course of this process, the state makes its own cost–benefit analysis on what size firm is worth monitoring and taxing to finance its mandate, and what size is not worth doing so, thus leaving the institutional space free for the kind of institutions discussed above. Tandler (2002), in fact, describes a “devil’s deal” in Brazil between the state and the informal sector implicitly on the boundaries of this space. More generally, there is an equilibrium where small firms find nonstate solutions to their needs, and the state occupies itself with firms above its enforcement threshold. In this view, informality is neither cause nor result of underdevelopment, nor is it necessarily pathological; rather it is a normal phase in the development process—a lack of formalization of enterprises and the dominance of local institutional systems that complement, compete with, or substitute for those of the state. It is not so much “exiting” as never really “entering.”

Three margins of informality

In all likelihood, all these agent–state interactions can be found among the phenomena lumped under what is clearly a very heterogeneous informal sector (see Cunningham

FIGURE 1.1

Margins of informality



Source: Maloney 2006.

and Maloney 2001; Fields 1990; and Henley, Arabsheibani, and Carneiro 2006). But policy demands that we identify the most germane interactions, and this, in turn, requires identifying the critical margins along which individuals and firms are making calculations about or facing constraints to becoming formal. Maloney (2006) sketches three such margins and suggest what types of agent–state interactions are at play (figure 1.1): (1) the *intrafirm margin* where firms are partly formal and partly not, (2) the *intersectoral margin* between informal and formal firms, and (3) the *intersectoral margin* of formal and informal workers operating through the labor market. These margins are not exhaustive, nor are they unrelated to each other. However, they do capture much of the relevant activity covered in the informality discourse and help us isolate the most relevant areas on which to focus.

The intrafirm margin

Firms across the size spectrum are often partially informal across several dimensions. Underreporting of sales is common globally. As chapter 6 will show, the investment climate surveys conducted by the World Bank find that the percentage of sales reported by Brazilian firms ranges between 60 percent in microfirms and 80 percent in very large firms. Survey and anecdotal evidence from Latin America suggests that medium-size to large firms will commonly have a substantial share of their operations, including workers, off the books. In Argentina, roughly

15 percent of workers will receive pay partly “*en blanco*” on the books and partly “*en negro*”—off the books—without the corresponding labor taxes paid by either worker or firm. As discussed later, firms can often be fully compliant in one dimension—perhaps paying taxes—and not in others—registering workers in social security.

The intersectoral margin (firms)

The growing informal firm is on the border of registering or complying with the labor or tax laws. To formalize the classic de Soto story, we can follow Lucas (1978) and think of steady-state firm size as determined by its underlying cost structure that reflects the ability of the entrepreneur, among other characteristics. The resulting heterogeneity of costs gives us the distribution of firms across the size spectrum, whether a mom-and-pop store or a Wal-Mart or Mexican Elektra. And, as Jovanovic (1982) argues, entrepreneurs have a rough idea of what their ability is, but only upon actually opening the business can they make that estimate precise. Some will find out they are unprofitable and quit; others will find their profits surprisingly high and seek to expand. At this point, the latter group may need the services of the state, or of collateral services that require being recognized by the state.

In the de Soto (1989) view, the costs of becoming formal are too high and firms are effectively excluded from the formal realm and forced to remain suboptimally small. Banerji and Jain (2006), following Rauch (1991), argue that lack of monitoring below a certain size threshold produces size dualism, where small firms take advantage of the wage differential, in the former case, to produce lower-quality goods for the poorer section of the consumption market.⁵

However, there are also firms whose underlying productivity is so low that they will never demand the services of the state. Going back to Geertz’s (1963) study of the bazaar economy in Indonesia, he describes entrepreneurs who, in fact, lack the organizational skills to function as a modern firm, let alone grow to a large size. Very poorly educated workers, many less than a generation away from subsistence farming, would also, on average, have low ability levels in running a firm. Alternatively, women may choose to operate as independents to better balance home and income-earning roles, with no plans to expand (Cunningham 2001). Further, we do not know how unsubstitutable formal inputs are for less formal ones.

Hence, there are two central open questions surrounding this margin. First, how many informal firms are actually

close to the margin of becoming formal? Alternatively put, how relevant is de Soto’s story in explaining informality? Second, how binding are the impediments at this margin? Chapters 7 and 8 will take up both these questions.

The intersectoral margin (workers)

The labor literature has long focused on the relationship and flows between workers in the formal sector, covered by labor legislation, and those in the informal microfirm sector who are not covered. The latter are often considered the most disadvantaged of the urban labor market, as they are precarious, often termed *subsistence*, and thought to be the rump end of the global value chain. A large informal sector has also been seen as evidence of a labor market segmented by acute formal sector rigidities arising from excessively high minimum wages or union bargaining (see, for example, Esfahani and Salehi-Isfahani 1989; Mazumdar 1976; and Rauch 1991).

However, evidence has been mounting that a sizable share of entrants into the informal sector do so because they will become better-off. The report will use the term *voluntary* to denominate entry yielding higher or equal levels of welfare. This does not imply that they are not poor or that they are happy—only that this is the better of two options, given their lower human capital and the low productivity of the economy. Returning to Lucas (1978), we may argue that, in fact, at low levels of aggregate productivity, the opportunity cost of becoming self-employed is such that more workers with a comparative advantage in operating a small business will actually do so. This may explain an important part of self-employment and microentrepreneurship in many countries in the region. The idea that entry into informality occurs for various reasons is not new. Hart (1973) never saw the informal sector as intrinsically bad, and Fields (1990) noted that there is an “upper tier” to informal employment that does very well. The critical empirical question is, What share of the sector corresponds to those who would prefer formal jobs versus those who are as well-off as they would be in the formal sector? Gregory (1986) and Maloney (1999, 2004) argued that, for Mexico, the evidence of segmentation is weak, the majority of the sector is “voluntary,” and the unprotected/exploited view of informality seems an inappropriate lens.⁶ The next three chapters confirm this finding for Mexico and the Dominican Republic, as well as for the majority of the informal self-employed workforce across the region. They also reveal that, in most countries

BOX 1.1

The ILO definition of informality

The three margins discussed in the text are fully consistent and useful for analyzing informality, both as the ILO traditionally defined it, based on what might be called the “productivity view” (rows in the table below) that focused on the type of production unit (rows); and the newer focus on *informal employment* defined according to the “social protection” or “legalistic” view by job status (columns). In the former definition, the informal sector enterprises are defined as production units operated by single individuals or households that are not constituted as separate legal entities independent of their owners and in which capital accumulation and productivity are low. This includes “family units” (those operated by nonprofessional own-account workers with or without contributing family workers) and “microenterprises” (productive units with no more than

five employees). As such, the table below shows that total employment in the informal sector includes self-employed (3); own-account workers, with or without family workers (5); microentrepreneurs (4); and their employees (6). Under this definition, understanding the logic of the production would have required focusing most on the second and third margins—how microfirms become formal and the nature of flows between those people working in such firms and those in the “modern” sector of the economy. The more recent shift to a “legal” definition of informality recognizes that “informal employment” can be found both within and outside the small-firm sector. Consequently, informal employment now includes informal contractual arrangement in firms that are otherwise formal, (1) and (2), and hence would now include the intrafirm margin.

ILO conceptual framework: informal employment

Production unit by type	Job by status in employment								
	Own-account workers		Employers		Contributing family workers	Employees		Members of producers' cooperatives	
	Informal	Formal	Informal	Formal	Informal	Informal	Formal	Informal	Formal
Formal sector enterprises					1	2			
Informal sector enterprises ^a	3		4		5	6	7	8	
Households ^b	9					10			

Source: Hussmanns 2004.

Note: Cells shaded in dark gray refer to jobs, which, by definition, do not exist in the type of production unit in question. Cells shaded in light gray refer to formal jobs. Unshaded cells represent the various types of informal jobs.

Informal employment: cells 1–6 and 8–10. Employment in the informal sector: cells 3–8.

Informal employment outside the informal sector: cells 1, 2, 9, and 10.

a. As defined by the Fifteenth International Conference of Labour Statisticians 1993 (excluding households employing paid domestic workers).

b. Households producing goods exclusively for their own final use and households employing paid domestic workers.

of the region, informal salaried workers appear to correspond more to the traditional queuing view.

Framing the informality “decision” as one occurring across three distinct margins focuses the diagnostic and policy discussion on the relevant set of individuals and considerations, and is general enough to encompass most existing frameworks as well. For instance, the ILO definition of informality, both traditional and more recent (box 1.1), is fully

compatible, with the latter spanning all three margins. Concerns with better enforcement of tax codes or “corruption,” more generally, are likely to focus primarily on the first margin. The World Bank’s Doing Business measures focus primarily on the second de Soto margin, as might those concerned with access to credit and informal microfirm productivity more generally. Traditional concerns with segmented labor markets focus primarily on the third margin.

Measuring the informal sector

The previous sections have suggested how difficult it may be to present a picture of informality, even if we had the data to see clearly the component elements—data that, almost by definition, we do not have. Which aspect are we interested in? Large firms evading taxes? Microfirms that will manage through family and community mechanisms and not bother with the state? Workers rationed out of formal jobs? Or women quitting jobs with benefits to stay with their children and work independently? This heterogeneity is clearly the manifestation of multiple social and economic phenomena that have given rise to a cacophony of characterizations and measurement attempts. The next sections discuss a subset of these measures, what they may or may not be capturing conceptually and in practice, and how they can foreshadow later discussions of the most relevant margins of informality on which policy makers should focus.

By definition, most economic activities that are classified as informal are not captured by national accounts and official statistics. One exception is informal employment, which can typically be measured or proxied using questions from household survey data on affiliation to social security, the mandated benefits workers receive, or the size of the firms they work for (in terms of the number of employees), or using a combination of those variables.

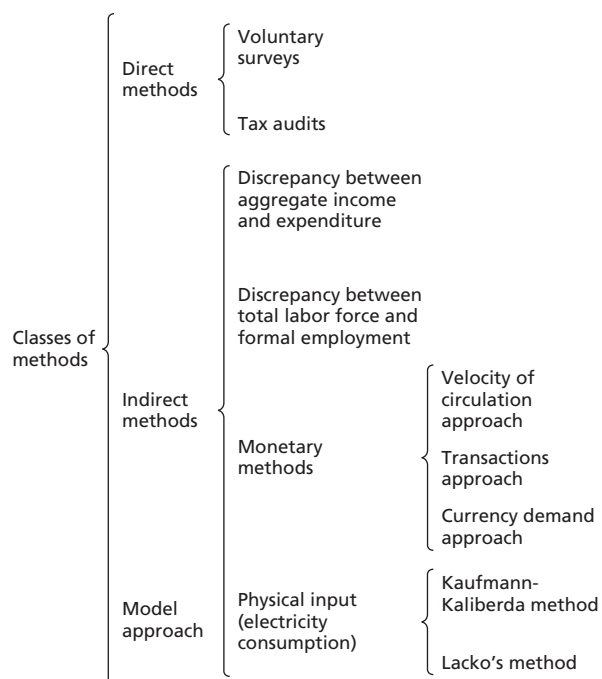
There are several methods that can be used to obtain estimates of the magnitude of the informal sector. These methods have been described in detail, and their strengths and weaknesses have been discussed extensively.⁷ These can be separated into three classes: (1) direct methods, (2) indirect methods or “indicator” approaches, and (3) the model approach (figure 1.2).

Direct approaches to measurement

The direct methods are microeconomic in nature and use either voluntary survey data or the results from tax audits to construct estimates of total economic activity and its official and unofficial (or measured and unmeasured) components. Voluntary surveys typically ask respondents to declare or reveal their incomes, labor status, or impressions of levels of tax compliance in their industry. This method has been criticized for its sensitivity to how the questions are posed, and its confidence in the respondents’ willingness to truthfully reveal their income. Tax audit-based measures define the magnitude of the informal economy as the

FIGURE 1.2

Methods for measuring the informal sector



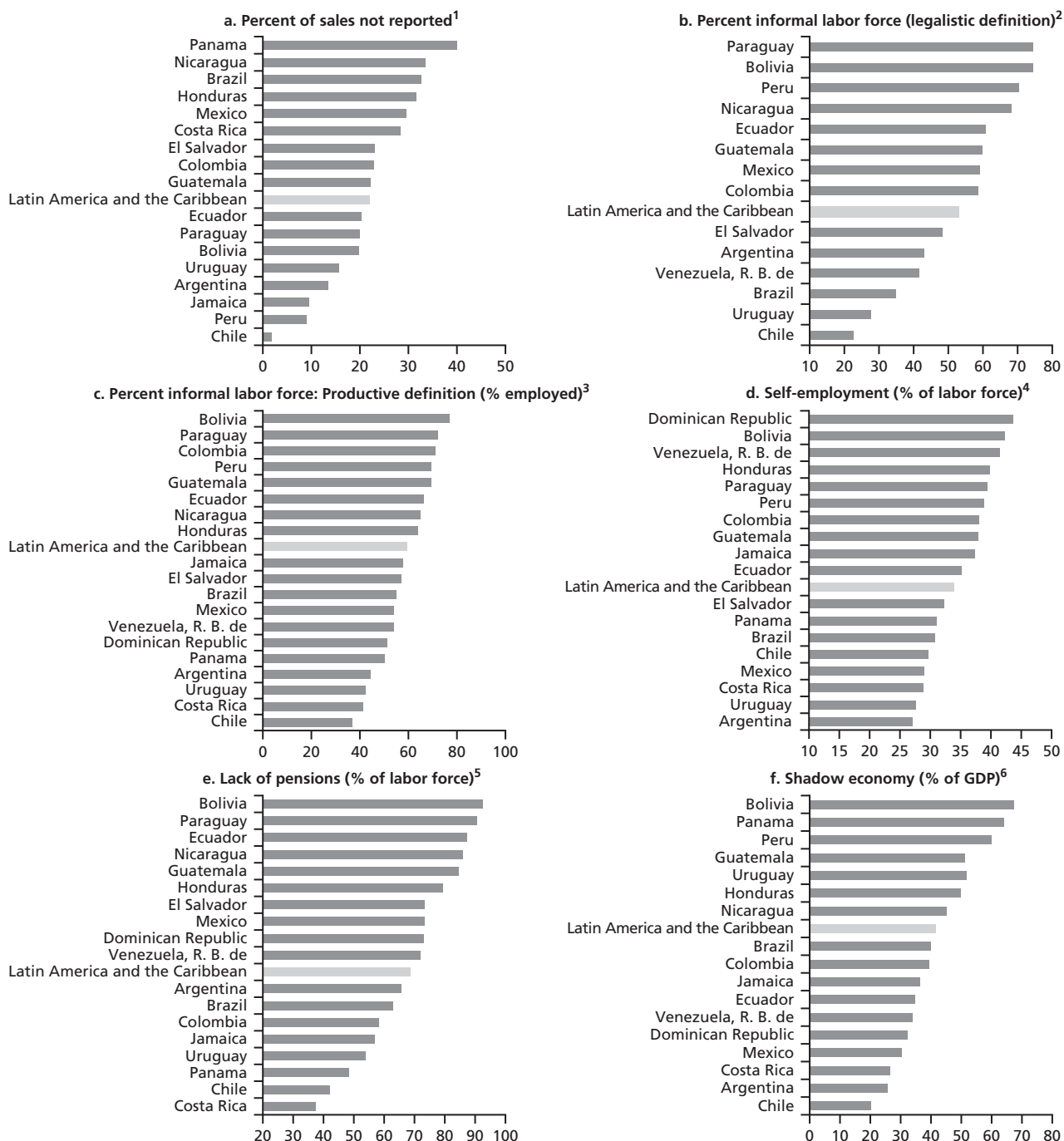
Source: García-Verdú 2007.

difference between the income declared in tax returns and the income actually found after an audit. A potential problem in extrapolating to the national economy is that audits are usually nonrandom and, hence, may not be representative. In both cases, the applications of these methods have been limited to a few developed countries because of the paucity of the available data.

For Latin America, we have three principal sources of data. One is the newly collected investment climate assessment that asks firms about the level of underreporting of income and workers. The World Bank Doing Business indicators use an analogous definition in their compilations. Figure 1.3 presents the unweighted average of these responses across firms in Latin America and the Caribbean and suggests a wide range of noncompliance, from less than 5 percent in Chile (similar to Organisation for Economic Co-operation and Development [OECD] levels) to over 40 percent in Panama.

Household and labor surveys provide the second and far more extensive source of direct data. The issue is finding the right definition of informality. As noted in box 1.1, the ILO has traditionally employed what might be called a

FIGURE 1.3

Selected measures of informality

Sources: Gasparini and Tornarolli 2006; investment climate surveys 2006; Loayza and Rigolini 2006; Schneider 2005; World Bank 2006b.

Note: 1. Informality is measured by the percentage of sales that businesses do not report for tax purposes (Investment Climate Surveys 2006).

2. "A salaried worker is informal if s(he) does not have the right to a pension linked to employment when retired" (Gasparini and Tornarolli 2006, 10).

3. "An individual is considered an informal worker if (s)he belongs to any of the following categories: (i) unskilled self-employed, (ii) salaried worker in a small private firm, (iii) zero-income worker" (Gasparini and Tornarolli 2006, 8).

4. "Self-employment is measured as the percentage of self-employed workers with respect to the total active population" (Loayza and Rigolini 2006, 15).

5. Share of the labor force not covered by a pension scheme (World Bank 2006b).

6. "The shadow economy includes all market-based legal production of goods and services that are deliberately concealed from public authorities for the following reasons: (1) to avoid payment of income, value added or other taxes, (2) to avoid payment of social security contributions, (3) to avoid having to meet certain legal labor market standards, such as minimum wages, maximum working hours, safety standards, etc., and (4) to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms" (Schneider 2005, 600).

In all cases, regional figures are unweighted averages.

“productive” definition, focused more on informal *firms*. Since these data are available on a global basis, they are commonly used in empirical work.⁸ The weighted average of 28 percent of the Latin American and Caribbean labor force is below those of Africa and South Asia, but above those of the OECD and Eastern Europe. Along these lines, the report tabulates a measure calculated by Gasparini and Tornarolli (2006) that is broadly consistent with the ILO measure for Latin America (see the annex) but adds in paid workers in those microfirms. For the purposes of this report, we will refer to the ILO definition as “self-employment” and to the Gasparini-Tornarolli definition as the “productive” definition.

A second definition, called by Saavedra and Chong (1999) the “legalistic” or “social protection” definition, focuses more on coverage of workers by mandated labor protections. It thus is more concerned with workers’ welfare per se (or perhaps with *job quality*) than with the nature of their employment; and when not including self-employed or owners, who are often not required to register with social security administrations, it captures compliance with labor laws. This is more consistent with the ILO’s (2002) more recent emphasis in its “Decent Work” report on noncompliance by either enterprises or workers with all or some of the rules and regulations in the body of national or local legislation, commercial, and/or labor legislation. This new focus implies expanding the definition to include informal contractual arrangements among otherwise formal entities (see box 1.1). Hence, this definition puts a greater emphasis on the division between *informal salaried workers* in any size firm and the *informal self-employed*; and, in fact, the report will show that there are substantial differences in the behavior of these two classes of workers that make the bifurcation of informal workers critical for analysis. Global data do not exist for this measure, but, again, Gasparini and Tornarolli (2006) have calculated consistent series for Latin America, and other authors providing background papers for this report use some close variant.

Overall, both the productive and the legalistic measures give broadly similar measures of the level of informality in Latin America and the Caribbean. However, the individuals under each measure may differ substantially.⁹ Column (ii) of table 1.1 tabulates the share of the labor force that, when classified as formal under the productive definition, is informal under the legalistic definition. In practice, it is a measure of what fraction of workers in firms with more than five employees probably legally should be covered by social

security but are not. The data suggest that evasion in larger firms is a relatively minor issue in Uruguay, for instance—around 10 percent—and quite large in Ecuador, Nicaragua, and Peru—more than 30 percent. Column (i)/(i)+(ii) captures a related measure—the percentage of workers classified as formal in the productive definition and as formal in the legalistic definition. The fit ranges from poor in Bolivia, Nicaragua, Paraguay, or Peru (at rates of 30–40 percent) to reasonably good in Brazil, Chile, Uruguay, and Colombia. The third column of the next panel suggests that workers classified as informal in small firms are, in fact, generally informal in the legalistic definition as well, with rates of overlap often greater than 90 percent.

Several points merit mention here. First, the substantial mismatch of classification of formal workers in the first panel, as captured in column (ii), reflects the different questions underlying the measures; the ILO focus on small firms in the productivity definition, by design, cannot capture evasion or coverage in large firms, while the legalistic definition is more informative in this respect. While, in the aggregate, the two measures are highly correlated, for particular questions the definition matters. For example, in 7 out of 12 countries in figure 1.4, the relative representation of women versus men in the informal sector depends on the particular measure used.

Second, informality measured along the labor dimensions in many countries is largely a small-firm phenomenon. In most countries, the share of uncovered workers in firms with more than 10 workers is a minority; and, as figure 1.5 shows for Argentina, Brazil, and Mexico, in firms of more than 10 workers, the share of workers plausibly not covered is small. Though there has been increasing informalization of the large-firm labor force in Argentina and metropolitan Brazil (and the reverse in Mexico), the relevant margins for understanding the *razón de ser* of the informal laborer seem more along the margin of small firms growing into formality, and the intrasectoral margin of worker flows among the formal and informal sectors. Hence, along this dimension of formality—that is, compliance with labor laws—the intrafirm margin with large firms in mind, while still employing an important share of informal workers in some countries, does not seem to be where the majority of the action is, overall. Understanding the decisions that employers and workers in small firms make on whether to register with the authorities becomes a central question to be taken up in later chapters.

A final direct measure capturing social protection is an index of pension coverage of the population that considers

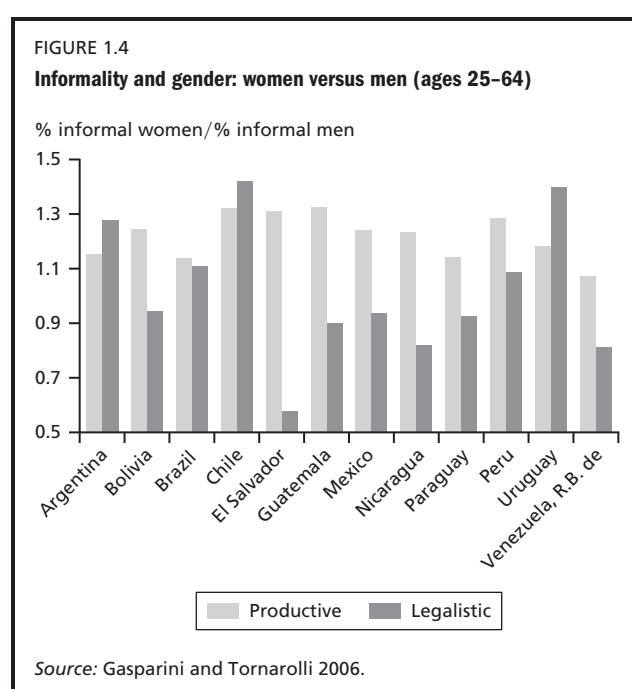
TABLE 1.1

Correspondence of the “productive” and “legalistic” definitions of informality

Sample			Formal productive			Informal productive			Total (i)+(ii)+(iii)+(iv)	(i)+(iv)
			Formal L (i)	Informal L (ii)	Formal P & L (i)/(i)+(ii)	Formal L (iii)	Informal L (iv)	Informal P & L (iv)/(iii)+(iv)		
Argentina	2004	Only salaried workers	50.6	15.6	76.5	6.1	27.8	82.1	100.0	78.4
Bolivia	2002	Only salaried workers	24.6	35.5	40.9	1.1	38.9	97.3	100.0	63.5
		All workers	7.6	15.5	33.0	0.8	76.1	98.9	100.0	83.7
Brazil	2003	Only salaried workers	53.3	10.6	83.4	11.8	24.2	67.2	100.0	77.6
		All workers	36.2	8.8	80.4	10.2	44.8	81.5	100.0	81.0
Chile	2003	Only salaried workers	67.0	11.6	85.3	10.7	10.8	50.1	100.0	77.7
		All workers	51.8	11.4	82.0	11.2	25.6	69.5	100.0	77.4
Colombia	1999	Only salaried workers	86.0	14.0	86.0				100.0	86.0
		All workers	13.6	10.6	56.1	2.8	73.0	96.3	100.0	86.6
Ecuador	1998	Only salaried workers	36.9	32.4	53.2	2.6	28.1	91.6	100.0	65.0
El Salvador	2003	Only salaried workers	49.9	20.8	70.6	1.9	27.4	93.6	100.0	77.3
		All workers	28.8	16.2	64.0	1.5	53.4	97.2	100.0	82.2
Guatemala	2002	Only salaried workers	37.8	24.1	61.0	2.3	35.7	93.9	100.0	73.5
		All workers	15.4	15.1	50.5	1.0	68.5	98.5	100.0	83.9
Mexico	2002	Only salaried workers	37.6	25.7	59.4	3.4	33.2	90.7	100.0	70.9
Nicaragua	2001	Only salaried workers	29.5	30.7	49.0	2.3	37.5	94.1	100.0	67.0
		All workers	14.9	20.4	42.2	1.5	63.3	97.7	100.0	78.1
Paraguay	2003	Only salaried workers	23.6	27.4	46.2	2.1	47.0	95.8	100.0	70.6
		All workers	10.5	17.1	38.1	1.4	71.0	98.1	100.0	81.5
Peru	2002	Only salaried workers	26.6	36.1	42.4	1.4	35.9	96.2	100.0	62.5
		All workers	11.4	21.6	34.6	1.4	65.5	97.9	100.0	77.0
Uruguay	2004	Only salaried workers	64.1	9.9	86.7	8.3	17.7	68.0	100.0	81.8
		All workers	49.3	8.4	85.5	10.0	32.3	76.3	100.0	81.6
Venezuela, R. B. de	2003	Only salaried workers	53.4	19.2	73.6	5.0	22.4	81.8	100.0	75.9

Source: Gasparini and Tornarolli 2006.

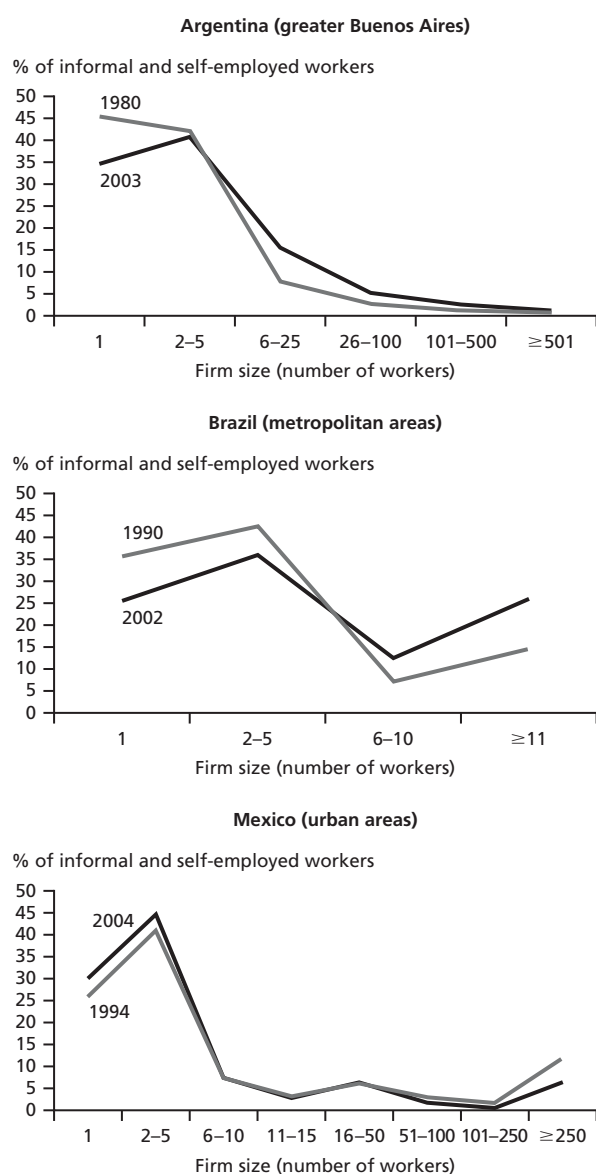
Note: Table shows what fraction of workers categorized as informal under the “productive” definition are also informal under the “legalistic” definition, and analogously for the formal. Formal L = formal using legalistic definition, informal L = informal using legalistic definition, and formal P&L = formal under both definitions.



both labor-related and universal pension schemes. This measure raises an important issue. If our concern is that families are covered by certain protections, it need not be the case that these protections are linked to the particular labor contract. Throughout the region, there has been an expansion of social protection programs that aim to provide a minimum safety net for families, regardless of labor market status. These are not captured in the pension measure, and cross-country comparisons are not yet available.

Finally, all these stock measures of informality obscure the fundamentally dynamic nature of the labor market. Chapters 2 and 4 document high flows between jobs with and without formal pension programs. Not only does this suggest a need to understand workers' choices among sectors and the constraints they face in making them, but also that the individuals in the informal population change substantially from one month to the next. Such high flows have implications for what coverage really means. As chapter 7 will show, in Mexico and Uruguay, poor workers flow

FIGURE 1.5

Informal salaried workers (legal definition) across firm size and time

Sources: Encuesta Nacional de Empleo Urbano 1994, 2004; Encuesta Permanente de Hogares 1980, 2003; Pesquisa Nacional por Amostra de Domicílios 1990, 2002.

so frequently in and out of covered jobs that, in practice, they will never accumulate enough years to gain a pension. They pay, but are de facto not covered. Further, such flows also raise the question (discussed at length in chapters 2 and 3) of what a worker's choice to be uncovered implies about the social protection and, to a lesser extent, pension definitions as measures of worker welfare. That is, if, as appears to be the case in the Dominican Republic and

Mexico, there is little evidence that either informal salaried or self-employed workers are less well-off than comparable workers in protected jobs, why should policy makers care about the particular bundle of money and benefits in which those workers are paid? The answer is complex but underscores the importance of understanding decisions across the margin of worker flows between the two sectors for the interpretation of informality measures.

The multidimensional continuum of informality

As suggested above, formality is multidimensional and continuous. At the microfirm level, there is a substantial gray area (see Tokman 1992) where firms comply with certain norms and not necessarily to the same degree. Gray areas prevail and, in the extreme, one can argue that a realistic—but one that is less easy to operationalize—description is that of a continuum from complete lack of integration with formal institutions to full compliance with all.

Compliance along any one particular dimension of formality is not discrete. For example, Robles et al. (2001) find that 16 percent of microfirms in Peru do not pay any tax, 83 percent pay some taxes, and 2 percent pay all taxes that they are required to pay. Figure 1.6 suggests that, in Mexico, firms of, say, seven workers may well have one worker registered but not all. Further, the registered–not registered dichotomy does not capture the gradations of protection that exist across labor contracts. Though not formally covered, workers nonetheless appear to benefit from some social norms of fairness. Souza and Baltar (1979) introduced the concept of the *efecto farol*, whereby the minimum wage ends up being an indexation mechanism valid also for the informal sector. In fact, Maloney and Nuñez (2001) and Cunningham (2007) show that, in many countries of Latin America, the minimum wage is most binding among the *informal* salaried, suggesting that salary norms are respected outside the realm of the official work contract. Chavez and Chacaltana (1994) show, in a study of microenterprises in Peru, that a large percentage of workers without a formal contract and without social protection benefits do enjoy vacations and the customary December bonus salary, much as dictated by labor legislation. Among larger firms, such norms may carry over to workers hired illegally (*bajo la mesa*) or subcontracted out, but on the same premises. In Argentina, however, there is an almost one-to-one correspondence between pension coverage and other labor benefits, suggesting substantial variance by country.

FIGURE 1.6

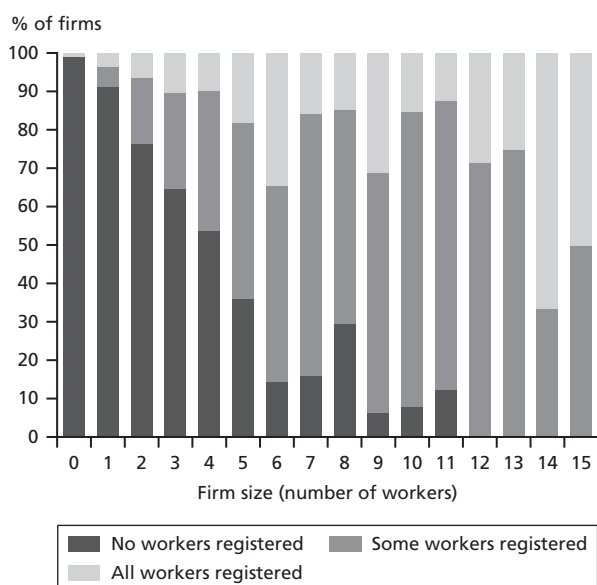
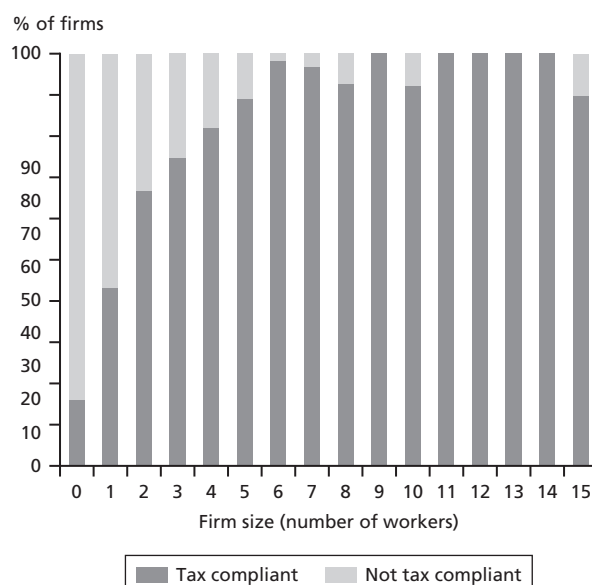
Distribution of manufacturing sector firms according to worker registration in MexicoSource: *Encuesta Nacional de Micronegocios 1992–2002*.

FIGURE 1.7

Distribution of manufacturing sector firms according to tax compliance in MexicoSource: *Encuesta Nacional de Micronegocios 1992–2002*.

The point remains, however, that many of those workers who are unprotected in large firms (identified in table 1.1) may, in fact, enjoy other elements of the standard labor contract.¹⁰

But it is also the case that the firms may be formal along one dimension, but not along another. Levenson and Maloney (1998), taking a broad view of formality as integration with not only government but also civil society more generally, find for Mexico that firms may pay taxes, but only pay labor taxes as they get larger, and engage with business associations only when they grow larger still. Figure 1.6, combined with figure 1.7, suggests that while most firms with a labor force of five workers are registered with tax authorities, compliance with labor law is far less complete. This is consistent with the findings of Robles et al. (2000) from Peru that the majority of firms pay the value-added tax, just over half pay municipal taxes, 45 percent pay income taxes, and 13 percent pay labor taxes. For larger firms, as later chapters will show, there tends to be partial formality that is relatively well synchronized across the tax and labor dimension. The logic here is clearer: larger, more frequently monitored firms will raise suspicions if they report half of their true product, but all of their workers producing it. Hence, the logic driving less easily monitored microfirms operating fuzzily

across the de Soto margin is likely to be somewhat different from that of larger firms partially evading obligations.

Indirect and modeling approaches to estimating aggregate informality

Indirect methods are macroeconomic in nature, and combine various aggregate economic variables and a set of assumptions to produce estimates of total economic activity (that is, measured and unmeasured, official and unofficial). Box 1.2 outlines several popular methods—in particular, those based on unexplained components of money demand or electricity consumption—and some of their drawbacks. By far the most common method is that of the Multiple Indicator–Multiple Cause (MIMIC) Model that imputes a level of underlying informality from a set of presumed causes of informality on one hand, and measurable consequences of it on the other. This exercise was first undertaken by Loayza (1996) for Latin America with a relatively tight, theoretically motivated set of input and outcome variables. However, recently, a more expansive estimate of the “shadow” economy as a share of gross domestic product (GDP) by Schneider and Enste (2000, 2002) has been tabulated globally and is presented in panel (f) of figure 1.3. As box 1.2 suggests, these have been

Indirect methods of estimating informality

One indirect method of estimating informality is to attribute the discrepancy between aggregate income and expenditure from the National Income and Product Accounts, which capture economic activity, to the informal sector. For this method to work, it is necessary to have measures of gross domestic product (GDP) obtained independently through the expenditure and the income approaches. Given that only one independent measure of GDP is typically available for most countries, in practice the application of this approach has been limited to a few developed countries.

Another indirect method commonly employed is the physical input (electricity consumption) approach. This method assumes that electricity consumption is the “single best physical indicator of overall [official and unofficial] economic activity” (p. 27). The method then defines the growth rate of the shadow economy as “the difference between the growth of official or measured GDP and the growth rate of electricity consumption” (p. 28). This method has been criticized on several grounds, all related to the assumption of a constant coefficient of use per unit of GDP. First, it does not consider technological progress, which reduces the amount of electricity consumption per unit of output. Second, it needs to assume a base year in which the magnitude of the informal economy is zero or negligible—an unrealistic assumption for most countries. Third, it does not consider the incorporation over time of new households to the electric grid, a fact that explains a large fraction of the increase in electricity consumption in developing countries.

A third indirect method that has also been commonly employed is the currency demand approach. This method begins by estimating a form of money demand equation in which the dependent variable is the ratio of cash holdings to current and deposit accounts ($M0/M2$). The equation controls for most known determinants of money demand; it also includes as covariates variables that are thought to be determinants of the shadow economy (for example, the tax burden). It then defines the growth rate of the shadow economy as the difference between the fitted values obtained using the estimated model and the observed values from actual data. Just like in the case of the physical input approach, this method has been criticized on several grounds. First, it assumes a common velocity of circulation of money between the official and unofficial economies. Second, transactions in the shadow

economy are assumed to occur only in cash. Third, it also assumes a base year in which the magnitude of the informal economy is zero or negligible—again, an unrealistic assumption for most countries.

Thus, both the physical input and the currency demand methods—two of the most widely used approaches—are somewhat arbitrary in the following sense. Depending on the assumption made about the base year in which the magnitude of the shadow economy is zero or negligible, one can obtain widely different estimates of the magnitude of the shadow economy.¹

The third group of methods is the model approach. The most popular among these is the Multiple Indicator–Multiple Cause (MIMIC) or structural equation model.² The MIMIC approach postulates that magnitude of the unofficial economy can be modeled as a latent or index variable. While this variable is unobservable, its causes (for example, an increase in the tax burden) and effects (such as an increase in the demand for currency) can be observed directly.

A system of equations forms the basis of this model: one set models the effects (or indicators) as a function of the latent variable; the other group models the magnitude of unofficial economy as a function of the causal variables. The parameters in this system of equations are estimated simultaneously, typically using maximum likelihood. The fitted values of the latent or index variable obtained from the reduced form equation are then used to produce an estimate of the unofficial economy.

The model approach has been criticized (see Breusch 2005) since it has been shown that its results are sensitive to transformations of the data, to the units of measurement, and to the sample used. Another criticism is that no theory is used in order to determine which variables to include as indicators or as causes. Moreover, the shadow estimates, while relying on the MIMIC Model to generate trends over time, appear to rely on traditional currency demand or the physical input methods for the initial levels, which makes it vulnerable to the criticisms of these two methods.

Notes

1. See Thomas (1993, 1999) for a more detailed description and criticism of these two methods.

2. For a detailed description and critique of these models, see Breusch (2005).

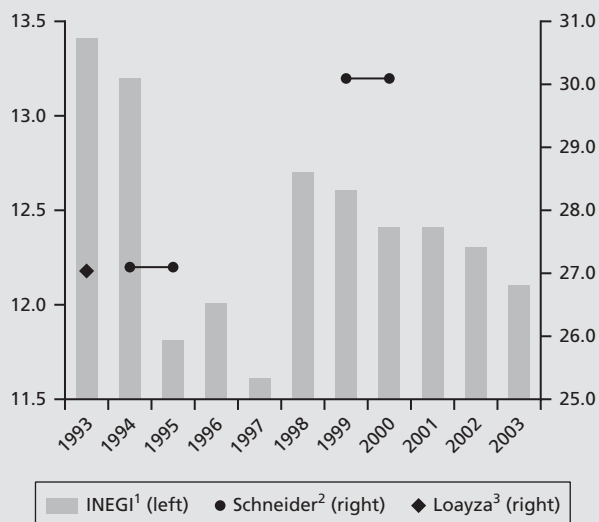
Source: García-Verdú 2007.

BOX 1.3

Schneider and Enste in the New World: Checking MIMIC estimates against Mexican data

Mexico is one of the few countries that calculate official statistics on the contribution of the informal sector to total value added. The National Statistical Institute (INEGI) employs the International Labour Organisation's official definition of the informal sector, and imputes its value added using a variety of sources. Over the period 1993–2003, the INEGI calculates that the informal sector averaged 12.4 percent of gross domestic product (GDP)—a share that remained relatively constant across the period (figure at right). This conflicts with the estimates from Schneider (2005) and Schneider and Enste (2000, 2002), which suggest that over the periods 1990–91 and 1999–2000, the shadow economy in Mexico increased from 24.1 to 30.1 percent of GDP.

Official and unofficial estimates of the informal economy, Mexico 1993–2003



Sources: 1. Cuenta Satélite del Subsector Informal de los Hogares. Sistema de Cuentas Nacionales de México. Instituto Nacional de Estadística, Geografía e informática (INEGI).
 2. Schneider (2005), using DYMIMIC and currency-demand methods.
 3. Loayza (1996), using MIMIC Method.

subject to substantial criticism for, among other reasons, their relatively atheoretical combination of different causal factors and indicators, and the difficulty in assuming that informality is the only thing linking the two. Both suggest that the shadow economy runs the risk of being a measure of an agglomeration of known size but unclear content. An exercise comparing MIMIC estimates with official estimates is presented for Mexico in box 1.3.

Correlations among measures and trends over time

As table 1.2 indicates, globally, the pensions, self-employment, and shadow measures show a modest degree of correlation generally of the expected sign. All are importantly negatively correlated with GDP. This makes certain sense since, as figure 1.8 shows, self-employment decreases sharply with development, from high levels in Latin America and the Caribbean region (60 percent in Peru) to near single-digit levels in the OECD. The close connection between self-employment and GDP per capita has already been documented by, among others, Blau (1987), Loayza and Rigolini (2006), and Maloney (2001) with a variety of

explanations, some of which will be touched on in this volume. Again, since most independent workers in Latin America are not covered by pensions, and many informal salaried workers are found working in these very small firms, we may expect lack of pension coverage to follow GDP closely as well. Globally, the shadow economy measure is moderately correlated with self-employment and pension coverage.¹¹ The limited correlation of the tax compliance measure with both GDP and the other measures is suggestive that, if it is reliable, it may be measuring a different phenomenon. As chapter 8 will show, tax evasion or elusion may be a relevant phenomenon across the firm-size spectrum and be related to levels of social norms or collective responsibility, and less to income levels. Labor informality measured as lack of compliance with legislation, however, may fundamentally be a small, low-productivity firm issue rather than a compliance issue per se.

In Latin America, the productive measure, including self-employed and all employees, and the legalistic/social protection measure are highly correlated (.8–.9) with each other and, not surprisingly, with the pension measure. It is perhaps not unexpected that all are somewhat less

TABLE 1.2

Correlations across measures of informality

All countries	Shadow economy	Self-employment	Sales nonreported	Lack of pensions	GDP pc PPP 05
Shadow economy	1				
Self-employment	0.58	1			
Sales nonreported	0.13	0.17	1		
Lack of pensions	0.60	0.81	0.43	1	
GDP per capita (PPP 05)	−0.69	−0.76	−0.30	−0.85	1

Latin America	Shadow economy	Self-employment	Sales nonreported	Lack of pensions	Informality (productive)	Informality (legalistic)	GDP pc PPP 05
Shadow economy	1						
Self-employment	0.35	1					
Sales nonreported	0.29	−0.06	1				
Lack of pensions	0.43	0.62	0.04	1			
Informality (productive)	0.60	0.70	0.11	0.80	1		
Informality (legalistic)	0.58	0.68	0.32	0.89	0.90	1	
GDP per capita (PPP 05)	−0.58	−0.75	−0.24	−0.66	−0.83	−0.75	1

Sources: Gasparini and Tornarolli 2006; investment climate surveys 2006; Loayza and Rigolini 2006; Schneider 2005; World Bank 2006b.

Note: GDP = gross domestic product; pc = per capita; PPP = purchasing power parity; PPP05 = data correspond to year 2005.

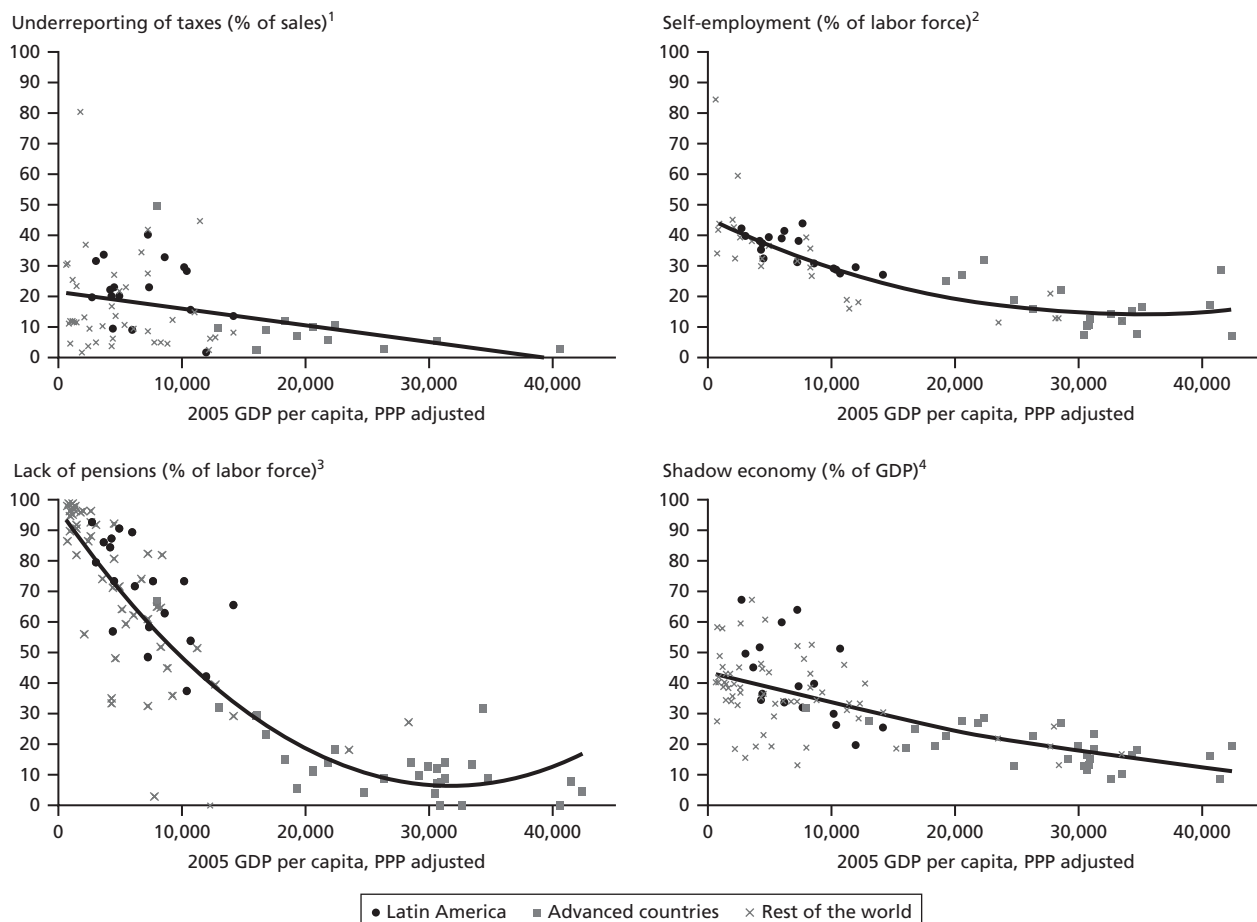
correlated with the ILO self-employment measure that omits salaried workers. The underreporting of sales measures is capturing something else and is poorly correlated with other measures. Finally, the shadow economy is not well correlated with the self-employment or underreporting of sales measure or pensions measure, but is moderately correlated with the productivity and legalistic measures. Its poor performance on the self-employment measure may be due to some odd values of the measure. For instance, Uruguay has relatively low levels of self-employment, but is among the highest in the size of its simulated shadow economy.

Trends across time

The underreporting of sales is available for only a handful of countries; however, we report the productive, legal/social protection, shadow economy, lack of pensions, and ILO self-employment for Latin America and the Caribbean where available (figure 1.9). With the exception of the self-employed definition, all measures suggest increases across the available sample period, although there are suggestive differences among them and even within individual country experiences. For instance, Argentina shows important increases in the legal and pension coverage variable, but not

in the productive measure. As will be documented in chapter 4, this arises from the fact that the increase happened through increasing informalization of large firms, not the emergence of new informal microfirms. Some anomalies—such as why Ecuador should have the largest increase in lack of coverage in pensions, but apparently little increase in legal informality—are likely due to the time spans covered. Also important to keep in mind are the subnational differences in trends. Though the legal definition for Brazil as a whole shows little increase, there has been a dramatic increase in *metropolitan* informality across the same period, and it will be explored in chapter 4. The shadow measure somewhat strangely shows global increases in informality for all countries of the world from 1990 to 2000—increases from 30 to 36 percent of GDP, an increase of 24 percent—and substantial increases in Latin America.¹² Perhaps the estimates are a bit too substantial, showing an average increase over the period 1990–2000 of 7.4 percentage points or approximately 21 percent. This is far above every other measure and is reasonable only if we assume that productivity in the informal sector rose substantially more than that in the formal sector across the period.¹³ Because of these odd results, and the theoretical concerns discussed above, the report does not rely significantly on this measure in its analysis.

FIGURE 1.8

Global correlation of measures of informality with GDP

Sources: Investment climate surveys 2006; Loayza and Rigolini 2006; Schneider 2005; World Bank 2006b.

Note: 1. Informality is measured by the percentage of sales that businesses do not report for tax purposes (Investment Climate Surveys 2006).

2. "Self-employment is measured as the percentage of self-employed workers with respect to the total active population" (Loayza and Rigolini 2006, 15).

3. Share of the labor force not covered by a pension scheme (World Bank 2006b).

4. "The shadow economy includes all market-based legal production of goods and services that are deliberately concealed from public authorities for the following reasons: (1) to avoid payment of income, value added or other taxes, (2) to avoid payment of social security contributions, (3) to avoid having to meet certain legal labor market standards, such as minimum wages, maximum working hours, safety standards, etc., and (4) to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms" (Schneider 2005, 600).

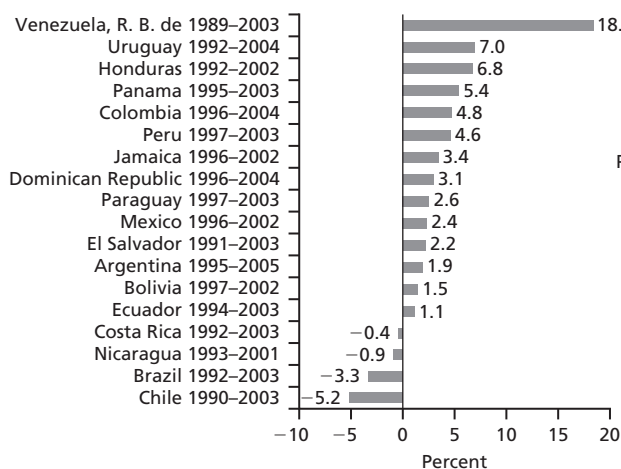
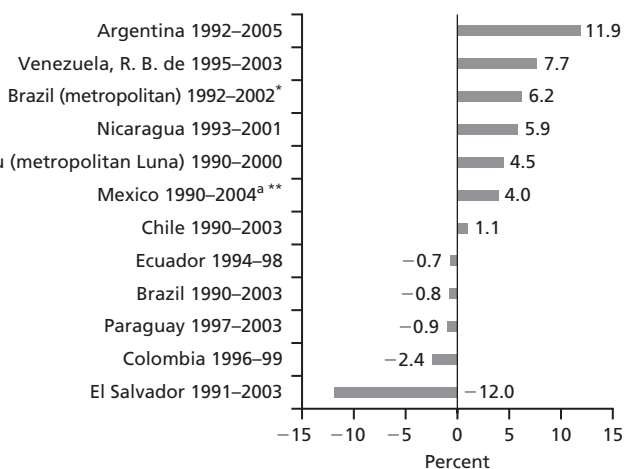
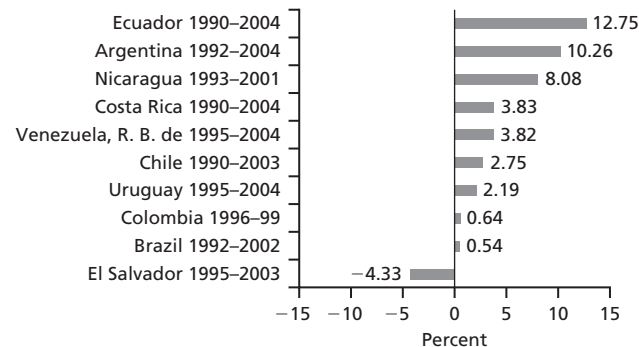
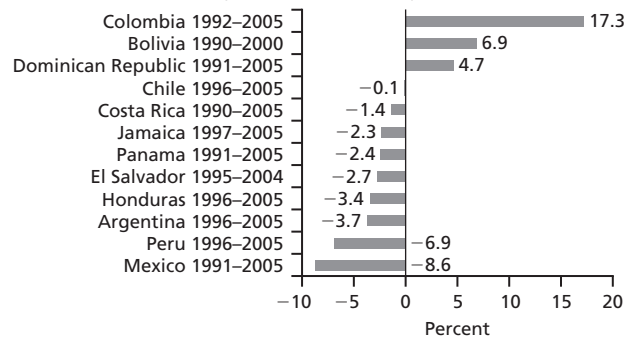
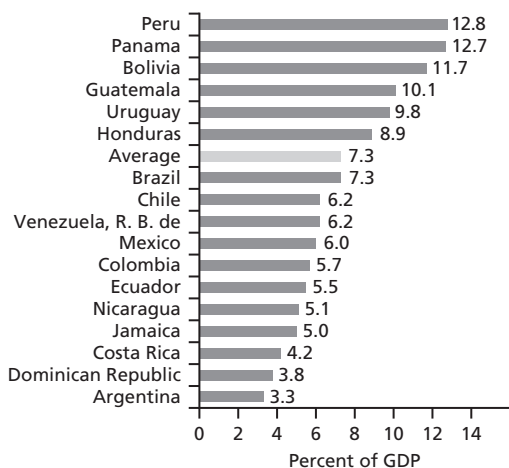
Nonetheless, the overall perception from the various measures and other sources, as well as the general "feeling" of the region, is that informality has risen. What will be dealt with in the next chapters are some of the forces driving the movements over the last decades and across which margins.

Conclusions

This chapter has sought to unpack our understanding of the term *informality*, why we may care about it, and, broadly speaking, what dynamics may be driving its elements. The

number of phenomena it encompasses and the limitations of its measures are manifold, raising some serious doubt about how useful the term really is. Yet two stylized facts remain: First, however measured, Latin America ranks high in its degree of informality. Not especially high for its level of development, but informality remains an important phenomenon. Second, several countries have experienced striking increases in informality across the last decades. Whatever adverse regulatory, poverty, growth, or social morale implications informality may have, they have become more relevant with time.

FIGURE 1.9

Trends in informality by various definitions**Informal labor force: Productive definition (% of employed workers)¹****Informal labor force: Legalistic definition (% of salaried workers)²****Lack of pensions (% of labor force)³****Self-employed: ILO (% of employed workers)⁴****Shadow economy 1990–2000 (% of GDP)⁵**

Sources: Gasparini and Tornarolli 2006; ILO 2006; Loayza and Rigolini 2006; Schneider 2005; World Bank 2006b.

Note: 1. "An individual is considered an informal worker if s/he belongs to any of the following categories: (i) unskilled self-employed, (ii) salaried worker in a small private firm, (iii) zero-income worker" (Gasparini and Tornarolli 2006, 8). 2. "A salaried worker is informal if s/he does not have the right to a pension linked to employment when retired" (Gasparini and Tornarolli 2006, 10). 3. Share of the labor force not covered by a pension scheme (World Development Indicators 2006). 4. "Self-employment is measured as the percentage of self-employed workers (employers, own account workers) and contributing family workers with respect to the employed workers" (ILO 2006, 15). 5. "The shadow economy includes all market-based legal production of goods and services that are deliberately concealed from public authorities for the following reasons: (1) to avoid payment of income, value added or other taxes, (2) to avoid payment of social security contributions, (3) to avoid having to meet certain legal labor market standards, such as minimum wages, maximum working hours, safety standards, etc., and (4) to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms" (Schneider 2005, 600).

*% of workers without *carteira* (work card). **Based on the balanced panel sample (common municipalities) for the period 1990–2004.

a. Author's calculations based on ENEU.

TABLE 1A.1

Comparisons of ILO and Gasparini-Tornarolli measures of self-employment

ILO Table: Proxied with Gasparini Data vs. ILO Data										
Country	Year	Informal sector								
		Independent workers		Domestic workers		Microfirms		Total		
		Gasparini	ILO	Gasparini	ILO	Gasparini	ILO	Gasparini	ILO	Difference
Argentina	2004	20.7	17.9	—	7.4	24.6	19.0	45.3	44.3	1.0
Bolivia	2002	47.4	44.6	—	4.3	18.2	17.8	65.6	66.7	–1.1
Brazil	2003	25.4	21.0	—	9.3	22.2	14.3	47.6	44.6	3.0
Chile	2003	21.3	21.5	—	6.2	15.7	11.1	37.0	38.8	–1.8
Colombia	2004	39.6	37.6	—	5.8	27.9	16.6	67.5	59.9	7.6
Costa Rica	2003	22.0	18.1	—	5.3	19.5	20.2	41.5	43.6	–2.1
Ecuador	2003	35.1	31.9	—	5.2	20.6	19.4	55.7	56.5	–0.8
El Salvador	2003	36.7	32.1	—	5.7	20.3	16.4	57.0	54.2	2.8
Honduras	2003	45.4	40.8	—	4.8	18.3	13.8	63.7	59.4	4.3
Mexico	2002	24.8	19.5	—	4.4	23.0	17.9	47.8	41.8	6.0
Panama	2003	34.8	24.7	—	7.1	15.4	10.7	50.2	42.5	7.7
Paraguay	2003	33.9	30.0	—	11.8	26.6	19.9	60.5	61.7	–1.2
Peru	2003	40.3	34.5	—	5.7	19.0	15.8	59.3	56.0	3.3
Rep. Dominicana	2004	39.1	32.5	—	5.7	12.2	11.3	51.3	49.5	1.8
Uruguay	2004	24.1	17.2	—	9.1	18.3	11.5	42.4	37.8	4.6
Venezuela, R. B. de	2003	39.4	33.0	—	3.0	14.7	17.6	54.1	53.6	0.5

Sources: Gasparini and Tornarolli 2006; ILO 2006.

Note: — = not available.

Annex

For most of the report, we require more detail than provided by the ILO tables. Hence, Gasparini and Tornarolli (2006) have replicated the ILO data with identical data and definitions. Although in this chapter we use measures that include informal salaried workers in our definition, we check here to be sure that the subcomponent of our measure that corresponds to self-employment is similar to that of the ILO. The right-hand column of table 1A.1 suggests that, with some exceptions—Colombia, Mexico, Panama—the two series are quite close.

Notes

1. This section draws heavily on Maloney (2006).
2. Without abandoning the central principle, a certain amount of *state-sanctioned evasion* may be optimal. Undertaking a cost–benefit analysis of monitoring and enforcement, the state may decide to leave its coverage incomplete. As an example, many countries collect no taxes below a certain level of income or have streamlined labor regulations for microfirms. In this case, the “informal” becomes simply the population that it wasn’t socially optimal to force to be formal.
3. The tax morale literature departs from the finding that, under normal estimates of individual risk aversion, the existing penalties for cheating and the probability of being caught are simply too low to explain the high rates of compliance in the advanced countries. See, for example, Graetz and Wilde (1985); Alm, McClelland, and Schulze (1992); and Frey and Feld (2002).

4. We also know that, within the formal sector, individuals choose among degrees of formal sector protection (for example, how much insurance and what kind to buy).

5. See Livingstone (1991), who argues that many of these goods and services are “appropriate” for lower-income consumers. In fact, these consumers are willing—given their preferences, information, rates of discount, and income—to eat in a cheaper, less hygienic restaurant in any Latin American capital or ride a less safe taxi or mototaxi.

6. Sethuraman (1981), in another ILO report, mentioned that there was no evidence that informal workers were en route to or queuing for a formal sector job, but that their activities were a permanent source of income. This may testify more to extreme barriers to entry than to voluntariness in the sector.

7. For two comprehensive surveys of these methods, see Thomas (1993) and Schneider and Enste (2000, 2002).

8. The version of the original ILO definition presented here considers an individual as an informal worker if she or he belongs to any of the following categories: (1) unskilled self-employed, (2) salaried worker in a small private firm, or (3) zero-income worker. It is important to note that labor market–related definitions include within the informal sector at least two types of workers, self-employed people and informal wage earners for whom the micro determinants and motivations to participate in formal or informal economic arrangement vary, as will be discussed in chapter 2.

9. For a discussion of Brazil, see Henley, Arabsheibani, and Carneiro (2006).

10. The gray area of informality may also exist in the public sector where rigid hiring and firing laws generate the need to hire public

employees through diverse contractual arrangements, ranging from low-skilled workers who work for a subcontractor to highly skilled professionals who work permanently in public institutions but are paid as consultants. In most of these cases, these workers are formal in terms of tax compliance, but informal from the labor viewpoint.

11. The documentation is not always clear on what variables are used as causes and indicators, although, in some articles by Schneider and Enste (2000, 2002), self-employment is mentioned as a cause.

12. In particular, the estimate of the shadow economy as a share of GDP for the group of African countries is approximately the same as the estimate for Latin America and the Caribbean, despite the fact that the group's average GDP per capita is only 38 percent of the average GDP per capita in Latin America and the Caribbean. In the same way, the group of countries in Eastern Europe and Central Asia has a significantly lower estimate of the shadow economy as a share of GDP than does Latin America and the Caribbean, despite the fact that the group's average GDP per capita is only 9 percent higher than the average GDP per capita in Latin America and the Caribbean.

13. For operational purposes, the ILO defines informal employment as consisting of self-employed or own-account workers (excluding administrative workers, professionals, and technicians), unpaid family workers, and employers and employees working in establishments with less than 5 or 10 persons employed, and excludes paid domestic workers. For Latin America, ILO self-employment suggests that the informality rate increased from 42.8 to 47.4, a rise of 11 percent. Using the productive definition of the ILO (2002), Gasparini and Tornarolli (2006) and Rofman and Luchetti (2006) suggest that labor-defined informality has indeed increased over time in the region. Gasparini and Tornarolli argue that informality increased in 13 countries, fell in 2, and remained constant in 3.

References

- Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. "The Colonial Origins of Development." *American Economic Review* 9 (5): 1369–98.
- Alesina, A., and E. La Ferrara. 2000. "Participation in Heterogeneous Communities." *Quarterly Journal of Economics* 114 (3): 847–904.
- Alm, J., G. McClelland, and W. D. Schulze. 1992. "Why Do People Pay Taxes?" *Journal of Public Economics* 48: 21–48.
- Banerji, A., and S. Jain. 2006. "Quality Dualism." Photocopy. Department of Economics, University of Virginia, Charlottesville.
- Bentolila, S., and A. Ichino. 2000. "Unemployment and Consumption: Are Job Losses Less Painful near the Mediterranean?" Discussion Paper 2539, Centre for Economic Policy Research, London.
- Blau, D. 1987. "A Time-Series Analysis of Self-Employment in the United States." *Journal of Political Economy* 95 (3): 445–67.
- Bond, P., and R. Townsend. 1996. "Formal and Informal Financing in a Chicago Ethnic Neighborhood." Federal Reserve Bank of Chicago *Economic Perspectives* 3–27.
- Breusch, T. 2005. "Estimating the Underground Economy using MIMIC Models." Photocopy. School of Economics, Faculty of Economics and Commerce, Australian National University, Canberra.
- Capp, J., H. Elstrodt, and W. Jones, Jr. 2005. "Reining In Brazil's Informal Economy." *McKinsey Quarterly*; available at <http://www.mckinseyquarterly.com>.
- Castells, M., and A. Portes. 1989. "World Underneath: The Origins, Dynamics and Effects of the Informal Economy." In *The Informal Economy: Studies in Advanced and Less Developed Economies*, ed. A. Portes, M. Castells, and L. A. Benton, 11–37. Baltimore: Johns Hopkins University Press.
- Centeno, M. A., and A. Portes. 2003. "The Informal Economy in the Shadow of the State." Photocopy. Princeton University, NJ.
- Chavez, E., and J. Chacaltana. 1994. "Cómo se Financian las Microempresas y el Agro." CEDEP, Ed. Stilo Novo SRL, Lima, Peru.
- Cunningham, W. 2001. "Breadwinner Versus Caregiver: Labor Force Participation and Sectoral Choice over the Mexican Business Cycle." In *The Economics of Gender in Mexico: World, Family, State, and the Market*, ed. E. G. Katz and M. C. Correia, 85–132. Washington, DC: World Bank.
- . 2007. "Minimum Wages and Social Policies: Lessons from Developing Countries." World Bank, Washington, DC.
- Cunningham, W., and W. F. Maloney. 2001. "Heterogeneity among Mexico's Microenterprises: An Application of Factor and Cluster Analysis." *Economic Development and Cultural Change* 50 (1): 131–56.
- de Ferranti, D., G. Perry, I. Gill, L. Servén, F. Ferreira, W. F. Maloney, and M. Rama. 2000. *Securing Our Future in the Global Economy*. Washington, DC: World Bank.
- de Soto, H. 1989. *The Other Path*. New York: Basic Books.
- Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1): 1–37.
- Esfahani, H., and D. Salehi-Isfahani. 1989. "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism." *Economic Journal* 99 (June): 818–36.
- Fields, G. S. 1990. "Labor Market Modelling and the Urban Informal Sector: Theory and Evidence." In *The Informal Sector Revisited*. Paris: Organisation for Economic Co-operation and Development.
- Frey, B., and L. Feld. 2002. "Deterrence and Morale in Taxation: An Empirical Analysis." Working Paper 760, Center for Economic Studies and the Ifo Institute, Munich.
- Friedman, E., S. Johnson, D. Kaufmann, and P. Zoido-Lobaton. 2000. "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics* 76: 459–93.
- García-Verdú, R. 2007. "Measurement of the Shadow Economy or Shadowy Measurement?" Photocopy. World Bank, Washington, DC.
- Gasparini, L., and L. Tornarolli. 2006. "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata." Photocopy. World Bank, Washington, DC.
- Geertz, C. 1963. *Peddlers and Princes: Social Development and Economic Change in Two Indonesian Towns*. Chicago: University of Chicago Press.
- Glaeser, E. L., D. Laibson, and B. Sacerdote. 2000. "The Economic Approach to Social Capital." Working Paper 7728, National Bureau of Economic Research, Cambridge, MA.
- Graetz, M., and L. Wilde. 1985. "The Economics of Tax Compliance: Facts and Fantasy." *National Tax Journal* 38: 355–63.

- Gregory, P. 1986. *The Myth of Market Failure: Employment and the Labor Market in Mexico*. Baltimore: Johns Hopkins University Press.
- Greif, A. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review* 83: 525–48.
- . 1998. "The Historical and Comparative Institutional Analysis." *American Economic Review Papers and Proceedings* 83: 80–84.
- Hart, K. 1973. "Informal Income Opportunities and Urban Employment in Ghana." *Journal of Modern African Studies* 11: 61–89.
- Henley, A., G. R. Arabsheibani, and F. G. Carneiro. 2006. "On Defining and Measuring the Informal Sector." Policy Research Working Paper 3866, World Bank, Washington, DC.
- Hirschman, A. O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Husmanns, R. 2004. "Defining and Measuring Informal Employment." Bureau of Statistics, International Labour Organization, Geneva.
- ILO (International Labour Organization). 2002. "Decent Work and the Informal Economy." Report VI for the 90th International Labour Conference, Geneva.
- . 2006. LABORSTA (labor statistics). ILO. Available at <http://laborsta.ilo.org>.
- INEGI (Instituto Nacional de Estadística, Geografía e Informática). 2004. Cuenta Satélite del Subsector Informal de los Hogares 1998–2003. Cuentas por Sectores Institucionales, Sistema de Cuentas Nacionales de México.
- Jovanovic, B. 1982. "Selection and Evolution of Industry." *Econometrica* 50 (3): 649–70.
- Levenson, A., and W. Maloney. 1998. "The Informal Sector, Firm Dynamics and Institutional Participation." Policy Research Working Paper, 1988, World Bank, Washington, DC.
- Livingstone, I. 1991. "A Reassessment of Kenya's Rural and Urban Informal Sector." *World Development* 19 (6): 651–70.
- Loayza, N. 1996. "The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129–62.
- Loayza, N., and J. Rigolini. 2006. "Informality Trends and Cycles." Policy Research Working Paper 4078, World Bank, Washington, DC.
- Loayza, N., L. Servén, and A. M. Oviedo. 2005. "The Impact of Regulation on Growth and Informality—Cross-Country Evidence." World Bank Policy Research Working Paper, World Bank, Washington, DC.
- Lucas, R. E., Jr. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (2): 508–23.
- Maloney, W. F. 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico." *World Bank Economic Review* 13: 275–302.
- . 2001. "Self-employment and Labor Turnover in Developing Countries: Cross-country Evidence." In *World Bank Economists' Forum*, ed. S. Devarajan, F. H. Rogers, and L. Squire. Washington, DC: World Bank.
- . 2004. "Informality Revisited." *World Development* 32 (7): 1159–78.
- . 2006. "Informality Deconstructed." Photocopy. World Bank, Washington, DC.
- Maloney, W. F., and J. N. Nuñez. 2001. "Measuring the Impact of Minimum Wages: Evidence from Latin America." Policy Research Working Paper 2597, World Bank, Washington, DC.
- Mazumdar, D. 1976. "The Urban Informal Sector." *World Development* 4: 655–79.
- North, D., J. Wallis, and B. Weingast. 2005. "The Natural State: The Political-Economy of Non-Development." Burke Center for International Relations, University of California, Los Angeles.
- O'Donnell, G. 1996. "Illusions about Consolidation." *Journal of Democracy* 7 (2): 34–51.
- Perry, G., O. Arias, H. López, W. Maloney, and L. Servén. 2006. *Poverty Reduction and Growth: Virtuous and Vicious Circles*. Washington, DC: World Bank.
- Portes, A., M. Castells, and L. A. Benton, eds. 1989. *The Informal Economy: Studies in Advanced and Less Developed Economies*. Baltimore: Johns Hopkins University Press.
- Rauch, J. E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35: 33–47.
- Robles, M., M. Torero, J. Saavedra, N. Valdivia, and J. Chacaltana. 2000. "Estrategias y Racionalidad de la Pequeña Empresas." ILO, Lima.
- Rofman, R., and L. Luchetti. 2006. "Pension Systems in Latin America: Concepts and Measurements of Coverage." Social Protection Discussion Paper 0616, World Bank, Washington, DC.
- Saavedra, J., and A. Chong. 1999. "Structural Reforms, Institutions and Earnings: Evidence from the Formal and Informal Sectors in Urban Peru." *Journal of Development Studies* 35 (4): 95–116.
- Schneider, F. 2005. "Shadow Economies around the World: What Do We Really Know?" *European Journal of Political Economy* 21 (3): 598–642.
- Schneider, F., and D. H. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38 (1): 77–114.
- . 2002. *The Shadow Economy: An International Survey*. Cambridge, UK: Cambridge University Press.
- Sethuraman, S. V. 1981. "The Urban Informal Sector in Developing Countries: Employment, Poverty and Environment." International Labour Organization, Geneva.
- Souza, P., and P. Baltar. 1979. "Salário mínimo e taxa de salários no Brasil." *Pesquisa e Planejamento Econômico* 9: 629–60.
- Stiglitz, J. E. 2000. "Formal and Informal Institutions." In *Social Capital: A Multifaceted Perspective*, ed. P. Dasgupta and I. Serageldin. Washington, DC: World Bank.
- Tendler, J. 2002. "Small Firms, the Informal Sector, and the Devil's Deal." *Institute of Development Studies Bulletin* 33 (3): 98–104.
- Thomas, J. J. 1993. *Informal Economic Activity*. Ann Arbor: University of Michigan Press.
- . 1999. "Quantifying the Black Economy: 'Measurement Without Theory' Yet Again?" *Economic Journal* 109 (456): 381–89.
- Tokman, V. 1992. "The Informal Sector in Latin America: From Underground to Legality." In *Beyond Regulation: The Informal Economy in Latin America*, ed. V. Tokman. Boulder, CO: Lynne Rienner.
- World Bank. 2006a. *Doing Business Indicators*. Available at <http://www.doingbusiness.org>.
- . 2006b. *World Development Indicators*. Available at <http://www.worldbank.org>.

CHAPTER 2

The *Razón de Ser* of the Informal Worker

SUMMARY: This chapter brings to bear a variety of conceptual and empirical approaches to flesh out the rationale (razón de ser) for the existence of the informal worker. The chapter argues that the traditional segmentation/exclusion view as the sole explanation of informality is seriously incomplete and offers evidence of a large voluntary entrance into informal work. Two types of informal employment emerge: one component corresponds to workers excluded from more desirable formal jobs, and the other to those driven by voluntary motives. The majority of independent or self-employed workers are voluntary, attaching significant value to the nonpecuniary benefits of autonomous work and choosing to “exit” formal social protection systems. In contrast, the majority of informal salaried workers appear to be excluded from more desirable jobs both as formal salaried and independent workers, although voluntary motives are significant for many workers (for example, youth) and predominant in some country contexts (such as the Dominican Republic and Mexico). In both cases there is substantial heterogeneity of motives and demographic characteristics with gradients and gray areas that cloud universal conclusions about what determines the relative size and welfare implications of each group.

CHAPTER 1 SUGGESTED THAT INFORMALITY in the labor market spans all three possible margins: firms of all sizes contracting some part of their workforce without mandated labor benefits; owners of small firms contemplating registering their workers with the labor ministries; and informal and formal workers weighing the advantages and disadvantages of jobs in the two sectors. What determines whether workers are formal or informal at each of these margins has been the focus of intense debate over the last four decades and the policy implications critically depend on the diagnosis. This and the next three chapters bring to bear a variety of analytical tools—motivational and self-rated welfare responses, labor market transitions, earnings comparisons, and labor force dynamics across the business cycle—to shed light on the significant drivers of decisions at these margins. Although a crisp, stylized image of the sector would be desirable, the picture that emerges is a complex one of a heterogeneous sector containing workers

with a range of motivations and constraints that spans the traditional view of workers excluded from the formal sector, as well as those stressing the voluntary “exit” of many workers from state-sponsored labor and social protection systems.

Informal work: Adding exit to exclusion

As noted in chapter 1, much of the debate over the nature and implications of a large informal sector can be traced broadly to two views of whether workers are driven to or pushed toward informal work. In the “exclusion” view, informal workers, either self-employed or salaried, are the disadvantaged class of a segmented labor market arising from economic dualism and institutional rigidities. Workers would prefer the presumed higher wages and benefits of formal work, but are rationed out. In contrast, the integrated view emphasizes the voluntary character of informal employment. Workers may choose informal jobs, given their preferences, skills, other means of social protection,

and their valuation of costs and the characteristics of informal and formal employment. The main rationale behind each view is discussed below. Since the traditional exclusion view is better known, there is a more detailed discussion of the integrated view.

The exclusion view: Segmentation and beyond

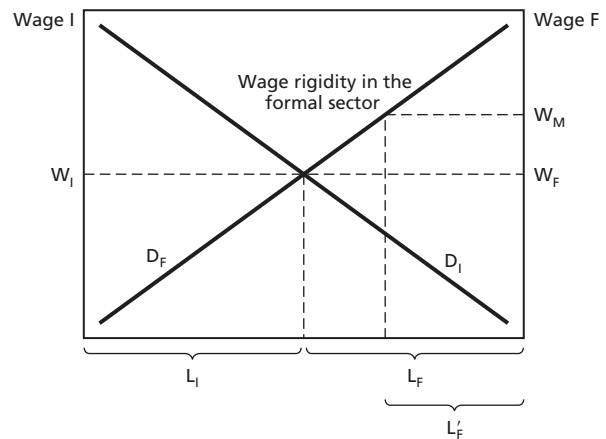
An extensive body of literature with its roots in Harris and Todaro (1970) has equated the informal sector with underemployment or disguised unemployment—the disadvantaged or excluded sector of a labor market segmented by rigidities in the formal or covered sector of the economy.¹ This view has mainly focused on the second margin, or labor flows between informality and the formal salaried sector: above-market-clearing wages force workers to queue for preferred formal jobs while subsisting in the informal microfirm sector, which is characterized by little capacity for capital accumulation (thus low productivity), an absence of labor benefits, irregular work conditions, high turnover, and lower rates of remuneration. The large proportion of women in the sector with lower earnings raises additional suspicions of discrimination in the allocation of “good” jobs. The steady rise in informal salaried work in larger firms in several countries, such as Argentina (see chapter 1), has also led to focus on the within-firm margin. Related to both is the Portes, Castells, and Benton (1989) structural articulation view that informal work is a means by which large firms resort to informal labor (directly or through subcontracting) to sidestep labor regulations and unions in order to adjust to increased global competition.

Figure 2.1 depicts one mechanism of how labor market segmentation may happen. A nominal wage floor, arising from a minimum wage (W_m) or union power, will expand informal employment and drive an earnings wedge between the two sectors, with informal workers occupying inferior (lower-pay) jobs. Thus, segmentation yields two key predictions: identical workers receive higher earnings in formal jobs; and, given the opportunity to move, informal workers would happily take a formal job. This and the next two chapters document numerous findings gauging these predictions, although, as argued below and in chapter 3, informal-formal earnings gaps cannot offer unambiguous tests of segmentation.

Poorly designed regulations may also impede firms to engage in as much formality as they may wish and, as a result, generate higher informal employment (Loayza 1996). Along the de Soto margin, since elaborated by the Doing

FIGURE 2.1

Relative sector sizes and wages under a nominal wage rigidity



Source: Authors' illustration.

Business Group at the World Bank, microfirms would like to formally register their businesses and workers, but burdensome or costly registration precludes their doing so, and mandated non-wage benefits make hiring/firing too costly. Thus, a size dualism is created in which small firms are forced to operate informally. More generally, if the state makes compliance with taxes and regulations too costly, even larger firms will be forced to operate as partially informal. In general, regulations that affect the profitability of firms in competitive markets can induce size and labor market dualism (Fortin, Marceau, and Savard 1997; Rauch 1991).

However, labor market dualism can result from factors other than rigidities in labor institutions or badly designed regulations. The high levels of quitting may induce formal firms to pay higher-than-market-clearing “efficiency wages” to secure the more productive employees (see Shapiro and Stiglitz 1984; Stiglitz 2000).² Since informal, small firms are able to monitor workers’ efforts almost without cost, they do not need to pay the efficiency wage (Esfahani and Salehi-Isfahani 1989). In his interviews with hundreds of business executives, labor leaders, and human resource personnel, Bewley (1999) found that there was an aversion to cutting wages of either current employees or new hires, even during economic downturns, under the concern that this would hurt workers’ morale and their support for internalizing the firm’s needs to enhance its competitiveness. Because of information asymmetries, young workers without a track record may find formal employers reluctant to hire them and hence may queue in the informal sector

while they gain experience. Finally, in many countries, social security contributions are collected with income taxes. Even when these are collected separately, the payment of social security contributions provides public information so that firms tacitly colluding to underreport output and tax liabilities will also resort to informal hiring. That is, part of labor informality may be a result of pure evasion of income and sales taxes.

The persistence of exclusionary factors is often associated with weak institutions and state capture by both elites and organized segments of middle classes, or, in Hirschman's (1970) terms, "the lack of voice of those excluded." The implications of this view can be disturbing. First, over half of Latin America's workers, far more in some cases, are without desired social protections. Second, the increase in informality is driven by either increasingly onerous labor, tax, or business regulations or increased global competition that is causing a race to the bottom in workers' rights and protections. Both informal workers and firms lack "voice" in the political system to change the exclusionary rules and processes that pushed them to informality. The next sections and chapters of the report show evidence that "exclusion" factors are indeed important, document their negative impacts on productivity and welfare, and discuss some of the reforms necessary to mitigate them.

The integrated market view: Opting out of or "exiting from" formality

However, there is a second and equally important lens through which to view informality that is more akin to Hirschman's voluntary "exit": workers and firms make implicit cost-benefit analyses about whether to cross the margin into formality, and they frequently decide against it.³ In this lens, much of the informal sector, in fact, offers jobs that are equally valued by workers to those they could get in the formal sector. Contrary to the predictions of the exclusion view, this implies that many informal workers are equally well-off (in broad welfare terms) as in other formal jobs fit to their skills; and, being "voluntarily" informal, they can move to the formal sector but choose not to. It is imperative to highlight that this does not imply that these workers are prosperous or happy—only that they would not be better-off in the formal sector jobs that workers with similar characteristics occupy.

Again, when Keith Hart coined the term *informality*, he surely never assumed the sector was necessarily bad, and neither did several subsequent authors. Some authors have

posited the idea of a two-tier sector—one of which is voluntary and prosperous—and, in fact, recent studies for numerous Latin American countries conclude that the majority of informal self-employed workers have chosen their occupation after undertaking a cost-benefit analysis of whether to be self-employed or salaried and then whether to be formal or informal.⁴ The literature on informal salaried workers is scant to date, but the same pair of considerations may enter both sectors' calculations: comparative advantage, and, as will be argued in chapter 7, the perceived inadequacy of the benefits of formality or their high opportunity cost and alternative self-insurance mechanisms or free social protection programs. As chapters 5 and 6 will elaborate, firms also appear, more generally, to choose the degree of participation in formal institutions according to their business needs. Finally, as will be posited in chapter 8, an important leitmotif underlying both worker and firm decisions is a concern with the overall competency and legitimacy of the state.

As will be discussed in more detail in chapter 3, workers' comparative advantage is central to their occupational choice. This arises from the matching of individual abilities or tastes to the different demands and characteristics of jobs. Lucas (1978) argued that individuals choose between salaried work and self-employment, depending on whether they are relatively more talented as an entrepreneur or as a salaried employee. Rosen (1981), Heckman and Sedlacek (1985), and, recently, Carneiro, Heckman, and Vytlačil (2005) argue that a worker's comparative advantage—emerging from a variety of work-related skills and characteristics—drives him or her to choose among distinct jobs the one that best matches his or her talents and tastes. Informal and formal jobs differ by more than labor protections, and formal benefits are just one ingredient in workers' calculations. In figure 2.1, workers equilibrate utilities—not just earnings—in choosing between jobs in the two sectors.

In addition to Lucas's (1978) discussion of comparative advantage based on relative entrepreneurial skill, informal jobs may offer an entry point to the labor market for youth and unskilled middle-age workers that partially remedies deficient or obsolete skills through on-the-job training unavailable to them in formal salaried jobs. For married women, informal jobs offer flexibility to better balance work and child rearing. For some talented workers, informality may offer better prospects for upward mobility than the formal sector.⁵

In addition, as Maloney (1999, 2004) notes, it is often critically overlooked that social protections are not free. Workers pay for them, either explicitly (through contributions) or implicitly in terms of lower wages, as has been documented, for example, in Ecuador, Peru, and Colombia (Kugler and Kugler 2003; MacIsaac and Rama 1997, 2001). Drawing on the industrialized countries' experience of predominantly formal salaried labor, "decent" jobs are generally considered to be those covered by labor legislation. However, the low quality of many formal services and high administrative overhead costs may cause many workers to see mandatory contributions to benefits programs as a disadvantage of formal salaried work. As the sociologist Bryan Roberts (1991) notes, based on his interviews with informal workers in Guadalajara, Mexico: "The absence of welfare coverage is a drawback, but, on the other hand, many informants cited the deductions made for welfare as a disadvantage of formal employment, particularly since the services they received were poor" (p. 50).

This misalignment of costs and benefits can lead workers to regard their contribution as a tax and to opt out of the system. In fact, a microentrepreneur, faced with borrowing constraints to expand a business, may be reluctant to hand over current resources for an uncertain promise of an old-age payment in the distant future. At worst, it means throwing money away if pension funds are raided or inflated away to finance the fiscal deficit. But even with stable macroeconomic times, Levy (2006) finds that the substantial rates of transition of many Mexican workers between jobs with formal pension contributions and those with none would, in all likelihood, leave them with insufficient months of benefits to qualify for a pension. As chapter 7 will discuss, this situation is relevant even in countries with low levels of informal employment, such as Uruguay.

Even in the case where benefits are well aligned with costs, workers may still opt out if they prefer higher cash compensation or if an alternative exists at lower cost or is better suited for their needs. The clearest example is that often an entire family is covered by medical benefits when any one member is formally employed, so the second formal sector worker sees zero return to paying labor taxes, explicitly or implicitly. Using cross-country employment surveys for the region, Galiani and Weinschelbaum (2006) find that secondary workers are indeed more likely to be informal if the household's primary worker has a formal job. Liquidity constraints may lead very poor workers to

consider saving for their retirement unaffordable or to resort to other informal arrangements, such as investing in a microenterprise that may be sold upon retirement. Youth, in particular, tend to have a high discount rate (or myopia) and be less concerned about having jobs with pension contributions. Further, poor workers who face temporary financial hardship will curtail savings to preserve current consumption so that incentives for opting out of benefits programs would be higher during crises. The work on Chile by Barr and Packard (2000) and Packard (2002) corroborates the significance of some of these considerations. They find that participation in the government's voluntary pension scheme, a private individual account in which workers receive what they contribute, is barely around 4 percent. Thus, independent workers are choosing to be "unprotected" even by the scheme that arguably best aligns costs and benefits in the region.⁶

Moreover, informal support networks may be able to partially substitute for unemployment and health insurance or retirement funds at lower transaction costs and with greater ease of monitoring those who intend to abuse the system. Morduch (1999) reviews numerous studies that find that informal insurance has a significant ability to offset income or other shocks, although it is far from perfect. Bentalila and Ichino (2000) argue that, in Italy and Spain, household consumption fell only 20 percent as much in response to unemployment of the household head as it did in Germany, where formal credit and social protections systems are more advanced. They attribute this to the greater influence of (informal) interhousehold transfers.⁷ Some workers may find settling for informal mechanisms justified by the other benefits of informality—for instance, flexibility. This is akin to strains of the social capital literature that see individuals optimizing their investment in informal networks with a view toward long-run returns broadly conceived.

In fact, not being formally enrolled with the social security system does not preclude being *de facto* covered by other labor protections. Recent work (Cunningham 2007; Maloney and Nuñez 2002) finds that minimum wages in several Latin American countries are actually most binding among salaried workers in *informal* firms. This "lighthouse" effect, as it is called in Brazil, suggests that there are norms on level of pay that informal employers follow even if they do not register their workers with social security administrations.

Similar considerations apply to firms approaching the *de Soto* margin contemplating registering their business or

workers with the authorities. Saavedra and Chong (1999) argue that, in deciding to register, firms consider whether the increased expansion of scale and access to other institutional supports that facilitate growth makes the registration expenses worthwhile. Levenson and Maloney (1996) show that formality along a broad set of dimensions (registering with tax and labor authorities, or joining business associations) can be considered “normal goods” in the production function. Entrepreneurs with the underlying ability to expand, eventually, need and seek these inputs. This is consistent with numerous studies showing that formality increases with firm size and time in business (see chapters 5 and 6).

In essence, low productivity may be the main reason why many small firms, and their employees, remain informal. The fact that most small firms operate in the informal sector (that is, size dualism) may just be a natural result of their lower managerial ability (Amaral and Quintin 2006; de Paula and Scheinkman 2006; Straub 2005; and Levenson and Maloney 1996). The observation that size and labor market dualism go hand in hand (that is, unskilled workers tend to be employed in informal small firms) can simply reflect the efficient allocation of more productive labor and entrepreneurship (Galiani and Weinschelbaum 2006). Sorting also leads to matching of employees and employers with higher (lower) propensities to evade. Thus, part of informal salaried employment could be a result of microfirms’ sole decisions to opt out of formal institutions for reasons unrelated to the presence of labor or other regulatory distortions, and some may reflect explicit opting-out collusions between owners and employees.

Given the very uneven enforcement of legislation in the region, workers and firms are able to choose the degree of participation in benefits programs or formal institutions. Of course, those operating informally need to weigh the benefits of being informal against the expected value of any sanctions if they are caught evading (Dávila Capalleja 1994; Hibbs and Piculescu 2006). But, in fact, a share of informality could be effectively mutually tolerated; the state may not find it cost-effective or equitable to levy penalties or force compliance on small enterprises, particularly if this could lead to bankruptcy (employment without social security contributions is deemed better than unemployment). The strength of the sense of reciprocity or, alternatively, the level of tax morale—that one should pay because others pay, or conversely that evasion is rampant and considered an acceptable business practice (discussed

later in chapter 8)—also enters here. In any case, the optimal choice for firms, workers, and the state itself would not obviously be the full formal sector package for all.

In sum, workers, in making their multidimensional evaluation of sectoral choice along the third margin—as well as microfirms contemplating a jump into formality along the de Soto margin—will weigh the costs and benefits of being formal, and we cannot assume, immediately, that all who are without protections would desire to be otherwise. As Maloney (1999) argues, and chapter 3 will investigate further, this insight also compounds the difficulties of establishing segmentation, and hence the degree of involuntary informality, by the traditional method of comparing earnings adjusted for human capital. In a market with no rigidities, informal earnings should be higher to compensate workers for the lost value of benefits and whatever risk they may be facing. On the other hand, they may be lower to compensate for taxes evaded, greater independence and flexibility, or, perhaps for young workers, on-the-job training.

Avoiding the extremes in the debate

The exaggerated size of the informal sector in the region raises the stakes of the debate dramatically: if over 50 percent of Latin American labor markets are thought of as disguised unemployment, then the implied labor market and regulatory distortions are indeed large. But if informality is largely opting out, as discussed above, then the concern with segmentation born of minimum wages, excessive union power, or excessive business regulations may be overstated.

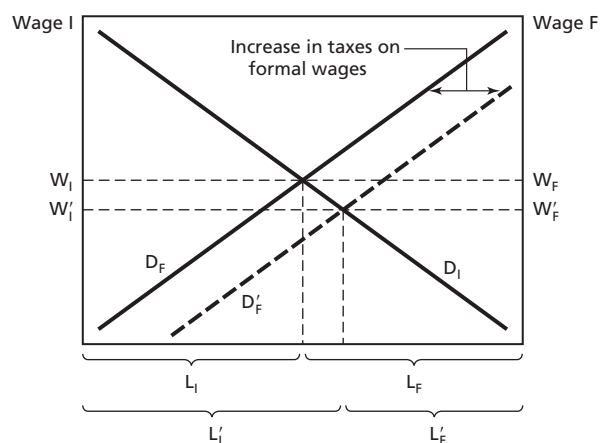
In fact, however, there is no need to be so Manichean in approaching the issue: the two views, exclusion and exit, are complementary rather than competing analytical frameworks. First, individual countries differ greatly in history, institutions, and legal framework; hence, exclusionary mechanisms may be more important in some, and exit mechanisms in others. Second, the informal sector is tremendously heterogeneous, and arguably, therefore, there is a continuum in the relative importance of exit and exclusion within countries. Third, in some cases the two can be virtually indistinguishable. A microentrepreneur concluding, through a cost–benefit analysis, that formality is not worth the exaggerated registration costs may be explicitly excluded or self-excluded, but the effect is much the same. A poor worker, excluded from health services by virtue of living in a remote, poor rural area or urban neighborhood, may see little point in being formal and paying labor taxes for services

to which he or she has no access. Finally, informality is a multidimensional concept in which agents interact with the state along some dimensions and not others, creating a large gray area between the extremes of full compliance and non-compliance. Exit and exclusion can play different roles across different dimensions: a microfirm owner unregistered due to the high costs of registration may feel de facto excluded from desired formal credit, while opting out from contributing to poorly designed state pension funds on behalf of his or her workers. The evidence presented below and across the next chapters suggests that each viewpoint has validity in some sectors, in some countries, at some points in time; both lenses can help understand and address the causes and consequences of informality in the region.

Furthermore, it is important to stress that saying that workers voluntarily choose to be unprotected does not imply that this is the best outcome for society as a whole. Ideally, we would like all families to be covered by at least the basic protections against income and health shocks, even if they, myopically, may not value, for instance, pensions, in the short run. That such protections should be linked to a characteristic of a job is not obvious, however, as will be discussed in chapter 7. Even further, saying that a worker is “voluntarily” informal means only that the worker would not be better-off in the jobs that comparable workers have in the formal sector, not that he or she is prosperous and happy, or would not prefer the kinds of formal jobs found in more advanced countries. However, it does have important ramifications for what we mean by “decent” work: if a worker is indifferent between a job with protections and one without, the jobs must, therefore, be equally decent in that worker’s judgment. Thus, the issue of increasing job quality concerns both formal and informal jobs, not just the latter.

Finally, a finding of equal levels of welfare across sectors or voluntary informality does not preclude the presence of large labor or business climate distortions. More straightforwardly, regulation-stunted growth prevents the generation of firms with potential for generating higher-paying jobs over the long run that would draw workers from the informal microfirm sector. Second, figure 2.2 shows how labor distortions—restrictions on firing, labor taxes—can shift in the demand curve for formal labor, leading to a larger share of informal workers without creating segmentation. Both arguments allow us to reconcile the findings that many informal workers are as well-off as their formal counterparts with those stressing how distortions affect the size of the informal sector (Botero et al. 2004; Djankov et al. 2002; Loayza 1996).

FIGURE 2.2

Relative sector sizes and wages in the presence of a labor tax

Source: Authors' illustration.

Note: L and W , respectively, are the labor allocated to and the wages paid in the informal (I) and the formal (F) sectors.

The next sections will bring to bear a variety of tools—motivational responses, sociological studies, and analyses of worker flows—to characterize the informal sector and sketch out the underlying motivations for being in it.

The sectors of informal labor: Characteristics and dynamics

Informal labor can broadly be broken into two subsectors: informal salaried and independent work. Table 2.1 shows the relative importance of the two sectors in the region using the social protection/legal definition where data permit. Formal salaried urban employment by this definition ranges from nearly 20 percent in Bolivia, Paraguay, and Peru, to roughly 60 percent in Chile—reaching roughly 40 percent for the region overall. Salaried informal work, including unpaid and domestic workers—mostly women—comprises roughly 33 percent of urban employment in the region, ranging from 17 percent in Chile to over 45 percent in Ecuador, Nicaragua Paraguay, and Peru. Informal independent (self-employed) workers, comprising single-person firms or owners employing other workers, represent roughly 24 percent of the regional urban workforce, ranging from 18 percent in Chile to over 35 percent in Bolivia, Colombia, the Dominican Republic, Peru, and the República Bolivariana de Venezuela. Clearly, the relative proportions depend on conventions for allocating workers—for instance, the International Labour Organization (ILO) categorizes domestic workers as independent—or whether unpaid workers are counted as “salaried.” However, the overall picture is one in

TABLE 2.1

Distribution of formal salaried, informal salaried (social protection/legal definition), and self-employed in urban areas of Latin America (percent)

Country	Year	Formal Salaried				Informal Salaried					Independent workers		
		Subtotal	Large firms	Small firms	Public workers	Subtotal	Large firms	Small firms	Domestic workers	Unpaid workers	Subtotal	Formal) (professional)	Informal (unskilled)
Argentina (GBA)	2005	45.33	27.96	4.03	13.34	30.36	10.83	10.87	7.65	1.01	24.32	9.31	15.01
Bolivia	2005	18.50	8.32	0.38	9.81	40.94	16.41	12.19	3.82	8.52	40.56	4.36	36.20
Brazil	2003	48.16	30.25	5.52	12.39	28.52	6.69	9.43	6.56	5.84	23.32	4.05	19.27
Chile	2003	60.81	43.94	4.60	12.27	17.20	7.64	4.26	4.03	1.27	21.99	4.45	17.54
Colombia	2006	30.64	21.25	1.80	7.59	28.01	7.99	12.06	4.65	3.31	41.35	2.85	38.50
Dominican Republic	2006	33.24	18.84	0.60	13.80	27.26	13.24	7.06	5.10	1.86	39.50	2.10	37.40
Ecuador ^a	1998	23.93	12.82	1.68	9.43	49.06	18.00	13.70	5.09	12.26	27.01	1.68	25.33
El Salvador	2003	38.78	26.54	1.13	11.10	33.78	11.25	11.43	4.31	6.79	27.45	0.72	26.73
Guatemala	2002	31.10	23.16	1.50	6.44	40.68	12.63	12.96	3.64	11.44	28.22	0.31	27.92
Mexico	2002	35.33	22.23	2.93	10.16	43.79	14.87	17.61	4.30	7.01	20.88	0.14	20.74
Nicaragua	2001	24.19	14.92	1.68	7.59	47.12	16.15	14.74	5.98	10.26	28.70	0.22	28.47
Paraguay	2003	17.93	6.15	1.05	10.74	48.06	15.74	15.79	11.79	4.74	34.01	0.59	33.42
Peru	2002	18.15	9.72	0.89	7.54	45.66	16.57	14.86	5.44	8.80	36.19	1.40	34.79
Uruguay	2004	51.85	29.26	3.98	18.61	21.81	7.08	6.69	6.65	1.40	26.34	5.75	20.59
Venezuela, R. B. de ^a	2003	38.77	21.80	2.37	14.60	22.55	8.24	9.47	2.06	2.77	38.68	3.46	35.22
Latin America		40.18	25.32	3.75	11.11	33.24	10.10	11.86	5.52	5.76	26.58	2.94	23.64

Sources: Gasparini and Tornarolli 2006; authors estimates.

Note: GBA denotes Greater Buenos Aires. ^aFormality among independent workers proxied by having a university degree.

which the informal salaried account for almost 60 percent of the informal, and independent workers slightly less.

As chapter 1 noted, firm size has been an important criterion in the past for categorizing the informal, and table 2.1 further breaks out each category by small firms of less than the traditional ILO definition of five workers and the rest as “large.” The characteristics of independent workers, particularly as entrepreneurs or firm owners, will be discussed at length in chapters 5 and 6 on firm dynamics, so here we

focus largely on salaried workers. Among these, formal employees are generally found in larger firms. The informal salaried appear to be found more or less equally in small and large firms, although table 2.2 suggests that, as discussed in chapter 1, most are concentrated in relatively small enterprises, if not necessarily in those employing fewer than five workers. In Mexico, 73 percent of paid informal salaried workers and 97 percent of unpaid workers are found in firms of fewer than 15, with the majority of those firms having

TABLE 2.2

Distribution of informal salaried, self-employed, and unpaid workers, by firm size (percent)

Argentina, 2003

Firm size	Informal salaried		Self-employed		Unpaid workers	
	Share of firm workers	Distribution	Share of firm workers	Distribution	Share of firm workers	Distribution
1	0.00	0.00	99.56	72.95	0.44	8.36
2 to 5	32.54	53.12	25.13	27.05	1.79	50.53
6 to 25	47.27	29.57	0.00	0.00	1.26	22.99
26 to 100	20.15	9.98	0.00	0.00	0.29	4.17
101 to 500	16.87	4.95	0.00	0.00	1.03	8.82
501+	13.35	2.38	0.00	0.00	0.99	5.13

Source: *Encuesta Permanente de Hogares*.

Note: Sample constrained to Gran Buenos Aires.

Brazil, 2002

Firm size	Informal salaried		Self-employed		Unpaid workers	
	Share of firm workers	Distribution	Share of firm workers	Distribution	Share of firm workers	Distribution
1	0.00	0.00	100.00	59.39	—	—
2 to 5	47.82	41.79	24.40	28.29	—	—
6 to 10	35.03	17.01	10.00	6.44	—	—
11+	17.25	41.20	1.85	5.87	—	—

Source: *Pesquisa Nacional Amostra Domicílios*.

Note: — = not available. Sample constrained to the metropolitan areas.

Mexico, 2004

Firm size	Informal salaried		Self-employed		Unpaid workers	
	Share of firm workers	Distribution	Share of firm workers	Distribution	Share of firm workers	Distribution
1	0.00	0.00	100.00	64.64	0.00	0.00
2 to 5	53.49	56.78	25.02	30.65	13.65	93.03
6 to 10	50.53	11.55	10.02	2.64	2.63	3.87
11 to 15	40.88	4.58	6.29	0.81	0.79	0.57
16 to 50	28.76	10.90	2.48	1.08	0.43	1.05
51 to 100	16.29	3.11	0.54	0.12	0.07	0.09
101 to 250	8.59	1.01	0.32	0.04	0.22	0.17
251+	10.47	12.06	0.01	0.01	0.17	1.23

Source: *Encuesta Nacional de Empleo Urbano*.

Note: Second quarter 2004.

TABLE 2.3

Characteristics of informal salaried employees in Mexican microfirms

Variable	Workers		Paid workers		Unpaid workers	
	Relatives	Nonrelatives	Relatives	Nonrelatives	Relatives	Nonrelatives
Observations (n)	1,317,330	994,045	372,032	964,275	945,298	29,770
% of Workers	56.99	43.01	27.84	72.16	96.95	3.05
Spouse	23.62		3.45		31.47	
Children	50.66		35.77		55.63	
Other relatives	25.72		60.78		12.89	
% Male	51.62	65.8	72.4	66.65	43.44	38.18
Average hours	33.17	44.09	42.47	44.58	29.51	27.97
Average age	28.15	30.17	28.64	30.38	27.96	23.03
% IMSS	4.31	51.7	13.02	53.28	0.90	0.30
% written contract	0.91	27.46	3.23	28.31	0.00	0.00

Source: *Encuesta Nacional de Micronegocios*, 1994.

Note: IMSS = Mexican Social Security Institute.

fewer than five workers. In Argentina, this is somewhat less the case, although 83 percent are found in firms of under 25 workers, which also encompass 80 percent of unpaid workers. Hence, understanding the nature of informal salaried work is, to an important degree, about understanding worker and firm decisions to become formal (the *de Soto* margin) and the determinants of worker flows between large formal firms and small informal firms. The evasion margin among large firms is more important in Argentina, where roughly 15 percent of salaried workers in firms of over 100 are informal, compared to roughly 10 percent in Mexico. Therefore, we are dealing with modestly differing mixes of phenomena across countries.

The Mexican microenterprise survey offers additional insight into the salaried and unpaid workers in small firms (under 16 workers). Several striking findings emerge from table 2.3. First, the microenterprise appears to be very much a family affair. In Mexico, over half of all workers and almost 30 percent of paid workers are directly related to the owner of the firm. This clearly complicates earnings comparisons, as family workers are likely to receive payment in kind—either food or shelter. Second, essentially all unpaid workers (97 percent) are family workers, and so probably should be seen as earning pay within the context of the family unit rather than truly “unpaid.” Whether their marginal product to the household is as high as it would be in the formal sector, or whether many of these workers are, in fact, underemployed, the data cannot tell us. Finally, even a large part of home-based work, which largely involves

women and is sometimes thought to be particularly exploitative, may also emerge from this dynamic (box 2.1).

Third, among family members, the verbal open contract is the rule (96 percent), while written contracts are reserved for the 28 percent of Mexican workers who are paid nonrelatives. This surprisingly small number suggests that most employment relations are ruled by informal mechanisms without recourse to any sort of formal employment contracts. As other dimensions of informality, this is very much a question of size of the firm, as noted in chapter 1 and will be further discussed in chapters 5 and 6 on formality decisions of firms. Relatively few firms with fewer than 3 workers have registered with the social security administration, while most firms of 10 or more have at least 1 worker registered. Plausibly, more formal contracting relationships will become more important as workers with fewer personal ties are hired.

Life-cycle patterns and dynamics

The informal self-employed and salaried labor show very different patterns of behavior both across the life cycle and in their flows through the labor market. To begin, figure 2.3 illustrates employment trajectories across the life cycle, plotting the share of the working-age urban employed population in the three categories. Young workers are most strongly represented in informal salaried employment in each of the three countries, peaking at around 20 percent for those in the workforce at around the age urban of 20. Thereafter, it falls steadily, although middle-aged and older workers still make

Home-based work: Exploitation or flexible work arrangement?

A particular modality of the informal microfirm that has received increasing attention is home-based work. Numerous authors (Arriagada 1998; ILO 1995; Prugl and Tinker 1997; Tomei 2000; WIEGO 2000) tend to see the subcontracting out of these workers as a way for large firms to maintain flexibility, quality, and global profitability by both avoiding benefits and transferring the risks of demand volatility to workers. The absence of an internationally accepted definition of home-based work, among other problems, means that the reported shares of the workforce, ranging from 1.5 to 20.0 percent (see table 2B.1), are probably not comparable (Chen, Sebstad, and O'Connell 1999).

Although some very famous firms, such as Italy's Benetton, started as a cluster of home-based workers, it is questionable how important international links really are. As with informal microenterprises more generally, the fraction that reports working as contract workers, perhaps for large companies, rather than selling their work directly in the local market, is generally small (0.7 percent in Brazil, 1.2 percent in Ecuador, and 1.6 percent in Mexico). Yet the sector is often over 75 percent women and those whose personal characteristics make outside work inaccessible (Carr, Chen, and Tate 2000; Cunningham and Ramos 2001; ILO 1995; Tomei 2000). The concern is that employers may take advantage of the situation of those whose work options are limited to the home by subjecting home-based workers to lower remuneration and labor standards than workers who can compete in the labor market (Krawczyk 1993; WIEGO 2000). Despite suggestive anecdotal evidence, statistical studies have yet to document it as a widespread phenomenon.

TABLE 2B.1

Home-based workers' participation (percent)

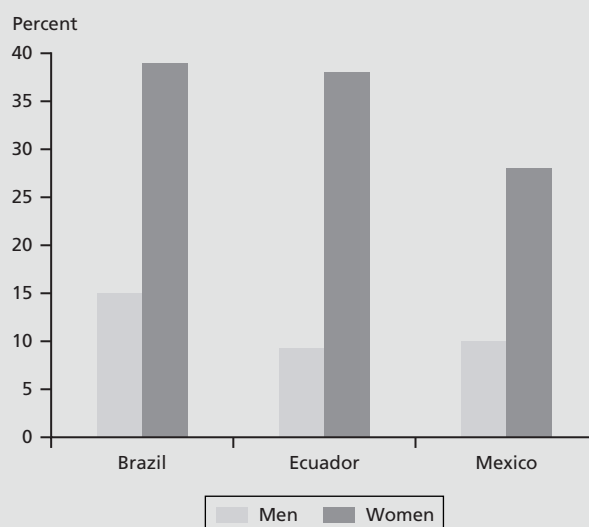
Country	Estimated proportion	Proportion female	Year of estimate
Algeria	3.3	97.0	1989
Australia	2.9	—	1989
Brazil	5.5	74.8	1999
Ecuador	17.3	73.6	1999
India	2.5	—	1981
Japan	1.6	93.5	1988
Mexico	4.4	64.4	1999
Philippines	23.0	—	1980s
Peru	10.5	—	1987
United Kingdom	2.3	70.0	1981
United States	7.53	—	1985

Sources: Cunningham and Ramos 2001; ILO 1995.

Note: — = not available.

As can be seen in figure 2B.1, home-based workers do earn less than other workers. Controlling for lower skill levels in the sector, home-based workers earn 22.8, 28.9, and 39.6 percent less than do workers outside the home in Mexico, Brazil, and Ecuador, respectively (Cunningham and Ramos 2001). But we cannot rule out that these lower wages are reflecting unobserved characteristics or the price of the flexibility that allows juggling other household responsibilities, not having to travel to the worksite, or other benefits that accrue to home care providers. This is partially borne out by the fact that home-based workers spend an average of 30 hours in productive activities each week, compared to more than 40 hours weekly among nonhome-based workers (Cunningham and Ramos 2001; Tomei 2000). This result is driven by women, especially those with young children and/ or spouses. In Brazil, Ecuador, and Mexico, home-based female workers spend one-third fewer hours on the job than do women who work outside the home. Home-based work may thus be a preferred work arrangement for those who have both home and market duties. Women with young children and/or who are married are more likely to engage in home-based work than are those without such household constraints, and less than half are household heads (Cunningham and Ramos 2001). Interviews with female home-based workers with children reveal that these women hope to work outside the home once their children have left home (Jelin, Mercado, and Wyczykier 2001).

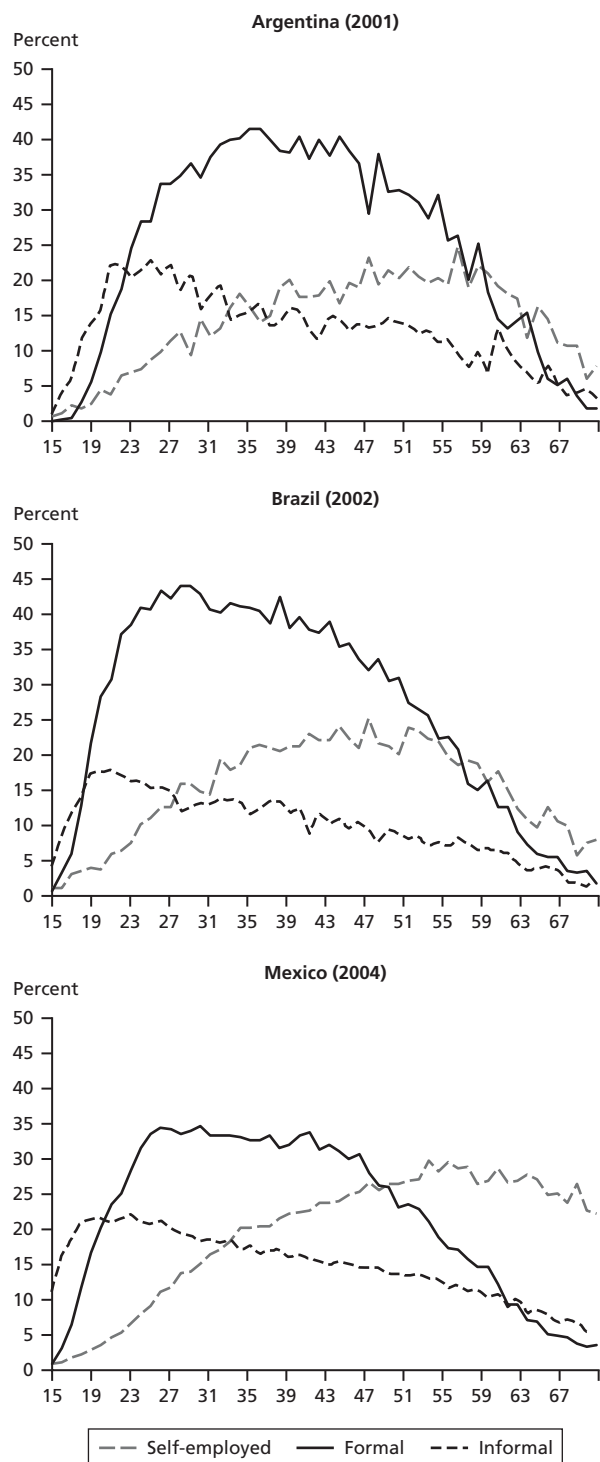
FIGURE 2B.1

Hourly wage premium of outside home work to home-based work

Source: Cunningham and Ramos 2001.

FIGURE 2.3

Rate of urban employment across sectors, by age



Source: Cunningham et al. 2007.

Note: Considers all self-employed and informal salaried uncovered by social protections in Argentina and Mexico, and the availability of a *carteria de trabajo* (work card) in Brazil.

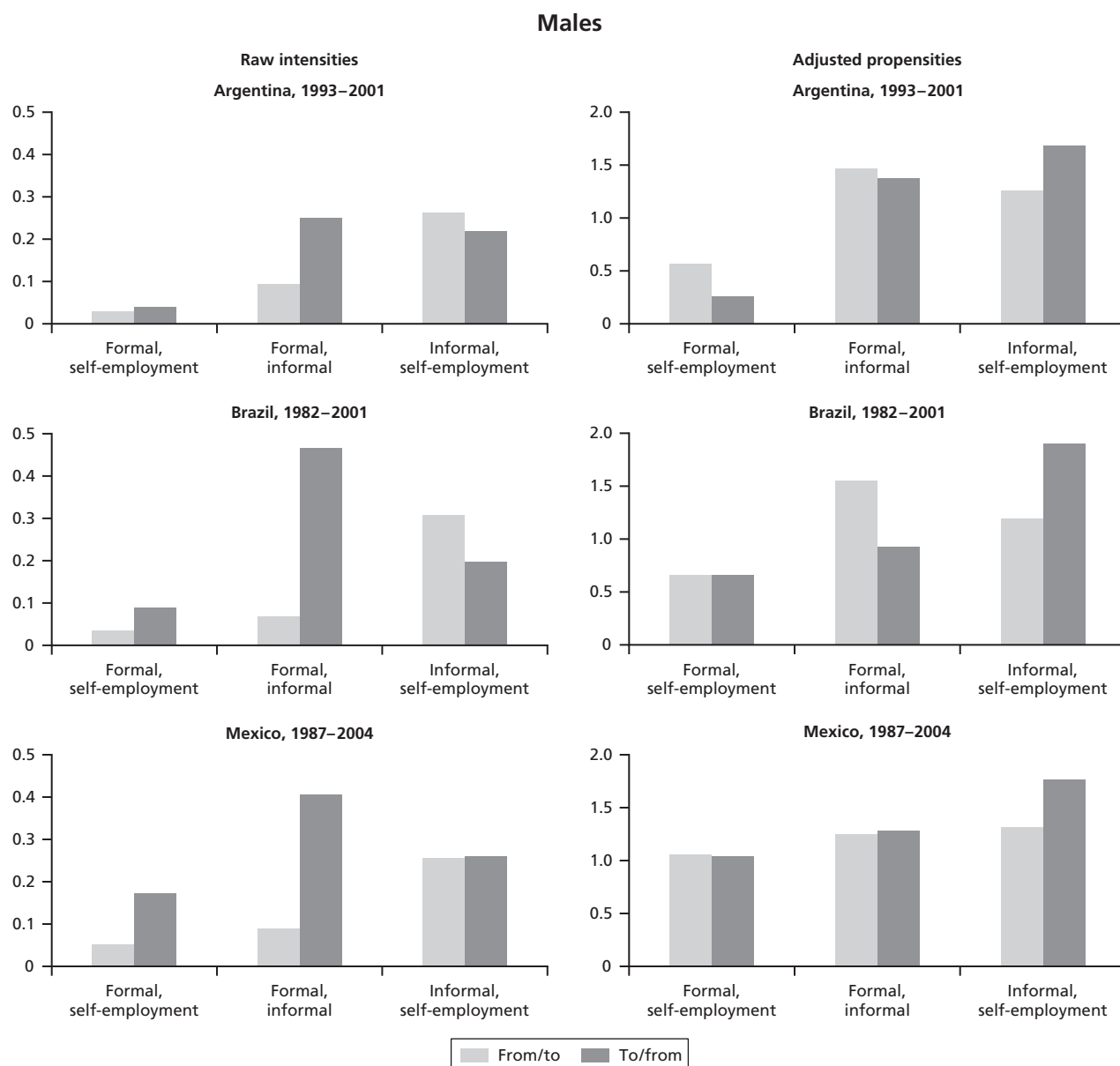
up around 40 percent (almost half in Argentina) of all informal employees. By contrast, self-employment is almost the mirror image, with virtually no representation among the very young, but then climbing steadily until late middle age. The formal sector is largest among prime-age workers. Older workers are as likely to be found in the formal sector or self-employment.

The commonalities across all three countries highlight a suggestive, distinct role for the two informal sectors. First, informal salaried work is a point of entry to the labor market for many of the young. This is consistent with the accumulation of skills enabling young workers to eventually find a job in the formal sector or realize any desire to be self-employed. Second, informal salaried work remains a source of jobs for older workers, around 10–15 percent of prime-age workers. They may lack the skills or capital to become self-employed or get a formal salaried job, or they may opt out of formality since they cannot accrue sufficient years to secure a meaningful pension. Finally, there is no inexorable evolution from informality to formality with age; informal self-employment and formal salaried work are equally common at the end of work life.

Further insight can be gained into the differences between the sectors and, more fundamentally, their respective roles in the labor market by looking at patterns of mobility of workers through types of work. As Maloney (1999) argues in the first analysis of this kind for Latin America, the segmentation or exclusion view would predict a predominantly unidirectional move toward formality: on average, workers leave school, enter informality to queue until they find a formal job, and eventually retire. In an integrated market, however, workers search among formal and informal jobs, treating them as different in type but not necessarily in quality, so that flows among sectors are more symmetrical.

Reliable panel data that permit constructing transition probabilities from actual movements of the same workers across sectors are available for three countries in the region: Argentina, Brazil, and Mexico (see box 2.2). Bosch and Maloney (2005) follow them through the three sectors of paid work—informal self-employed, informal salaried, and formal salaried—as well as unemployment and being out of the labor force (OLF). Such transitions represent “reduced forms,” capturing disposition as well as ease of entering a sector. Without further relating them to the dynamics of earnings and the business cycle, they need to be seen as more suggestive than conclusive. Such structured analysis of the determinants of entry into sectors, as well as the analysis of cyclical labor force dynamics, is taken up in chapter 4.

FIGURE 2.4a

Probabilities of transition among sectors of employment

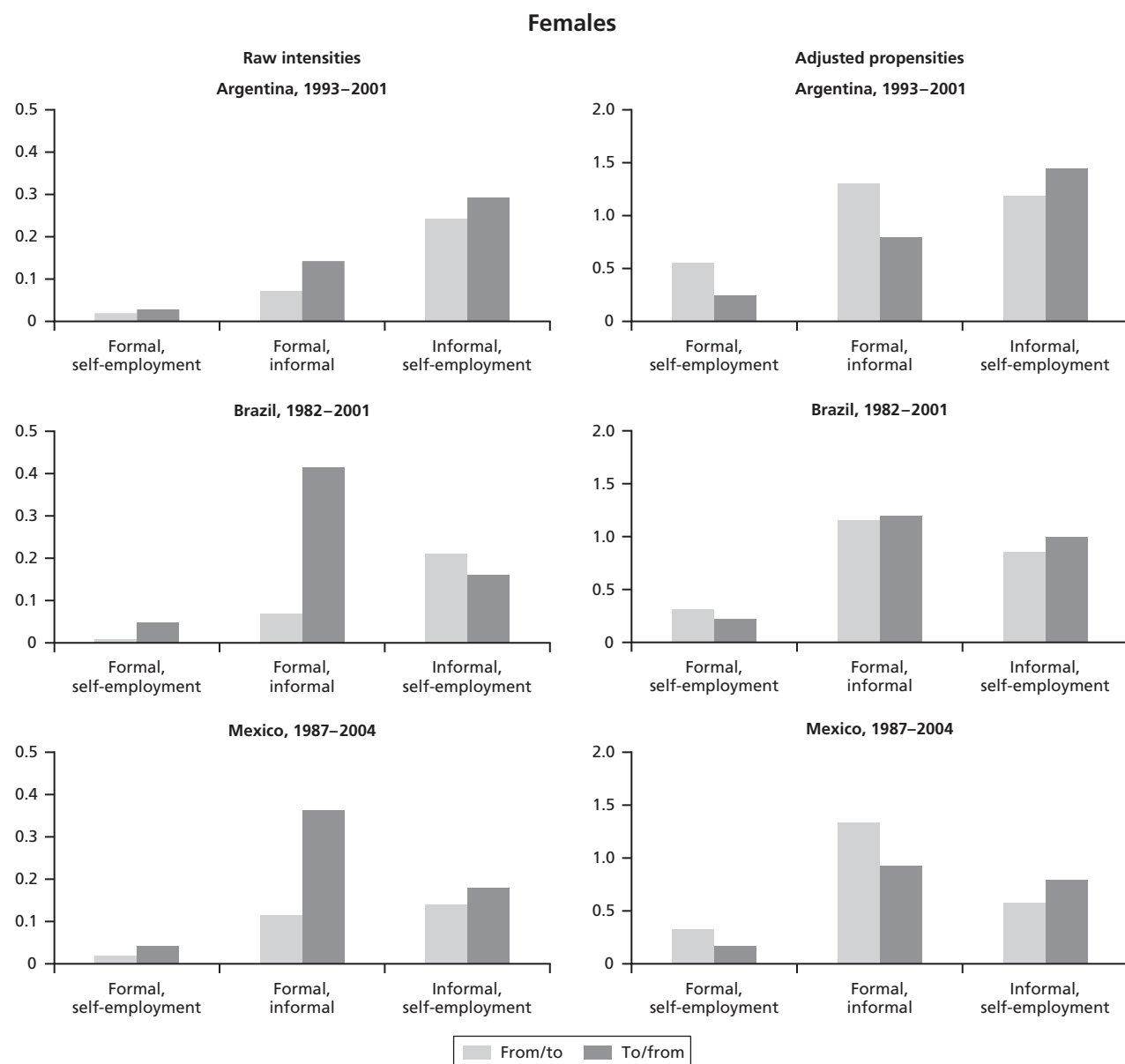
Sources: Bosch and Maloney 2005, 2007.

Note: Probabilities in the second column are standardized by final sector job turnover and destination sector size.

The first column in each panel of figures 2.4a and b presents the probability of moving from one sector of work to another,⁸ and two points emerge as noteworthy. First, the three labor markets show a high degree of commonality in most transitions. The relative mobility between any given pairing of sectors is broadly similar across the three countries, although with some notable exceptions. Second, the level of mobility is relatively high. For example,

in both Mexico and Brazil, the probability that a worker found in informal salaried work moves to formal work across a one-year period is 40–50 percent. The relatively smaller reverse probabilities would seem consistent with the conventional one-way flows out of informality. However, as Maloney (1999) argues, such observations need to take into account both the likelihood of a worker separating from the previous job, regardless of destination sector, and

FIGURE 2.4b

Probabilities of transition among sectors of employment

Sources: Bosch and Maloney 2005, 2007.

Note: Probabilities in the second column are standardized by final sector job turnover and destination sector size.

the availability of positions to move into the destination sector. The first differs greatly across sectors. Figure 2.5 shows the inverse of the probability of leaving—the average amount of time spent or duration in a sector. Again, the similarities across countries are far more striking than the differences: durations are shortest in unemployment, followed by informal salaried, self-employment, or out of the labor force; and longest in formal salaried employment.

Given the low rates of turnover in formal salaried and self-employment, it is not surprising that flows between these sectors should be the lowest in figure 2.5. This turnover, combined with the size of the destination sector, is also related to the number of positions into which a worker can move.

The second column in each panel of figures 2.4a and b standardizes, by both turnover and destination, available

BOX 2.2

Data from rotating panels in Argentina, Brazil, and Mexico

In order to construct continuous time transition matrices to study worker flows across sectors, Bosch and Maloney (2005) employ three different surveys that compile information about the labor status of workers and other relevant information. They employ one year as the time unit to analyze labor mobility dynamics, mainly as a common sampling interval for the three countries.

Argentina

For Argentina, in the *Encuesta Permanente de Hogares* (Permanent Household Survey) a panel covers the area of the Federal District and surroundings (Gran Buenos Aires), which accounts for approximately 60 percent of Argentina's urban employment. Until 2003, the survey was conducted every six months (April/May and October) with a 25 percent rotation of the panel. As a consequence, each household is followed for two years at sampling intervals of six months. The authors employ panels from May 1993 to October 2001. The sample is notably smaller than the Mexican and Brazilian surveys, and only 29,000 transitions (13,900 for men, 15,100 for women) can be studied.

Brazil

The *Pesquisa Mensual de Emprego* (Monthly Employment Survey) follows monthly employment indicators. Households are interviewed four months in a row, and then reinterviewed eight months later, so that 25 percent of the sample is renewed every month. Given this panel structure, we can construct four yearly employment status transitions for each individual. The authors put together nine consecutive panels starting in February 1982. Each panel consists of 12 consecutive cohorts

covering approximately two years through the period 1982–2001. The total number of transitions is 2,520,000 (1,190,000 for men, 1,330,000 for women).

Mexico

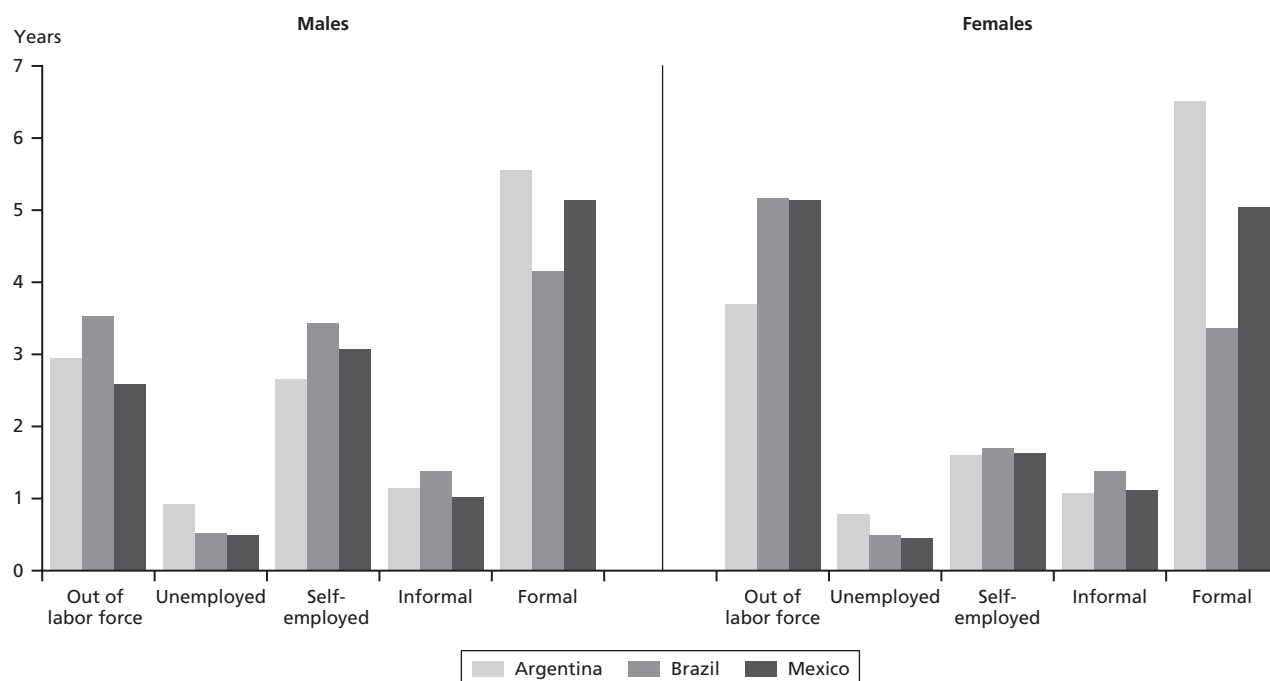
The *Encuesta Nacional de Empleo Urbano* (ENEU, National Urban Employment Survey) conducts extensive quarterly household interviews in the 16 major metropolitan areas. The questionnaire is extensive in its coverage of participation in the labor market, wages, hours worked, and so on that are traditionally found in such employment surveys. The ENEU is structured so as to track a fifth of each sample across a five-quarter period. The authors concatenated panels from the first quarter of 1987 to the fourth quarter of 1999. Each individual contributed two transition pairs (from the first quarter to the fourth, and from the second to the fifth), giving rise to approximately 1,785,000 transitions (810,000 for men, 975,000 for women).

“Informal” refers here to self-employed (including small firm owners) and informal salaried workers in firms with fewer than 16 employees who lack social security or medical benefits. Formal salaried workers are defined as those enjoying labor protections. To focus on the traditional first margin of informality, Bosch and Maloney (2005) do not consider unprotected workers in large firms. The remainder of the sample is divided into two groups: those out of the labor force and the unemployed. Unfortunately, the Brazilian survey lacks information on firm size, and informal status is given by whether the worker holds a signed work card guaranteeing access to benefits in Brazil. Mexican and Argentine surveys contain very similar questions about benefits and firm sizes.

positions in the final sector. This gives a measure of whether a separating worker from one particular sector is more or less likely to move into an available job in another sector than are all workers entering that sector from all other sectors (see Duryea et al. 2006). Bosch and Maloney (2007) also show that this standardization can be interpreted as a measure of revealed comparative advantage of a worker: in the absence of barriers to entry, what sectors do the individual characteristics and personal constraints predispose workers to enter?

What is striking is that there is a reasonably high degree of symmetry in these adjusted probabilities, with some exceptions: flows from formal salaried to informal salaried are not so different from the reverse—or, put differently, if an informal salaried worker has a comparative advantage in formal salaried work, the reverse is also likely to be true. Further, the symmetry of the flows also suggests that, overall, there is not a one-way flow from informality into formality, with some exceptions. For instance, Brazilian formal salaried

FIGURE 2.5

Absolute mean duration of labor force status

Source: Bosch and Maloney 2005.

men move relatively more into informal salaried work than the reverse. The same is true for Mexican females. In the case of Brazil, this may suggest the reduction in the availability of formal salaried jobs across this period, described in detail in chapter 4. It is also notable that, for all cases except Mexico, formal workers do not seem to have a comparative advantage to enter self-employed work and vice versa, although these flows are generally symmetrical, except for Argentine females.

As chapter 4 will show, in fact, these averages hide substantial changes in relative fluidity across the business cycle. Further, as discussed below, this rough symmetry hides important demographic differences, consistent with the previous patterns of participation across the life cycle. However, it seems safe to say that, on average, workers do not graduate unidirectionally from informal work to available jobs in the formal sector.

More generally, the exercise suggests that static summary statistics of the allocation of workers obscure important aspects of the dynamism of worker flows among sectors. On average, informal salaried workers do not remain so for long periods of time; their durations are very short. As is argued later, this does not necessarily mean that informal salaried labor is a transitory state that the majority of workers even-

tually leave behind—there may be many instances of subsequent return to informal salaried work in later transitions. However, what is clear is that, first, we are dealing more with “quantum” informality where there is a large amount of movement among sectors, rather than patient queuing by most of the informal salaried to enter a formal salaried job permanently. Second, the high frequency of job changes, even for formal workers, is consistent with the Levy (2006) observation that the rotation in and out of formal pension schemes implies that many “quantum formal” workers will never accrue enough years in the system to earn a pension, thus creating incentives for opting out. Exploring the dynamic demographic signature of the two informal sectors over the next sections will further help in understanding their role.

Patterns of mobility of the self-employed

The patterns of transition into informal self-employment by age and education suggest that this sector behaves broadly similarly to its counterpart in advanced economies and rather less like a queue for formal sector jobs (figure 2.6). Evans and Jovanovic (1989) argue that the observed increasing probability of entry with age, despite a presumably lower level of

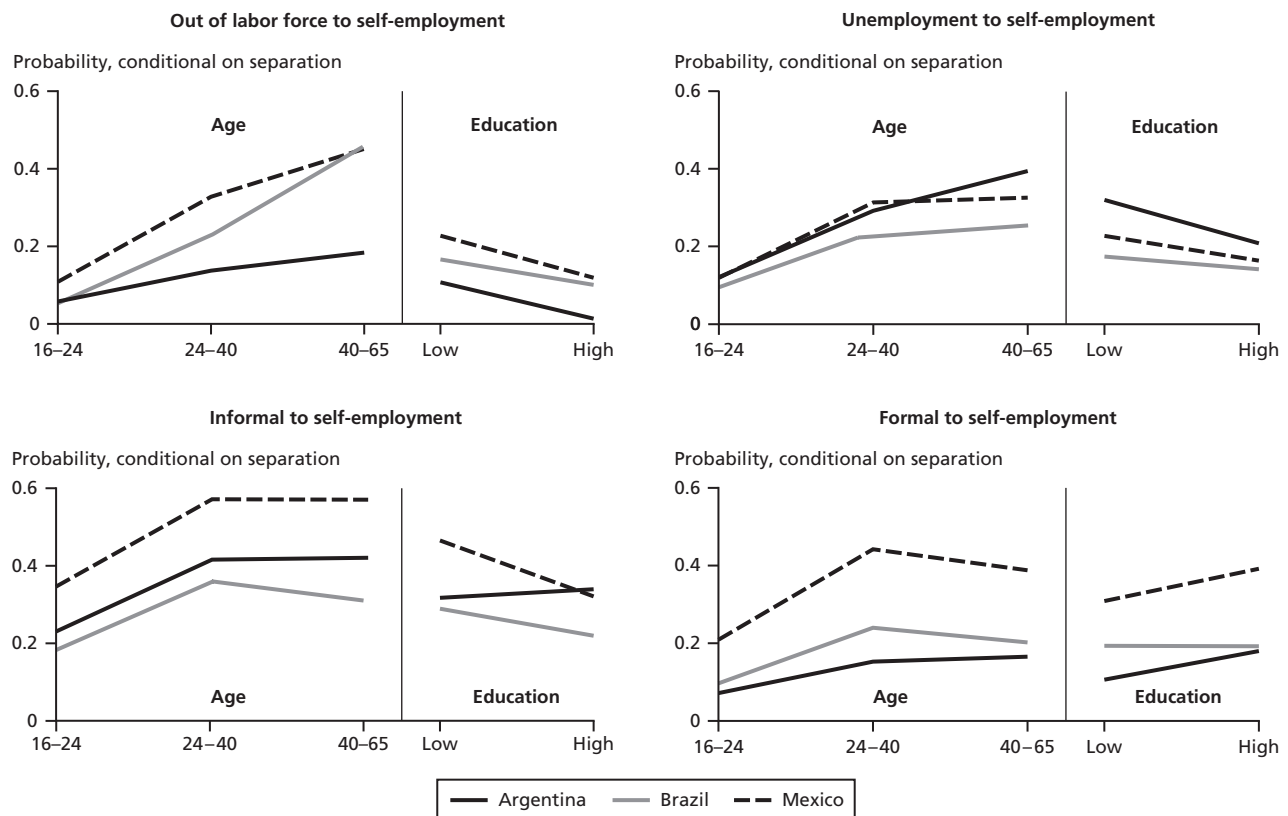
risk aversion among the young, is consistent with the existence of credit constraints that dictate that potential entrepreneurs must accumulate the necessary physical and working capital. These credit constraints may be exacerbated in developing countries, where not only do credit markets function poorly, but weak educational systems make the accumulation of human capital difficult as well.

Strikingly consistent with this view, and with our earlier examination of life-cycle employment trajectories, self-employment is not a port of entry into work. In all three countries, the probability of entry for young workers from OLF or unemployment is a mere fraction of that for older workers. For less than 10 percent of young workers (aged 16–24) leaving the OLF sector and unemployment, self-employment appears as the entry point to the labor market, approximately three times less than for middle-aged workers. The propensity to enter the sector from informal or formal salaried work is also, in all cases, often over double for older workers than for younger workers. For instance, 35 percent (20 percent) of younger

Mexican workers leaving the informal (formal) salaried sector become independent, compared with the 57 percent (44 percent) of their older counterparts. Similar results are found in Argentina and Brazil, where roughly 20 percent (10 percent) of exiting young informal (formal) salaried workers find a self-employment opportunity, compared to 40 percent (20 percent) among older workers.⁹ These low rates of entry among the young confirm the pattern found in figure 2.4 that self-employment is the least viable source of employment for the young in all three countries.

Again, the evidence is consistent with a life-cycle model in which many workers enter formal employment to accumulate both human and financial capital, and then become self-employed or open a small business—a pattern identified in the OECD literature (see Blanchflower and Oswald 1998; and Evans and Jovanovic 1989). The sociologists Balán, Browning, and Jelin (1973) argue precisely for such a life-cycle model in Mexico, where workers with entrepreneurial motives enter into salaried work; accumulate

FIGURE 2.6

Propensity to move to self-employment from different sectors

Source: Bosch and Maloney 2005.

knowledge, capital, and contacts; and then quit to open their own informal businesses (cited in Maloney 1999).¹⁰

A further disaggregation of figure 2.6 (not shown) suggests that older and better-educated workers spend longer in self-employment (Bosch and Maloney 2005). This is consistent with the mainstream firm dynamics literature that suggests that young firms, which, *ceteris paribus*, are more likely to be opened by younger workers, have very high failure rates (Evans and Leighton 1989; Jovanovic 1982). The higher tenure for better-educated workers is similar to that found in the formal sector and the opposite of the pattern for the informal salaried. This topic will be analyzed further in chapter 5.

It may be argued that both the pattern of late entry and longer duration with age are also consistent with exclusion: a middle-aged worker who loses his or her formal sector job is unable to find a new one, and therefore is forced into self-employment as a safety net. This dynamic may have been of particular relevance during the economic restructurings of the 1990s, when older displaced workers may have found their skills obsolete and of little demand in the emerging sectors. The features of the observed transitions mitigate against this view as being the entire or even a large part of the story. The rate of transition into the sector in all three countries seems broadly linear in age until middle age. That is, there is not a sudden spike in propensities among older workers, but rather a gradual increase with an inflection point at prime age when the propensity to enter often *decreases*. Overall, this pattern seems more consistent with the life-cycle pattern found in more advanced economies, although coexisting with some older workers with little access to the formal sector market once fired.¹¹

Even if the life cycle well describes the patterns of entry into self-employment, it is still the case that this is a risky business. Not only are most independent workers out of the formal circuit of employment protection, but also operating a small business is intrinsically riskier everywhere, with many firms not lasting through their first year (box 2.3). Again, the sociological literature provides striking confirmation of this insight when Balán, Browning, and Jelin (1973) argue that “although self-employment is a goal for many Mexican workers . . . [b]ecoming self-employed involves a large risk, especially for those men who had stable and secure jobs” (p. 216). In fact, chapter 5 will show that comparisons of patterns of entry and exit derived from the Mexican and Nicaraguan microenterprise surveys appear to be broadly similar to those derived for the United

States by Evans and Leighton (1989), so the considerations of these potential entrepreneurs are similar to those found everywhere.¹² Nonetheless, since small firms have high mortality rates, both their owners and their workers in the region are more likely to find themselves without employment and searching for a new job, without the broad coverage of safety nets common in most developed countries.

Further insights are gained by examining transitions by gender. Figure 2.7 suggests a more dynamic corridor of mobility between being out of the labor force and informal self-employment for women as compared to men. While at first this may suggest discrimination against women in formal sector jobs, further examination reveals that this is largely driven by married women. Figure 2.8 shows that while it is true that single mothers have high levels of participation in informal self-employment and wives are overrepresented in the unpaid sector, single women without children have the highest rate of participation in *formal* jobs of any group, male or female. Moreover, replicating this analysis for a cohort of single women in Argentina and Mexico (the two surveys with a marital status variable), Bosch and Maloney (2005) show that the transitions of single women are very similar to those of men, and this similarity holds up when disaggregated by duration and propensity. The high mobility is partly due to the fact that women have far lower tenure in self-employment, suggesting that these jobs are not intended to be careers, but rather transitory supplements to family income.

In a view with lineage to Becker's (1991) work stressing structural determinants of employment patterns, Cunningham (2001a) argues that Mexican women's patterns of participation—and particularly their gravitation toward self-employment—are driven by their need to balance their other responsibilities in the household; child-raising requires greater job flexibility than the salaried sector offers. Women with young children are often more prone to be self-employed rather than formal sector employees (Cunningham 2000; Cunningham and Ramos 2001; see also chapter 3), and interview data from Geldstein (2000) for Argentina and Chant (1991) from Mexico suggest that women may more easily balance their productive (market) and reproductive (home care) roles if they work for themselves than if they are employees.¹³

Patterns of mobility of the informal salaried

The patterns of transitions for informal salaried workers in small firms offer a more nuanced picture. As noted earlier, the

BOX 2.3

Informal self-employment: Risky and informal, or risky *because* informal?

Many of the “precarious” characteristics associated with informality are natural by-products of the fact that the informal self-employed resemble the fundamental characteristics of the owners of a small firm. The industrial country literature on firm behavior offers two important findings about such firms. First, there is a wide range of sizes among long-standing firms determined by factors such as how efficient or hardworking an entrepreneur is, how well placed is his or her firm, and the logic of the production process. This means that the existence of many small firms does not necessarily imply failure of either labor or credit markets. The reason that 80 percent of microfirms have only one or two employees and tend to be family based may reflect a logic rooted in the tradition of the family farm, or possibly reflects the sustainable reach of informal contracting relations.

A second finding about small firms everywhere is their extraordinarily high rate of failure. Seeking to explain the U.S. case, Jovanovic (1982) argues that this is due to the fact that entrepreneurs cannot know how good their location is or how competent they will be as entrepreneurs until they actually start the business. Very soon after starting, many enterprises find that they are not viable and fail. Rough calculations from the Mexican microenterprise survey suggest that these firms show high failure rates, but not particularly higher than those in the United States.

Thus, most of the characteristics of the sector may be intrinsic and not imply any inferiority or undesired precariousness. Levenson and Maloney (1996) treat “formality”

as participation in the numerous formal institutions: federal and local treasuries, governmental programs such as social security, the legal system, the banking system, health inspection, firm censuses, trade organizations, and civic organizations. These, of course, have costs in terms of compliance with legal norms, which very small firms can easily avoid in developing countries. Small firms are anchored in social networks of family and immediate neighborhood that allow them to enforce implicit contracts and insure against risks. However, as firms grow, they increasingly need to secure property rights or permit formal contracting mechanisms, pool risk, and gain access to credit. De Soto (1989) offers a striking example in which informal street vendors in Peru tried to *pay* their taxes since this would guarantee them some property rights over their pitch and hence offer some security to investments they wanted to make. Chapters 5 and 6 will show that firms do become more formal with age and size.

Combining the two characteristics of microfirms and this notion of formality implies that small firms will have higher costs, are likely to be informal, and will have very high failure rates. Although this corresponds exactly to the standard picture of the stagnant, precarious, unprotected informal worker familiar in the literature, it emerges naturally from potential entrepreneurs trying their luck (risk taking), often failing, and not engaging in the formal institutions until they grow. Thus, the existence of the microfirm informal sector may be largely unrelated to questions of labor market dualism or even credit market distortions.

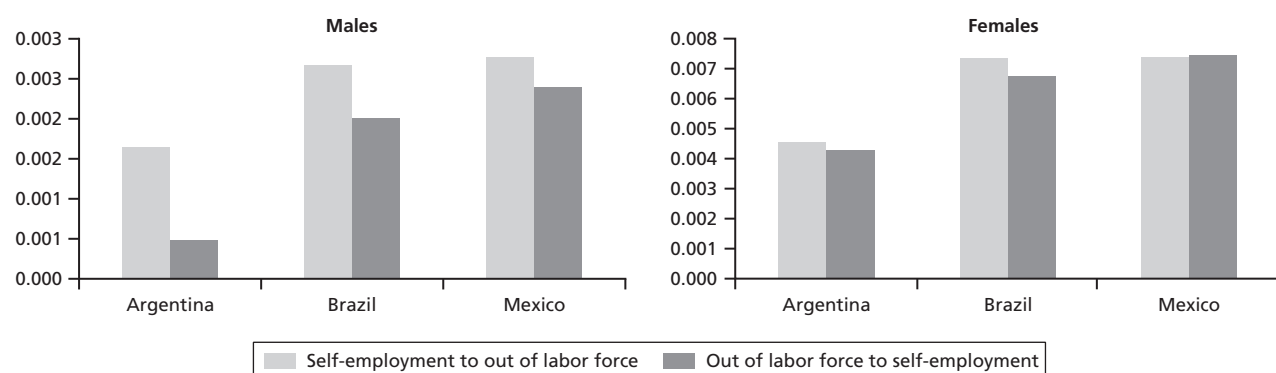
age trajectories of informal salaried workers sketched in figure 2.4 are almost the mirror opposite of the self-employed; it is the sector with the highest share of very young workers and shows a constant decline across all subsequent ages. In Mexico, for instance, this translates into the average age of the informal salaried being 29, five years below that of formal salaried workers, and 14 years below that of the self-employed. Still, 40 percent to almost half of the sector is composed of older workers in the three countries. Furthermore, figure 2.9 shows that in contrast to self-employment, entry into the sector from either unemployment or OLF decreases with age, except in Brazil, as well as education,

except in Argentina. Thus, the sector may represent both an important element of disguised unemployment and a port of entry into the workforce, particularly for youth, older unskilled workers, and some women.

As noted before, workers appear to spend relatively little time in this sector—durations are just over a year. The brevity of tenure is similar to that found for young workers in Brazil by Sedlacek et al. (1995), and is similar to the United States, where the median tenure for young workers 16–24 years of age is only 1.4 years, and for workers aged 25–34, the median tenure is 3.4 years (BLS 1991). More detailed studies of transitions also show a high degree of

FIGURE 2.7

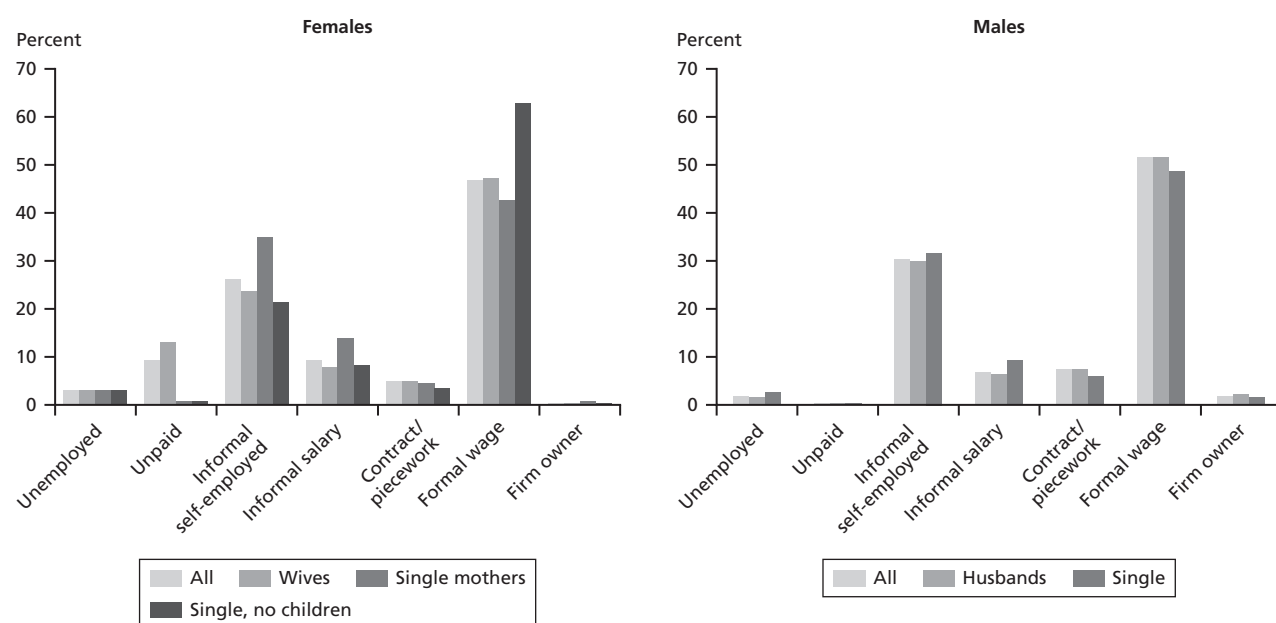
Transition between self-employment and out of labor force, by gender



Source: Bosch and Maloney 2005.

FIGURE 2.8

Employment sector allocation by gender, marital status, and parental status in Mexico



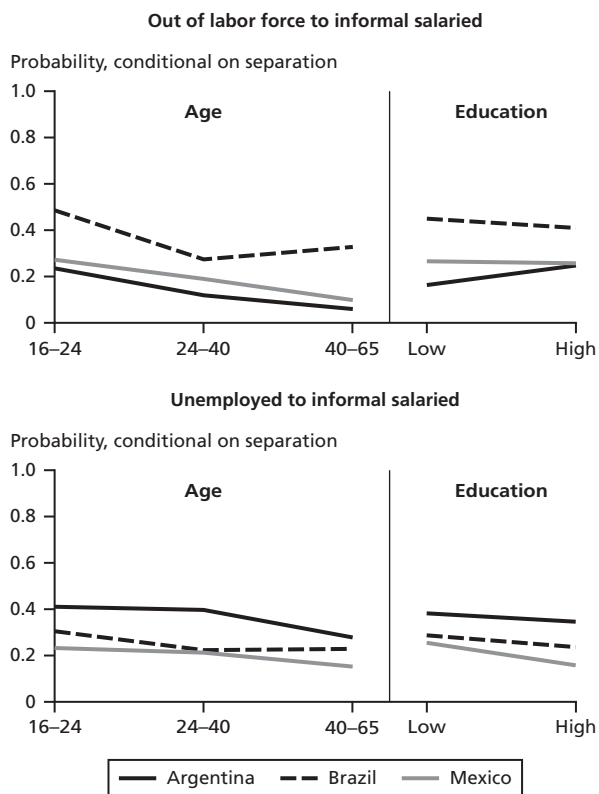
Source: Cunningham 2001a.

mobility between school, unpaid work, and, to a lesser extent, unemployment that suggests a pool of generally poorer workers not yet tracked into regular employment (see World Bank 2007).¹⁴

In this sense, informal salaried work has an important component that corresponds to the very high levels of unemployment found among youth in the OECD that are often double or triple those of prime-aged males 25–54 years of age (Jimeno and Rodríguez Palenzuela 2003). The causes

of this are numerous, including the relative size of the pool of incoming young people, the difficulty of signaling the appropriateness of an applicant for a particular job, the difficulty of dismissing a new worker if the match appears poor, or exclusionary factors related to deficient schooling or discrimination of unskilled and minority youth. These factors translate to the Latin American context; thus, even in the absence of nominal rigidities that might segment the market, both very restrictive labor legislation and poor

FIGURE 2.9

Propensity to move to informal salaried status from different sectors

Source: Bosch and Maloney 2005.

mechanisms for resolving information asymmetries can impede young workers' entry into the formal sector. As chapter 4 will show, in Argentina and Brazil, the rise in informal salaried workers had a disproportionate impact on young entrants to the workforce.

One scenario is that while in school—and just after leaving school—many students work at a family business, with or without pay, before moving on to formal salaried jobs.¹⁵ As noted further below, in the case of Mexico, fully half of those working for microenterprises are children of the owner. Even for those who are not family members, informal small enterprises may train more apprentices and workers than the formal education system and the mostly government job-training schemes together (Hemmer and Mannel 1989). The informal salaried experience may constitute continued schooling with an attendant lower wage. Still, the education–occupational profiles indicate that graduation to formal salaried work is unlikely for youth who drop out of school before completing at least a full course of secondary education.

For the sizable fraction of the informal salaried who are prime age or older, both exclusionary and exit rationales apply. While the arguments of greater flexibility and higher earnings are less applicable, as noted before, the informal salaried in several countries do earn the minimum wage and often receive other benefits (such as vacation and extra salaries), and may prefer not to contribute to social security, especially if they are past the age at which they cannot accrue enough years in the system to earn a pension. However, the case for exclusion is also strong, particularly for the unskilled and those fired former formal sector workers whose skills have been rendered obsolete. As chapter 4 will show, much of the trend in informality in Argentina, the metropolitan area of Brazil, and over the last five years in Mexico, is due to the increase in informal salaried work affecting, in particular, groups that previously were covered by formal work contracts, including prime-aged and older workers.

Finally, the low tenures observed among informal salaried workers do not mean that this is a transitory state that the majority of workers eventually overcome. The actual spells of informal salaried work experienced by many workers (especially low-skilled youth and older workers) may be much longer, since many spells end with unemployment or dropping out of the labor force, and may revert to informal salaried employment.¹⁶ The actual time these workers can take to find a formal salaried job, if desired, may be quite long. In fact, the lifetime occupational history data in table 2.4 show that around half and 84 percent of prime-aged and older informal salaried workers in Argentina and the Dominican Republic, respectively, have never held formal salaried employment. In fact, one-third of 46–55-year-old informal salaried workers report being informal salaried their entire work-life in Argentina, and nearly 60 percent in the Dominican Republic. Whether this is voluntary or involuntary, the data alone cannot tell; in fact, both the motivational and job satisfaction responses analyzed below and in chapter 3 show consistency with both hypotheses.¹⁷

Motivations for participation in informal work

As noted above, informal employment may reflect workers opting out of benefits programs voluntarily or exclusion to employment of last resort given the lack of better alternatives. This section introduces another tool for distinguishing between the two: asking workers directly what drives their occupational choices. Employment surveys in a handful of countries of the region have done this in the 1990s, although until recently only Mexico and Brazil did so with

TABLE 2.4

Informal employment and work-life occupational history, Dominican Republic and Argentina (percent of workers)

Have ever held a job as . . .	Dominican Republic (2006)			Argentina (2005)		
	Status in current job is . . .					
	Formal salaried	Informal salaried	Independent	Formal salaried	Informal salaried	Independent
Workers aged 15–25						
Formal salaried only	76.5	5.1	7.8	66.3	7.5	11.3
Informal salaried only	15.9	88.7	16.9	25.1	84.4	23.3
Independent only	5.9	5.1	72.8	2.2	7.2	62.4
More than one category	1.7	1.0	2.5	6.4	0.9	3.0
Total	100.0	100.0	100.0	100.0	100.0	100.0
Workers aged 26–45						
Formal salaried only	65.9	12.8	12.2	56.0	27.9	29.5
Informal salaried only	17.9	69.4	23.2	25.7	50.5	19.2
Independent only	11.3	14.1	57.4	9.3	9.8	31.0
More than one category	4.9	3.7	7.2	9.1	11.8	20.4
Total	100.0	100.0	100.0	100.0	100.0	100.0
Workers aged 46–65						
Formal salaried only	64.9	12.5	13.3	58.2	37.5	50.7
Informal salaried only	14.6	59.1	21.1	15.4	33.2	8.8
Independent only	16.0	24.1	61.8	16.4	7.8	23.9
More than one category	4.5	4.3	3.8	10.1	21.6	16.7
Total	100.0	100.0	100.0	100.0	100.0	100.0

Sources: Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), based on household survey, 2005, 2006 data.

Note: Each cell gives the percentage of workers currently in each occupational group who have held employment in the other occupational groups some time during their work-lives.

broad coverage. Below we report the findings from studies using these surveys.

In preparation of this report, new special surveys on informal employment were collected in the main urban areas of Bolivia, Colombia, and the Dominican Republic, building on a 2005 survey conducted in Greater Buenos Aires (see box 2.4, World Bank 2007b). The surveys included, among other things, questions on the motivations or reasons for workers to be salaried or self-employed and to participate or not in labor benefits programs (box 2.4). Workers were asked questions such as, “If you were able to choose, would you rather be a salaried or independent worker?”; “What is(are) the main reason(s) why you are a salaried employee rather than independent?” (and the converse question); and “What is the main reason why you do not have a job with contributions to pension and health insurance benefits?” The responses were structured around key proposed drivers of informality in the literature. Despite some variations in questionnaire design, the responses are largely comparable across the four countries. Therefore, below we discuss the evidence arising from the Brazilian and Mexican surveys, followed by the other four countries. Tables 2.5 and 2.6 present the main

results; these data offer fresh *prima facie* evidence on the significance of the voluntary or involuntary nature of informal employment in a variety of country contexts.

Self-employment

Both the sociological and economics literatures offer reasons why many workers might actually prefer self-employment over salaried work. In addition to the already discussed greater flexibility, autonomy, and entrepreneurship motives, the literature also talks about risk taking, family tradition, and mobility opportunities. Blau (1987) and Maloney (1999) argue that in poor countries, lack of human capital specific to an occupation or firm and the low formal sector productivity mean that the opportunity cost of self-employment is low. Contrary to these “exit” factors are, of course, the view of the sector as employment of last resort for workers unemployable in the formal sector.

The evidence below suggests that the majority of the sector corresponds more to exit motives and the entrepreneurial view in the advanced country literature, although a significant component corresponds to involuntary workers.¹⁸

BOX 2.4

Special informal employment surveys: What can we learn from them?

The World Bank, in partnership with line ministries and the national institutes of statistics in Argentina, Bolivia, Colombia, and the Dominican Republic, supported the implementation of special modules on informality attached to regular household surveys in these countries. These modules were collected in the fall of 2005 in Argentina and Bolivia, and in the fall of 2006 in Colombia and the Dominican Republic. The analysis presented in this volume is just a launching point for new work that can be carried out with the rich data in these surveys.

Depending on the country, the surveys covered four main sets of issues: (1) questions on motivations for employment category choice and participation in social security and health insurance plans; (2) questions related

to the informality of the productive units and the motivations to participate or not participate in formal institutions; (3) questions on informal-formal occupational transitions through all other categories except the current one (“When was the last time you’ve been . . . independent worker/registered worker/unregistered worker?”);¹ and (4) questions on access to social protection mechanisms through private means or publicly provided benefits programs other than social security.

Note

1. The third set of questions will allow analyzing long-term labor market dynamics, which has been impossible with available short-term panel data.

Table 2.5 shows the data from the most recent surveys in Greater Buenos Aires, and urban areas of Bolivia, Colombia, and the Dominican Republic. Except for Colombia, over 70 percent of independent workers are voluntary, in the sense that they would rather be independent if they were able to choose their job. In Colombia, by a different measure, only 41 percent of urban independent workers can be considered voluntary; they reportedly would not take a formal salaried job with earnings equal to the earnings in their current job. When asked if they would take the same formal job but with lower earnings (a stricter standard), 71 percent of the Colombian informal self-employed said they would not. These findings are also remarkably in line with those for Mexico and Brazil based on very different surveys, as well as the findings

of the sociological literature. For instance, more than two-thirds of the Brazilian informal self-employed in the early 1990s reported that they would not take a formal salaried job, and less than 20 percent in Mexico reported involuntary reasons (see below).¹⁹ Moreover, over half of *salaried* workers in Bolivia and the Dominican Republic and close to one-third in Argentina and Colombia have intrinsic preferences for independent work, consistent with Blanchflower and Oswald’s (1998) findings for the United Kingdom, the United States, and Germany. The informal salaried show somewhat stronger preferences for independent work than the formal salaried, especially in Bolivia.

A comparison of the reported desires of self-employment with comparable international data shown in figure 2.10

TABLE 2.5

Preference for independent employment (percent of workers)

Country	All salaried	Formal salaried	Informal salaried	All independent	Formal independent	Informal independent
Argentina	37	38	43	70	86	60
Bolivia	55	41	62	73	65	74
Colombia	36	34	40	41	46	41
Dominican Republic	53	51	57	75	85	75

Sources: Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), based on household survey, 2005, 2006 data.

Note: Independent workers include self-employed and the small fraction of microfirm owners in each country. Except for Colombia, percentages are based on responses to the question, “If you were able to choose, would you rather be a salaried or independent worker?” For Colombia, the figures are based on (1) independent workers reporting they would not take a salaried job with benefits and the same earnings they get as independent; and (2) salaried workers reporting the main reason for being in their current job was their inability to become independent.

gives further credence to these findings.²⁰ First, the rates of desired self-employment are not out of line with the average in this sample of countries. Second, the four Latin American and Caribbean countries considered (the only ones with such data) are among those with the lowest divergence between desired and actual self-employment rates. That is, self-employment rates in these four Latin American and Caribbean countries are largely in line with workers' inclinations to independent work. In a sense, many more workers become self-employed than what one would expect, given the credit and human capital constraints of these countries. This may be a result of a host of factors, including the low opportunity cost of salaried work (formal or informal) and rigidities that create rationing of salaried jobs. In countries such as Colombia and Bolivia, this may lead many workers who lack the qualities to succeed as entrepreneurs to end up in self-employment; as noted earlier, the importance of tradition and the family business enters as an important consideration here.

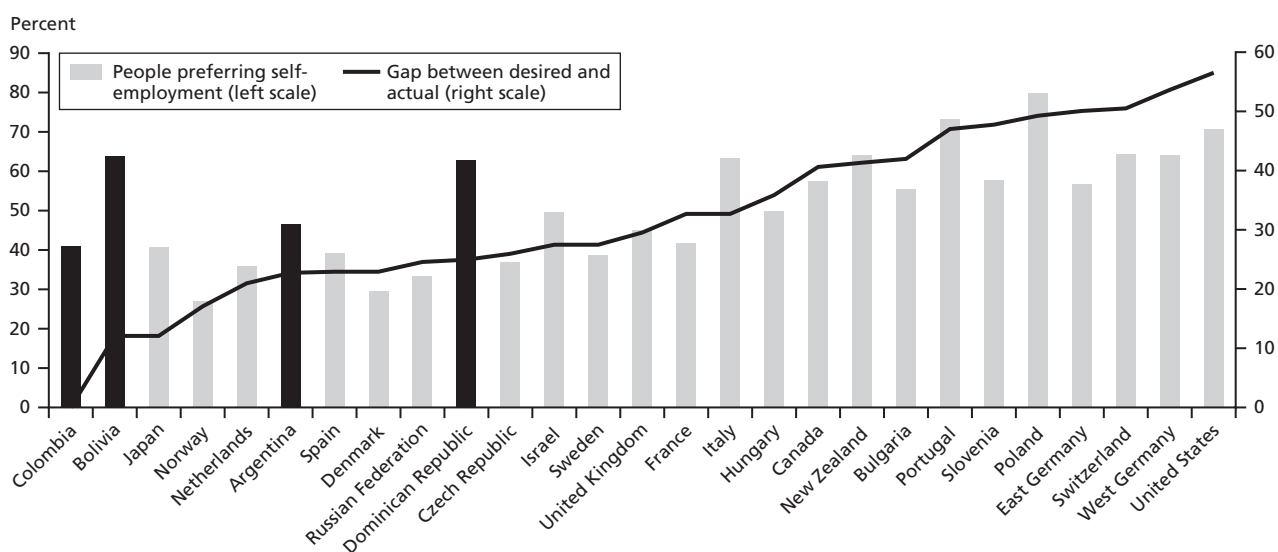
A noteworthy pattern of the data is the lack of a clear correlation between desires of self-employment and their materialization with level of development. While the United States has one of the highest rates of want-to-be-but-immateralized entrepreneurship, Japan appears in the other extreme. In fact, using these data, Blanchflower (2004) and Blanchflower, Oswald, and Stutzer (2001) find

few individual socioeconomic and demographic characteristics to be good predictors of self-employment desires. This suggests that entrepreneurship motives have a large idiosyncratic component, such as the reasons enumerated above, and that most workers found in this sector are likely to show strong preferences for its characteristics.

Nevertheless, table 2.5 shows that a significant fraction of the self-employed is involuntary and conforms to the traditional view of the sector. In particular, the fraction of informal independent that would rather be employees is roughly 40 percent in Argentina and 25 percent in Bolivia and the Dominican Republic, while a sizable 59 percent of the Colombian informal self-employed would rather take a salaried job with benefits over their current job. Again, by similar measures, in Brazil roughly one-third of the informal self-employed (45 percent among women) are involuntary.²¹

The results are further corroborated by the reported motivations for actually engaging in independent employment. Table 2.6 first presents the aggregate responses of formal and informal independent workers in the most recent surveys. Where workers were asked to report more than one reason, the table also shows the second-most important reason or the relative frequency of total responses. The fraction of informal independent workers who report loss or difficulties

FIGURE 2.10

International comparison of desired and actual self-employment rates

Sources: Based on data from Blanchflower 2004; Blanchflower, Oswald, and Stutzer 2001; and author's estimates for Latin American and Caribbean countries, using labor survey data, 2005 and 2006.

in obtaining a salaried job as their primary motivation is 25 percent in Bolivia, and 44 percent in the Dominican Republic, although it emerged as the second-most important reason for only 8 percent of cases in the latter. Lack of salaried employment was cited among the main reasons by 59 percent of the informal self-employed in Argentina and by 55 percent in Colombia, and accounted for 43 and 35 percent of their responses, respectively. All the remaining responses correspond to voluntary reasons in nature, such as higher earnings or mobility opportunities, greater autonomy, flexible hours, family tradition, or having the opportunity to become independent, and a variety of other reasons (for example, marriage, studying, the job being one's trade), and age (in Colombia). The rates of involuntary motivations among the formal self-employed compared to the informal are much lower in Argentina, somewhat lower in the Dominican Republic, and similar in Bolivia and Colombia.

The results are further corroborated by the earlier surveys in Mexico and Brazil. Tables 2.7 and 2.8 show the responses in these countries further disaggregated to gauge

differences in motivations by gender and age. Most workers in the two countries report reasons that imply the sector is a source of valued jobs, including higher earnings, greater flexibility, or family and personal motivations. In Mexico, 44 percent of males said they entered the sector for higher earnings, and just under 17 percent said they entered because they either lost a previous job or could not find a salaried formal job. Only 7 percent of women appear to be rationed out of salaried jobs. In Brazil, almost two-thirds of men and 44 percent of women reported that they were happily in the sector.

These findings are consistent with smaller sociological surveys that track workers across time. For instance, in Monterrey, Mexico, Balán, Browning, and Jelin (1973) find that being one's own boss was well regarded and that movements into self-employment from salaried positions often represented an improvement in job status. Figure 2.11 lends credence to these responses. Tracking Mexican workers from formal salaried jobs into informal self-employed jobs, on average voluntary movers gained around 15 percent while those who entered because of losing a formal job did not experience a statistically significant change in income.

TABLE 2.6

Distribution of the motivations/reasons for being in the current job as an independent worker (percent)

Motivation/reason	Argentina		Bolivia	Colombia		Dominican Republic	
	Cited as relevant	Percent of all responses ^a	Main reason	Cited as relevant	Percent of all responses ^a	Main reason	Secondary reason
Formal independent							
Could not find job as salaried	19.6	9.6	30.8	41.4	35.0	32.5	9.0
Autonomy/no boss	29.5	6.6	13.5	20.3	10.9	16.1	31.5
Flexible hours/less responsibility	34.3	7.0	8.3	23.9	17.6	10.7	24.5
Family tradition	21.5	11.3	6.8	2.3	2.5	3.7	4.1
Earns more	n.a.	n.a.	n.a.	20.3	5.2	26.4	5.3
Better mobility/benefits/prospects	59.4	31.3	n.a.	14.9	4.3	8.4	20.4
Able to become independent	66.8	25.8	24.9	n.a.	n.a.	n.a.	n.a.
Age	n.a.	n.a.	n.a.	21.1	16.5	n.a.	n.a.
Other	10.2	8.5	15.8	7.7	8.0	2.3	5.3
Informal independent							
Could not find job as salaried	58.8	43.0	25.3	55.3	34.5	44.3	8.1
Autonomy/no boss	16.5	3.2	9.5	22.2	16.4	16.2	33.5
Flexible hours/less responsibility	30.1	6.1	13.3	21.7	11.5	10.2	22.0
Family tradition	11.7	4.2	17.4	4.4	2.8	3.8	9.2
Earns more	n.a.	n.a.	n.a.	11.5	9.7	14.0	7.8
Better mobility/benefits/prospects	29.6	15.0	n.a.	7.4	5.7	7.8	12.4
Able to become independent	43.0	15.9	26.0	n.a.	n.a.	n.a.	n.a.
Age	n.a.	n.a.	n.a.	23.2	13.5	n.a.	n.a.
Other	15.5	12.6	8.5	10.8	5.8	3.7	6.9

Sources: Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), based on household survey, 2005, 2006 data.

Note: n.a. = not applicable. In Argentina, individuals could mark all relevant response options; in Colombia, they were asked for the two most important reasons, but without distinguishing first and second, unlike the case of the Dominican Republic.

a. Computed as the ratio of the frequency with which each reason is reported to the total number of valid responses in the sample.

TABLE 2.7

Reported reasons to be informal self-employed in Mexico (percent of workers)

Reason	Male					Female				
	All ^a	15–18 years	19–24 years	25–54 years	55–70 years	All ^a	15–18 years	19–24 years	25–54 years	55–70 years
Family tradition	8.9	6.0	17.4	8.3	10.0	3.6	9.8	8.4	2.8	4.5
Complement family income	18.0	40.2	13.9	17.1	20.9	65.1	21.3	59.4	64.5	74.7
Higher income than salaried	43.7	15.2	41.8	47.0	31.5	12.2	7.5	9.9	13.1	6.9
No salaried job	12.4	17.9	8.3	11.2	18.3	6.1	22.4	5.0	5.7	8.0
Flexible hours	3.8	9.4	9.1	3.4	3.3	4.2	11.5	9.0	4.6	1.2
Fired or lost job	4.0	0.3	0.8	4.2	5.5	0.6	0.0	0.3	0.8	0.2
Other	9.1	11.1	8.7	8.8	10.6	8.1	27.5	8.0	8.6	4.5
Percent of sample	68.1	0.7	4.6	49.4	11.3	31.9	0.2	2	23.6	5.1
Percent of gender	100.0	1.0	6.7	72.6	16.5	100.0	0.7	6.2	73.8	16.0

Source: Author's estimates, based on *Encuesta Nacional de Micronegocios* 1994.

a. Considers all self-employed (including people under 15 years of age and above 70) who respond to this question.

TABLE 2.8

Reported reasons to be informal self-employed in Brazil (percent of workers)

Would you like to leave [your current] job for a job with a signed work contract? (Self-employed)										
Response	Male					Female				
	All ^a	15–18 years	19–24 years	25–54 years	55–70 years	All ^a	15–18 years	19–24 years	25–54 years	55–70 years
No	67.9	29.9	52.6	68.3	80.6	55.5	24.6	41.8	55.2	71.2

Motivations to prefer an unprotected job (Self-employed)										
Reason	Male					Female				
	All ^a	15–18 years	19–24 years	25–54 years	55–70 years	All ^a	15–18 years	19–24 years	25–54 years	55–70 years
Earn more in current job	18.0	13.4	17.6	21.0	9.6	10.6	5.1	13.3	12.1	3.9
Needed to care for home	0.2	0.0	0.0	0.1	0.4	26.9	15.3	22.8	27.5	28.8
Need time for other activities	2.9	7.5	3.2	2.5	3.2	6.7	6.8	7.9	6.6	6.8
Happy in current job	64.9	69.5	68.3	64.0	67.6	44.1	59.3	47.7	44.8	39.0
Did not want the commitment	10.1	8.0	9.8	10.0	10.6	7.5	8.5	7.6	6.9	9.8
No answer	4.0	1.6	1.1	2.2	8.6	4.1	5.1	0.8	2.2	11.6
Percent of sample	73.5	0.8	5.0	51.3	13.7	26.5	0.3	1.7	20.0	3.9
Percent of gender	100.0	1.1	6.8	69.8	18.6	100.0	1.0	6.3	75.4	14.8

Source: Author's estimates, based on *Pesquisa Nacional por Amostra de Domicílios* 1990.

a. Considers all self-employed (including people under 15 years of age and above 70) who respond to this question.

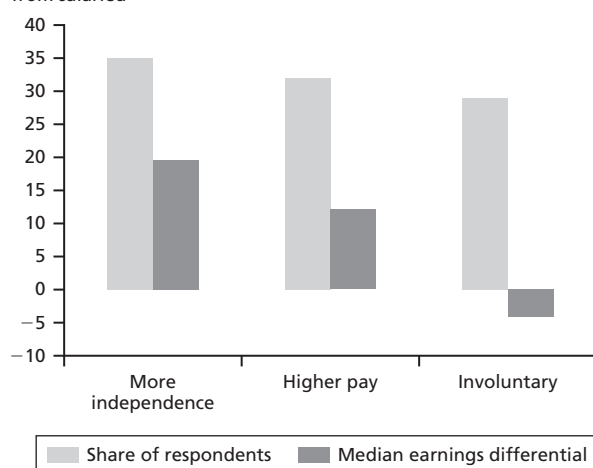
The results also confirm the advantages of self-employment for many women. In Mexico, almost two-thirds of women in self-employment reported the need to supplement family income, and only 7 percent appear to be rationed out of salaried jobs. Unlike men, Mexican women show a sharp increase in responding that they have entered self-employment to complement family income after the age of

19, above 60 percent compared to roughly 15–20 percent for men. In Brazil, one-third of women alluded to the need for flexibility to attend to family and personal activities. Roughly 25 percent of Brazilian women over the age of 19 say they entered self-employment to care for the home, compared to virtually none of the men. This correlation of rising self-employment with moving into marriage age is also

FIGURE 2.11

Earnings gain from voluntary movement to informality in Mexico

% hourly wage gain for males entering informal self-employment from salaried



Sources: *Encuesta Nacional de Micronegocios*, linked to *Encuesta Nacional de Empleo Urbano*, 1992; Maloney 1999.

consistent with the effect of having coverage through their spouses, as suggested by the results of Galiani and Weinschelbaum (2006). Again, a large part of home-based work may also emerge from this dynamic. In no way does the evidence here intend to validate the division of labor in the household, but simply to suggest that women's patterns of employment are consistent with their pressing need to balance family and work-life—whether or not this is largely a socially ascribed gender role.

Furthermore, tables 2.7 and 2.8 offer supporting evidence to the life-cycle nature of self-employment. Of the very few (1 percent) of male self-employed in the age range of 15–18 in Mexico, most are entering to complement family income and almost 30 percent cite lack of salaried jobs. For Brazil, of the similarly small fraction of self-employed in the same age group, about 70 percent would prefer a formal salaried job, unlike the older age groups that increasingly prefer self-employment.²² In sum, self-employment is not a viable or good alternative for most youth to the extent they still lack the experience and capital to have a chance to succeed as entrepreneurs.

Finally, the motivations of Brazilian and Mexican self-employed offer mixed evidence on the rationing out of middle-aged and older workers from formal salaried jobs. In Mexico, only 9 percent of prime-age males 19–24 years old report entering the sector because they either lost or

could not find a formal job. But this steadily rises with age to 15 percent for males aged 25–54 and finally to 24 percent for male workers aged 55–70. Involuntary entrants continue to represent a minority of entrants, but there is some evidence of the sector serving increasingly as a sector of last resort for older workers. In Brazil the trend is exactly the reverse, however; the oldest age group (55–70) reports being least interested in finding a formal salaried job at 81 percent. This is consistent with many of these workers not interested in having a covered job if they can no longer assure sufficient years of formal experience for a pension. The age and gender breakdowns of the survey responses for the other four countries (not shown) studied are broadly consistent with the results for Mexico and Brazil, with different gradients across countries.

Informal salaried work

Although many of the considerations for why workers might prefer to be informal apply to the informal salaried as well, overall this sector shows a much larger component of involuntary entry consistent with the exclusion view. Until now, the information on the informal salaried was much scarcer; the informality modules conducted in the four countries above included special components designed precisely for them.

Table 2.9 suggests that, in contrast to the self-employed, the majority of informal salaried workers appear to be involuntarily in their jobs, although not necessarily queuing for formal salaried employment. The inability to find a better job constitutes a much higher fraction of the reported reasons for being in informal salaried jobs than for the formal salaried: 48.4 versus 22.4 percent in Argentina; 64 versus 32 percent in Bolivia; 43 versus 16 percent in Colombia; and 40 versus 22 percent in the Dominican Republic. These are consistent with responses from the Brazilian informal salaried in 1990 (table 2.10) that roughly 70 percent would rather have had a formal salaried job, although some care is in order since, as chapter 4 will show, this was a time of extreme segmentation in the Brazilian labor market.

Although the traditional queuing view enjoys substantial support, some clarifications are in order. While higher than among the formal salaried, the inability to find a better formal salaried job is not the overwhelming majority response among informal salaried workers. For instance, in the case of the Dominican Republic, only 17 percent cite this when asked about the second main reason. This explains why, as will be shown in chapter 3, the 18-percentage-point

TABLE 2.9

Distribution of the motivations/reasons for being in the current salaried job (percent)

Motivation/reason	Argentina		Bolivia	Colombia	Dominican Republic	
	Cited as relevant	Percent of all responses ^a	Main reason	Main reason	Main reason	Secondary reason
Formal salaried						
Could not find another job	50.7	22.4	31.7	16.2	21.5	9.4
Independent work is unstable	42.7	18.9	11.3	13.3	24.4	15.7
Less responsibility	n.a.	n.a.	n.a.	0.7	1.0	5.0
Earns more than as independent	29.3	12.9	15.1	24.8	16.8	26.9
Better mobility/benefits/prospects	69.3	30.6	n.a.	8.5	16.7	22.8
Unable to become independent	25.5	11.2	15.0	33.7	17.4	17.3
Other	8.9	3.9	26.9	2.9	2.3	3.0
Informal salaried						
Could not find another job	87.3	48.4	64.2	43.0	39.8	16.8
Independent work is unstable	23.5	13.0	4.5	6.4	20.0	13.1
Less responsibility	n.a.	n.a.	n.a.	0.8	1.6	8.6
Earns more than as independent	17.9	9.9	4.8	3.9	5.9	13.7
Better mobility/benefits/prospects	17.4	9.6	n.a.	2.3	7.5	13.9
Unable to become independent	26.1	14.5	15.7	40.0	22.3	29.1
Other	8.1	4.5	10.8	3.7	3.0	4.9

Sources: Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), based on household survey, 2005, 2006 data.

Note: n. a. = not applicable.

a. Computed as the ratio of the frequency with which each reason is reported to the total number of valid responses in the sample.

TABLE 2.10

Reported reasons to be informal salaried in Brazil (percent of workers)

Would you like to leave [your current] job for a job with a signed work contract? (Informal salaried)										
Response	Male					Female				
	All ^a	15–18 years	19–24 years	25–54 years	55–70 years	All ^a	15–18 years	19–24 years	25–54 years	55–70 years
No	30.3	22.8	25.0	31.8	55.3	37.4	25.5	28.5	43.5	70.1
Motivations to prefer an unprotected job (Informal salaried)										
Reason	Male					Female				
	All ^a	15–18 years	19–24 years	25–54 years	55–70 years	All ^a	15–18 years	19–24 years	25–54 years	55–70 years
Earn more in current job	6.2	3.1	7.6	10.6	2.4	2.2	1.7	3.0	2.6	0.8
Needed to care for home	0.2	0.2	0.0	0.1	0.0	18.2	4.7	9.5	25.7	24.3
Need time for other activities	7.1	8.5	5.1	3.5	4.4	6.7	8.7	5.3	5.4	3.7
Happy in current job	67.1	74.1	74.7	68.7	57.1	60.3	73.6	71.6	55.6	47.6
Did not want the commitment	9.8	6.6	8.3	10.8	16.6	7.8	6.7	7.7	7.1	13.7
No answer	9.7	7.6	4.3	6.2	19.5	4.8	4.6	2.9	3.6	9.8
Percent of sample	56.0	9.7	9.6	20.3	5.9	44.0	6.1	6.8	22.3	4.2
Percent of gender	100.0	17.3	17.2	36.2	10.5	100.0	13.8	15.4	50.8	9.4

Source: Author's estimates, based on *Pesquisa Nacional por Amostra de Domicílios* 1990.

a. Considers all informal workers (including people under 15 years of age and above 70) who respond to this question.

difference does not prevent these workers from reporting similar levels of welfare and job satisfaction than formal employees. Further, while similar responses are not available for Mexico, rates of on-the-job search, a measure of dissatisfaction, are only somewhat higher than for the largely voluntary self-employed.²³

More important, there are numerous indications that while a majority of informal salaried workers in Argentina may be in fact queuing for formal salaried employment, in the other countries many lean more toward independent employment, most of which is informal. As shown in table 2.9, the inability to become independent accounts for 15 percent of their reported reasons for being in their current job in Argentina and Bolivia, similar to the formal salaried. In contrast, in Colombia and the Dominican Republic, these represent about 40 percent of all responses, about seven points higher than for the formal salaried. Moreover, table 2.5 shows that the informal salaried have higher rates of overall preference for self-employment than the formal salaried: in Bolivia (62 versus 41 percent); the Dominican Republic (57 versus 51 percent); and in Argentina (62 versus 57 percent). Finally, the majority of “other reasons” reported by the informal salaried for being in their current job are related to the exercise of their trade, apprenticeships, and flexibility.

Finally, there are clear differences across demographic characteristics in the view of being in the sector. Table 2.10 breaks down motivations by age and gender for the Brazilian informal salaried. Younger workers show the higher discontent with being informal salaried at about 80 percent. Meanwhile, half of male workers and 70 percent of females close to retirement (55 and above) show no inclination to move to a formal salaried job. This may arise because of quality of life issues—formal jobs being excessively demanding for older individuals (Gonzalez de la Rocha 1994; Lorenzen, Selby, and Murphy 1990), the realization of the limited use in further contributing to pensions, or perhaps now being able to take advantage of the family benefits enjoyed by their prime-age children in the formal sector.

Even among prime-age workers, 44 percent of females age 25–54 are inclined to stay as informal salaried and among these, again flexibility to attend family life gains importance in marriage age (25 percent). Another logic applies to domestic workers, a main component of this group, although with nagging tradeoffs. Interviews with poor working-class Argentine women reveal that it is not uncommon for unskilled women to leave formal sector jobs

that may be better paid to enter (or return to) domestic service, citing flexibility, in terms of work hours and pay schedules, and security in terms of a source of emergency income, networks to other sources of employment, and a general feeling of connection with the employer (Geldstein 2000). However, it is also reported that this is a vulnerable job, in which employer abuse, isolation, few chances for organization, and limited opportunities for career advancement are a reality.

The dissatisfaction of younger informal salaried workers does reveal that few see the sector as a career job. Only 2 percent of informal salaried workers in Colombia, 5 percent in Bolivia, roughly 10 percent in the Dominican Republic, and 17 percent in Argentina consider their job as preparation for formal salaried employment or as offering better opportunities for social mobility. These are broadly corroborated by the responses from the Dominican Republic of unpaid workers where roughly 22 percent see their unpaid work, for 73 percent their first job, as preparation for the future. This is only partially consistent with Hemmer and Mamel’s (1989) asserted role of training in the sector. In any case, consistent with previous results for Mexican enterprises, table 2.11 shows that most unpaid workers are related to the owner and paid in kind, and that helping out the family is the overwhelming motivation for working without pay.

In sum, while the informal salaried have a majority component of discontented workers corresponding to the more classic queuing view of the sector, it is also a heterogeneous sector, and the magnitude of each subcomponent would vary across countries and time. All the observations made earlier are relevant here. The sector comprises young workers

TABLE 2.11

Principal reason for working without monetary compensation in the Dominican Republic (percent of workers)

Reason	Distribution
Learning a profession	8.5
Helping family	62.4
Getting experience to get a better job	13.4
Paying a debt to the owner of the firm	1.3
Compensated with food, housing, or apparel	5.1
Compensated with partial or full payment for studies	2.8
Other	6.5

Source: Authors’ estimates, based on employment survey data, 2006.

possibly rationed out by labor and regulatory distortions, and those entering the labor force, shopping around for jobs and being frustrated by the lack of a track record and information asymmetries in the labor market, and some women who find more flexibility in microfirm salaried jobs. Again, labor or regulatory distortions that curtail formal salaried employment are not the only factors that may be behind the largely involuntary nature of the sector. The final section of this chapter examines more closely the reasons that cause workers in both informal sectors to be without old-age and health protections in their jobs.

Workers' motivations to lack old-age and health protection

When asked explicitly about the reasons for not having pension and health insurance benefits, a large fraction of informal salaried workers report relatively external constraints. As shown in table 2.12, when the informal salaried in Argentina were asked whether they lack social security benefits because their employers did not offer them or because they reached a consensual agreement to obtain higher earnings, 95 percent said it was the former reason. In Bolivia, faced with a larger number of response options, about

42 percent of the informal salaried cited their unawareness about how private pensions work as the prime reason for being out of the social security system; one-quarter said their employers did not offer the benefit, and another quarter said their income was too low. In Colombia, 30 percent reported that their employers did not offer the benefit or that most available jobs are like this; 56 percent said their income was too low, and only 10 percent reported reasons that reflect an explicit opting out. In the Dominican Republic, 42 percent reported similar employer or labor market constraints (including 5 percent unaware of their pension rights); 23 percent alleged insufficient incomes, but one-third actually cited voluntary reasons. In all four countries, only a small fraction (5–7 percent) said that the main reason for not being affiliated was low expected benefits or lack of trust in the social security system.

Likewise, most of the informal self-employed alluded to low incomes or lack of knowledge as the main reason for not contributing to the pension system. Insufficient income to afford a pension plan is reported by three-fourths of independent workers in Argentina and Colombia, half in the Dominican Republic and one-quarter in Bolivia. Meanwhile, 55 and 20 percent of the informal self-employed in Bolivia and the

TABLE 2.12

Main reasons why the informal do not contribute to social security (percent of workers)

Reason	Argentina		Bolivia		Colombia		Dominican Republic	
	Salaried worker	Independent	Salaried worker	Independent	Salaried worker	Independent	Salaried worker	Independent
Too young, too far in the future	n.a.	n.a.	n.a.	n.a.	3.8	2.4	5.0	3.4
Prefers higher earnings	4.5	1.7	n.a.	n.a.	1.7	1.3	11.8	1.8
Income too low	n.a.	76.0	25.4	26.0	55.8	74.8	22.9	51.2
Not worth it	n.a.	3.0	n.a.	n.a.	n.a.	n.a.	2.9	1.9
Employer only offered this job	95.5	n.a.	23.7	n.a.	11.8	0.6	21.0	n.a.
Lack of trust	n.a.	4.3	6.9	15.4	2.0	6.2	2.4	5.6
Jobs are like that	n.a.	n.a.	n.a.	n.a.	4.6	1.8	15.7	n.a.
Unaware of right/ could contribute	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	5.2	19.4
Employer does not require it	n.a.	n.a.	n.a.	n.a.	13.4	1.3	n.a.	n.a.
Children, spouse will look after	n.a.	n.a.	n.a.	n.a.	0.3	2.1	0.6	10.1
Prefers own savings	n.a.	n.a.	n.a.	n.a.	0.8	2.5	4.3	2.3
Don't know how system works	n.a.	n.a.	42.1	55.3	n.a.	n.a.	n.a.	n.a.
Other	n.a.	15.0	1.9	3.3	5.8	6.9	8.3	4.5

Sources: Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), based on household survey, 2005, 2006 data.

Note: n.a. = not applicable.

Dominican Republic, respectively, alleged lack of knowledge about the functioning of private pensions as the prime reason for not contributing to social security. Among the 20–30 percent of the informal self-employed who reported explicit voluntary reasons, only 6–8 percent in Colombia and the Dominican Republic and 15 percent in Bolivia alluded to low expected benefits or lack of trust in the social security system as the main reason not to be affiliated. Most voluntary reasons reported in Colombia and the Dominican Republic were relying on their spouses or children to cover their living expenses in old age, being too young to be concerned about this, and a preference to earn more now or save on their own. However, a small fraction (10–30 percent) of those in jobs not contributing to pensions in all four countries reported that they were using other means (e.g., savings, investment in assets) to prepare for old age.

Some caution is in order in approaching these responses. First, it is hard to distinguish true credit constraints from classic myopia or inertia in the mainstream pension literature. Turner and Verma (2007) show that of those who were eligible to participate in 401K plans in the United States but did not, 40 percent responded that they could not afford to. Yet these are workers earning several orders of magnitude the income of Latin American workers. Further, automatic enrollment substantially increased their participation; many allegedly income-constrained workers did not opt out, thus suggesting that for many the contributions were, in fact, quite manageable. The explanation for this is some mixture of inertia and myopia. Whether, in fact, informal workers in the countries surveyed here or elsewhere in the region would do the same, or whether they would still opt out into informality, we cannot know. However, in several countries of the region, including Chile, Uruguay, and to a lesser extent today Argentina, many low-income formal salaried workers are able to contribute to pension plans. In fact, as will be shown in chapter 3, the Dominican self-employed have average earnings similar to those of formal salaried workers, suggesting that, in principle, contributions to the system should be as affordable. Thus, the response of “lacking income” in these surveys cannot be taken at face value and should be plausibly viewed as voluntarily opting out of old-age savings, combining a mixture of unaffordability due to economic hardship, inertia, and myopia.

Second, although low expected returns and trust in the pension system do not appear prominently as main reasons, this does not negate that many workers in these countries might in fact have concerns with the quality and reliability

of the benefits and services mandated in formal labor contracts. As will be discussed in chapter 8, other survey evidence reveals that trust and satisfaction with public institutions are very low, and in some cases are declining, in these and other countries in the region, and that this issue may cultivate a social norm of noncompliance with taxes and regulations. While the survey responses indicate that a perceived inadequacy of benefits is not a first-order driver of workers’ informal status, they cannot tell us whether this factor could be a secondary but significant driver of informality (for example, if income and informational constraints were not binding).

Table 2.13 shows that the reported reasons why the informal salaried lack health insurance coverage in Colombia and the Dominican Republic (the only two countries that collected such data) reflect a much larger degree of voluntary choice. About 82 percent of salaried workers without pension benefits in Colombia and three-fourths in the Dominican Republic do not have health insurance coverage on their own, in contrast to the almost full coverage of the formal salaried. In the two countries a small fraction of informal salaried workers (20 percent in Colombia, 11 percent in the Dominican Republic) reported the employer not offering the benefit as the main reason, and again, a significant fraction (23 and 46 percent, respectively) cited very low incomes as the main reason.

Among those reporting clear opting-out reasons, 44 percent in Colombia said they have the benefit through another family member or the public subsidized regime, and 30 percent in the Dominican Republic reported a host of reasons ranging from preferring to cover health costs on their own (10 percent), having the benefit through another family member (6 percent), and relying on the public health system (7 percent). In both countries, less than 5 percent reported low coverage or quality of health services as the main reason not to be affiliated.

Low incomes and the preference for other alternatives also emerged as the main reported reasons of independent workers in Colombia and the Dominican Republic to lack health insurance coverage. Between 12 and 14 percent of independent workers without pension plans in Colombia and the Dominican Republic had other types of health insurance coverage, compared to three-fourths and two-thirds of the formal self-employed, respectively. Around half of the informal self-employed in the Dominican Republic and one-third in Colombia said they do not have health insurance because they could not

TABLE 2.13

Main reasons why workers do not contribute to health insurance (percent of workers)

Reason	Colombia			Dominican Republic		
	Informal salaried	Formal independent	Informal independent	Informal salaried	Formal independent	Informal independent
Health service inefficient/ low coverage	0.4	1.1	1.0	1.1	1.0	1.0
Health service low quality	0.4	4.3	0.6	1.6	0.0	1.1
Beneficiary through relative	44.0	57.9	53.7	6.0	19.0	5.6
Prefers higher earnings	1.3	0.5	0.5	2.6	0.0	1.1
Employer does not offer it	9.9	0.4	0.6	6.7	1.1	0.4
Employer does not require it	10.0	0.5	1.2	4.7	0.0	0.8
Don't know how to enroll	0.6	0.1	1.0	4.0	11.5	7.7
Too low incomes	23.4	11.7	33.2	45.5	41.5	49.1
Temporary situation	2.3	0.4	0.9	4.6	1.8	2.8
Prefers self-coverage	1.1	7.6	1.9	10.2	12.4	14.2
Has private insurance	0.3	4.8	0.7	n.a.	n.a.	n.a.
Attend public hospitals	n.a.	n.a.	n.a.	6.7	7.1	11.5
Insurance companies rejection	n.a.	n.a.	n.a.	0.5	0.0	0.5
Other	6.2	10.9	4.6	5.8	4.6	4.2

Sources: Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), based on household survey, 2005, 2006 data.

Note: n.a. = not applicable.

afford it, compared to 42 and 12 percent of formal self-employed, respectively.

About 54 percent of the Colombian informal self-employed and 58 percent among formal cited having the benefit through another family member or the public subsidized regime. The Dominican informal self-employed reported a host of reasons, including preferring to cover health costs on their own (14 percent), reliance on the public health system (12 percent), lack of information on how to enroll (8 percent), and being a beneficiary of another family member (6 percent). These are very similar for the formal self-employed, except for their heavier reliance on coverage through relatives. In both countries barely 2 percent said they were not affiliated due to low coverage or quality of health services.

Again, considerations similar to the above apply to the interpretation of the “low income” responses and the low fraction of workers citing overriding concerns with the quality of health services. In this case, risk aversion behavior becomes relevant. The fact that many workers resort to self-coverage or are happy with the protection offered by universal health services suggests that these are judged to be adequate as compared to privately provided health insurance plans considering the cost of the latter. In fact, as chapter 3 will show, the lack of pension and health cover-

age does not prevent informal workers in the Dominican Republic from reporting levels of job satisfaction equal to those of formal salaried employees.

The policy implications of the reported significance of low incomes for low pension and health insurance coverage depend on the alternative interpretations. If workers are unable to save for their retirement because they are forced to spend their incomes entirely on necessities, forcing them to save (through mandatory savings programs such as social security) may reduce their welfare during working years, and, therefore, well-targeted social insurance pillars would be more advisable. However, to the extent that this behavior partially reflects myopia or inertia in the savings behavior of some workers for retirement, there is an important role for programs that mandate or provide incentives for voluntary savings. In any case, whether workers end up with insufficient savings for old age because they consume “excessively” during their earnings years or because they have low long-term income potential, there is some justification for some government intervention to influence individuals’ savings and retirement decisions or to provide minimum protection from poverty at old age by pooling efforts through the tax system. These issues will be discussed in detail in chapter 7.

Furthermore, the significance of information failures and the availability of alternatives to substitute for mandated

benefits, either publicly provided or through private means, require policy attention. It is imperative to ensure a good design of social protection and social assistance systems to minimize distorting individuals' employment behavior, as will also be discussed in chapter 7.

Finally, to the extent that the decisions of firms to operate informally are important drivers of the informality of salaried workers, as will be discussed in chapters 4 through 6, improved design and enforcement of labor, tax and business regulations, and incentives for firms to find benefits in participating in formal institutions have an important role to play in addressing informal employment.

Conclusions

This chapter first laid out several conceptual and empirical approaches to understanding the *razón de ser* of the informal worker. It then brought to bear two sets of empirical tools to begin to distinguish among these approaches. The first approach moved beyond simple tabulations of stocks of workers in different sectors to studying the dynamics of their movements through the labor market, and what these sectoral transitions can tell us about why workers are where they are. The second analysis relied on the reasons workers report for being in their current job with or without social protections.

In the cases of both the informal salaried and self-employed, there is evidence of substantial heterogeneity of motives and demographic characteristics. However, the characters of the two sectors emerge as very distinct, lending credence to both the exclusion and voluntary nature of informal employment.

Evidence from workers' patterns of mobility and reported motivations to be in their current occupations lends support to the view that the majority of independent workers are largely voluntary and attach significant value to the nonpecuniary benefits of autonomous work (although an important fraction conforms to the traditional view of the sector as employment of last resort). Self-employment is concentrated among older workers and is not an entry point of work.

Meanwhile, the majority of informal salaried workers appear to be involuntary, although there is also significant heterogeneity of characteristics (most are young) and motives. This may not necessarily imply that they are queuing for preferred, better-quality formal salaried employment; although this seems to be the case for a majority in Argentina, a significant fraction of the informal

salaried in the other countries leans more toward independent employment, most of which is informal, as a desired destination.

The next chapter revisits this question and the more general question of the role of the different sectors by studying whether workers' stated voluntary or involuntary motivations to be in informal employment conform to their earnings performance and, ultimately, their welfare. Chapter 4 on labor dynamics will introduce additional evidence exploiting transitions and patterns across the business cycle to more fully flesh out the nature of the informal sector.

Notes

1. The Harris-Todaro (1970) model is perhaps the traditional statement of this view. Also see Fields 1990; Peattie 1987; Portes and Schaffler 1993; Tokman 1992; and Turnham and Eröcal 1990.

2. See also Krebs and Maloney (1999) for an efficiency wage model applied to Mexico.

3. Schneider and Enste (2000) linked Hirschman's "exit" and "voice" to the causes of the informal economy, but their emphasis was on opting out driven by the desire to avoid dealing with excessive state burdens, rather than on the perceived inadequacy of the benefits that formality entails.

4. Fields (1990) argued for two tiers, with a minority upper tier that is voluntary and prosperous. Yamada (1996), Maloney (1999), Saavedra and Chong (1999), and Mondino and Montoya (2002) have argued with evidence from Peru, Mexico, Brazil, and Argentina that the majority of the informal self-employed are voluntary.

5. The interviews by sociologists Balán, Browning, and Jelin (1973) in Mexico suggest that the paucity of openings for promotion on the rigid *escalafón* (seniority-driven job ladder) can make informal self-employment the remaining outlet for further advancement.

6. In fact, even in the United States, some workers leave large firms to open their own, taking on the responsibility for health care and pensions through private or independent means.

7. The fraction of households receiving financial help from relatives in other households is 38 percent in Spain and 11 percent in Italy versus 1 percent in the United Kingdom and 6 percent in the United States.

8. Maloney (1999) standardized probabilities both on initial turnover and on final sector size. Bosch and Maloney (2005) standardized on initial separations.

9. As one example, broadly replicated in the other surveys, the National Urban Employment Survey (ENEU) from Mexico shows that transitions into self-employment from the other paid sectors occur four to six years later than transitions into formal or informal salaried work, leaving the mean age eight years higher than the next closest sector (see Maloney 1999; Maloney and Aroca 1999).

10. See Lopez-Castaño (1990) in Colombia, and Fields (1990) and Peattie (1982), who find a tendency for employees of large firms to leave and open their own businesses. Aroca and Maloney (1997)

find confirmation in a logit analysis for Mexico with rotating panel data.

11. This mixed story is put forward by Guadalajara sociologist Gonzalez de la Rocha (1994), who argues that, for many older workers, self-employment provides a safety net by offering insecure occupations (such as the services) in which their age is not a limitation after they have been kicked out of the formal manufacturing or formal services, while also suggesting some degree of voluntary movement in that older men may also find the pace of industrial (formal) work too arduous and leave such jobs. The voluntary take is stressed by anthropologists Selby, Murphy, and Lorenzen (1990) who note the “surprising desirability of informal sector employment as the basis for a household earning strategy, particularly for poorer, older households with lower educational qualifications” (p. 144).

12. Fajnzylber, Maloney, and Rojas (2006) find that Mexican firms have mortality rates of the same order of magnitude as those found in the United States by Evans and Leighton (1989); the 18 percent exit to wage work (measured as the fraction of self-employed workers moving to wage work) compares to 14 percent across a year. This is consistent with much higher entry rates as a fraction of wage workers, averaging 8 percent in Mexico compared to the 4 percent found by Evans and Leighton (1989) for the United States. Among young workers 20–28 years old, failure rates are roughly equivalent at around 50 percent, while for prime-age workers, 31–35, failure rates are roughly equivalent at 20 percent. See also Levenson and Maloney (1996).

13. While the above findings do not preclude discrimination against married women or those with children if, for example, an employer fears that the women may be likely to be absent for work for long periods (Chant 1992), further evidence favors the idea that many of them choose self-employment on the basis of its desirable characteristics. In fact, evidence from Argentina, Mexico, Costa Rica, and Brazil shows that, when disaggregating the self-employment sector, a clear subgroup of young mothers emerges, along with the subgroups of new entrepreneurs, family business owners, and those who cannot find work elsewhere. The motivational responses discussed later lend more credence to these conclusions.

14. Balán, Browning, and Jelin (1973) note that, in Mexico, this period of life for young workers is one of “shopping around” and trying out various possible life choices.

15. This is consistent with findings that the largest determinant of whether a girl reports beginning to work in Mexico is whether her mother opened a microenterprise (Cunningham and Maloney 2001).

16. This issue is very similar to the arguments and findings of the literature on the importance of re-incidence of unemployment for unemployment durations. See, for example, Clark and Summers (1979) and Akerlof and Main (1980) for the United States and Galiani and Hopenhayn (2003) for Argentina.

17. Note that the proportion figures in table 2.4 are not directly comparable to the “raw” transition probabilities presented in the first bar on each pair of bars in figure 2.4 because the former are computed taking the total number of workers in each occupational and age group as denominators, whereas the latter divides by the total number of workers.

18. These results are very close to Gottschalk and Maloney’s (1985) finding that roughly 70 percent of U.S. job changes are voluntary. Put differently, the implied rates of involuntary entry into self-employment would be normal by the standards of a flexible industrial country market.

19. Survey data for other countries are also broadly consistent. For instance, Itzigsohn (2000) reports that most informal entrepreneurs in Costa Rica and the Dominican Republic preferred their jobs to being a formal employee in a special export zone. In Paraguay, only 28 percent of those in the informal sector stated a desire to change occupations (DGE 1998). In Guatemala, only 31 percent of informal independent workers were willing to become formal employees (CIEN 2006).

20. These data come from surveys with the same or very similar questions on preferences for self-employment over salaried work to the adult (employed) population of the countries.

21. Interestingly, the fraction of formal independent workers who are involuntary is only 15 percent in Argentina and the Dominican Republic, but 35 percent in Bolivia and 54 percent in Colombia.

22. This disproportionately high level of youth reporting complementing income makes sense, given that around 2.5 percent of 15–18-year-old males are heads of their households, while 30 percent are in the next category.

23. Likewise among Paraguayan informal salaried workers, the percent looking for a better job was only slightly higher than the self-employed, at 32 percent (DGE 1998).

References

- Akerlof, G., and B. Main. 1980. “Unemployment Spells and Unemployment Experience.” *American Economic Review* 71 (5): 885–93.
- Amaral, P., and E. Quintin. 2006. “A Competitive Model of the Informal Sector.” *Journal of Monetary Economics* 53 (7): 1541–53.
- Arriagada, I. 1998. “The Urban Female Labour Market in Latin America: The Myth and the Reality.” Report LC/L 1034-P/I, Economic Commission for Latin America and the Caribbean, Santiago, Chile.
- Arias, Omar, and Monserrat Bustelo. 2007. “Profiles and Dynamics of Informal Employment in Latin America.” Photocopy. World Bank, Washington, DC.
- Arias, Omar, Fernando Landa, and Patricia Yáñez. 2007. “Movilidad Laboral e Ingresos en el Sector Formal e Informal de Bolivia.” Documento de Trabajo, UDAPE, La Paz, Bolivia.
- Aroca, Patricio, and William Maloney. 1997. “Logit Analysis in a Rotating Panel Context and an Application to Self-employment Decisions.” World Bank Working Paper 2069, World Bank, Washington, DC.
- Balán, J., H. L. Browning, and E. Jelin. 1973. “Men in a Developing Society.” Photocopy. Institute of Latin American Studies, University of Texas, Austin.
- Barr, A., and T. Packard. 2000. “Revealed and Concealed Preferences in the Chilean Pension System: An Experimental Investigation.” Discussion Paper 53, Department of Economics, Oxford University, UK.
- Becker, G. S. 1991. *A Treatise on the Family*. Enlarged ed. Cambridge, MA: Harvard University Press.

- Bentolila, S., and A. Ichino. 2000. "Unemployment and Consumption: Are Job Losses Less Painful near the Mediterranean?" Discussion Paper 2539, Centre for Economic Policy Research, London.
- Bewley, T. F. 1999. *Why Wages Don't Fall during a Recession*. Cambridge, MA: Harvard University Press.
- Blanchflower, D. G. 2004. "Self-employment: More May Not Be Better." *Swedish Economic Policy Review* 11 (2): 15–74.
- Blanchflower, D. G., and A. J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16 (1): 26–60.
- Blanchflower, D. G., A. J. Oswald, and A. Stutzer. 2001. "Latent Entrepreneurship across Nations." *European Economic Review* 45: 680–91.
- Blau, D. 1987. "A Time-Series Analysis of Self-Employment in the United States." *Journal of Political Economy* 95 (3): 445–67.
- BLS (Bureau of Labor Statistics, U.S. Department of Labor). 1991. *Bureau of Labor Statistics News* 92–386.
- Bosch, M., and W. F. Maloney. 2005. "Labor Market Dynamics in Developing Countries: Comparative Labor Markets Analysis Using Continuous Time Markov Processes." Policy Research Working Paper 3583, World Bank, Washington, DC.
- . 2007. "Comparative Analysis of Labor Market Dynamics and Informality in Developing Countries." Photocopy. World Bank, Washington, DC.
- Botero, J., S. Djankov, R. La Porta, and F. Lopez-de-Silanes. 2004. "The Regulation of Labor." *Quarterly Journal of Economics* 119 (4): 1339–82.
- Carneiro, P., J. Heckman, and E. Vytlacil. 2005. "Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education." Photocopy. University of Chicago.
- Carr, M., M. Chen, and J. Tate. 2000. "Globalization and Home-based Workers." *Feminist Economics* 6 (3): 123–42.
- Chant, S. 1991. *Women and Survival in Mexican Cities: Perspectives on Gender, Labour Markets and Low-Income Households*. Manchester, UK: Manchester University Press.
- , ed. 1992. *Gender and Migration in Developing Countries*. London: Belhaven Press.
- Chen, M., J. Sebstad, and L. O'Connell. 1999. "Counting the Invisible Workforce: The Case of Home-based Workers." *World Development* 27 (3): 603–10.
- CIEN (Centro de Investigaciones Económicas Nacionales). 2006. *Economía Informal: Superando las barreras de un Estado excluyente*. Guatemala City, Guatemala.
- Clark, K., and L. H. Summers. 1979. "Labour Market Dynamics and Unemployment: A Reconsideration." *Brookings Papers on Economic Activity* 1: 13–60.
- Cunningham, W. 2000. "Mexican Female Small Firm Ownership: Motivations, Returns, and Gender." Photocopy. LCSPPR-Gender, World Bank, Washington, DC.
- . 2007. "Minimum Wages and Social Policy: Lessons from Developing Countries." Photocopy. World Bank, Washington, DC.
- Cunningham, W., R. García-Verdú, L. McGinnis, and C. Tesliuc. 2007. *The Promise of Youth: Policy for Youth at Risk in Latin America and the Caribbean*. Washington, DC.: World Bank.
- . 2001a. "Breadwinner Versus Caregiver: Labor Force Participation and Sectoral Choice over the Mexican Business Cycle." In *The Economics of Gender in Mexico: Family, Work, State, and Market*, ed. E. G. Katz and M. C. Correia, 85–132. Washington, DC: World Bank.
- . 2001b. "Sectoral Allocation by Gender of Latin American Workers over the Liberalization Period of the 1990s." Policy Research Working Paper 2742, World Bank, Washington, DC.
- Cunningham, W., and W. F. Maloney. 2001. "Heterogeneity among Mexico's Microenterprises: An Application of Factor and Cluster Analysis." *Economic Development and Cultural Change* 50 (1): 131–56.
- Cunningham, W., and C. Ramos. 2001. "The Home as the Factory Floor: Home-based Work in Latin America." Photocopy. LCSPPR-Gender, World Bank, Washington, DC.
- Dávila Capalleja, E. 1994. "Regulaciones laborales y mercado de trabajo en México." In *Regulación del Mercado de Trabajo en América Latina*, ed. G. Marquez. San Francisco: Centro Internacional para el Desarrollo Económico.
- de Paula, A., and J. Scheinkman. 2006. "The Informal Sector." Photocopy. University of Pennsylvania, Philadelphia.
- de Soto, H. 1989. *The Other Path. The Invisible Revolution in the Third World*. New York: Basic Books.
- DGE (Dirección General de Estadística, Encuestas y Censos). 1998. "Sector Informal." *Encuesta Integrada de Hogares, 1997/1998*, Paraguay.
- Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1): 1–37.
- Duryea, S., G. Marquez, C. Pages, and S. Scarpetta. 2006. "For Better or for Worse? Job and Earnings Mobility in Nine Developing and Emerging Economies." Photocopy. Inter-American Development Bank, Washington, DC.
- Esfahani, H., and D. Salehi-Isfahani. 1989. "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism." *Economic Journal* 99 (June): 818–36.
- Evans, D. S., and B. Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97 (4): 808–27.
- Evans, D. S., and L. S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79 (3): 519–35.
- Fajnzylber, P., W. F. Maloney, and G. V. Montes Rojas. 2006. "Microenterprise Dynamics in Developing Countries: How Similar Are They to Those in the Industrialized World? Evidence from Mexico." *World Bank Economic Review* 20 (3): 389–419.
- Fields, G. S. 1990. "Labor Market Modelling and the Urban Informal Sector: Theory and Evidence." In *The Informal Sector Revisited*, ed. D. Turnham, B. Salome, and A. Schwartz. Paris: Organisation for Economic Co-operation and Development.
- Fortin, B., N. Marceau, and L. Savard. 1997. "Taxation, Wage Controls and the Informal Sector." *Journal of Public Economics* 66: 293–312.
- Galiani, S., and H. Hopenhayn. 2003. "Duration and Risk of Unemployment in Argentina." *Journal of Development Economics* 71 (1): 199–212.
- Galiani, S., and F. Weinschelbaum. 2006. "Modeling Informality Formally: Households and Firms." Photocopy. World Bank, Washington, DC.

- Gasparini, L., and L. Tornarolli. 2006. "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata." Photocopy. World Bank, Washington, DC.
- Geldstein, R. 2000. "Non-labor Market Coping Strategies in Argentina." Photocopy. LCSPP-Gender, World Bank, Washington, DC.
- Gonzalez de la Rocha, M. 1994. *The Resources of Poverty: Women and Survival in a Mexican City*. Oxford, UK: Blackwell.
- Gottschalk, P., and T. Maloney. 1985. "Involuntary Terminations, Unemployment, and Job Matching: A Test of Job Search Theory." *Journal of Labor Economics* 3 (2): 109–23.
- Harris, J. R., and M. P. Todaro. 1970. "Migration, Unemployment, and Development: A Two-Sector Analysis." *American Economic Review* 60 (1): 126–42.
- Hart, K. 1972. *Employment, Income and Inequality: A Strategy for Increasing Productive Employment in Kenya*. Geneva: International Labour Organization.
- Heckman, J., and G. Sedlacek. 1985. "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market." *Journal of Political Economy* 93 (6): 1077–125.
- Hemmer, H., and C. Mannel. 1989. "On the Economic Analysis of the Urban Informal Sector." *World Development* 17 (10): 1543–52.
- Hibbs, D. A. Jr., and V. Piculescu. 2006. "Tax Toleration and Tax Evasion: Why Firms Enter the Unofficial Economy." Working Paper in Economics 173, Department of Economics, Göteborg University, Sweden.
- Hirschman, A. O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- ILO (International Labour Organization). 1995. *World Employment*. Geneva.
- Itzigsohn, J. 2000. "Immigration and the Boundaries of Citizenship: The Institutions of Immigrants' Political Transnationalism." *International Migration Review* 34 (4): 1126–55.
- Jelin, E., M. Mercado, and G. Wyczykier. 2001. "Home Work in Argentina: Boosting Employment Through Small Enterprise Development." Working Paper 6, Series on Homeworkers in the Global Economy, International Labour Organization, Geneva.
- Jimeno, J., and D. Rodriguez Palenzuela. 2003. "Youth Unemployment in the OECD: Demographic Shifts, Labour Market Institutions and Macroeconomics Shocks." Working Paper 19, European Network of Economic Policy Research Institutes, Brussels.
- Jovanovic, B. 1982. "Selection and Evolution of Industry." *Econometrica* 50 (3): 649–70.
- Krawczyk, M. 1993. "Women in the Region: Major Changes." *ECLAC Review* 49: 7–19.
- Krebs, T., and W. F. Maloney. 1999. "Quitting and Labor Turnover: Microeconomic Evidence and Macroeconomic Consequences." Policy Research Working Paper 2068, World Bank, Washington, DC.
- Kugler, A., and M. Kugler. 2003. "The Labor Market Effects of Payroll Taxes in a Middle-Income Country: Evidence from Colombia." IZA Discussion Paper 852, Institute for the Study of Labor, Bonn, Germany.
- Levenson, A. R., and W. Maloney. 1996. "Modeling the Informal Sector: Theory and Empirical Evidence from Mexico." Photocopy. University of Illinois, Chicago.
- . 1998. "The Informal Sector, Firm Dynamics and Institutional Participation." Policy Research Working Paper 1988, World Bank, Washington, DC.
- Levy, S. 2006. "Productividad, Crecimiento y Pobreza en México: ¿Qué sigue después de Progresar-Oportunidades?" Photocopy. World Bank, Washington, DC.
- Loayza, N. 1994. "Labor Regulations and the Informal Economy." Policy Research Working Paper 1335, World Bank, Washington, DC.
- . 1996. "The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129–62.
- Lopez-Castaño, H. 1990. "Inestabilidad Laboral y Ciclo de Vida en Colombia." *Coyuntura Económica* 20 (1): 173–91.
- Lucas, R. E., Jr. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (2): 508–23.
- MacIsaac, D., and M. Rama. 1997. "Do Labor Market Regulations Affect Labor Earnings in Ecuador?" *Journal of Labor Economics* 15 (3): 36–65.
- . 2001. "Mandatory Severance Pay: An Assessment of Its Coverage and Effects in Peru." Policy Research Working Paper 2626, World Bank, Washington, DC.
- Maloney, W. F. 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico." *World Bank Economic Review* 13: 275–302.
- . 2004. "Informality Revisited." *World Development* 32 (7): 1159–78.
- Maloney, W. F., and P. Aroca. 1999. "Logit Analysis in a Rotating Panel Context and an Application to Self-employment Decisions." Working Paper 2069, World Bank, Washington, DC.
- . 2007. "Labor Market Dynamics in Latin America: Evidence from Argentina, Brazil and Mexico." Photocopy. World Bank, Washington, DC.
- Maloney, W. F., and J. N. Nuñez. 2002. "Measuring the Impact of Minimum Wages, Evidence from Latin America." In *Law and Labor Markets*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- Mondino, G., and S. Montoya. 2002. "The Effects of Labor Market Regulations on Employment Decisions by Firms. Empirical Evidence for Argentina." In *Law and Employment: Lessons from Latin America and the Caribbean*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- Morduch, J. 1999. "Between the Market and State: Can Informal Insurance Patch the Safety Net?" *World Bank Research Observer* 14: 187–207.
- Packard, T. 2002. "Pooling, Saving and Prevention: Mitigating Old Age Poverty in Chile." Policy Research Working Paper 2849, World Bank, Washington, DC.
- Peattie, L. 1982. "What Is to Be Done with the 'Informal Sector'? A Case Study of Shoe Manufacturers in Colombia." In *Towards a Political Economy of Urbanization in Third World Countries*, ed. H. Safa. Delhi: Oxford University.

- . 1987. "An Idea in Good Currency and How It Grew: The Informal Sector." *World Development* 15 (7): 851–60.
- Portes, A., M. Castells, and L. A. Benton, eds. 1989. *The Informal Economy: Studies in Advanced and Less Developed Economies*. Baltimore: Johns Hopkins University Press.
- Portes, R., and R. Schauffler. 1993. "Competing Perspective on the Latin American Informal Sector." *Population and Development Review* 19 (1): 33–59.
- Prugl, E., and I. Tinker. 1997. "Microentrepreneurs and Homeworkers: Convergent Categories." *World Development* 25 (9): 1471–82.
- Rauch, J. E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35: 33–47.
- Roberts, B. D. 1991. "Employment Structure, Life Cycle, and Life Chances: Formal and Informal Sectors in Guadalajara." In *The Informal Economy: Studies in Advanced and Less Developed Economies*, ed. A. Portes, M. Castells, and L. A. Benton. Baltimore: Johns Hopkins University Press.
- Rosen, S. 1981. "The Economics of Superstars." *American Economic Review* 71: 845–58.
- Saavedra, J., and A. Chong. 1999. "Structural Reform, Institutions and Earnings: Evidence from the Formal and Informal Sectors in Urban Peru." *Journal of Development Studies* 35 (4): 95–116.
- Schneider, F., and D. H. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38 (1): 77–114.
- Sedlacek, G. L., R. Paes de Barros, and S. Varandas. 1990. "Segmentação e Mobilidade No Mercado de Trabalho: A Carteira de Trabalho em São Paulo." In *Pesquisa e Planejamento Econômico*, 20(1): 87–104.
- . 1995. "Segmentação e Mobilidade no Mercado de Trabalho Brasileiro." Unpublished manuscript.
- Selby, H., A. Murphy, and S. Lorenzen. 1990. *The Mexican Urban Household, Organizing for Self-Defense*. Austin: University of Texas Press.
- Shapiro, C., and J. Stiglitz. 1984. "Equilibrium Unemployment As a Worker Discipline Device." *American Economic Review* 74: 433–44.
- Stiglitz, J. E. 2000. "Formal and Informal Institutions." In *Social Capital: A Multifaceted Perspective*, ed. P. Dasgupta and I. Serageldin. Washington, DC: World Bank.
- Straub, S. 2005. "Informal Sector: The Credit Market Channel." *Journal of Development Economics* 78: 299–321.
- Tokman, V. 1992. *Beyond Regulation: The Informal Economy in Latin America*. Boulder, CO: Lynne Rienner.
- Tomei, M. 2000. "Home Work in Selected Latin American Countries: A Comparative Overview." Photocopy. International Labour Organization, Geneva.
- Turner, J., and S. Verma. 2007. "Why Some Workers Don't Take 401(K) Plan Offers: Inertia Versus Economics." Working Paper 56/07, Center for Research on Pensions and Welfare Policies, Moncalieri, Italy.
- Turnham, D., and D. Eröcal. 1990. "Unemployment in Developing Countries, New Light on an Old Problem." Technical Paper, Development Center, Organisation for Economic Co-operation and Development, Paris.
- WIEGO (Women in Informal Employment: Globalizing and Organizing). 2000. "Women in the Informal Economy," available at <http://www.wiego.org>.
- World Bank. 2007. *Argentina: Poverty Assessment. Informal Employment in Argentina: Toward Understanding Its Causes and Consequences*. Washington, DC: World Bank.
- Yamada, G. 1996. "Urban Informal Employment and Self-employment in Developing Countries: Theory and Evidence." *Economic Development and Cultural Change* 44 (2): 289–314.

CHAPTER 3

Informality, Earnings, and Welfare

SUMMARY: Based on an examination of differences in earnings and self-rated welfare, this chapter lends credence to both the “exclusion” and the “voluntary” nature of informal employment suggested in chapter 2 by workers’ reported motivations to be informal. Earnings and welfare assessments support the view that the majority of independent workers are largely voluntary and attach significant value to the nonpecuniary benefits of autonomous work, while the majority of informal salaried workers tend to be excluded from more desirable jobs both in the formal salaried and independent sectors. However, the data caution against simple generalizations. Both groups exhibit important heterogeneity, and data for the Dominican Republic show that both the informal self-employed and informal salaried are as well-off as formal salaried workers, while in Colombia both informal groups exhibit much lower levels of satisfaction with current jobs. While there are seemingly two tiers of informal employment, gradients and gray areas cloud universal conclusions about who they are, what determines their relative size, and the welfare implications of each tier.

CHAPTER 1 ILLUSTRATED THE considerable heterogeneity of informal workers in the region and how, on average, they tend to face a significant earnings disadvantage. Chapter 2 discussed the various hypotheses that aim to explain the *razón de ser* of the informal worker and their implications regarding the relative quality of formal and informal jobs. This chapter delves into a more in-depth investigation of the characteristics associated with the propensity of workers to be informal, and the variation in the earnings and welfare differentials associated with informal work, making a clear distinction between informal salaried and independent workers.

Who are the informal (salaried and independent) workers? Does informal employment in itself imply lower earnings—that is, do informal workers receive lower pay for equal work? Are all informal jobs “bad jobs” in terms of carrying lower earnings for a worker’s skills, or do some individuals find niches in informal employment that are more suitable for their skills, preferences, and situations? Does informal work lead to lower individual welfare, or do the nonpecuniary

(for example, flexibility, autonomy) characteristics of informal jobs partially compensate workers for any lower earnings? This chapter examines these questions, relying heavily on analyses of household surveys containing the relevant data from Argentina, Bolivia, Colombia, and the Dominican Republic to illustrate the issues in a variety of country contexts.

The chapter first describes the characteristics of workers, their jobs, and the firms that affect the propensity to informal employment (proxied by lack of pension coverage). It then examines whether informal workers receive equal pay for equal work along the entire remuneration scale, and what fraction of informal-formal earnings gaps result from differences in worker characteristics and in how these are rewarded in the labor market, including differences in the returns to human capital (schooling or work experience). Finally, the chapter examines workers’ self-rated welfare (propensity to consider themselves poor) and job satisfaction responses as more complete indicators to assess welfare differences between workers in formal and informal jobs, as well as the significance of nonpecuniary job characteristics

(that is, flexibility, autonomy, stability/mobility) in the sectoral participation of various workers (for example, women/men, the skilled/unskilled).

Compensating differentials, comparative advantage, and informal work

As discussed in chapter 2, the nature of informal employment has been examined mainly through two lenses: the “exclusion” and “voluntary” views. In the exclusion view, workers would prefer the benefits of formal jobs but are rationed out due to segmentation of the labor market. The latter is traditionally linked to institutional rigidities, but—again—can arise from economic dualism, barriers to labor mobility (geographical or informational), efficiency wages, or coordinated evasion of corporate, sales, and payroll taxes. In contrast, the voluntary view suggests that informal jobs largely reflect workers’ implicit choices given their preferences, skills, the costs and benefits of formality, and the availability of other means of social protection, the last in turn being a function of the quality of existing country policies and institutions. As noted before, both views have potential explanatory power to understand the nature of informality in the Latin American and Caribbean region.

The labor literature on compensating differentials and occupational choice based on workers’ comparative advantage provides a framework that encompasses these two seemingly irreconcilable views. The basic idea, first advanced by Adam Smith (1876, p. 111), is that the wages paid to various types of labor must, in general, equalize total *advantages and disadvantages*, pecuniary and nonpecuniary, and that workers select occupations that yield the highest net advantage for their tastes and skills. This is the cornerstone of the modern literature on comparative advantage in the labor market as a determinant of occupational choice, human capital investments, and earnings performance (Carneiro, Heckman, and Vytlačil 2001; Heckman and Sedlacek 1985; Rosen 1981).

According to this, individuals choose the jobs that better fit their preferences and specific ranges of talents, including cognitive, social, and mechanical skills. Jobs that are more desirable (due to amenities such as fringe benefits, stability, safety, autonomy, and flexibility) or that require relatively abundant skills should have *lower-than-average* wages while jobs that are less desirable or demand scarce skills should pay *higher* wages. A competitive labor market determines an implicit (hedonic) wage for each type of labor and equilibrates when labor mobility leads to a set of relative

wages that makes workers indifferent toward the various types of jobs. The differences in these implicit wages are called *compensating differentials*. Given heterogeneity in worker preferences and skills, both supply and demand for particular jobs determine the size of the compensating differential between jobs with different working conditions. Thus, the labor market comprises a set of interrelated markets for different labor types whose wages are set by supply and demand.

The consideration of major labor-demand constraints in the framework can bring to bear the concerns of the exclusion view of informality. In particular, with high unemployment, an oversupply of some types of labor, or in a situation when workers are in “noncompeting groups,” the market pressure to pay compensating differentials can vanish. In this case, more workers would be willing to take the less desirable jobs at wages below those in the more desirable-but-scarce jobs. Therefore, when the demand of some labor types is very constrained for whatever reason, including market segmentation, the room for workers’ choice is greatly diminished and their wages are set primarily by the employing firms.

This framework has three key implications for the nature of informal employment. First, in choosing between informal and formal employment, workers weigh the advantages and disadvantages of each potential job, subject to the availability of jobs with their desired attributes; therefore, comparative advantage could make the informal sector a better match for many of them. As noted in chapter 2, informal and formal jobs tend to require varying degrees of skills, and the various amenities of informal and formal jobs may be valued differently by some groups. Some workers might find advantageous niches to their observed and unobserved skills in occupations with a higher propensity to be informal (such as those in construction).

Others may be willing to forgo some of the benefits of formality in exchange for the nonpecuniary benefits of informal jobs. As an example, for low-skilled youth and older workers, informal salaried jobs may offer an entry point to the labor market that partially allows them to remedy deficient schooling or the obsolescence of skills through on-the-job training unavailable to them in formal salaried jobs. For women, the demands for flexibility to balance work and family responsibilities that arise with child rearing may render the greater flexibility and autonomy of informal jobs a better match. Furthermore, employers could attract

enough workers without offering a compensating wage differential if a large subset of workers does not regard a given job attribute (such as not having social security) as a disadvantage. In some cases, amenities and disadvantages may cancel each other out without giving rise to compensating differentials.

Second, there is no clear a priori presumption that informal employment should carry lower, equal, or higher earnings than formal jobs. On the one hand, informal jobs should command higher earnings to compensate workers for the value of lost fringe benefits such as social protection for health risks and old age (net of payroll contributions), and their greater risk of unemployment and income shocks. On the other hand, earnings may be lower to adjust for the value of nonpecuniary amenities such as flexible hours, autonomous work, training, and tax savings, particularly for independent workers but much less so for the informal salaried workers. Thus, differences in earnings cannot be used to test segmentation in the labor market (Magnac 1991; Maloney 1999).

Moreover, since workers may sort into formal and informal jobs acting on the returns to their skills and competencies in each sector, direct average earnings comparisons are flawed. Frequently, high-paying skilled jobs also offer better fringe benefits. It is necessary to control for differences in observed and unobserved skill level (leading to productivity differences reflected in wages) in order to measure correctly informal-formal compensating differentials. Furthermore, average earnings gaps do not fully characterize the earnings gains or losses from formal employment for any given worker. Some workers could emerge as top earnings performers for their skills in informal jobs, but if they were to move to a formal job they may actually lose out in the earnings rankings. Section 3 of the chapter presents empirical results in this vein.

Finally, more generally, the relative qualities of informal and formal jobs and the ensuing differences in workers' welfare have to be assessed more broadly to take into account nonpecuniary job attributes. The conditional informal-formal earnings comparisons could be seen as providing orders of magnitude of the potential net value of the amenities and disadvantages of informal jobs. Establishing whether amenities fully compensate for any lower informal earnings so that workers are equally well-off in terms of welfare faces the challenge of limited data. The last section of the chapter presents new analyses of subjective welfare and job satisfaction data that attempt to provide some evidence on this key issue.

The profile of participation in informal employment

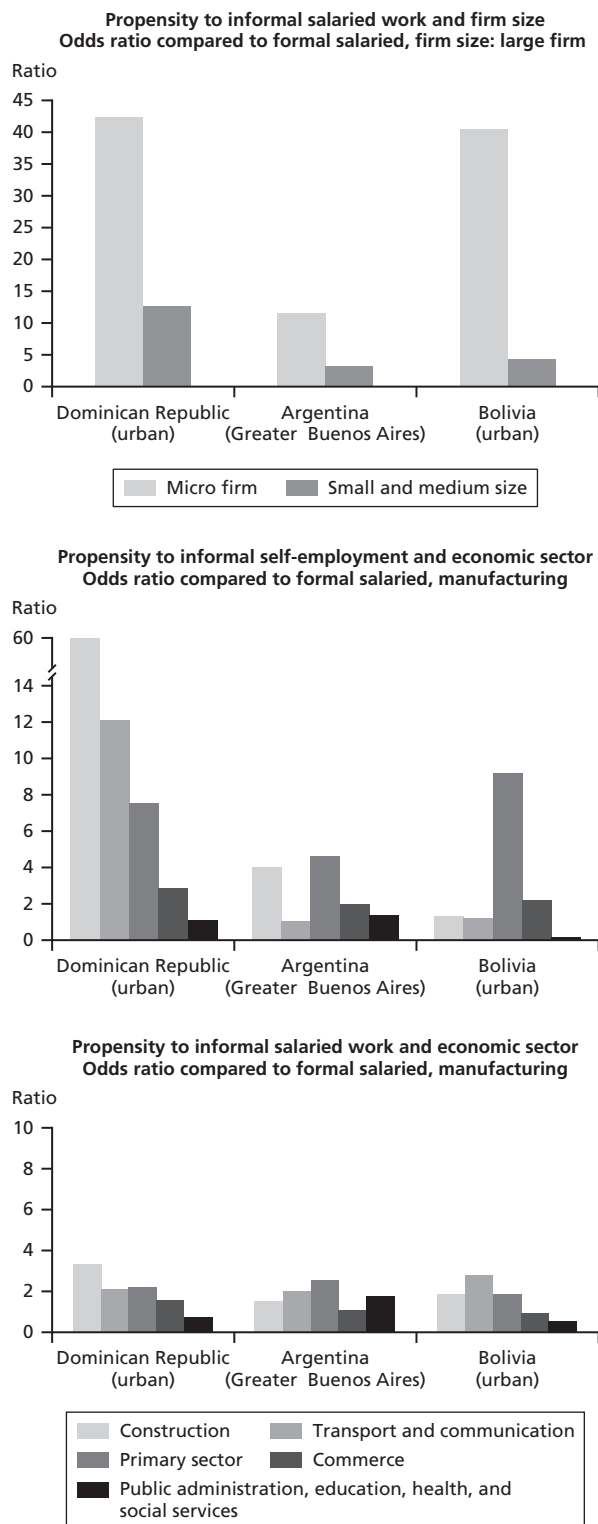
As discussed in chapters 1 and 2, informal employment encompasses a diverse range of people. Using contributions to social security as a criterion to define informal employment, we can distinguish four employment categories: informal and formal independent, informal and formal salaried. The informal independent generally comprise relatively better-paid small firm owners, self-employed professionals (for example, doctors, lawyers, teachers), semiskilled workers with some technical competencies (for example, artisans, handymen, construction laborers, taxi drivers), and unskilled workers in precarious employment or under a semi-dependent work relationship (street vendors, small artisans under subcontract). The informal salaried largely comprise domestic employees, a variety of young apprentices, and older unskilled workers in small enterprises (for example, sales clerks, beauticians, salaried artisans), but also the workers in larger firms who are under informal labor arrangements.

Regression analysis can be used to isolate the characteristics that best predict the propensity that a given worker engages in the four employment categories. The analysis is carried out for Argentina, Bolivia, and the Dominican Republic, considering a host of characteristics of workers, their jobs, and the firms for which they work. Figures 3.1 to 3.3 present the odds ratios or relative propensities that a worker with a given characteristic belongs to each of the two informal groups rather than to the formal salaried, holding other characteristics constant. An odds ratio of 1 represents no effect of a variable; a ratio greater than 1 indicates that the characteristic increases the odds of that employment category compared to the formal salaried group, and a ratio less than 1 indicates that it diminishes the odds. For instance, an odds ratio of 0.75 signifies that the chance that a worker with the given characteristic gets a given job type is 75 percent of the chance of being formal salaried. An odds ratio of 1.5 implies a 1.5 times greater chance of being in the given category than formal salaried. The larger the odds ratio, the stronger the relationship, so these provide an implicit ordering of workers with different education levels, age groups, sectors, firm sizes, and so on across the four occupational groups.

Adjusting for workers' personal characteristics, firm size has the strongest negative association with informal work, and sector of employment also has an important independent effect. The odds of being informal salaried in microenterprises relative to large firms are over 40 to 1 in Bolivia and the Dominican Republic, and over 11 to

FIGURE 3.1

Propensity to informal employment by firm size and economic sector, 2005-06

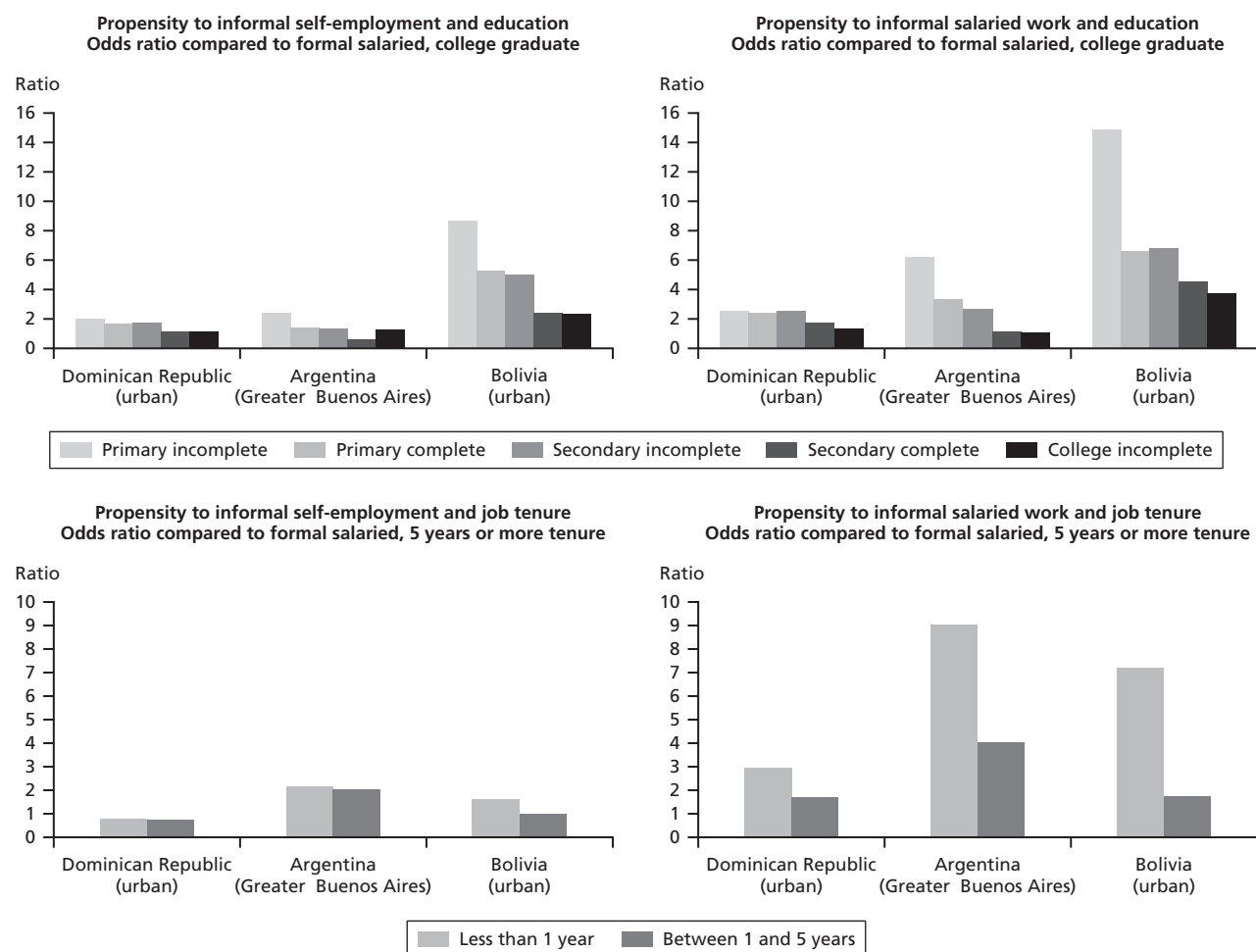


Source: Based on Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), and author's estimates using household survey 2005, 2006 data.

1 in Argentina.¹ In these countries, the odds fall significantly for employees of medium-sized companies, although remaining at 12 to 1 in the Dominican Republic. Employment industry also significantly affects informality propensities, with no universal patterns across countries. Adjusting for firm size and other characteristics, informal independent workers are more likely to be found in construction and transport as well as in agricultural activities and commerce in the Dominican Republic, and in construction and commerce in Argentina. Informal salaried employment is more likely in the same sectors in the Dominican Republic, while it is more prevalent in transport, social, and personal services in Argentina. Informal employment propensities are quite similar across sectors in Bolivia, except for greater odds of independent work in agriculture and informal salaried employment in transport. These differences across and within countries in sectoral propensities to informality may reflect differences in the incentives entailed by existing regulations and in their enforcement across sectors and countries, as well as an intrinsic inclination toward informality of workers with occupational traits fit to sectors such as construction and commerce.

Among personal characteristics, education and tenure are the strongest predictors of informal employment (figure 3.2). The informal employment propensities decline progressively with education, especially with full completion of secondary schooling. Compared to a college graduate, the odds that a worker with incomplete primary or no education is informal salaried or informal independent are 15 and 8 times greater in Bolivia, respectively, about 2 to 1 in the Dominican Republic, while in Argentina the odds are 6 to 1 and above 2 to 1, respectively. Holding constant other characteristics, those with a high school diploma face equal chances of informal employment as the college-educated in Argentina, 4-to-1 odds in Bolivia, and still close to a 2-to-1 chance in the Dominican Republic. The significant fall in the informality propensity with the completion of high school suggests that this diploma may serve as a signal to employers that a worker has the minimum skills required to justify the cost of a formal labor contract. This may be more important in the presence of job rationing due to labor market rigidities or other sources of job segmentation. This is consistent with the differences observed between Argentina and Bolivia, on one hand, and the less rigid Dominican labor market on the other. Interestingly, results (not shown here for brevity) indicate that in all three

FIGURE 3.2

Propensity to informal employment by education and tenure, 2005–06

Source: Based on Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), and author's estimates using household survey 2005, 2006 data.

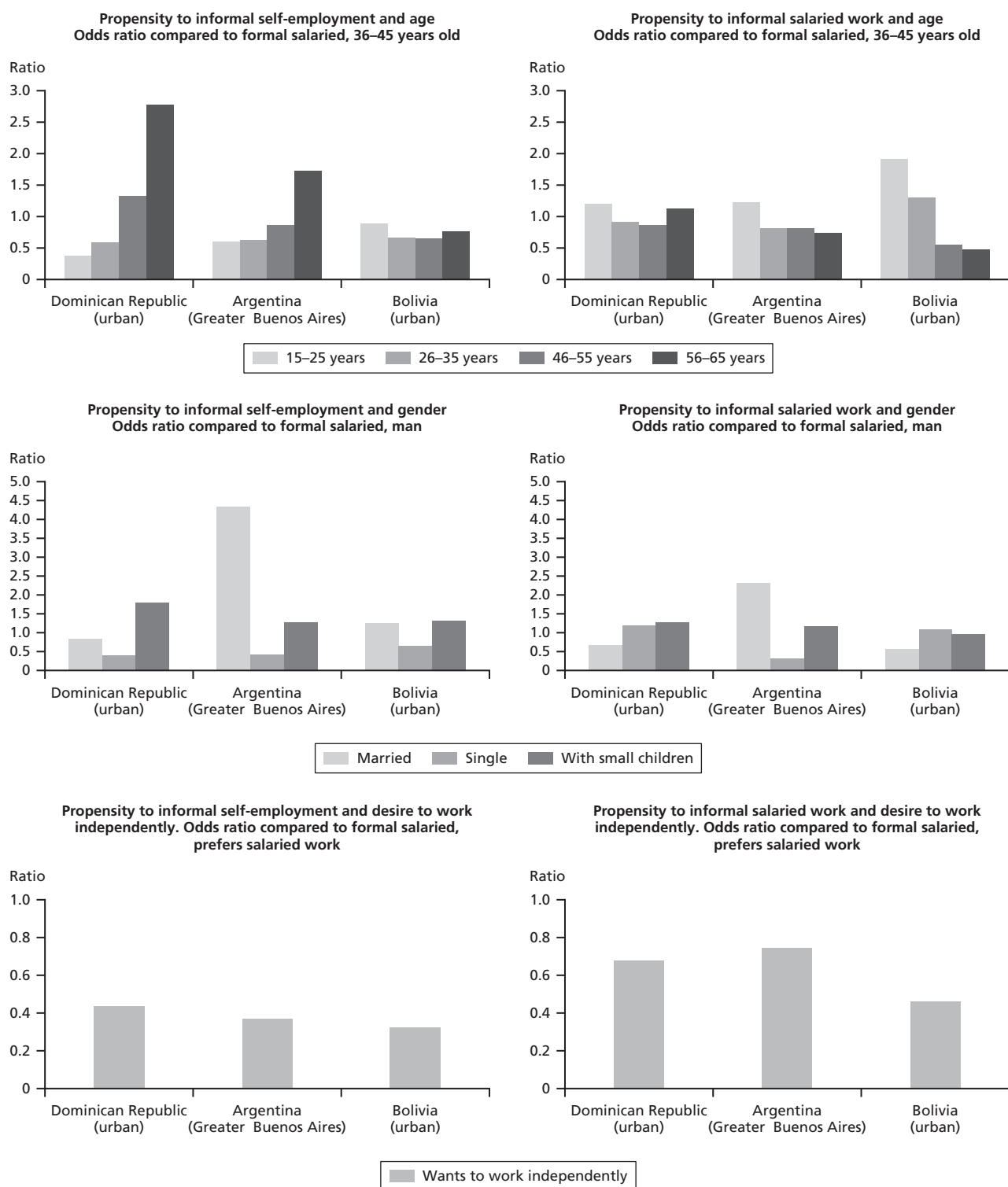
countries the college-educated are generally equally or more likely to be formal independent than formal salaried, adjusting for other characteristics.

Finally, workers with less than one year in their current occupation have a much larger risk of being informal salaried than those with longer tenures. For instance, compared to those with five or more years in their occupation, the odds are 9 to 1 in Argentina, 7 to 1 in Bolivia, and 3 to 1 in the Dominican Republic. Short tenures also correlate with greater chances of being informal independent relative to formal salaried in Argentina and Bolivia (4 to 1 and 2 to 1, respectively), while the odds are similar in the Dominican Republic.

Individual and family demographics show an independent but second-order correlation with informality

propensities (figure 3.3). Although far from universal, some common patterns emerge by age and women's marital status. As shown in chapter 2, younger workers experience much higher rates of informal salaried work. The regression analysis shows that this is largely a result of youth's much shorter tenures and higher rates of employment in small firms. In fact, holding these and other factors constant, the odds ratio of informal to formal salaried work is constant across all age groups in Argentina and the Dominican Republic, although they still fall systematically with age in Bolivia, where youth are 50 percent more inclined to informal over formal salaried work than middle-aged workers. The age profile of informal independent work is consistent with life-cycle theories of

FIGURE 3.3

Propensity to informal employment by age, gender roles, and work preferences, 2005–06

Source: Based on Arias and Bustelo (2007), Arias, Landa, and Yanez (2007), and author's estimates using household survey 2005, 2006 data.

self-employment, although it also varies across countries. The odds of independent (formal or informal) employment to formal salaried work increase with age in Argentina and the Dominican Republic, but are unrelated to age in Bolivia. The latter finding suggests that the accumulation of savings and occupation-specific skills may be less important for the ability to engage in the type of independent employment predominant in very low-income countries.

The gender patterns of participation in informal work are broadly consistent with the findings reported in chapter 2 for Mexico and other middle-income countries, although with some country variation. For instance, Argentine married women are more inclined than men to be independent (4 to 1 informal, 2 to 1 formal) and informal salaried (2 to 1) than formal salaried, regardless of family structure. The same is true for Dominican married women with small children. Conversely, single Argentine and Dominican women have greater odds than men of engaging in formal salaried employment. These results are consistent with Galiani and Weinschelbaum's (2006) findings; using data for a large sample of Latin American countries, they found that, *ceteris paribus*, a spouse is more likely to work informally if the head of household has a formal job. Meanwhile, Bolivia shows no gender differences in informal participation, regardless of the presence of small children (or elderly) in the family. Overall, there is no systematic evidence of discrimination forcing women into informal employment; the pattern of female employment that allows them to balance work and family life varies with economic and social conditions, including the availability of protection through other family members or publicly provided benefits programs, in their countries.

With regard to other personal characteristics, there is strong evidence that workers act on their tastes for independent or salaried work, and little evidence of an independent correlation of informal employment with migration or ethnicity. The bottom panel of figure 3.3 shows that the chance that workers who report a preference for salaried over independent work are actually employed as independent is only 30 to 40 percent of their chance of being formal salaried. Interestingly, the chances that these workers are informal salaried are 50 percent (Bolivia) to 70 percent (Argentina and the Dominican Republic). This could reflect the fact that formal salaried workers have a stronger

preference for the job amenities they enjoy—or, conversely, it could be an expression of their higher satisfaction with salaried work. Other results (not shown) indicate that, controlling for all characteristics, urban migrants (coming from rural or other urban areas) in all three countries and foreign migrants in Argentina and the Dominican Republic have similar employment propensities as nonmigrants. The same is true for indigenous workers in Bolivia. Thus, the observed higher incidence of informal employment in these groups is largely a result of their having personal characteristics or jobs in sectors or firms with a higher propensity to informality, rather than to their migrant or ethnicity status, *per se*.

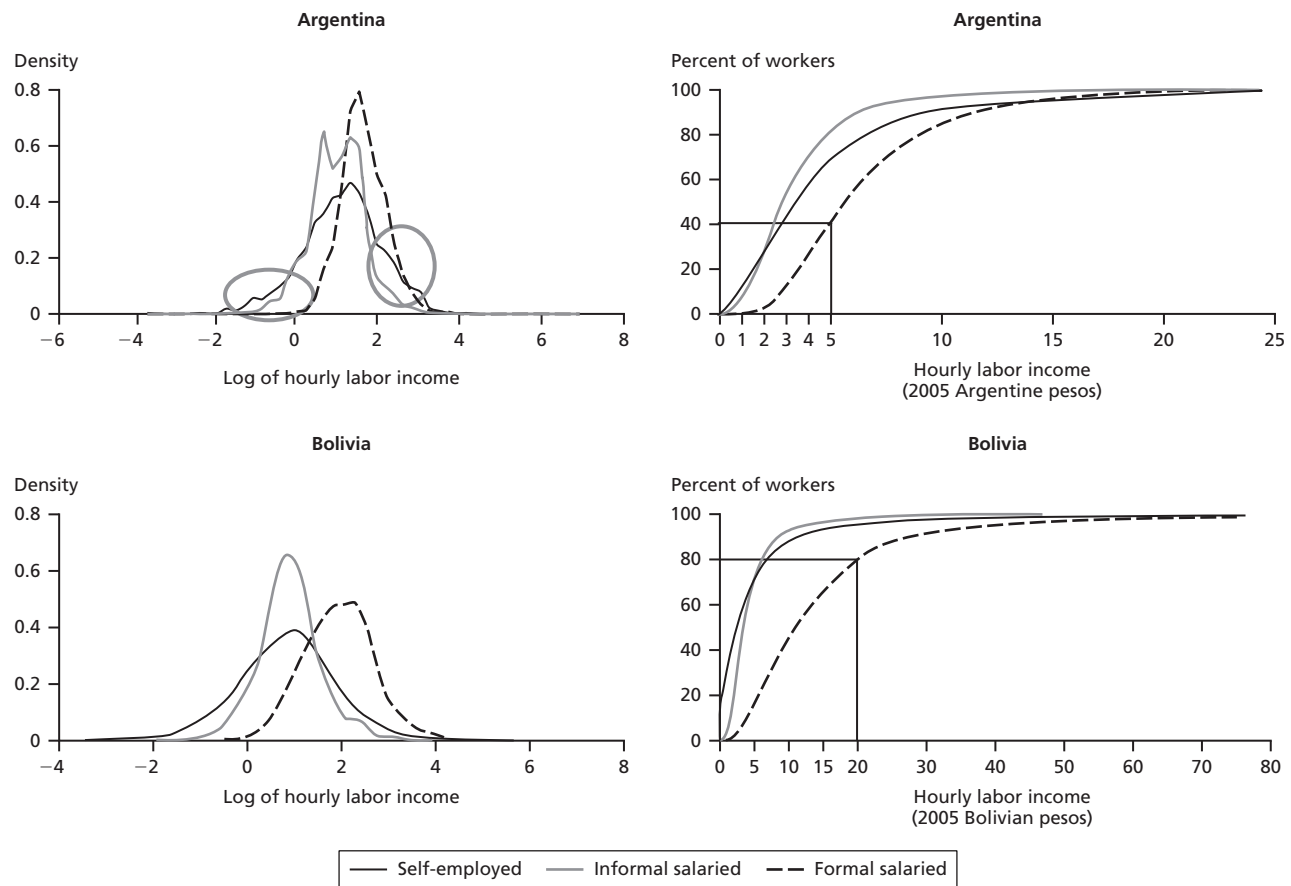
In summary, the results show that firm size, education levels, tenure, and sector of employment—variables related to the productivity of workers and firms and the ease of government enforcement—are the most important predictors of informal employment in Argentina, Bolivia, and the Dominican Republic. There is a similar pattern of impacts of these characteristics on the propensity to be informal salaried or informal self-employed in all three countries. However, the orders of magnitudes differ markedly, which suggests that the underlying mechanisms determining informal employment vary with level of economic development, economic structure, and the nature of public policies and institutions.

The next section examines empirically whether the informal-formal characteristics of jobs have an impact on workers' earnings in Argentina, Bolivia, and the Dominican Republic. The analysis considers the ample spectrum of remunerations and worker prototypes in the labor market and demonstrates that workers' sectoral participation is affected by the expected returns to their skills and the perceived qualities of informal and formal jobs.

The question of equal pay for equal work in the informal and formal sectors

The heterogeneity in the prototype of informal workers is reflected in a wide earnings variation. Figure 3.4 illustrates this with a comparison of the distributions of hourly earnings for formal, informal salaried, and independent workers in urban areas of Argentina and Bolivia.² No distinction is made between the informal and formal self-employed, as the latter are a small fraction of employment and their separation would complicate the econometric analysis. The left panels contain the frequency distributions (that is,

FIGURE 3.4

Distribution of hourly earnings for workers in urban areas of Argentina and Bolivia, 2005-06

Source: Based on Arias and Khamis (2007), Arias, Landa, and Yanez (2007), and author's estimates using household survey 2005, 2006 data.

Note: Incomes measured in 2005 local currency.

smoothed histograms) showing the earnings where most workers are concentrated in each group (that is, near the peaks). The right panels present the cumulative distributions of hourly earnings which show the fraction of workers at or below any given level of hourly earnings (that is, the percentiles). The horizontal distance between the distributions yields measures of the gap in informal-formal earnings at different earnings percentiles. For instance, in Argentina the difference in hourly earnings between the formal and the two informal groups at the 40th percentile is \$2.5 ($= \$5 - 2.5$), while in Bolivia the difference at the 80th percentile is \$15 ($= \$20 - 15$).

There are three main points worth noting. First, although the formal salaried have an earnings advantage at any point of the pay scale (their distribution is further to the right), informal-formal earnings gaps are larger for workers in low-earnings jobs (the distributions are further apart at

the bottom). Second, the informal salaried have a slight earnings advantage over independent workers in the bottom 40 percent of the distribution (their distribution lies below and closer to the formal salaried), but this is reversed in the top of the distribution (the high-earnings jobs). Third, the earnings distribution for independent workers is more dispersed, consistent with the international evidence of higher earnings variations in independent activities.³ Moreover, there are two clear tiers of low- and high-earnings independent workers (illustrated in circles for Argentina), consistent with the heterogeneous composition of the group.

Table 3.1 quantifies the raw (unconditional) gaps in hourly earnings for jobs of low, median, and high remuneration within each occupational category for the three countries considered.⁴ Clearly, there are significant differences in the pattern of earnings gaps across and within the countries. On average, informal salaried workers earn between 40 to

TABLE 3.1

Unconditional hourly earnings gaps (percent difference in earnings) for formal employees, informal employees, and independent workers in urban areas in Argentina, Bolivia, and the Dominican Republic, 2005

Worker status by country	Low-earnings jobs	Median-earnings jobs	High-earnings jobs	Average-earnings jobs
<i>Argentina</i>				
Informal/formal salaried	–53.3	–47.8	–45.1	–43.1
Independent/formal salaried	–58.0	–37.3	–19.3	–22.3
Informal salaried/independent	11.1	–16.7	–32.0	–26.8
<i>Bolivia</i>				
Informal/formal salaried	–60.0	–64.0	–68.0	–67.0
Independent/formal salaried	–69.6	–61.0	–51.0	–60.0
Informal salaried/independent	32.3	–7.3	–31.8	–18.0
<i>Dominican Republic</i>				
Informal/formal salaried	–44.0	–46.5	–46.0	–45.9
Independent/formal salaried	–9.4	11.0	5.8	5.7
Informal salaried/independent	–38.2	–51.8	–49.0	–48.8

Source: Based on Arias and Khamis (2007), Arias, Landa, and Yanez (2007), and author's estimates using household survey 2005, 2006 data.
Note: The high, medium, and low pay correspond, respectively, to the 80th, 50th, and 20th percentiles of earnings for each group; that is, they represent the top 20 percent of workers with the highest earnings, the bottom 20 percent with the lowest earnings, and workers with median earnings. Average pay = mean earnings.

66 percent less than formal salaried employees in all countries, while independent workers earn as little as 60 percent less in Bolivia and 28 percent less in Argentina but have a slight advantage in the Dominican Republic.

The differences in average hourly earnings understate the disadvantage of informal salaried workers in low-earnings jobs in Argentina, but not as much in Bolivia and the Dominican Republic. Independent workers face a larger earnings disadvantage in the low-earnings jobs, which falls progressively for workers at the top of their pay scale and actually turns into an earnings advantage in the Dominican Republic. The low-earnings informal salaried are in fact slightly better off than low-earnings independent workers in Argentina and Bolivia, but earn less in the top-paying jobs. In these countries low-earnings workers face a more precarious employment situation whether they are salaried or independent. The distinct patterns across countries suggest that the relative earnings of formal and informal workers probably vary with economic structure, overall productivity, and other conditions specific to their countries.

However, these earnings gaps do not prove that formal sector jobs are superior, per se. As we saw before, informal workers tend to be younger and less educated, working in smaller firms and sectors such as commerce and construction, all of which are cause for lower earnings. Their lower earnings may be due to their disadvantage in productive attributes (both observed and unobserved) rather than to

having an unregistered job. More generally, as noted earlier in the chapter, differences in earnings can result from labor market segmentation or barriers that prevent workers from accessing the best possible job for their skills, or can simply reflect the compensated earnings differentials from differences in the amenities of informal and formal jobs.

To determine whether informal workers receive equal pay for equal skills, one must compare earnings of workers with similar observed and unobserved characteristics who differ only in having a formal or an informal job. This is best accomplished through regression analysis to purge earnings gaps of spurious correlations induced by observed and unobserved worker characteristics that affect earnings and cause selection (either by choice or rationing) into formal, informal salaried, or independent work. For instance, the most talented individuals may be more likely to obtain formal salaried employment because of better prospects for mobility in a career as wage earner. On the contrary, individuals with more entrepreneurial ability are more likely to succeed as independent workers. Those with low work attachment and little adherence to authority or rigid work schedules may be excluded from formal salaried employment or voluntarily seek the flexibility of self-employment even at the cost of lower earnings. Many women may forgo the higher earnings of being employees in exchange for flexibility in informal employment. The end result is selection into occupations based on returns and tastes.

Even among workers with the same credentials, experience, and line of work, some earn more than others due to, for example, their innate abilities, or because they landed jobs in more productive firms. That is, there are low-paying and high-paying jobs among doctors, teachers, carpenters, and salespeople, as well as in construction, manufacturing, and large and small firms. Gaps in average earnings between informal and formal workers with the same observed characteristics may misrepresent the situation of those with earnings below or above the average pay for their skills; some workers may be paid differently or find advantageous niches across job categories. Informal jobs may carry a lower reward in manufacturing where small-scale production tends to be less productive, while pay may be similar across all categories in service jobs where economies of scale are less important.

More important, the returns to observed skills (such as education or work experience) may be different. Schooling returns may be higher for the formal salaried, since education is generally more productive in activities employing more skills and physical capital. However, informal occupations prevalent in small-scale activities require fewer cognitive skills and may offer low-educated workers compensatory practical skills that enhance the productivity of their schooling. Returns to education might be higher for independent workers, since they can optimize the use of their credentials in their economic activity with the flexibility of small size and lesser division of labor. The returns to work experience can also differ for similar reasons.⁵ Although many of these differences are hard to ascribe to segmentation or to voluntary sorting of workers, they have implications on life-cycle incomes.

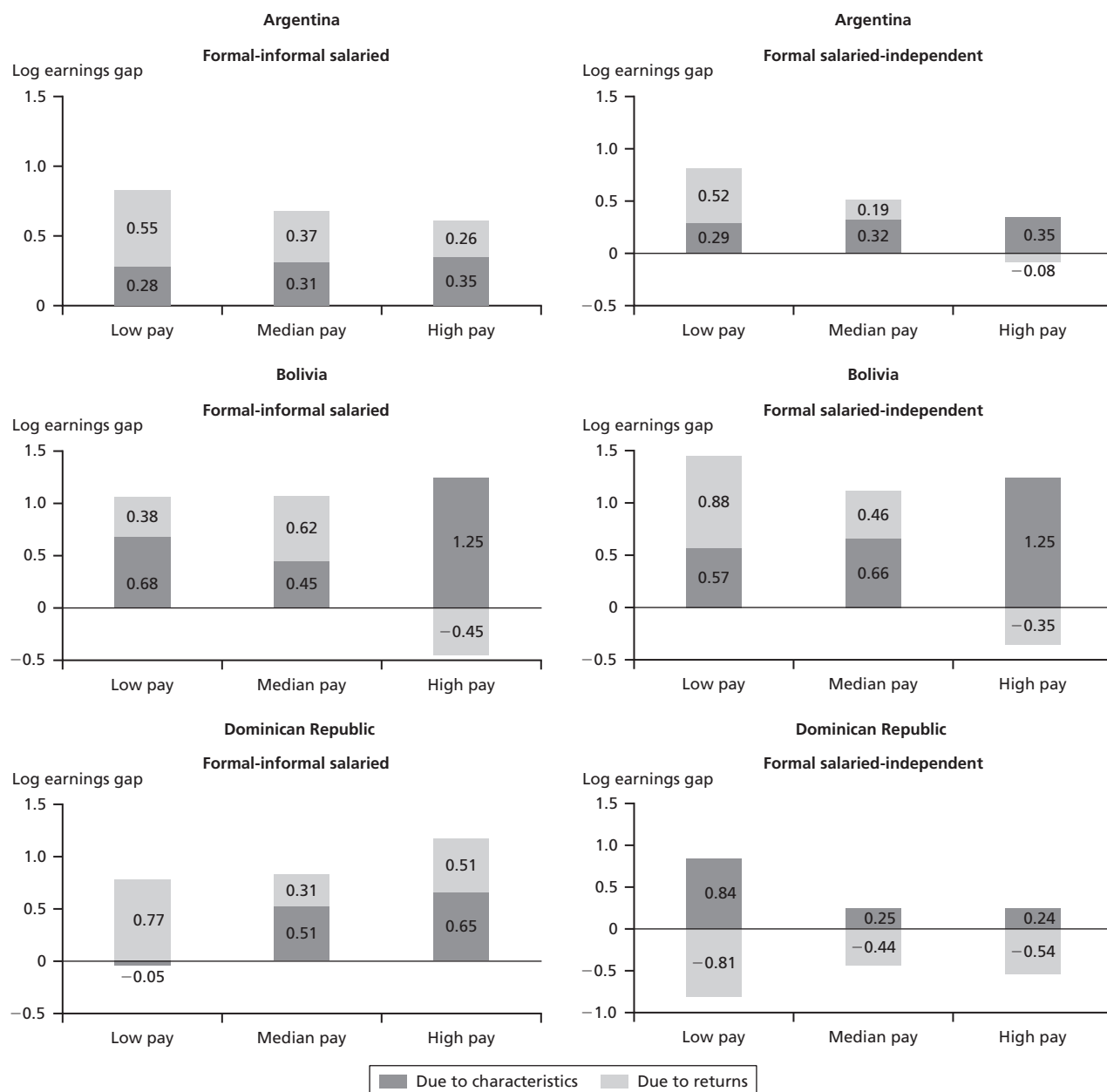
In light of all this, earnings regressions are estimated separately for each of the three employment categories at different percentiles of the group-specific earnings distributions. This allows measurement of earnings gaps between formal and informal workers who have the same observed characteristics (for example, education, work experience, gender) and other unmeasured characteristics and in the same positions of the pay scale fit to their skills and job categories. For instance, we measure the earnings gap between the best-paid formal salaried and the best-paid self-employed with the same skills set using quantile earnings analysis.⁶ To deal with selection biases with cross-section household survey data, we proxy unobserved factors by estimates of the relative probabilities that individuals work as formal, informal salaried, or independent and include these in the earnings regressions. This is a modification of conven-

tional Heckman-selectivity corrections. The extent to which this procedure corrects fully or partially for self-selection biases depends on the strength of variables that affect sector participation but not earnings (see the studies by Arias [2007, 2005] and the coauthors for details).

We decompose the estimated earnings gaps into the portion due to between-group differences in characteristics (endowments) and the part due to differences in how these characteristics are remunerated (returns) across job categories. The latter differences in returns are more adequate measures of the compensated earnings differentials between informal and formal jobs.⁷ Figure 3.5 shows the results of this decomposition for the three countries considered. The predicted earnings gaps (in logs) at low-, median-, and high-paying jobs are given by the height of the bars. The shaded areas reflect the portion attributed to informal workers' less favorable characteristics; the rest represents our measures of the compensating earnings differentials among formal and informal jobs, which also encapsulate any differences in earnings arising from possible labor market segmentation due to macroinstitutional rigidities.

The results provide a different picture across countries. In Argentina the bulk of the informal earnings gap reflects unequal pay for similar skills, especially at the bottom of the earnings scale. Approximately two-thirds of the earnings difference between formal and informal Argentine employees in low-paying jobs is due to the overall lower remuneration to the skills of the latter in the labor market. This fraction falls to 54 and 42 percent in the median- and high-paying jobs.⁸ The contribution of unequal pay is similar for independent workers in the low-paying jobs, falling to 38 percent in median-pay jobs, and turning insignificant at the top. That is, all of the earnings advantage of the best-paid formal salaried workers over the best-paid independent workers in Argentina is due to differences in their characteristics rather than unequal pay. In fact, the absolute level of both informal earnings gaps falls as we move up to compared workers in the best-paid jobs of each of the three sectors. The results are consistent with high-earnings independent workers having primarily voluntary motivations to be independent in Argentina, to the extent that they cannot get better pay for their skills in formal salaried jobs. The gaps at the bottom of the earnings distribution are consistent with compensating earnings differentials in favor of those in formal salaried jobs (for example, to compensate for lower flexibility/autonomy), and/or labor market segmentation against the self-employed.

FIGURE 3.5

Earnings cost of informality in urban areas of Argentina, Bolivia, and the Dominican Republic

Source: Based on Arias and Khamis (2007), Arias, Landa, and Yanez (2007), and author's estimates using household survey 2005, 2006 data.

Note: See table 3.1 for definitions.

The results for Bolivia show the existence of unequal pay for informal jobs, but with an opposite pattern for the informal salaried and independent workers. Differences in characteristics account for the bulk of earnings gaps between the formal and informal salaried, especially in the low- and high-pay jobs, while differences in returns are more important in the average-pay jobs (accounting for close to 60 percent of

the gaps). The differences in returns actually favor the informal salaried in high-pay jobs. That is, workers in the high-pay jobs of the informal salaried sector receive better remunerations for their skills than workers in the high-pay formal jobs. The pattern is similar for the self-employed and is close to the results for Argentina. The returns to their skills become increasingly similar to the formal salaried as

we move up the earnings scale, and in fact they are remunerated slightly better than the formal salaried at the best-paying jobs for their skills set. As in Argentina, all of the earnings advantage of the best-paid formal salaried workers over the best-paid independent workers arises from their more favorable characteristics rather than unequal pay, and yet the absolute level of the earnings gaps is less monotonic along the pay scale.

Finally, the Dominican Republic shows a mixed pattern for the informal-formal salaried earnings gaps and a very different situation for independent workers. Close to 60 percent of the earnings gaps between formal and informal employees in jobs of median pay or above are due to differences in their characteristics, but the latter explain little of the gap in the low-pay jobs. That is, the lower earnings of informal salaried workers in the bottom of the earnings distribution are entirely due to lower remunerations to their characteristics. However, contrary to Argentina and Bolivia, they are not as worse off in an absolute sense, since the level of the informal-formal salaried earnings gaps is actually much lower at the bottom than at the top of the distribution. Meanwhile, consistent with the simple earnings differentials shown above, the results actually show that independent workers are in a favorable situation regardless of their position in the earnings scale. In fact, their advantage in earnings arises from better remunerations to their characteristics, which fully compensate for their less favorable personal and job characteristics, especially in the low-earnings jobs. If they had a similar set of average characteristics as formal salaried workers, they would actually earn even more than their formal salaried counterparts.

In sum, all three countries show a clear earnings disadvantage of informal salaried workers that cannot be accounted for by lower observed productivity, and is indeed much larger in the low-paying jobs for any skills set. The situation is similar for independent workers in Argentina and Bolivia, and markedly different in the Dominican Republic. There is substantial heterogeneity in earnings performance, particularly among independent workers. The similarities and differences in performance across countries may be related to differences in country policies and institutions as well as the dynamism of the informal economy. For instance, the good performance of the self-employed in the Dominican Republic has been associated with the growing and well-performing tourism sector and construction booms that the country has experienced, in the context

of a relatively flexible labor market. Labor rigidities, however, are known to be relatively high in Argentina and Bolivia, although other factors may also be at play—for example, excessive costs of registration in Bolivia, lax enforcement, and coordinated general tax and payroll evasion in Argentina.

Again, these “unexplained” earnings gaps cannot be interpreted as evidence of segmentation in the labor market. They provide an order of magnitude of the pecuniary value that informal job amenities may have for informal salaried and independent workers net of the implicit value of the foregone benefits of formal salaried jobs. Yet, they provide inconclusive evidence to discriminate against the segmentation and comparative advantage hypotheses.

Moreover, the results may still suffer from biases arising from the failure to fully account for unobserved productivity differences between workers. In fact, even if it accounts for self-selection into the sectors, the analysis still assumes that workers’ positions in the earnings distribution remain intact were they to switch sectors. For example, it presumes that an informal worker positioned at the top of the informal earnings distribution would remain at the top of the formal salaried distribution were he or she to take a formal salaried job. However, when comparative advantage is present, some workers could emerge as top earnings performers for their skills in the informal sector, but if they were to move to a formal job they may actually lose out in the earnings rankings if their unobserved skills are less rewarded in that sector. The decision to participate in the formal sector could depend on the expected return for the individual to both observed and unobserved characteristics. In this case, the above comparisons of high-earnings workers in the informal and formal sectors would not give a true measure of the potential change in earnings that either would derive by moving across sectors. This complicates the estimation through conventional regression methods.

Arias and Khamis (2007) apply recently developed econometric methodologies to Argentinean data (Heckman, Urzua, and Vytlačil 2006; Heckman and Vytlačil 2001, 2005) to deal with these issues and properly analyze the relevance of labor market comparative advantage in the participation and earnings performance of workers in the formal and informal sectors. To correct for selection biases, they use variations in the enforcement of labor legislation across Argentine provinces and workers’ reported preferences for salaried or independent work as factors that affect employment sector participation, but not the earnings returns to

TABLE 3.2

Impact of informality on earnings in Argentina, estimated log-earnings differences and treatment parameters

Treatment type	Formal salaried versus informal salaried [SE]	Formal salaried versus self-employed [SE]	Informal salaried versus self-employed [SE]
Treatment on the treated	2.088* [0.187]	0.187 [0.443]	−0.449 [0.426]
Treatment on the untreated	1.892* [0.204]	−0.122 [0.245]	−0.989** [0.510]
Average treatment effect	2.002* [0.105]	0.105 [0.291]	−0.600* [0.244]

Source: Arias and Khamis 2007.

Note: SE = standard error.

* $p < .005$.

** $p < .001$.

formality or informality. The methods yield a whole range of earnings differentials as a function of workers' propensity to formal or informal employment, by comparing workers who are at the margin of indifference between jobs fit to their skills and preferences in the three sectors. A summary of the results is presented in table 3.2. This presents a distinct set of summary parameters to answer different policy questions: the average treatment effect (ATE), that is, the mean earnings gain from formality for a randomly selected worker; the treatment on the treated (TT), that is, the mean earnings gain from formality derived by those with characteristics similar to workers currently in formal jobs; and the treatment on the untreated (TU), that is, the mean earnings gain (or loss) for those in informal (salaried or independent) jobs were they to switch to formal salaried jobs.⁹ These are alternative measures of the mean earnings gain from having a formal occupation for workers with the same set of observed and unobserved characteristics, who are indifferent between a formal and an informal job and are found participating in different sectors.

The results corroborate the mixed view of the Argentine labor market and support the importance of comparative advantage in workers' selection into formal salaried and self-employment but also a role for segmentation. On the one hand, the results reveal little difference in the earnings of formal salaried and independent workers once one fully accounts for the sorting of workers based on preferences and the returns to their observed and unobserved skills. All three treatment parameters are statistically insignificant. The ATE is positive; however, this reflects a combination of a positive earnings effect for workers whose characteristics make them more prone to formal salaried work (TT) and a

negative effect for those workers who have more self-employed-like characteristics. That is, workers with independent-like characteristics (observed and unobserved) would receive lower earnings were they to move to formal salaried jobs, although the effect is not statistically significant. The results are again consistent with workers' primarily voluntary motivations to be independent in Argentina. These results are in contrast to the findings from quantile earnings estimation and suggest that the sorting of workers based on differences in tastes for various types of work is of major importance for those entering self-employment. This again underscores the importance of considering differences in the nonpecuniary qualities of independent work as revealed by the motivational analysis of chapter 2. The absence of compensating differentials suggests that the perceived amenities (for instance, independence and flexibility) and disadvantages (for example, lack of benefits and risk) of self-employment tend to cancel each other out.

On the contrary, for informal salaried workers all the parameter estimates are positive and large, so that informal salaried work implies very high earnings penalties when compared to formal salaried work. Although the full results suggest some heterogeneity in the full comparison of earnings of formal and informal salaried workers, the summary treatment parameters are very similar, so that workers with informal-like characteristics (observed and unobserved) would experience roughly similar earnings gains were they to move to formal salaried jobs. To the extent that these are derived from comparing identical workers at the margin of indifference between the two sectors, they provide measures of differences in earnings arising from nonpecuniary characteristics of jobs across sectors or from labor market

disequilibria or segmentation. The magnitude of earnings gaps seems very large to arise from compensating earnings differentials and suggests the presence of segmentation between informal and formal salaried employment. This is consistent with the evidence cited above on the involuntary nature of individual reasons for being informal salaried in Argentina. The next section presents an analysis of subjective response data that attempts to establish more definitely whether the value of nonpecuniary job characteristics fully compensates for the lower earnings of informal workers in Argentina as well as other countries.

Informality and self-rated welfare

One way to assess whether the informal-formal earnings disparities reflect actual differences in welfare is by considering individuals' perceptions of their own welfare. Some household surveys in a number of countries have collected data on subjective well-being by asking individuals questions such as the following: "Do you consider yourself or a member of your family to be: very poor, poor, nonpoor, or rich?" We use available surveys for Argentina, Bolivia, and the Dominican Republic to examine the link between informal-formal occupational status and expanded notions of welfare beyond incomes. In Argentina and the Dominican Republic, the surveys include data that allow alternative measures of informal employment including pension coverage, firm size, and the temporary or permanent nature of the labor contract.¹⁰ Regrettably, the Bolivian survey does not allow a clear distinction between informal and formal salaried workers, other than through proxies of whether the worker is temporary and employed in a blue-collar occupation. Workers are classified according to two groups: those answering "rich" or "nonpoor" and those answering "poor" or "very poor." Except for Bolivia, the surveys contain sufficient income data that allow classification of workers also in terms of their income poverty. Regressions are estimated relating the subjective responses to a large number of individual and family characteristics including the job category variables, and comparing these with the results from similar regressions of income-poverty status of the individual according to conventional poverty-line analysis.

Before discussing the results, it is useful to establish some facts on the data. Table 3.3 first presents a comparison of the self-rated and income poverty classifications in Argentina and the Dominican Republic. Despite the markedly different processes involved in classifying workers as income-poor and self-rated poor, there is a high level

TABLE 3.3

Correlation between income and self-rated poverty

Country		Income			Coincidence
		Nonpoor	Poor	Total	
Argentina					
Self-rated	Nonpoor	45.1	13.8	58.8	64.8
	Poor	21.5	19.7	41.2	
	Total	66.5	33.5	100.0	
Dominican Republic					
Self-rated	Nonpoor	46.4	19.0	65.4	62.2
	Poor	18.9	15.7	34.6	
	Total	65.3	34.7	100.0	

Source: Based on Arias and Lucchetti (2007).

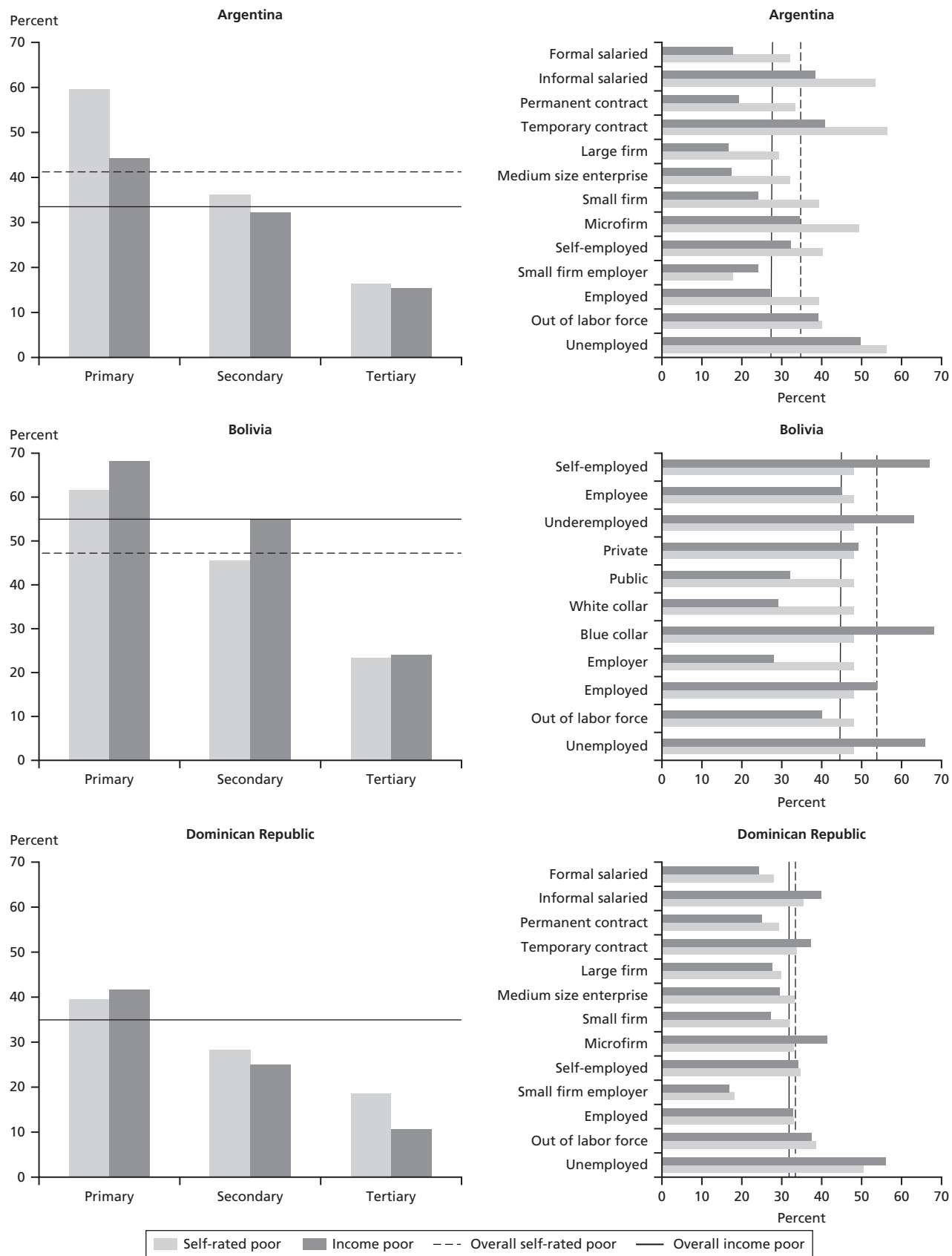
of coincidence in the classifications: in both countries, approximately 65 percent of individuals self-rate their status similarly to what conventional income poverty measures dictate.

Figure 3.6 also shows that self-rated poverty measures follow fairly closely the pattern of income poverty classifications by education level and labor market status of respondents. In the case of Bolivia, income poverty rates are calculated with the 1999 national living-conditions survey. There is a strikingly similar correlation in the case of education; both poverties are lower among the well-educated. There is also considerable coincidence in the case of occupational status, although some noteworthy divergences are already apparent. For example, the gap between the self-rated and income poverty rates of the self-employed is much smaller than among the informal and formal salaried. Moreover, small-business owners tend to self-rate less poor than their income would suggest.

Figure 3.7 and table 3.4 present the main results of the multivariate regression analyses for Argentina and the Dominican Republic. The coefficients in the figure represent the impact of alternative measures of informal employment on the propensity of individuals to consider themselves poor (figures in bold indicate that the effect is zero at conventional statistical significance), and have a similar interpretation to the odds ratios discussed earlier in the chapter. The main conclusions are summarized below. The coefficients in the table are standardized for proper comparisons of the effect of the variables on self-rated (binary indicator) and income poverty (a continuous variable).

FIGURE 3.6

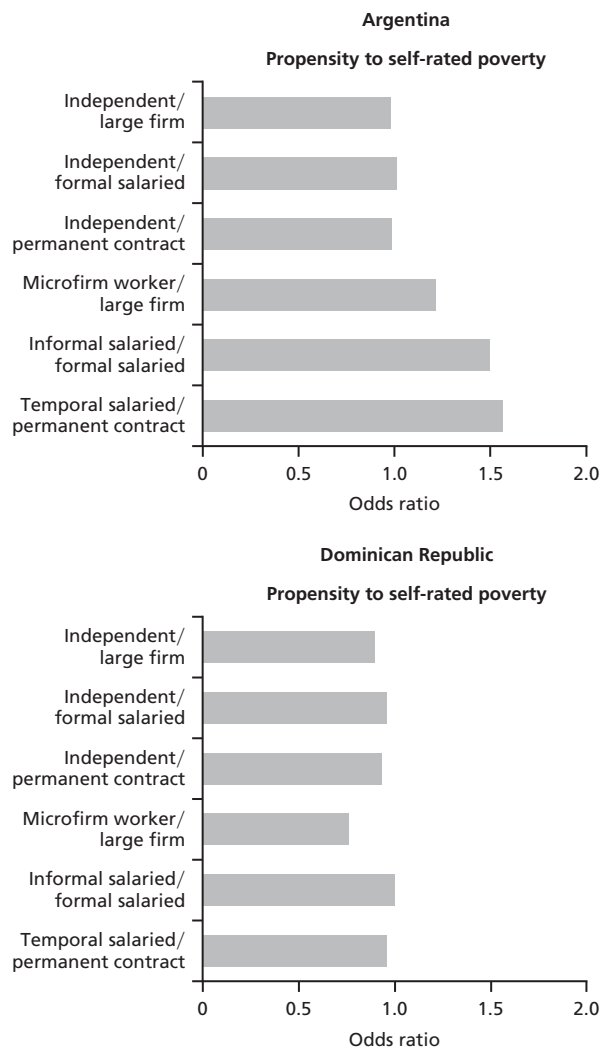
Fraction of workers who are income-poor and self-rated poor, by education and occupational group



Source: Based on Arias and Lucchetti (2007).

Note: Self-rated poverty is based on individual responses. Income poverty is based on national poverty lines and per capita family incomes.

FIGURE 3.7

Impact of informality on self-rated poverty

Source: Based on Arias and Lucchetti (2007).

Note: Self-rated poverty is based on individual responses.

The impact of being an informal salaried employee on self-rated poverty is identical to its impact on income poverty in Argentina, but is quite opposite that in the Dominican Republic. The informal salaried Argentines consider themselves poorer than formal salaried workers, holding other factors constant, even when we control for their own labor income and the incomes of other family members. Moreover, the informal salaried in Argentina tend to self-rate as poor as their family income indicates, all other factors held constant. Thus, they have a lower level of self-rated welfare that cannot be explained on the basis of their individual and family characteristics (proxies for

wealth, living conditions, and so on). However, the Dominican informal salaried consider themselves as poor as formal salaried workers, despite being more likely to be income-poor. Thus, incomes are not a good proxy for the actual welfare situation of the Dominican informal salaried workers. As can be seen in figure 3.7 and table 3.4, these results are strikingly robust to different definitions of informal salaried employment. In the case of Bolivia, the regression results of Arias and Sosa-Escudero (2005) find that employees, temporary salaried workers, and blue-collar workers, most of whom are likely to be informal, also have lower self-rated welfare than the formal salaried.

Contrary to the results for the informal salaried, independent workers in both countries consider themselves as poor as formal salaried workers. This is at odds with their being more likely to be income-poor in Argentina, all other factors constant, but is consistent with their lower income poverty in the Dominican Republic. Thus, in these two countries, independent workers report having similar levels of welfare as the formal salaried workers. This was also found to be the case in Bolivia, where the self-employed are as likely to be self-rated poor as formal salaried workers, holding other factors constant (Arias and Sosa-Escudero 2005).

An important finding also shown in figure 3.7 and table 3.4 is that access to social protection correlates with higher self-rated welfare. In both Argentina and the Dominican Republic, individuals who live in families where some member has social security benefits are about 17 percent less likely to self-rate as poor (see figure 3.8). This result also holds if access to benefits is proxied by employment in large firms or by having a permanent labor contract. Thus, workers seem to attach an important welfare value to having access to social protection.

The results that Dominican informal workers report the same level of welfare as formal salaried employees are corroborated by the recently collected data on job satisfaction as part of the special informal survey in this country. Table 3.5 presents the responses to the question “How satisfied are you with the following aspects of your current job?” for the four occupational groups. These questions refer to workers’ general satisfaction with their jobs and a variety of specific working conditions including incomes, working hours, benefits, flexibility, and mobility opportunities. As can be seen, there is strikingly little difference in the levels of satisfaction reported by informal salaried and informal independent workers relative to formal salaried workers

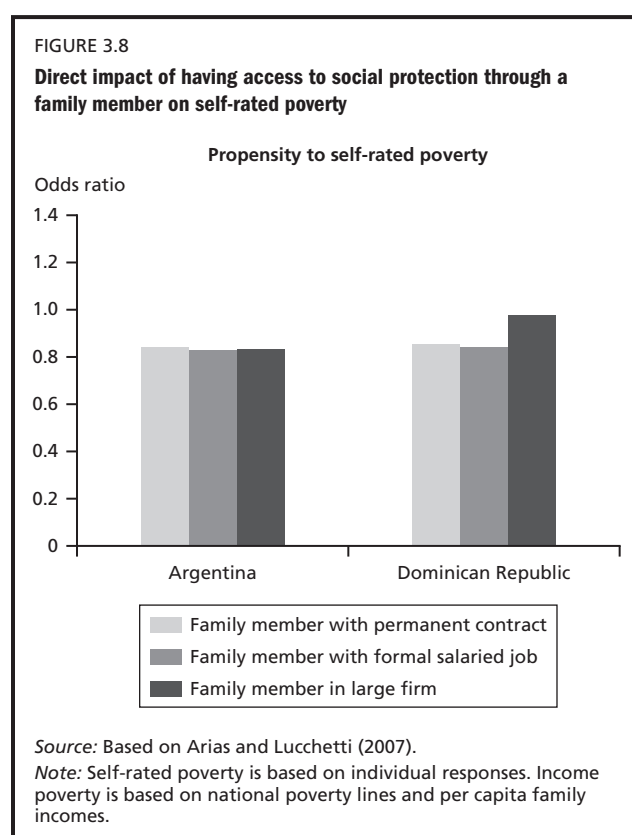
TABLE 3.4

Impact on the propensity to self-rate poor and be income poor (normalized logit regression coefficients)

Informality proxy	Argentina		Dominican Republic	
	Self-rated	Income	Self-rated	Income
<i>Informality proxied by firm size</i>				
Independent worker	0.06	0.27	-0.07	-0.12
Microfirm worker	0.17	0.23	-0.13	0.06
Family member in large firm	-0.11	-0.34	-0.02	-0.39
<i>Informality proxied by pension</i>				
Independent worker	0.06	0.24	-0.01	-0.06
Informal salaried worker	0.26	0.27	0.04	0.10
Family member formal salaried	-0.16	-0.41	-0.11	-0.45
<i>Informality proxied by type of contract</i>				
Independent worker	0.04	0.23	-0.04	-0.06
Temporal salaried worker	0.27	0.29	0.00	0.09
Family member with permanent contract	-0.15	-0.41	-0.16	-0.35

Source: Based on Arias and Lucchetti (2007).

Note: Coefficients are normalized so that they measure the relative weight of each explanatory variable on the total variation explained. Coefficients in bold are not statistically significant at the 5 percent confidence level. Regressions control for other individual and family characteristics.



(informality proxied by pension contributions). Interestingly, the formal self-employed report the higher levels of job satisfaction. The fact that the overall job satisfaction rate is only 5–7 percent higher for the formal salaried than for

the informal is remarkable given that the differences in characteristics between the groups are not controlled. The only exception is with regard to the benefits offered by the job which, as can be expected, are regarded as inferior by the informal, as well as mobility opportunities (which may reflect between-group differences in skills). The lower satisfaction with respect to incomes among the informal salaried is exclusively due to family workers being included in this group, when the latter are excluded, the rates become similar to those for the formal salaried. Moreover, in terms of flexible work schedules, the informal self-employed Dominicans report higher levels of job satisfaction.

Regrettably, the recent Argentine informal survey did not include questions on job satisfaction; however, a comparison with the responses to a similar question in Colombia offers a useful polar case to benchmark. These are presented in table 3.6. In sharp contrast to the Dominican Republic, the informal salaried and informal self-employed workers in Colombia are clearly more dissatisfied with their jobs in all of the surveyed dimensions. In several cases, the differences are fairly large, although still only a quarter of informal workers report being unsatisfied with their jobs.

The similarities in responses between the informal salaried and the informal self-employed in Colombia are quite remarkable. The results are consistent with the overall lower degree of voluntariness directly revealed and implicit in the reasons to be in the current occupation for both

TABLE 3.5

Job satisfaction and informal employment in the Dominican Republic

What is your overall level of satisfaction with your economic activity or main job?	Formal salaried	Informal salaried	Formal independent	Informal independent
Very satisfied	9.3	5.7	25.2	8.8
Satisfied	69.3	67.6	55.0	62.4
Unsatisfied	20.0	25.0	14.7	25.0
Very unsatisfied	1.4	1.7	5.1	3.9
Total	100.0	100.0	100.0	100.0
What is your overall level of satisfaction with the following aspects of your main job?	Formal salaried	Informal salaried	Formal independent	Informal independent
<i>The number of hours you work</i>				
Very satisfied	5.3	4.5	12.7	5.9
Satisfied	73.2	70.7	75.4	71.6
Unsatisfied	19.1	22.4	10.4	20.0
Very unsatisfied	2.5	2.4	1.5	2.5
Total	100.0	100.0	100.0	100.0
<i>The level of income received (earnings and monetary pay)</i>				
Very satisfied	3.9	2.4	7.2	3.7
Satisfied	48.5	40.8	61.9	49.2
Unsatisfied	40.1	39.9	22.3	36.8
Very unsatisfied	7.5	8.5	8.6	10.4
Not applicable/don't have	0.0	8.4	0.0	0.0
Total	100.0	100.0	100.0	100.0
<i>Additional benefits received (health insurance, pension, paid vacations, and so on)</i>				
Very satisfied	5.0	1.0	1.9	0.1
Satisfied	70.7	26.1	1.3	1.0
Unsatisfied	18.9	33.8	0.5	0.3
Very unsatisfied	5.5	18.9	0.0	0.1
Not applicable/don't have	0.0	20.2	96.3	98.5
Total	100.0	100.0	100.0	100.0
<i>Flexible work schedules</i>				
Very satisfied	6.1	5.4	15.1	8.8
Satisfied	70.2	72.7	80.1	78.8
Unsatisfied	21.7	19.5	4.1	10.9
Very unsatisfied	2.0	2.3	0.7	1.5
Total	100.0	100.0	100.0	100.0
<i>Opportunities for economic mobility or to be promoted</i>				
Very satisfied	6.1	3.6	9.5	5.2
Satisfied	58.0	46.7	61.7	46.4
Unsatisfied	32.2	39.4	15.3	35.3
Very unsatisfied	3.8	10.3	13.5	13.1
Total	100.0	100.0	100.0	100.0

Source: Author's estimates, based on household survey data, 2006.

Note: Numbers are rounded.

groups presented in chapter 2. These results for Colombia and the Dominican Republic caution against simple-minded generalizations across countries and highlight that voluntariness or levels of job satisfaction can vary among both the informal self-employed and the informal salaried sectors across and within countries.

Finally, in order to have a quantitative notion of the potential value attached to the amenities of informal and

formal work, figure 3.9 presents some calculations derived from the self-rated poverty regressions for Argentina. These reflect the difference between the actual average labor incomes of various groups of workers and the income level required to make them borderline of self-rating poor. That is, these are measures of the "excess income" above the amount individuals in each group need to reach a level of welfare that puts them with a 50:50 chance of self-rating

TABLE 3.6

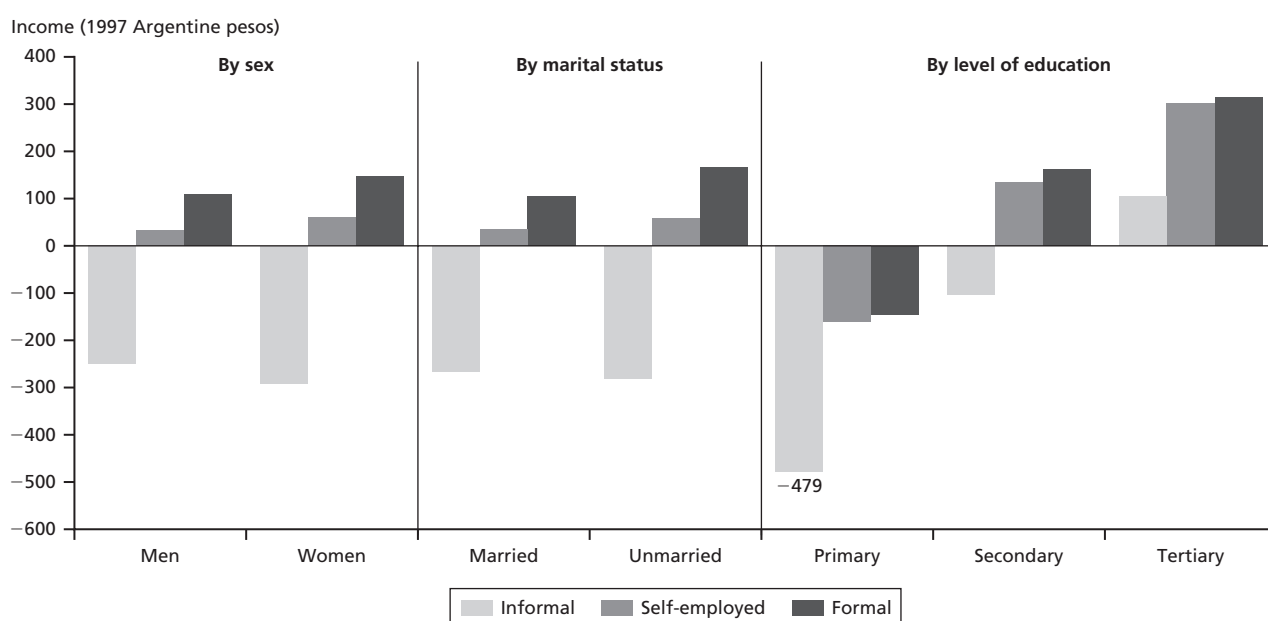
Job satisfaction and informal employment in Colombia

What is your overall level of satisfaction with our economic activity or main job?	Formal salaried	Informal salaried	Formal independent	Informal independent
Very satisfied	13.2	3.0	14.9	4.2
Satisfied	79.3	69.9	71.1	70.2
Unsatisfied	7.4	27.1	14.0	25.6
What is your overall level of satisfaction with the following aspects of your main job?	Formal salaried	Informal salaried	Formal independent	Informal independent
<i>The number of hours you work</i>				
Very satisfied	8.0	2.0	10.0	2.3
Satisfied	78.8	67.4	69.6	67.4
Unsatisfied	13.2	30.6	20.4	30.3
<i>The level of income received (earnings and monetary pay)</i>				
Very satisfied	6.8	1.3	8.9	1.8
Satisfied	62.5	45.3	55.4	45.3
Unsatisfied	30.7	53.3	35.7	53.0
<i>Additional benefits received (such as health insurance, pension, and paid vacations)</i>				
Very satisfied	8.3	0.9	5.1	1.0
Satisfied	73.7	28.7	43.5	29.4
Unsatisfied	18.0	70.4	51.4	69.6
<i>Flexible work schedules</i>				
Very satisfied	5.9	1.2	8.5	1.7
Satisfied	81.1	68.1	71.7	68.3
Unsatisfied	13.0	30.7	19.8	30.0
<i>The application of your knowledge at work</i>				
Very satisfied	17.6	6.4	22.3	8.7
Satisfied	75.7	76.1	66.8	75.3
Unsatisfied	6.7	17.5	10.9	16.0

Source: Author's estimates, based on household survey data, 2006.

Note: Numbers are rounded.

FIGURE 3.9

Differences between actual labor incomes and the level of income needed to avoid self-rating poverty in Argentina

Source: Based on Arias and Lucchetti (2007).

Note: Estimates based on the logit regression of self-rated poverty. Positive (negative) values indicate that average labor incomes exceed (fall short of) the self-rated poverty threshold of individuals.

poor. The negative values for the informal salaried indicate that they actually have an income deficit—that is, they earn less than what is needed to be equally likely to self-rate poor or nonpoor. The fact that independent workers have lower values than the formal salaried means that they derive a similar level of self-rated welfare with less income.

Since the regressions hold other observed individual and family characteristics constant, these differences may partially reflect the significance of the above-mentioned nonpecuniary aspects of employment for the two informal groups. Although the data do not provide sufficient grounds to estimate the imputed value of formal benefits in relation to the cost of contributions and of the nonpecuniary amenities of informal work, the figures are suggestive of the relative significance of these factors for different groups of workers. The implications for antipoverty and social protection policies will be discussed in detail in chapter 7.

Conclusions and policy implications

This chapter lends credence to both the “exclusion” and “voluntary” nature of informal employment suggested by workers’ reported motivations to be in their current occupation presented in chapter 2. Evidence from workers’ reported motivations to be in their current occupation, the sources of earnings differentials, and self-rated welfare assessments lends support to the view that the majority of independent workers are largely voluntary and attach significant value to the nonpecuniary benefits of autonomous work. Meanwhile, informal salaried workers tend to be excluded from more desirable jobs in either formal salaried or self-employment.

The existence of a sizable earnings differential between informal and formal salaried workers, unrelated to compensating differentials, like those found in Argentina and Bolivia, has implications for the functioning of labor markets. This can reflect “queues” for formal salaried sector jobs, given that they are comparatively better paid across the spectrum of low- and high-paid jobs in the labor market and have social benefits. It may, for instance, be a product of the labor market not being flexible enough to equalize earnings through arbitrage. However, as discussed in chapter 2, it may be related to other sources of labor segmentation. For instance, it may reflect a serious problem of general evasion of income and value-added taxes that must be addressed (at least in part) with tighter enforcement and improvements in the structure of these taxes. Thus, policies

to address the barriers to more desirable formal salaried jobs are of first-order importance, and may include removal of labor market frictions, tighter enforcement and improvement of labor and tax laws, and improvement of the role of unions in achieving better overall employment, wages, and productivity growth.

As noted in chapter 2, there is a two-tier divide among independent workers: a majority of workers who choose this sector voluntarily and conform closer to entrepreneurship motives, and an important but lower fraction who use it as a safety net. The lower tier faces significantly higher earnings disadvantage in the low-paying jobs, but the best-paid independent workers enjoy remunerations similar to the best-paid formal salaried workers with similar characteristics. The evidence for Argentina is consistent with workers sorting into formal salaried and self-employment occupations according to labor market comparative advantage. That is, some workers find advantageous niches for their observed and unobserved skills and tasks in sectors or occupations where jobs have a different propensity to be exercised as formal salaried or independent. Moreover, the results on self-rated welfare differentials suggest that nonpecuniary factors may be to a large extent compensating independent workers for any lower earnings. This is consistent with chapter 2’s survey-reported motivations for being independent and not contributing to social security.

However, the Dominican Republic (as well as Mexico, as will be argued in chapter 4) and Colombia provide exceptions that caution that this distinction between the informal salaried and the informal self-employed need not be true in every country. Overall, the nature and determinants of informal employment would depend on country-specific contexts, particularly on cross-country and over time variation in formal sector productivity, the demographic and skills composition of the labor force, and the incentives to comply with tax and labor regulations (including participation in the social security system).

The fact that a fraction of independent workers enjoys higher—although presumably more variable—incomes but are not contributing to social security poses the question: How much will they be willing to pay for the social protection benefits that formal wage earners enjoy (which, as indicated above, are positively valued by workers)? The issue of employment protection and old-age security merits an analysis that considers the incentives to participate in the social security system of workers with

different preferences regarding job flexibility, with different concerns as to their futures, with different intertemporal discount rates, and with respect to who derives different levels of welfare from a particular benefit package. Workers may have a differing willingness to pay or accept lower take-home earnings in exchange for such benefits depending on their individual preferences, the cost and quality of the services (real and perceived) provided by the public and private sectors, and the characteristics of alternative sources of services and benefits not related to the labor contract (for example, informal insurance, social networks, and such). There also may be a role for flexible benefits plans. Chapter 7 analyzes these issues in more detail, while the next chapter summarizes the main lessons for labor market policies emerging from the existing literature for Latin America and this report.

Notes

1. This effect cannot be identified for independent workers, because the majority of them are in microfirms and one-person enterprises.
2. Earnings of formal salaried are net of labor tax contributions, while independent earnings are computed netting out the costs and returns to capital as much as labor surveys allow.
3. Maloney and Mendez (2003) find evidence that the influence of minimum wage norms is far stronger on the more concentrated informal earnings distribution in Argentina.
4. A low-earnings formal job gets the earnings ceiling of the bottom 20 percent formal workers (the 20th percentile of the formal salaried distribution), while a high-earnings formal job gets the earnings floor of the top 20 percent formal workers (the 80th percentile).
5. The economics literature offers various rationales for lower returns to experience in independent work. While pay raises based on seniority are used in salaried work to induce the best workers to remain in the firm, the self-employed have less incentives to shirk in the job (or quit). Also, since independent workers often have a sunken investment at the start-up of their businesses, they may not be able to move out quickly from a poor-performing activity.
6. The chapter loosely uses the term “low- and high-paying jobs” to refer to the conditional percentiles of the earnings distribution.
7. For more on the empirical testing of labor market segmentation and compensating differentials theories, see Magnac (1991). The earnings gap decomposition technique was first presented by Oaxaca (1973). For more on earnings decompositions, see Oaxaca and Ramson (1994).
8. The results here differ from the findings of Pratap and Quintin (2006), who found no evidence of systematic differences in earnings between formal and informal salaried workers. They use 1993–95 *Encuesta Permanente de Hogares* (EPH) data and propensity scores methods to deal with the problem of bias due to self-selection of low productivity workers into the informal sector. While the

methods used here are different, the results might reflect the difference in time periods. The informal-formal salaried earnings gap has widened considerably in the last 10 years.

9. As shown by Heckman and Vytlačil (2001, 2005), these parameters can be derived from an estimate of the marginal treatment effect using local instrumental variables (LIVs) and other approaches. The results shown here are based on semiparametric estimation but also robust to different empirical specifications and alternative estimation approaches. (See Arias and Khamis [2007]).

10. The surveys used are the 1997 Encuesta de Desarrollo Social in Argentina, the 1999 Encuesta Nacional de Aspiraciones y Prioridades de Desarrollo Humano in Bolivia, and the Encuesta de Condiciones de Vida in the Dominican Republic.

References

- Arias, O., and M. Khamis. 2007. “Comparative Advantage and Informal Employment.” Photocopy. World Bank, Washington, DC.
- Arias, O., and W. Sosa-Escudero. 2005. “Subjective and Objective Poverty in Bolivia.” Background paper for the 2005 World Bank Bolivia Poverty Assessment. World Bank, Washington, DC.
- Arias, Omar, Fernando Landa, and Patricia Yáñez. 2007. “Movilidad Laboral e Ingresos en el Sector Formal e Informal de Bolivia.” Documento de Trabajo, UDAPE, La Paz, Bolivia.
- Arias, Omar, and Monserrat Bustelo. 2007. “Perfiles y Dinámicas del Empleo Informal en América Latina.” Photocopy. World Bank, Washington, DC.
- Arias, Omar, and Leonardo Lucchetti. 2007. “Informal Employment and Self-Rated Welfare: Measuring Compensating Differentials.” Photocopy. World Bank, Washington, DC.
- Carneiro, P., J. Heckman, and E. Vytlačil. 2001. “Estimating the Rate of Return to Education When It Varies among Individuals.” Paper presented at the Royal Economic Society annual meeting. Durham, UK, April.
- Galiani, S., and F. Weinschelbaum. 2006. “Modelling Informality Formally: Households and Firms.” Photocopy. World Bank, Washington, DC.
- Heckman, J., and G. Sedlacek. 1985. “Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market.” *Journal of Political Economy* 93 (6): 1077–125.
- Heckman, J., and E. Vytlačil. 2001. “Policy-relevant Treatment Effects.” *American Economic Review* 91 (2): 107–11.
- . 2005. “Econometric Evaluations of Social Programs.” In *Handbook of Econometrics*, Vol. 5, ed. J. Heckman and E. Leamer. Amsterdam: North-Holland.
- Heckman, J., S. Urzua, and E. Vytlačil. 2006. “Understanding Instrumental Variables in Models with Essential Heterogeneity.” *Review of Economics and Statistics* 88(33): 389–432.
- Magnac, T. 1991. “Segmented or Competitive Labor Markets?” *Econometrica* 59 (1): 165–87.
- Maloney, W. F. 1999. “Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico.” *World Bank Economic Review* 13: 275–302.

- Maloney, W., and J. N. Mendez. 2003. "Measuring the Impact of Minimum Wages: Evidence from Latin America." Working Paper 9800, National Bureau of Economic Research, Cambridge, MA.
- Oaxaca, R. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14: 693–709.
- Oaxaca, R., and M. Ramson. 1994. "Discrimination and Wage Decomposition." *Journal of Econometrics* 61 (1): 5–21.
- Pratap, S., and E. Quintin. 2006. "Are Labor Markets Segmented in Argentina? A Semiparametric Approach." *European Economic Review* 50 (7): 1817–41.
- Rosen, S. 1981. "The Economics of Superstars." *American Economic Review* 71: 845–58.
- Smith, A. 1876. *The Wealth of Nations*. Chicago: University of Chicago Press.

CHAPTER 4

The Informal Labor Market in Motion: Dynamics, Cycles, and Trends

SUMMARY: *This chapter takes an aggregate view of the informal sector, asking what determines the size of the sector, both across the business cycle and across longer periods of time. In the process, we shed more light on the razón de ser of the informal sector, the overall functioning of labor markets in least-developed countries (LDCs), and what legal and regulatory factors affect the allocation of workers across sectors. The chapter begins with a discussion of gross labor flows—the movements of workers across types of work—to establish the dynamic relationships among sectors. The chapter then takes a step back and locates the labor market in the context of a standard small-country macroeconomy model to explain pro-cyclical movements in informality. Finally, it examines determinants of the longer-term trends observed in LDC labor markets.*

CHAPTER 3 DISCUSSED THE DECISION-making process of workers choosing between the formal and informal sectors. This chapter takes a more aggregate view, asking what determines the size of the sector, both across the business cycle and across longer periods of time. In the process, we shed more light on the *razón de ser* of the informal sector, the overall functioning of LDC labor markets, and what legal and regulatory factors affect the allocation of workers across sectors between the formal and informal sectors.

We begin the chapter with a discussion of gross labor flows—the movements of workers across sectors of work and unemployment—to establish the dynamic relationships among sectors. This approach has two advantages. First, it offers a complement to the traditional comparisons of wages adjusted for human capital that, because of their inability to account for unobserved job characteristics and nonpecuniary welfare effects, turn out to be *faux amis* in the quest to establish segmentation in the market or relative inferiority or job quality of sectors. Second, it allows us to draw on the recent advanced country literature to offer an

updated view of how the labor market adjusts across the business cycle. Overall, the gross labor flow analysis offers further evidence that, as a first approximation, informal jobs—particularly those in the informal self-employed sector—are not obviously worse than formal sector jobs for many workers. This does not imply that there are no involuntary components, or that there are no distortions in the business climate, but it does offer an importantly different lens through which to view these issues.

The chapter then takes a step back and locates the labor market in the context of a standard small-country macroeconomy model. This approach allows for a richer exploration of the interaction of macroeconomic shocks and the labor market. More specifically, it offers an explanation for the pro-cyclical movements we sometimes see. In addition, it shows how looking at the relative size and relative earnings of the informal sector can offer a diagnostic of the state of the labor market.

Finally, in the last section we explore the long-run determinants of the size of the informal sector and, in the process, examine some explanations for the expansion of informality in some countries over the 1990s.

Informality through the lens of gross labor force dynamics

The recent availability of panel data sets that permit following workers across employment states allows us to revisit several long-standing issues in the study of developing-country labor markets and, in particular, the role of the informal sector. As noted in the previous chapter, a substantial body of literature with intellectual roots in the Harris-Todaro (1970) model sees informal workers as constituting the disadvantaged component of a labor market segmented by wage rigidities. This general view has implications for the cyclical adjustment of the labor market during recessions: downward rigidities prevent wages from adjusting to adverse shocks to the formal sector, leaving the informal sector to absorb workers who would be unemployed in societies where workers could afford to be so. As will be discussed in the next section, on average, it does appear, although with important and frequent exceptions, that the informal sector shows a countercyclical behavior consistent with this view and especially during crises. Figure 4.1 suggests that, especially during the Tequila crisis of 1995 in Mexico and the 1999 Brazilian crisis, formality fell along with the increase in unemployment.

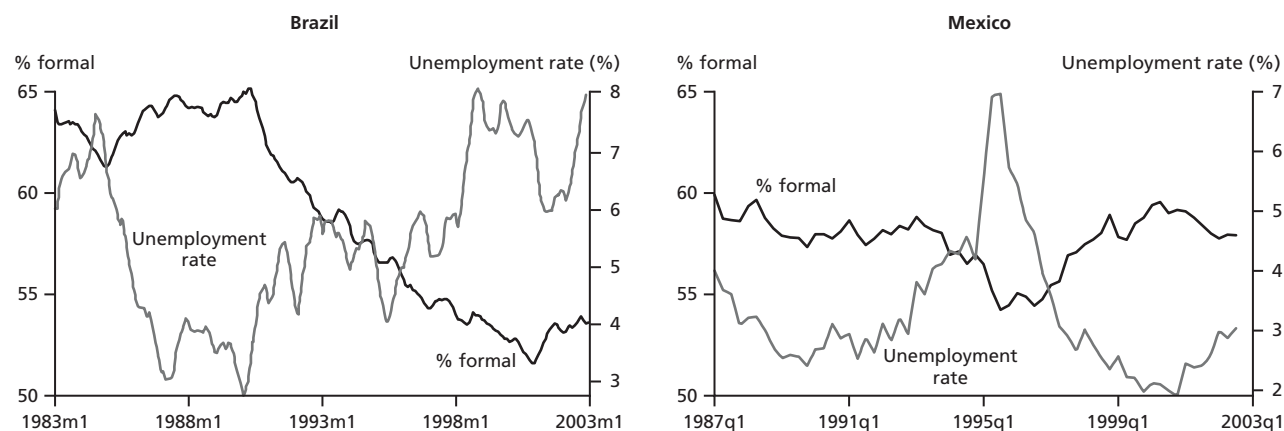
Worker flows provide the movie to the snapshot provided by the simple stock indicators presented in chapters 1 and 2 and further support the idea that much of the informal sector, and particularly the self-employed, is not

primarily an employment sector of last resort. More generally, studies of labor market dynamics have moved to the center of the debates about how advanced-country markets adjust and offer important lessons for Latin America. In the United States, recent work by Davis and Haltiwanger (1992, 1999), Hall (2005), and Shimer (2005a, 2005b, 2005c) have documented the huge amount of churning of workers among sectors. For instance, Davis, Faberman, and Haltiwanger (2005) show that 10 percent of U.S. workers separate from their employers *each* quarter, some moving directly to a new job with a new employer, some becoming unemployed, some exiting the workforce. This literature has set off a debate about the sources of unemployment during downturns, whether caused by the destruction of jobs (Davis and Haltiwanger 1992, 1999) or by the cessation of hiring (job finding) (Hall 2005; Shimer 2005a). It has also introduced several new tools and uncovered striking findings that, as we will show, are directly applicable to the developing-country context and, in particular, explain the cyclical movements of the informal sector.

Bosch and Maloney (2006) and Bosch, Goñi, and Maloney (2006) draw on two rotating panels constructed from employment surveys that permit estimating the probabilities of transiting among labor market states and how these change across time in Brazil and Mexico. Box 4.1 details the theory and some relevant empirical context.

FIGURE 4.1

Formal share of the labor force and unemployment, Brazil and Mexico



Sources: Bosch, Goñi, and Maloney 2006; Bosch and Maloney 2006.

BOX 4.1

Conceptual issues in gross worker flows

Over the last 20 years, the study of the labor market has departed from the Marshallian view of instantaneous adjustments, where prices (wages) are the only allocation mechanism, toward a view of trade in the labor market as an uncoordinated, time-consuming, and costly process for both firms and workers (Pissarides 2000). Agents in the labor market have to spend time and resources in order to achieve the desired outcomes. Workers are in search of jobs offered by the firms while firms are looking for workers to fill their vacancies in order to start production—and even if they agree on the wage, there is a chance that they may not find each other.

The process by which firms and workers find each other has been the subject of researchers trying to understand the functioning of the labor market. One of the dominant views is that this process is governed by a “matching function,” a sort of production function that brings together vacancies and workers, and whose inputs are the number of vacancies and the number of job searchers, $m = m(u, v)$. This matching function is normally considered to be increasing in both arguments and has constant returns to scale. The ratio of vacancies to workers determines what the literature calls market tightness, v/u .

This matching function is the workhorse of an immense amount of research aimed at understanding the flows between employment and unemployment in an equilibrium framework. From it, one can obtain what is the probability of an unemployed worker finding a job by dividing the number of matches by the unemployment rate $m(u, v)/u$.

The relevant question is how this process works in a country with a division between protected or formal and unprotected or informal jobs. This question is actually not very different from the distinction between skilled and unskilled jobs that has been widely studied in a developed-country framework (Acemoglu 2003; Albrecht and Vroman 2002; and Dolado, Jansen, and Jimeno 2002). It requires making a number of assumptions about how the labor market works. Do firms post both types of vacancies? Do workers look for both types of jobs? The theoretical discussion of how the search process works in developed economies is by no means a settled

question. It does, however, highlight potentially important directions for the modeling of labor markets with informal sectors.

First, empirical evidence has shown that informal workers concentrate their activities in the provision of nontradables, whereas formal firms operate mainly with tradables. Hence, it would be plausible to assume an economy where there are formal firms posting vacancies and, in parallel, there are informal firms posting vacancies in a different sector (or, equivalently, self-employment opportunities arise). Workers may then direct their search toward one of the sectors or search randomly. The first case gives rise to two different market tightness conditions and to an arbitrage condition. That is, the expected return from the search must be equal in both sectors. This actually implies that if wages are higher in the formal sector, the average waiting time must be shorter in the informal sector. If workers search randomly, then the probability of finding a formal job is given by the ratio of formal to informal vacancies.

Albrecht, Navarro, and Vroman (2006) propose even another mechanism to generate an informal sector in a searching and matching framework. Assume that workers are heterogeneous. The highly educated workers always work in the formal sector, whereas the uneducated always work in the informal sector. The middle-class workers shift between the formal and the informal sectors, depending on the incentives provided by policy makers. Similarly, Bosch (2006) also employs worker heterogeneity to generate an informal sector. Firms post vacancies, initially undefined; when the match between the firm and the worker has been established, the firm has to decide what type of contract it wants to sign (formal or informal). This decision depends on the idiosyncratic productivity of the match and on labor cost in the formal sector. This approach implies that firms decide a threshold level of formality. If the match is sufficiently good, the firm is willing to bear the higher labor cost of a formal job. If not, it will hire the worker informally.

Labor markets in developing countries most likely contain some or all elements discussed here. Understanding how searching and matching occur in the presence of informal labor markets is essential to determine the effect of public policies in the labor market.

Cyclical patterns in gross labor flows

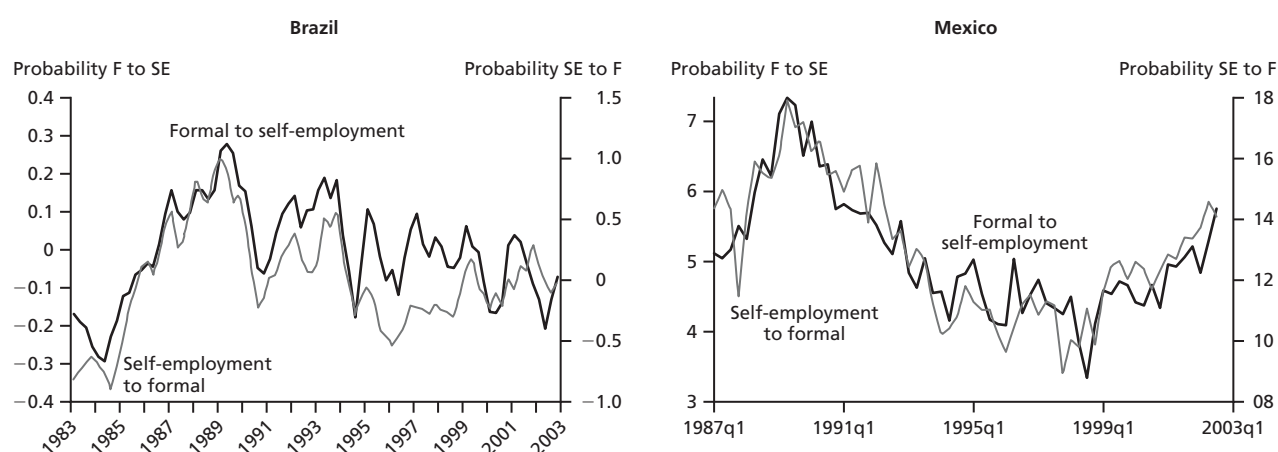
Transitions between formality and informality

Figure 4.2 offers additional evidence in favor of a more integrated view of the labor market where informal jobs offer distinct, but not necessarily inferior, employment opportunities, particularly self-employment. The figure plots the probabilities of transition among formal and informal sectors, the raw transitions in Mexico and in Brazil, in the latter case detrended to account for the longer-term rise in informality across the period that will be discussed later. In both

countries, flows among the formal sector and the two informal sectors are remarkably symmetrical and appear highly correlated across the business cycle, rising in upturns and falling in downturns. That is, it is not the case that, as the economy recovers, we see fewer individuals being thrown out of formality into the informal sector and more informal individuals being able to find jobs in the formal sector. Such asymmetric behavior does characterize flows between employment and unemployment in the Organisation for Economic Co-operation and Development (OECD) and, in

FIGURE 4.2a

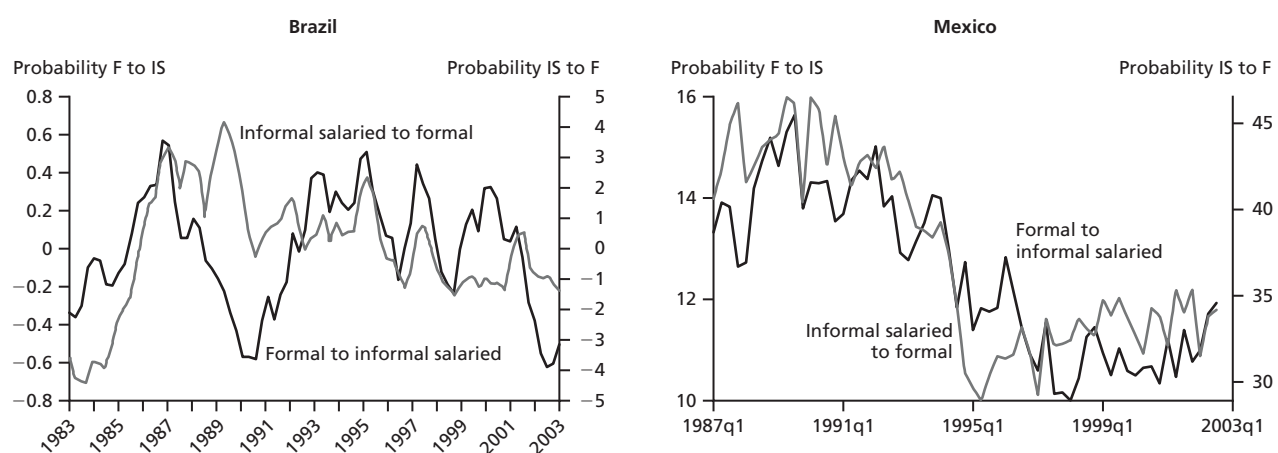
Probability of transition between formal salaried and self-employment, Brazil and Mexico



Sources: Bosch, Goñi, and Maloney 2006; Bosch and Maloney 2006.

FIGURE 4.2b

Probability of transition between formal salaried and informal salaried, Brazil and Mexico



Sources: Bosch, Goñi, and Maloney 2006; Bosch and Maloney 2006.

Note: Magnitudes differ between Brazil and Mexico due to technique used to detrend Brazil.

Brazil and Mexico, *between all sectors of employment both formal and informal, and unemployment*. But, among sectors of employment, what we see are flows in both directions increased during upturns and decreased during downturns, exactly consistent with the pro-cyclical patterns in job-to-job flows observed in U.S. literature (Shimer 2005b) that are generally attributed to workers searching for and finding better jobs in tighter job markets.

The results are most striking for Mexico, but the story is broadly supported with some caveats in Brazil as well. In both countries, the gross flows among the informal, particularly the self-employed, and formal sectors do not behave as if the former were a kind of unemployment, but rather as an alternative job. This is consistent with response data presented in chapters 2 and 3 that most of the self-employed choose to enter the sector and show equal levels of welfare as those in the formal sector. For Mexico, figure 4.3 shows that 75 percent report voluntary entry from both formal and informal salaried work, even during the crisis, although, predictably, the share declines substantially then.

An absence of complementary data for workers entering informal salaried work leaves us to rely on less direct evidence. If we assume that employed workers searching for another job (“Have you been looking for a job over the last two months?”) are “less voluntarily” employed, then it appears from figure 4.4 that in the early 1990s and early 2000s, effectively, job satisfaction was relatively similar across sectors. In the peak of the recession in 1996, both

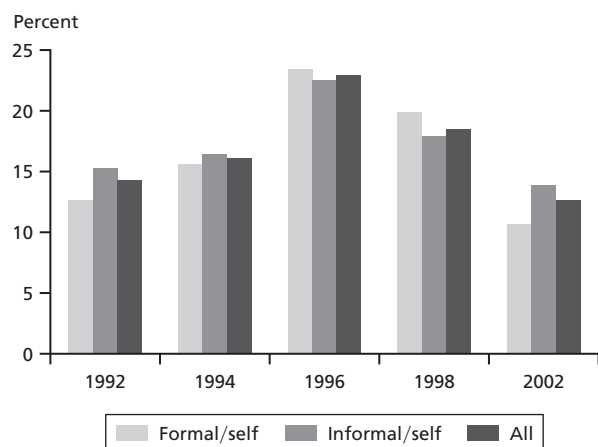
informal sectors are substantially higher than the formal, but relatively close to each other at a remarkably low 6 percent. Broadly speaking, then, in Mexico, sentiments about job quality seem similar among the self-employed and informal salaried, consistent with the high correlation across the cycle of gross flows across all sectors.

However, as chapter 3 notes, this differs strikingly across countries. In the Dominican Republic, self-reported happiness among informal salaried does not suggest that either is worse-off than formal salaried. However, in several countries, the informal salaried clearly are. The motivational responses of Brazilian informal salaried workers suggest that a large fraction are queuing there as well. This may account for the much lower cyclical correlation observed (0.26 for I/F-I transitions versus 0.85 for SE-F/F-SE transitions).

Perhaps paradoxically, a high degree of voluntary entry does not rule out some degree of segmentation for certain workers, nor distortions more generally. First, at any time, there are workers seeking jobs in each sector. Introducing a binding minimum wage or other formal sector wage rigidity will reduce the probability of finding a job in the formal sector and make the available informal jobs relatively less attractive, thereby introducing a progressive asymmetry to the flows and a lower degree of voluntary entry, as rigidities become more severe, especially during downturns.

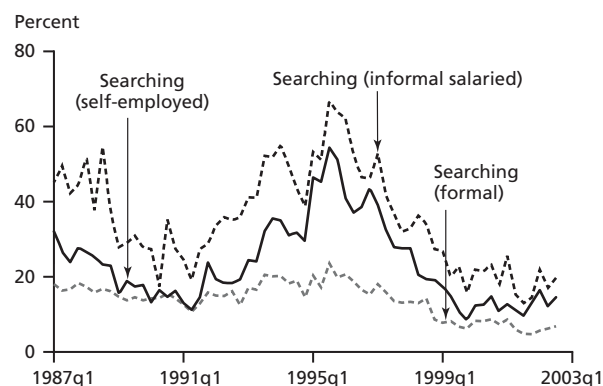
Second, a broad class of distortions—excessive labor taxes or firing restrictions, for example—may not induce

FIGURE 4.3

Involuntary transition to self-employment in Mexico

Source: Author's estimations, based on *Encuesta Nacional de Micro Negocios*, several years.

FIGURE 4.4

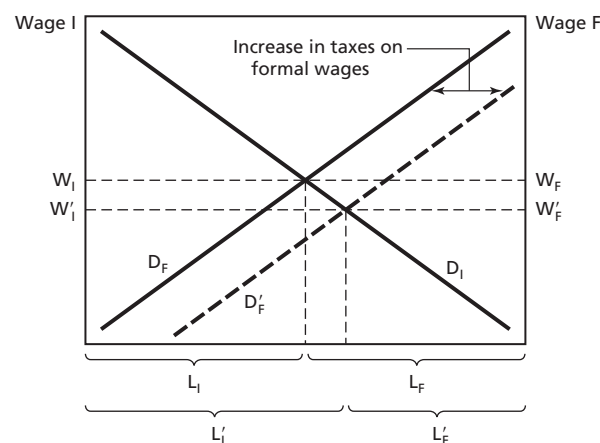
Searching while employed, Mexico

Source: Quarterly data from the National Urban Labor Survey 1987:Q1 to 2002:Q4.

Note: Searching refers to the proportion of employed workers in the sector who claim to be looking for a new job and have not changed employment status in the previous quarter.

segmentation, but will nonetheless reduce the vacancies opened in the formal sector. Effectively, this will be analogous to a shift in the formal sector demand for labor, and will lead to lower earnings in both sectors; but the vast majority of workers entering informality may still declare that they do so voluntarily, given earnings in the formal sector (see figure 4.5). This point is important: it is entirely possible for all workers found in the informal sector to be as well-off as they would be in the formal sector, in spite of very high, nonsegmenting distortions in the formal sector. At the limit, we could imagine an oppressive business climate that prevented growth and the creation of new, modern sector jobs, but where workers freely chose between the extant menu of vacancies. A discussion of models of the impacts of regulations on worker flows and the size of the informal sector is found in box 4.2.

FIGURE 4.5

Decreased availability of formal sector jobs without segmentation

Source: Authors' calculations.

BOX 4.2**Simulated effects of labor market legislation on the size of the informal sector**

Several models have been developed to examine the plausible impact of policies on labor market variables. These are generally calibrated with a generic economy in mind and, hence, should be taken only as suggestive.

Albrecht, Navarro, and Vroman (2006) and Pries and Rogerson (2005) calibrate matching models and simulate the effects of labor market regulation on labor force turnover and composition. Although Pries and Rogerson work only with employment and unemployment, Albrecht, Navarro, and Vroman (2006) explicitly model the informal sector as well.

The Pries-Rogerson model combines variants of two benchmark models from the literature: the Jovanovic (1979) learning model and the Pissarides (1985) matching model. Job flows are driven by idiosyncratic shocks to job productivity, and worker turnover (in excess of job turnover) is driven by a process of accumulation of information about the quality of the match. Both parties to a match observe a signal about the match's true quality prior to deciding whether to form a match, and matches form only if they judge that the match quality exceeds a threshold value. True quality is revealed over time, but only if a match is formed. Labor market regulations have an impact by influencing hiring practices—specifically, the level of the threshold.

Table 4B.1 shows the simulations of three changes in labor market policies: a 15 percent rise in minimum wage, the introduction of dismissal costs worth three months' wages, and an increase of unemployment insurance benefits equivalent to 20 percent of the wage funded by a 15 percent tax on output. In all cases, we see a reduction in the probability that workers will find a job vacancy and substantial changes in worker turnover (rather than job destruction) that drive the results. All policies lead to a reduction in the employment rate of between 1 percentage point (dismissal costs) and 4 percentage points (minimum wages) and similar levels of "formal" sector output. The welfare costs are high only for the minimum wage. Under a combination of policies, the interactions roughly double the impact of all single policies.

Albrecht, Navarro, and Vroman (2006) extend Mortensen and Pissarides (1994) by adding an informal sector and allowing for heterogeneity among workers. They assume that workers differ in their maximum productivities (low, medium, or high productivity) in formal sector jobs. The decision about whether to accept an informal sector job thus depends on a worker's type. All workers have the option to take up informal sector opportunities as these come along, and all workers are equally productive in that sector, but workers who are most

TABLE 4B.1

Effects of varying minimum wages, dismissal costs, unemployment insurance, and taxes

Affected variable	Differences with respect to benchmark of equilibrium			
	Policy			
	Minimum wage	Dismissal cost	Unemployment insurance	Taxes
Probability to meet a vacancy	-0.13	-0.04	-0.14	-0.13
Unemployment duration	3.02	1.32	1.50	1.06
Annual job destruction (%)	-0.20	0.00	-0.30	-0.30
Annual worker turnover (%)	-6.00	-4.3	-1.80	-0.50
Employment rate	-0.04	-0.01	-0.02	-0.02
Output	-0.04	-0.01	-0.03	-0.02
Welfare loss (%)	1.40	0.18	0.29	0.29

Source: Pries and Rogerson 2005.

TABLE 4B.2

Effects of varying severance taxes and payroll taxes

Affected variable	Differences with respect to benchmark of equilibrium		
	Policy		
	Severance tax	Payroll tax	Both taxes
Labor tightness (vacancies/unemployment)	-0.360	-0.090	-0.230
Low productivity (informal)—medium productivity cut-off	0.053	0.063	0.065
Medium productivity—high productivity (formal) cut-off	0.048	0.101	0.085
Unemployment	-0.018	0.003	-0.005
Productivity ^{formal}	-0.061	0.036	-0.003
Wage ^{formal}	-0.069	-0.070	-0.070
Output	-0.019	-0.004	-0.008

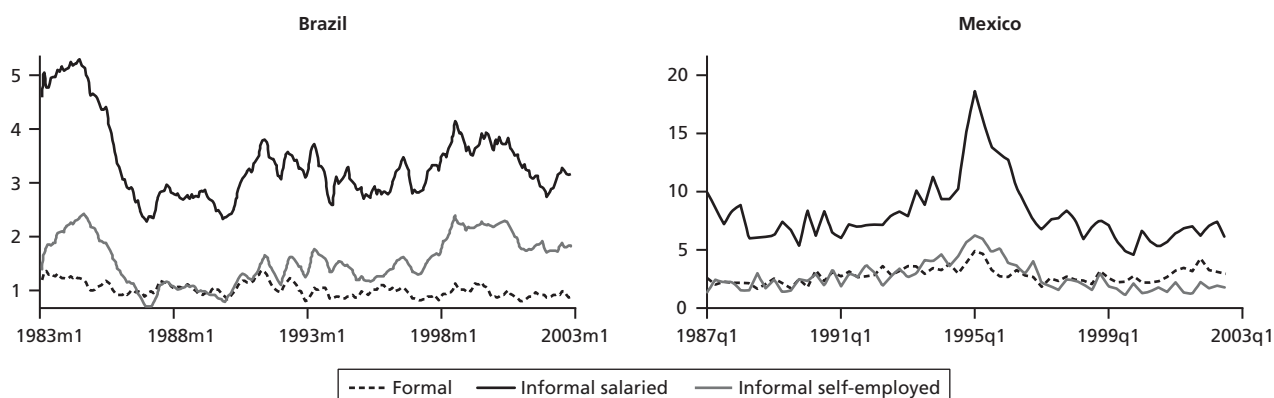
Source: Albrecht, Navarro, and Vroman 2006.

productive in the formal sector will reject informal work in order to wait for a formal job. Similarly, the least productive workers are not hired in the formal sector. The cut-off values that determine how the different types of workers are allocated across sectors are endogenous and are influenced by formal sector labor market policy. A policy change can disqualify some workers from formal sector employment; similarly, some workers accept informal sector work, although they would not have done so earlier.

Effects of raising severance taxes and payroll taxes under this framework are reported in table 4B.2. The first column shows that an increase in severance payments from 0 to 20 percent of the wage makes creating vacancies less attractive for formal sector employers. This should, in principle, increase unemployment. However, a severance tax also encourages firms to keep low-productivity

matches, reducing formal productivity and pushing the unemployment rate down. In this case, the latter effect dominates and unemployment falls. Overall, as fewer formal vacancies are open, more workers are forced into accepting informal jobs. A rise in payroll taxes from 0 to 20 percent has similar effects on the reduction of vacancy creation and the compositional shift toward informal jobs. However, contrary to the severance tax, the introduction of payroll taxes makes the firms keep only highly productive matches. This increases average formal productivity but also increases the unemployment rate. The last column of the table reports the effects of applying both policies simultaneously (each tax is set to be equal to 10 percent). Both taxes make vacancy creation less attractive, employment duration in the formal sector increases (that is, the severance tax effect dominates), productivity falls in the formal sector, and net output decreases.

FIGURE 4.6
Probability of transition to unemployment (separation rate), Brazil and Mexico



Sources: Bosch, Goñi, and Maloney 2006; Bosch and Maloney 2006.

The cause of unemployment: Separations from the informal sectors

A second provocative finding emerges from the flows in and out of unemployment that suggest some rethinking about the informal sector serving as sort of disguised unemployment. Figure 4.6 suggests first that, at any time, the probability of separation from the informal salaried sector is much higher than from the formal sector. In both Mexico and Brazil, informal salaried workers are roughly three times more likely to transit to unemployment, compared with their formal counterparts. Second, job separation probabilities present much higher volatility in the informal sector, especially for the informal salaried. Simulations for Mexico (Bosch and Maloney 2006) and Brazil (Bosch, Goñi, and Maloney 2006) suggest that, in fact, the increases in unemployment during recessions are generated primarily by workers separating from the informal sectors.

A number of possible mechanisms may be at play here. First, workers in the informal sector may be occupying low-skill jobs for which training or experience is not excessively important. During a downturn, these workers are easily disposed of since they are not protected by labor regulations. Further, informal workers are employed normally by small firms, which may suffer disproportionately during recessions and, hence, destroy more jobs. Bosch and Maloney (2006) explore this issue for Mexico and find that job separations are, in fact, largely driven by firm size, with the aggregate finding of acyclical firings in the formal sector being

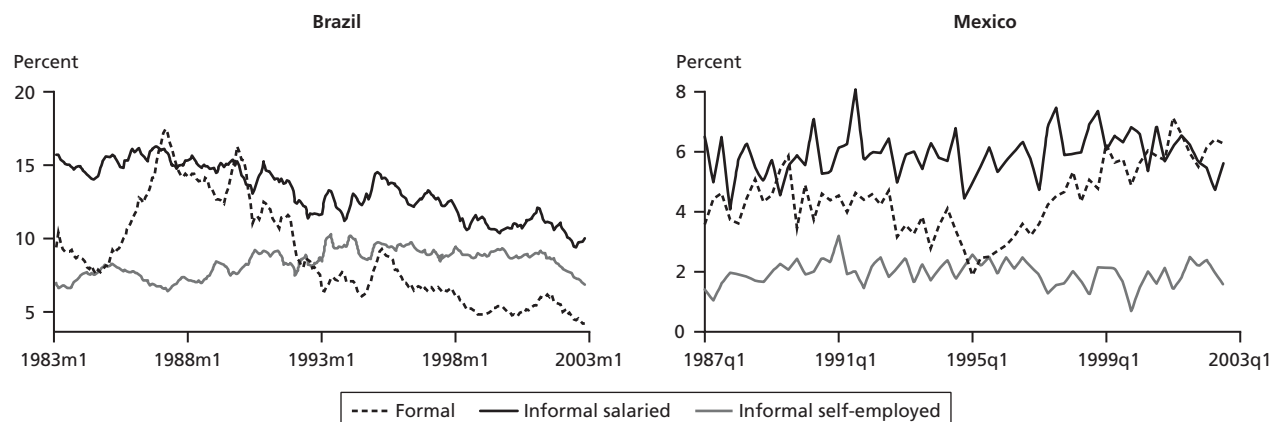
driven by the fact that most workers in formal firms are in large firms. Smaller and medium-size formal firms are extremely countercyclical in their separations. This is, perhaps, consistent with smaller firms being more likely to fail in a bad economy (Jovanovic 1982; Hopenhayn 1988) or perhaps it is because a lack of access to credit to smooth over difficult times leads to smaller firms being unable to hang on to valuable workers while larger firms can.² However, it is also the case that informal workers in firms of all sizes in Mexico show extraordinarily high separation probabilities. This would seem to suggest that, again, it is the type of contractual relationship that is determining firing behavior. However, an alternative possibility is that, though these workers report that they are working in large firms, in fact they are contracted by intermediaries that are smaller firms and that hire and fire as such.

Why the informal sector generally expands in downturns

That said, why does figure 4.1 suggest that the informal sector expands during downturns, much as traditional views would suggest? Indeed, as Gasparini, Haimovich, and Olivieri (2006) and Loayza and Rigolini (2006) have shown, this appears, on average, to be the case.

The advanced-country literature on understanding unemployment through the lens of gross worker flows offers important insights and suggests that we focus on two sets of flows: those into sectors of employment—the *job-finding rate*—and those out of those sectors—the *separation rate*.

FIGURE 4.7

Probability of transition from unemployment (job-finding rate), Brazil and Mexico

Sources: Bosch, Goñi, and Maloney 2006; Bosch and Maloney 2006.

Figure 4.7 shows the job-finding rates of the formal and informal sectors for Brazil and Mexico over the period 1983–2003. The main stylized fact emerging from these figures is the strong pro-cyclicality of the job-finding rate in the formal sector, compared with the relative stability of the job-finding rate in the informal sector. The 1995 crisis in Mexico made the formal job-finding probability oscillate between 30 and 15 percent, while the job-finding rate in the two informal sectors fluctuated only around 4 percentage points. Evidence from Brazil corroborates this finding. The 1984 crises decreased formal job-finding rates dramatically, while the informal counterparts were relatively stable. Moreover, after the reforms at the end of the 1980s, the job-finding rate saw major decreases from 15 to 5 percent. Thus, the job-finding probability in the formal sector appears to be an important adjustment variable.³

These differential rates of job finding across the cycle are critical to explaining the frequent expansion of the informal sector in downturns. As the economy slows, formal sector hiring falls sharply but informal hiring falls much less, and algebra dictates that the relative size of the informal sector is likely to rise. Though we previously showed that separations from informal salaried were the most volatile, net flows into the sector during crises are positive.

This raises the question of why the job-finding rates are so distinct between the two sectors, an issue with striking resonance in the recent puzzle about the high volatility of job-finding rates in the developed economies. In the

United States, the job-finding rate is also strongly procyclical, and Shimer (2005a) and Hall (2005) argue that the magnitude of these fluctuations cannot be well explained by state-of-the-art search and matching models. Both Shimer and Hall argue that one explanation for the excess volatility arises from wage rigidities. In most matching models (see box 4.2), an adverse shock to the economy will lead to a decline in the productivity of potentially hired workers and to their wages. However, if wages cannot fall, then the likely profitability of hiring a new worker falls even more and, as a result, job vacancies will fall much more than if wages could fall.

The relevance to the LDC case is clear. If formal sector wages cannot fall while informal earnings can, then this might well explain why formal job finding fluctuates so much more than informal job finding does. In a sense, then, the focus on gross labor flows could be seen as simply putting the Harris and Todaro (1970) vintage insight in new bottles. However, the evidence to date suggests that the question should be left open. First, Bosch and Maloney (2006) for Mexico and Bosch, Goñi, and Maloney (2006) for Brazil find that certainly for self-employment and, to a lesser degree, for informal salaried work, wages are more flexible than in the formal sector, but not always compellingly so. For example, in the 1988–92 recovery where formal hiring showed a disproportionately large gain, formal salaried and informal salaried earnings stayed effectively equal.

More fundamentally, the mainstream literature has not converged on wage rigidities as being responsible for the common pro-cyclical patterns in job finding that we see in both the Brazilian and Mexican formal sectors and in the United States. Mortensen and Nagypal (2005), for instance, even adding wage rigidities and several other modifications to the Mortensen-Pissarides (1994) model, can simulate only 40 percent of the volatility of the job-finding rate, and they conclude that “in sum, the dilemma [of high employment volatility] persists” (p. 24). This suggests that the distinctive job creation behavior in the two sectors may involve far more than the traditional focus on formal sector rigidities. Further, as discussed in chapter 2, there is substantial evidence of binding minimum wages in the informal sector, although this may be less the case during downturns where “norms” may prove more flexible than legal dictates.

Second, even if wage rigidities turn out to be the critical difference, the mechanism driving them may or may not include the usual considerations of unions, minimum wages, or other distortions. Kennan (2005) and Menzio (2005) both stress asymmetric information in the wage-bargaining process as inducing rigidities in wages with respect to productivity shocks. More generally, in the United States, for instance, Bewley (1999) argues that, for reasons of staff morale, firms would prefer to fire workers, the collective memory of whom will soon fade, rather than reduce all workers’ wages, the resentment about which will not drop. In short, wage rigidities, the traditional driver of segmentation, are part of the story but only one candidate for explaining the sharp fall in formal hiring rates.⁴ Each of the elements of this debate surrounding “the dilemma” of excess volatility in the job-finding rate in the United States applies to understanding the behavior of the formal sector job-finding rate in Latin America.

The broader macro-context

Working more in the de Soto (1989) tradition of firms deciding to be formal or informal, Loayza and Rigolini (2006) develop a model based on Rauch (1991) that offers an alternative perspective on possible drivers of cyclical behavior of the informal, in this case measured as the self-employed. Here, the cyclical movements are not driven by the net effects of worker flows among sectors, but rather across the de Soto margin. Effectively, individual firms weigh the benefits of being formal against the costs and, in

turn, their productivity is affected by their access to the benefits of formality. The size of the informal sector is a function of the productivity differential between formal and informal workers among sectors. This differential is, in turn, determined by the cost of becoming and remaining formal and the distribution of skills in the workforce. Specifically, the productivity differential has a worker-driven component, given by the worker’s individual skills, and a sector-related component, given by the relative informal-formal regulatory burden, the strength of enforcement, and the access to productivity augmenting public services. The size of informal employment is then given by the proportion of workers whose skills fall below a threshold level where the worker is indifferent between the two sectors. When regulation decreases or enforcement increases, the formal sector becomes relatively more attractive for them and more firms join it. Moreover, regulation is a fixed cost that all formal firms have to bear. Therefore, when overall productivity increases, the cost of regulation becomes proportionally smaller so that more firms join the formal sector, and the reverse is true in downturns.

Empirically, Loayza and Rigolini (2006) find that, on average, gross domestic product (GDP) growth is negatively correlated with the size of the informal sector—defined as the share of self-employed—indicating that the informal sector is, on average, countercyclical. Further, the degree of countercyclicality appears smallest for the countries with the very largest informal sectors, Peru and Bolivia, behaving relatively acyclically (see figure 4.8). Consistent with their model, they find improvements in law and order strongly reduce the cyclicity of informal employment.

Explaining pro-cyclical movements in informality

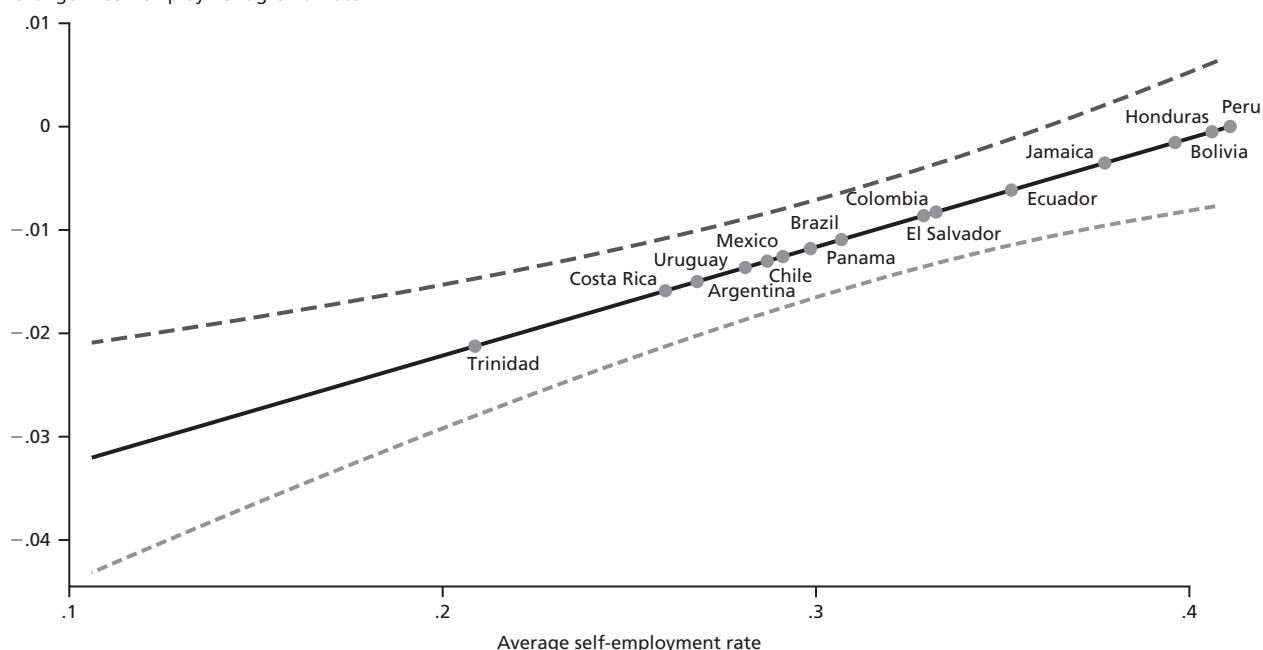
Both of the previous models offer explanations for countercyclical patterns of informality. However, a closer look at figure 4.1 suggests that this is not the whole story. In Mexico, from 1988 to 2001, unemployment fell at the same time that informality rose. That is, informality is *pro-cyclical* across this period. While gross flows increased in both directions, net flows were toward self-employment.

Fiess, Fugazza, and Maloney (2002, 2006) offer an explanation by introducing distinct shocks to the formal and informal sectors. More specifically, they argue that because informality is largely concentrated in the nontradables sector and formal jobs in the more tradables sector, standard

FIGURE 4.8

Informal employment reaction to the business cycle

Change in self-employment growth rate



Source: Loayza and Rigolini 2006.

Note: The graph simulates the change in informal employment growth due to a 5 percentage point change in the GDP per capita growth rate. The simulation is based on a short-run regression like the one shown in table 4, col. 2, but with a sample resulting from applying the constraint of min. obs. = 2. The dashed lines are the 90% confidence bands. Only countries in Latin America and the Caribbean are highlighted. From left to right: Trinidad and Tobago, Costa Rica, Argentina, Uruguay, Chile, Mexico, Panama, Brazil, El Salvador, Colombia, Ecuador, Jamaica, Bolivia, Honduras, and Peru.

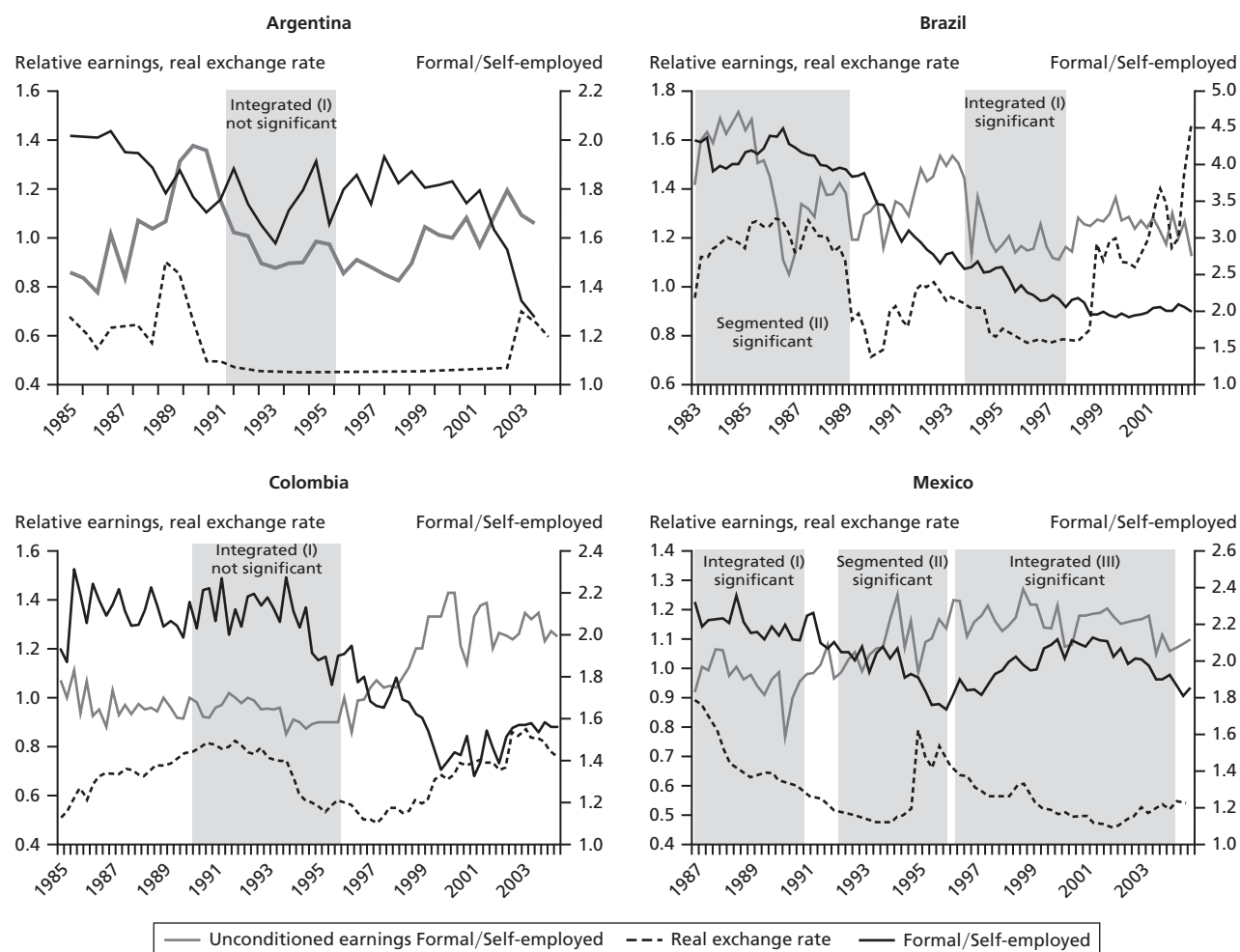
small-economy models that focus on the response of these two sectors to cyclical shocks (see Obstfeld and Rogoff 1996) offer insights into the cyclical behavior of the informal sector. A boom in construction, for instance, opens opportunities for informal contractors and others, and could lead to a pro-cyclical expansion of the sector. Loosely linking the discussion back to the matching models in the previous section, although increases in job-finding rates in the formal sector may generally exceed those in the informal sector, a strong enough stimulus to the nontradables/informal sector can outweigh those increases, leading to pro-cyclical net job creation in the informal sector. On the other hand, a negative shock to the formal sector, perhaps due to a major crisis, combined with downward rigidities in earnings, can lead to countercyclical behavior of informality identified previously, on average. The nature of shocks and the degree of rigidities of earnings, simultaneously determine movements of the exchange rate. The correlation of these variables informs what is driving

movements in the labor market and, in the reverse way, movements in the labor market can inform what is driving movements in the macroeconomy.

Figure 4.9 plots these three variables for Argentina, Brazil, Colombia, and Mexico. Again, the early Mexican period shows a relative increase in the size of the informal self-employed sector and relative self-employed earnings, at the same time that the real exchange rate appreciated sharply. More generally, we find three types of regimes. First, co-integration tests show that, in Brazil during 1994–97 and in Mexico in the period 1987–91, we find a positive co-movement of the two labor series and an appreciation of the exchange rate. Even in Argentina, from 1991 to 1996, and in Colombia, from roughly 1990 to 1996, visual inspection suggests a similar pattern. This suggests that the expansion of informality across this period may have been due to either a productivity or a demand shock to the nontradables sector. A very probable candidate is that the liberalization of the capital account and other reforms taking place around these

FIGURE 4.9

Relative sector shares and earnings, real exchange rate



Sources: Ffies, Fugazza, and Maloney 2006; and author's estimates, based on *Pesquisa Mensal de Emprego*, *Instituto de Pesquisa Econômica Aplicada*, *Encuesta Nacional de Empleo Urbano*, and International Financial Statistics.

Note: Formal/self-employed earnings capture relative reported earnings in the formal and informal sectors. Formal/self-employed captures the relative size of these sectors as a ratio of employed population. "Significant" means a significant cointegrating relationship found. I, II, and III refer to regime (see text).

periods led to an increased demand for nontradables and an expansion of the informal sector.

In a second regime, Brazil (1983–89 and 1998–2002) and Mexico (1992–96) correspond to the case of a negative shock to the formal/traded sector in the presence of some barriers to adjustment of formal sector wages. This is also clearly the case in Colombia after 1996 and for much of the period in Argentina, although the small sample sizes do not permit testing subperiods. The pattern suggests the classic informal/nontradable sector adjusting to take in labor no longer absorbed in the formal sector, discussed previously. This is historically plausible, given that all

four countries experienced deep recessions across these periods.

Finally, in the third regime, Mexico (1997–2003), we find a negative shock to the formal sector in the absence of downward formal sector rigidities that leads to an expansion of the relative size and remuneration in the informal sector and to a depreciation of the currency.

Drivers of the increase in informality

As chapter 1 (figure 1.10) noted, several measures of informality suggest substantial increases in informality across the last decades in Latin America. As noted, using a social

protection definition, urban salaried informality increased in Argentina from 1992 to 2003 by around 9 percentage points and by 17 percentage points in Greater Buenos Aires from 1980–2003. Although, nationally, there was effectively no change in informality in Brazil, in a single decade (1990–2000), informality in the *metropolitan* areas rose by 10 percentage points. In Colombia, figure 4.9 suggests that, over a longer period of time, informality rose by perhaps 6 percentage points. And Mexico, despite relatively stable rates of informality across the 1990s, saw a sharp 4 percentage point inflection in informal salaried work from 2000 to 2004.

Diagnosing across which margin these changes are occurring and why is critical to forming appropriate policy toward the sector. This, however, is made difficult by the fact that the 1990s were years of far-reaching reforms on diverse fronts, as well as the final push on the unfinished agenda of macroeconomic stabilization. Teasing out the impact of each policy is perilous and compounded by the fact that research on the determinants of informality remains relatively inchoate. The aims of this section are thus relatively modest—to use several country cases to highlight viable hypotheses and to present some evidence that informs the emerging discussion. In turn, we discuss issues of macroeconomic stabilization, trade liberalization, labor market reforms, supply side effects, and, briefly, reforms to facilitate formalization of microfirms.

Macroeconomic stabilization and capital account opening

The previous section noted that substantial movements in informality could be generated by distinct shocks to the informal and formal sectors. At the beginning of the 1990s numerous countries in the region fixed the exchange rate as a means of controlling inflationary expectations while liberalizing the capital and current accounts. Virtually all countries saw sharp appreciations of the exchange rate that, in many cases, were seen as arising from backward indexing of earnings in the context of falling inflation. However, the previous section suggests that these countries experienced the Regime I pattern of a rise in relative size and earnings of the informal self-employed along with the exchange rate appreciation. This is consistent with a demand shock to the nontradable informal sector and views stressing borrowing against expectations of future income that had been sharply revised upward in light of the improving macroeconomic situation and reforms. Thereafter, however, in the lead-up to the crises, macroeconomic

deterioration, and the lack of flexibility of nominal formal wages and the exchange rate, led to a more classic pattern of relative contraction of formal sector employment, expansion of relative earnings, and appreciation that is more consistent with traditional queuing models. Hence, a pretty straightforward story in the context of macroeconomic adjustment can explain substantial movement in informality across this period.

However, dramatic as these movements have been—terminating in crises in several of the countries—in Argentina, Brazil, Colombia, and perhaps, in the early 2000s, Mexico, the data suggest longer-term trends that seem to go beyond medium-term macroeconomic adjustments and point to longer-term evolutions in underlying determinants of informality.

Teasing out the impact of these distinct effects is extremely difficult. To begin, global consensus on the impact of many of these reforms on labor markets remains elusive and the Latin American literature remains in its infancy. To date, the strongest effort to study the impact of labor legislation has been undertaken by Heckman and Pages (2004), although even here, where the direction of impact has been identified, often the magnitudes are still up for discussion. Relatively little of the work on trade liberalization has focused on the impact of informality, and perhaps even less on the impact of stabilization. Second, this report has stressed the potential importance of, particularly, informal independent employment as an exit option for workers evading social protection coverage that costs them far more than the benefits they receive from it. Changes in the relative attractiveness of formal versus informal work can lead to shifting the supply curve of labor to the formal sector and, hence, change the allocation of labor. Some of the reforms of the period had potentially far-ranging effects in this area, but the efforts to quantify them, presented here, are rudimentary at best. Finally, all the reforms happened at once, thus making isolation of one effect from the others especially difficult.

Demographic and structural factors

A natural place to begin is in demographic and structural changes in the economies over the last two decades. Female labor market participation has risen sharply and, as chapter 2 documents, women at marrying age tend to be disproportionately informal: Galiani and Weinschelbaum (2006) argue that the inflow of secondary workers to the Argentine workforce, particularly of female workers who choose to be

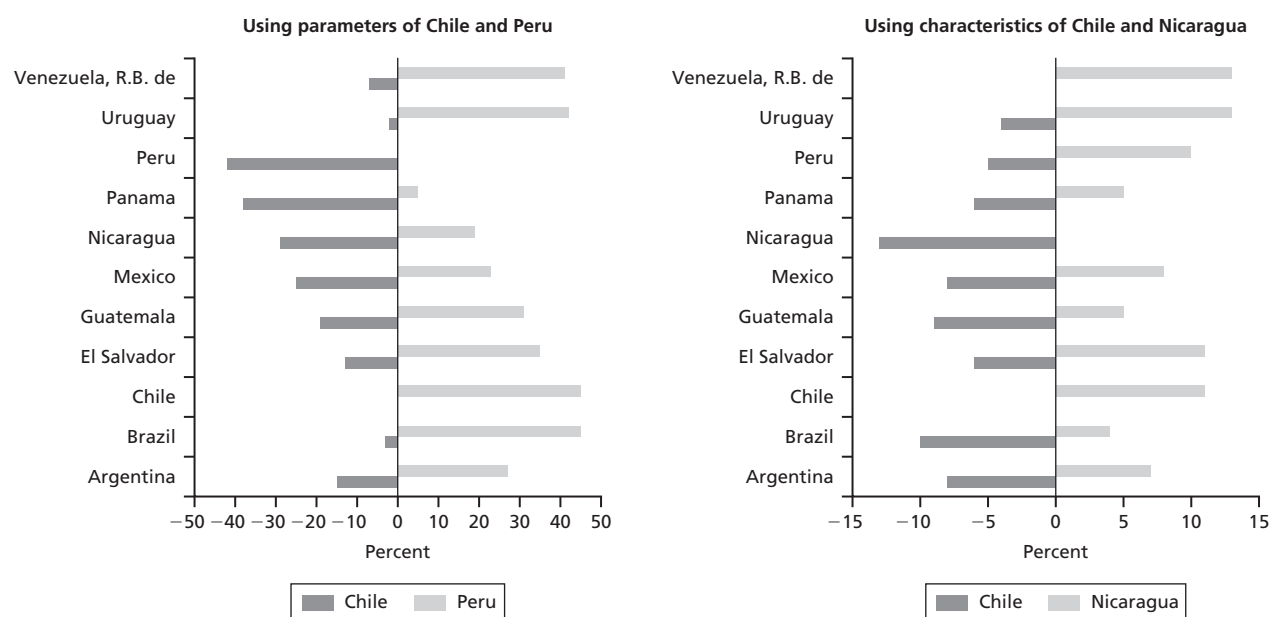
informal because their husbands are already covered by benefits, has led to a 13 percent increase in informality from 1974–76 to 1999. Jimeno and Rodriguez Palenzuela (2003) argue that trends in youth participation in the labor market are partly responsible for the rise in unemployment in the OECD, which, as chapter 2 suggested, also maps to more informal salaried work. However, youth participation has broadly stabilized as students spend more time in school and, in turn, education appears often as a determinant of informality (see chapter 3)—with some models, such as Krebs and Maloney (1999) and Loayza and Rigolini (2006), showing that more educated workers or workforces are less likely to be informal. In terms of economic structure, numerous studies have shown that different sectors have predispositions to informality—services more, manufacturing less (see Goldberg and Pavcnik 2003, and Bosch, Goñi, and Maloney 2006).

To explore how important such changes may be, Gasparini and Tornarolli (2006) undertook two sets of decompositions, identifying for numerous countries the relative impact of various demographic and structural factors on the level of informality. For each country, they estimate the probit coefficients that suggest how much a given

characteristic affects the probability of being informal. Figure 4.10 plots how much informality would change if, using the set of estimated country-specific coefficients of the impact of any particular demographic or sectoral characteristic, we assign to the country the set of characteristics that tends to lead to both the highest (Nicaragua) and lowest (Chile) levels of informality in the sample. As is clear using both the social protection and productive definitions, characteristics alone can account for a difference of more than 10 percentage points in level of informality. How a given set of characteristics translates into informality appears even more important. Using the parameters that create the highest level of informality (Peru) and those that yield the lowest (again, Chile) yields large differences: Peru's structure, using Chile's coefficients, would lead to a 25–50 point reduction in the size of the informal sector. Clearly, it is critical to understand how the economy translates demographic characteristics into informality.

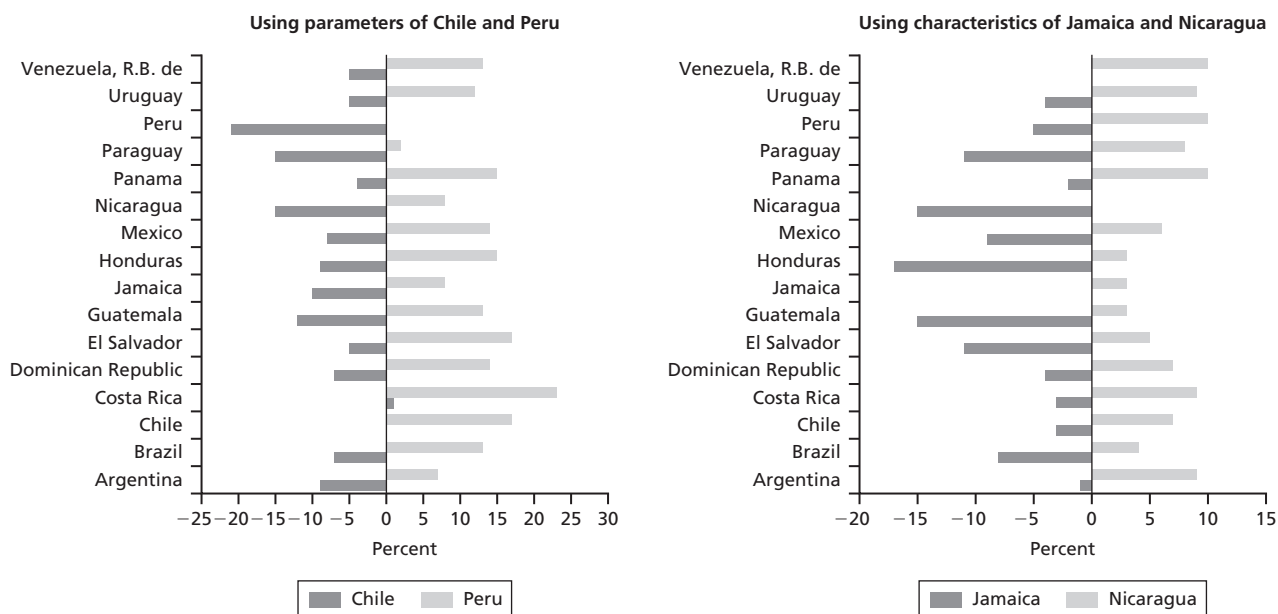
This translation also turns out to have been far more important to explaining the increase in informality than changes in structure or demography. Figure 4.11 shows that for Argentina (1995–2003), Brazil, Chile, El Salvador, Paraguay, and Uruguay, structural changes

FIGURE 4.10a

Variation of informality rate (social protection definition)

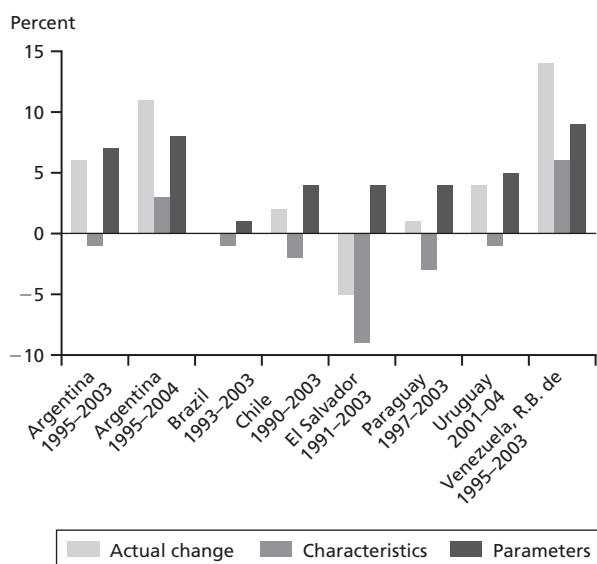
Source: Gasparini and Tornarolli 2006.

FIGURE 4.10b

Variation of informality rate (productive definition)

Source: Gasparini and Tornarolli 2006.

FIGURE 4.11

Decomposition of changes in informality (social protection definition) for urban salaried workers

Source: Gasparini and Tornarolli 2006.

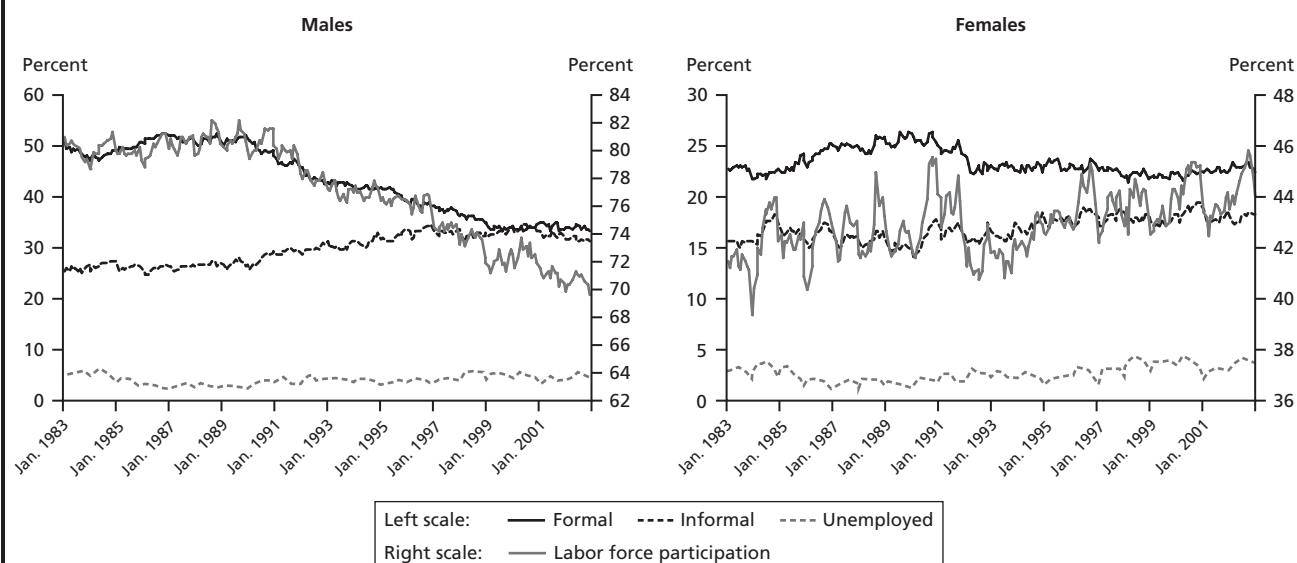
would lead to a *decrease* in informality if their impact had stayed the same. Only for a particular time period for Argentina and the República Bolivariana de Venezuela is there an informality-expanding effect. That said, an

independent decomposition by the Argentina Poverty Assessment (World Bank 2007a) concludes that changes in such structural factors were not, overall, responsible; rather, there was a generalized tendency toward noncompliance with labor legislation. Similarly, an independent exercise for Chile (Contreras, Puentes, and Sanhueza 2006) broadly concurs with Gasparini and Tornarolli (2006), arguing that changes in worker characteristics over the last 40 years would have led to a much larger decrease in informality than they actually observe.

We might argue that the changes in economic structure arising from reforms, away from formality-intensive sectors into those less so, might be critical. At this point, these do not seem a first-order explanation. Ramos (2002) found that, across the 1990s in Brazil, there was an expansion into services, but that this would explain a maximum of 25 percent of the increase in informality. Goldberg and Pavcnik (2003) for Brazil and Colombia; Bosch, Goñi, and Maloney (2006) for Brazil; and Bosch and Maloney (2006) for Mexico show that most secular changes in informality have arisen *within* sectors, not across them.

This point also applies to the Galiani and Weinschelbaum (2006) argument about the importance of the gender composition of the workforce. Looking closely at Brazil and Colombia (figure 4.12), we see a complex

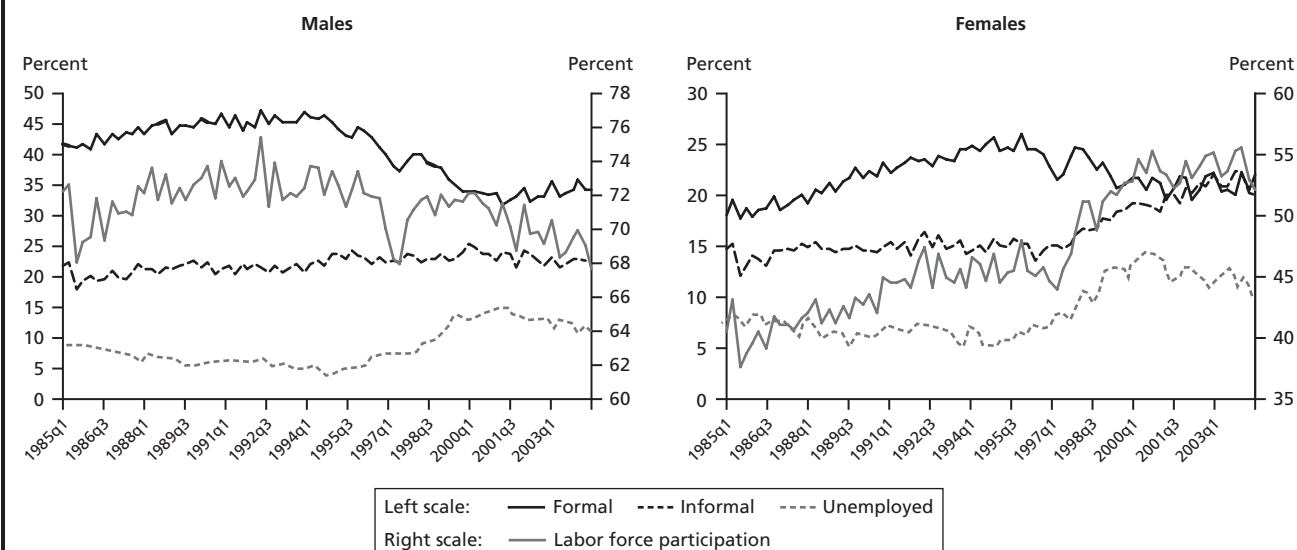
FIGURE 4.12a

Labor force participation, rate of formality, and unemployment, Brazil

Source: Author's estimations, based on *Encuesta de Calidad de Vida* and *Pesquisa Mensal de Emprego*.

Note: All aggregates are a share of total population and hence sum to unity.

FIGURE 4.12b

Labor force participation, rate of formality, and unemployment, Colombia

Source: Author's estimations, based on *Encuesta de Calidad de Vida* and *Pesquisa Mensal de Emprego*.

Note: All aggregates are a share of total population and hence sum to unity.

evolution between and across genders but with a heavy emphasis on informalization within genders.

In Colombia, the share of men in formal employment as a share of the population has fallen sharply since 1995

while the male share in informal work rose less strikingly, with the lost formal workers appearing in the sharply rising unemployment of the time. Female participation rose sharply beginning in 1996 and went entirely to informality.

This represents a break with past trends where the gently rising female participation from 1985 to 1995 left the share of females who were informal relatively constant. The apparent correlation of female participation with rising unemployment would seem consistent with an added worker effect, documented in Mexico and Argentina (Cunningham 2001a, 2001b): women—particularly married women—enter the workforce if their spouses become unemployed. Hence, adverse macroeconomic shocks are driving the increase in female participation and the rise in informality both overall and within genders.

Brazil saw much less pronounced increases in unemployment, female participation, and the share of women who were informal. Rather, what emerges there is a replacement of male formal positions with informal positions concomitant with an unexplained fall in male labor market participation. In both cases, the rising informality *within* gender contributes dominantly to the overall evolution of informality in the economy.

Finally, the region's overall gain in income would seem to mitigate against the rise. As noted in chapter 1, numerous studies now have shown that informality, measured as the share of the workforce, falls with development; and, overall, in spite of the crises, the region made some small progress (see Blau 1987; Gollin 2002; Loayza and Rigolini 2006; and Maloney 2001).⁵ However, as Loayza and Rigolini (2006) and others have shown, this broad trend with development explains perhaps 60–75 percent of the variance in the size of the self-employed sector. The residual 25–40 percent is broadly consistent with the simulations above that suggest that similar countries can have very different levels of informality. We now look at changes in trade regimes and labor legislation as possible explanations.

Trade reforms

The far-reaching trade reforms of the 1990s are a logical suspect, although, theoretically, the effects on informality are ambiguous. On the one hand, cheaper imports (or the appreciation of the currency that accompanied the trade reforms) may introduce pressure on domestic prices, driving local firms out of business, reducing their incentives to open new positions, or pushing them toward cheaper means of production in the informal sector. In the Fiess-Fugazza-Maloney (2006) model, this could be seen as a negative productivity shock to the formal/traded sector, the adjustment to which would depend on the degrees of rigidity in the formal sector, but in any case would lead to a

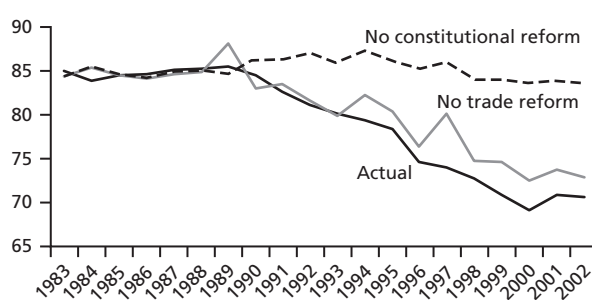
decline in formal sector employment. The increase in salaried informality could be made manifest through a couple of channels. The negative shift of the demand curve for formal labor would lead to lower employment and earnings in the formal sector. Part of the fall in earnings could occur through lower benefits, an effect that might be exacerbated if wages were relatively rigid. The same scenario would lead to hiring workers without benefits or subcontracting tasks to lower-paid external workers.

However, lower tariffs may also foster the import of technology and capital from abroad, thereby increasing the demand for complementary skilled labor that, in the long run, tends to greater formality. Generally speaking, industries that are more exposed to trade tend to pay higher wages and be more formal (see de Ferranti et al. 2001), given the human capital of their workers. In addition, the availability of higher-quality or lower-cost intermediate inputs in essence, constitutes a productivity shock increase to the formal sector, which, as shown earlier, lends to lower informality.

Empirical evidence of openness to trade on levels of informality is mixed, but generally suggests small effects. Goldberg and Pavcnik (2003) find a very modest impact of trade reforms in Colombia and none in Brazil. Bosch, Goñi, and Maloney (2006), revisiting the Brazilian case through the lens of job creation and destruction, find a positive, but again small, impact. Figure 4.13 plots the predicted values using the Shimer (2005c) methodology and suggests that, in the absence of trade liberalization, formal employment may have been 10 percent higher. The evidence from Mexico does not suggest a huge impact either. As noted by García-Verdú (2007), among others, given the dramatic unilateral liberalization beginning in

FIGURE 4.13

Actual and predicted size of the industrial formal sector, Brazil



Source: Bosch, Goñi, and Maloney 2006.

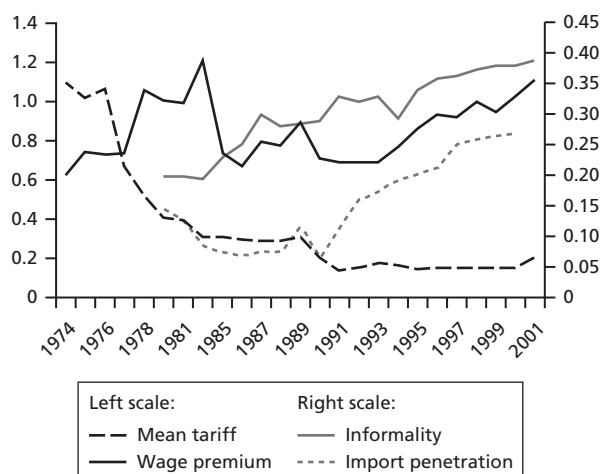
1987 and then continuing through the North American Free Trade Agreement, there is little trend in informality. Aleman-Castilla (2006), broadly following the Goldberg and Pavcnik (2003) methodology, finds that those industries more exposed to trade saw higher increases in the rate of formality. The reason, he argues, was that the impact on product prices was minor, while the reduction in import prices raised the productivity of the tradables sector and, hence, expanded the demand for formal labor overall.

However, suspicions remain. As Mondino and Montoya (2002) and World Bank (2007a) have shown, a very large increase in the share of informal salaried workers began in the early 1980s (figure 4.14). Though the last round of trade liberalization began only in 1990, the reforms begun under Argentina's Minister of the Economy José Alfredo Martínez de Hoz in the late 1970s radically lowered tariffs and led to an appreciated exchange rate. Galiani and Sanguinetti (2003) and Porto and Galiani (2006) find that the decreased protection had some effect on both the absolute level of wages and the gap between skilled and unskilled labor. To the degree that part of the downward pressure on unskilled wages came through the reduction of benefits, or subcontracting, it seems possible that trade liberalization had an impact. However, preliminary analysis replicating the Goldberg-Pavcnik (2003) exercise⁶ for

Argentina suggests that the impact of trade reform per se has a magnitude similar to that in Brazil, although there may be significant additional impacts from the various periods of sustained currency overvaluation.

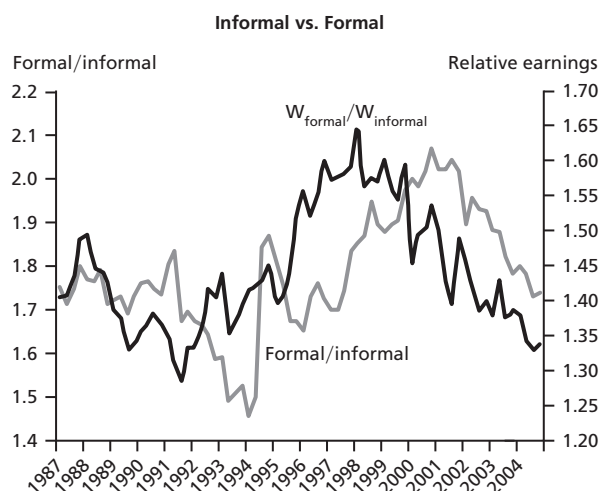
Recent trends in Mexico also seem plausibly related to international exposure. The sharp increase in both self-employment (shown above) and informal salaried work (see figure 4.15) after 2000 has occurred concomitantly with the entry of China as a major competitor in some areas of Mexico's comparative advantage. Hanson and Robertson (2006) argue that, had China's growth in export capacity remained unchanged after 1995, Mexico's annual export growth rate of Chinese-substitutable goods would have been 1.5 percentage points higher in the late 1990s and 3.0 percentage points higher than the 1.9 percent it experienced going into the new millennium. This does suggest that international competition is putting a constraint on the expansion of some export jobs. On the other hand, Lederman, Olarreaga, and Soloaga (2006), using estimations of the gravity model of trade, argue that there is little evidence that Mexican (and Central American) nonfuel exports overall were affected. It is also noteworthy that the sharp increase seems to occur with the relaxation of restrictions on Chinese textiles and apparel imports in the United States, which was not one of Hanson and Robertson's affected sectors. As an alternative explanation, the overall

FIGURE 4.14

Informality rate for salaried workers in Greater Buenos Aires

Sources: Porto and Galiani (2006) for mean tariff and wage premia; Gasparini (2002) for informal salaried work; and Olarreaga, Castro, and Saslavsky (2006) for import penetration.

FIGURE 4.15

Informal salaried versus formal salaried, Mexico, relative earnings and sector size

Source: Author's calculations.

reduction in exports due to the U.S. recession may have had a straightforward impact through, effectively, a reduction in tradables productivity that, in the absence of wage rigidities, puts us exactly in Regime III—a depreciation of the currency concomitant with a rise in relative sector size and relative earnings (figures 4.9 and 4.15). There may have been nothing more exotic going on than that, given the slowdown in the U.S. economy, the opportunities became better in the informal microfirms than in the maquilas across this period.

An additional table and graph comparing Argentina, Brazil, and Mexico are potentially informative here. Table 4.1 shows the change in the allocation of informal sector workers across sector sizes over the relevant period of

increase in informal work. In both Argentina and Brazil, the striking fact is the shift of informality toward larger firms, as the rate of informality in those large firms has almost doubled. This casts some doubt on the Beccaria, Carpio, and Orsatti (1999) conjecture that the rise in informality in Argentina was driven by increased subcontracting to smaller firms.

However, the reverse appears to happen in Mexico where medium-size and large firms are becoming more formal over time. Table 4.2 suggests that this is not due to increased subcontracting relationships with the firms of fewer than 16 employees covered in the microenterprise survey. Though the 2002 data are not exactly comparable with earlier data, there does not appear to have been an important increase in those selling their products or services to or buying inputs from large clients. Putting together the story, we might argue that the shifts in informality measured here are due to the increased relative attractiveness of opening or working for a microfirm over the last quinquennium, and not to greater subcontracting or within-large-firm informality due to trade opening.

Second, figure 4.16 suggests that, for Mexico, the small changes in the distribution of formal employment across ages over the last 20 years of trade liberalization have been minimal, with some loss of formality among prime-age males and perhaps older workers in the 1987–96 period that was absorbed both in informal salaried and independent work. There have been no substantial changes in the 1996–2004 period. In Brazil, however, the 1990–2002 period brought a shifting down of formal employment of roughly 10 percentage points across the whole age spectrum, with a fall of 20–30 percentage points for young workers. In Argentina, a similar pattern has prevailed, although the similarly dramatic losses of formal jobs

TABLE 4.1

Evolution of informality across firm size and time: Distribution of informal and self-employed workforce

a. Argentina (GBA)

Firm size	1980	2003	Variation (%)
1	45.35	34.60	–10.74
2 to 5	42.11	40.75	–1.36
6 to 25	7.82	15.55	7.72
26 to 100	2.71	5.25	2.54
101 to 500	1.25	2.60	1.35
501+	0.76	1.25	0.49

Source: *Encuesta Permanente de Hogares*.

b. Brazil

Firm size	1990	2002	Variation (%)
1	35.70	25.52	–10.18
2 to 5	42.52	35.99	–6.53
6 to 10	7.15	12.47	5.32
11+	14.63	26.02	11.39

Source: *Pesquisa Mensal por Amostra de Domicílios*.

c. Mexico

Firm size	1994	2004	Variation (%)
1	25.9	30.0	4.2
2 to 5	41.0	44.7	3.7
6 to 10	7.4	7.4	0.1
11 to 15	3.2	2.8	–0.4
16 to 50	6.1	6.3	0.2
51 to 100	3.0	1.7	–1.2
101 to 250	1.7	0.6	–1.1
251+	11.9	6.5	–5.4

Source: *Encuesta Nacional de Empleo Urbano*.

TABLE 4.2

Subcontracting in Mexico, 1992–2002 (percent)

Input/client	1992	1994	1996	1998	2002 ^a
Big inputs	37.78	39.8	41.6	46.99	41.79
Big clients	5.31	3.85	3.77	4.14	3.19
Maquila	2.28	1.42	1.31	2.98	0.07

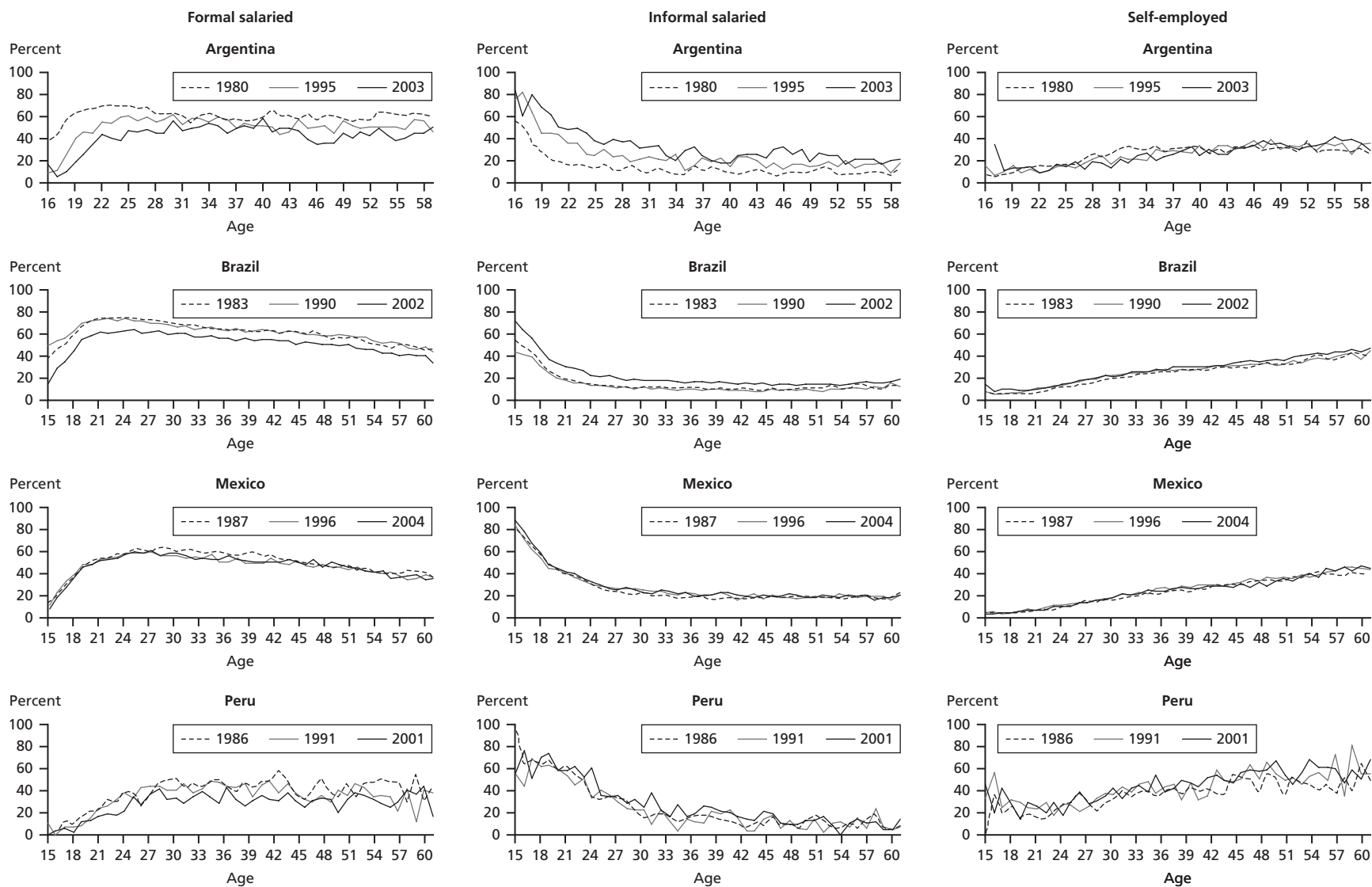
Source: *Encuesta Nacional de Micro Negocios*, several years.

Note: Firms are considered to be subcontracting if there is a big provider or a big client, or if the firm does maquila work.

a. Includes exports and imports as category in inputs/clients; no separate question for maquila, included as category of input/client.

FIGURE 4.16

Evolution of informality across age groups



Source: Author's calculations.

among workers in the early part of the life cycle extend to those in their late 20s, unlike in Brazil where the really big losses level out at about 20. Further, there is also a marked decrease in formalization among workers over 45 years of age that is roughly double that of the prime-age males.

In light of this, there can be no wonder about the concern in Argentina about the increased informalization of the workforce. The richest country in Latin America once had a constant and high formal employment rate of almost 70 percent, where 17-year-old workers had the same access to formal sector jobs as prime-age males. It now looks closer to Mexico, except the large firms in Mexico are relatively more formal. At the very least, this represents different experiences with trade liberalization. But the summary picture is striking. Mexico's far-reaching trade liberalizations, begun in 1987, had minimal impact on informality or its allocation across age groups or firm sizes. Meanwhile, Brazil has experienced an evolution in magnitude of informality and its allocation across both dimensions similar to that in Argentina. Yet, as figure 4.13 shows, little of this can be attributed to its far-reaching trade liberalization. At this point, we begin to look for other explanations.

Regulation and distortion

A substantial body of literature sees the size of the informal sector to be determined substantially by regulatory distortions or corruption (see Djankov et al. 2002; Friedman et al. 2000). Looking particularly at self-employment in cross-section, Loayza and Rigolini (2006) find that variations in business flexibility account for 16 percent of the variance; variations in government expenditure (a measure of monitoring intensity) account for 7 percent; and enforcement of contracts, prevalence of the rule of law, and the efficiency of the policy and judicial system account for 26 percent. GDP still accounts for 61 percent.⁷ Loayza, Oviedo, and Servén (2006), using the Schneider and Enste (2000) shadow economy estimates, find similar correlations with regulatory measures, although, as discussed at length in chapter 1, it is unclear what this measure is actually capturing. In general, such effects account for about 30 percent of the variation of informality across countries.

A long tradition sees distortionary regulation, particularly in the labor market, as leading to increased informality for a variety of reasons and across all three margins. Box 4.2 discusses three modeling exercises within the context of matching type models discussed in the first part of this chapter. Together, these simulations suggest that modest

increases in firing costs, labor taxes, and nominal wage rigidities can decrease the size of the formal sector by a few percentage points. Ulyssea (2006), modeling both barriers to entry and labor regulations, found the latter to be of small importance, although, as he makes clear, the assumption of perfect wage flexibility means that many distortions are passed on to workers and, hence, have little effect on the demand for labor. A large body of literature (for example, Botero et al. 2003; Heckman and Pages 2004; and Nickell and Layard 2000) concludes that more stringent labor market regulations hamper productivity growth. Holmes (1998) and Besley and Burgess (2004), exploiting regional variation for the United States and India, respectively, find important impacts of labor regulation on output, and the latter on employment. Following their general approach, Almeida and Carneiro (2005) find that Brazilian states with lower enforcement have higher productivity, wages, and employment.

Documenting the impact of legislation on formal sector demand in Latin America has been the subject of several major efforts in the region. Of particular importance, the National Bureau of Economic Research volume edited by Heckman and Pages (2004) collected some of the most serious attempts made to that date to quantify the impacts of various regulatory changes, and the literature is well reviewed in the introduction and chapters in that book. For the region and several countries, a credible case can be built that labor legislation had a substantial impact on the size of the formal sector (see Heckman and Pages 2004; and Saavedra and Torero 2004).

The Mexican case has attracted little academic work since there were significant changes in neither the size of the informal sector nor, perhaps relatedly, labor legislation. However, the fall in formal employment in the municipal areas of Brazil has received formidable attention by the analysts⁸ and the present discussion can only graze the tip of the iceberg. However, increased labor market rigidities are among the prime candidates to explain the fall in formal employment not accounted for by trade liberalization in figure 4.12. The constitutional reform of 1988 included several important changes to the labor legislation (see Paes de Barros and Corseuil 2004 for a summary). First, several measures increased labor costs and reduced employer flexibility. Maximum working hours per week were reduced from 48 to 44, the maximum daily work day in selected industries was reduced from 8 to 6 hours, overtime remuneration was increased from 1.2 to 1.5 times the

normal wage rate, vacation pay was raised from one to one and a third of the monthly wage, and maternity leave increased from 90 to 120 days. Second, union power was expanded: unions were no longer required to be registered and approved by the Ministry of Labor; decisions to strike now rested purely on the union's decision, and the required advance notification to the employer fell from five to two days; strikes in certain strategic sectors were no longer banned. Previous work by Menezes-Filho (1997) documenting the impact of the reduction in union power across the 1980s on firm profitability and earnings suggests that this channel can be important. Finally, firing costs were raised. The penalty levied on employers for "unjustified" dismissal, a category encompassing most legitimate separations for economic reasons in the United States, increased by four times, from 10 to 40 percent of the accumulated separation account (*Fundo de Garantia por Tempo de Serviço*).

To date, the most comprehensive work relating these changes to the functioning of the labor market was undertaken by Paes de Barros and Corseuil (2001) who find that separation rates decreased after the constitutional changes for short employment spells and increased for longer spells, but find an ambiguous overall effect. However, again, matching models discussed in the previous section suggest that several of these reforms would lead to a reduction in hiring (job-finding) rates as opposed to the separations that Paes de Barros and Corseuil study. By exploiting cross-industry variation in proxies related to these reforms, Bosch, Goñi, and Maloney (2006) find suggestive evidence that proxies for restrictions on hours worked, increases in union power, and increased firing costs appear to have a strong impact on decreasing formal job creation across the 1990s. Simulations suggest that, had there been no change in the constitution, formal hiring rates would have been 40 percentage points higher and overall levels of formality would have fallen negligibly.

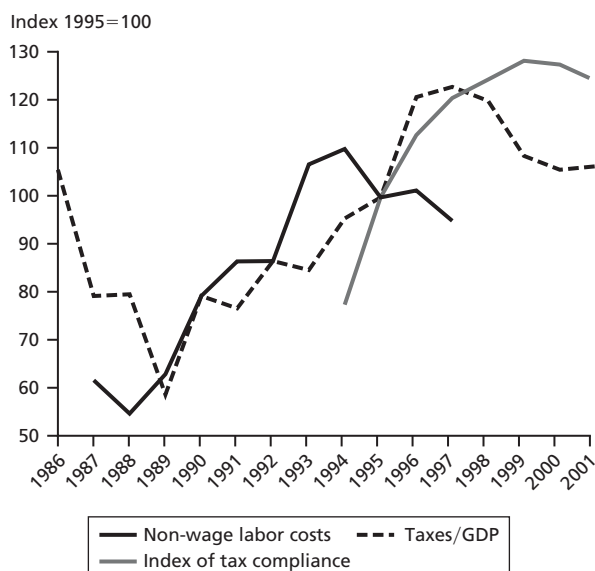
The impact of minimum wages interacted with other labor reforms plays an important role in explaining the evolution of informality in Colombia. Figure 4.9 shows that, around 1996, the labor market began a period of severe segmentation captured by the fact that relative earnings in the formal sector rose while the relative size of the sector fell sharply. As Arango and Pachón (2004), Kristensen and Cunningham (2006), Maloney and Nuñez Mendez (2003), and Santamaria (2000) show, Colombia's minimum wage is among the highest and most binding in Latin America, and it was raised in real terms in the depths of the crisis of the late

1990s. In the context of the Fiess-Fugazza-Maloney model, this corresponds well to a negative formal sector productivity shock in the presence of a nominal rigidity and, in fact, the empirical evidence supports that interpretation. In addition to this and consistent with it, other policy changes may have led to shifting the demand for formal sector workers that interacted with the minimum wage. Kugler (2000) argues that the 1990 reforms that reduced firing costs induced greater worker flows in and out of unemployment and may have contributed to 10 percent of the reduction in unemployment. However, they may also have deepened the fall in formal employment by reducing the costs of firing. Perhaps more directly, Kugler and Kugler (2003) argue that the sharp rise in payroll taxes in Colombia—over 10 percent from 41 to 51.5 percent—between 1989 and 1996 may also have had important employment effects. As they note, both nominal rigidities and the weak link between taxes and benefits lead to these higher costs being only 1.4–2.3 percent of those taxes shifted to workers, the rest leading to a decline in formal employment of 4–5 percent. Cardenas and Bernal (2003) similarly find high estimated wage elasticities that would suggest that such an increase in labor taxes would have led to a substantial drop in labor demand.⁹

Peru's large increase in informality of roughly 10 percentage points from 1986 to 2001 in metropolitan Lima also seems partly due to the increased regulatory burden. As Saavedra and Chong (1999) note, during the early 1990s Peru implemented dramatic trade liberalizations, fiscal reforms, deregulation of the goods markets, extensive privatization, and labor market reforms—the respective impacts of which are difficult to disentangle. The results were the substantial downsizing of many firms, and the productivity increase achieved by the reforms was concentrated in select sectors that created few additional formal jobs. On the one hand, the labor reform reduced employment rigidities, firing costs were reduced substantially,¹⁰ red tape on hiring temporary workers was minimized, and the use of training contracts for younger workers was facilitated. Further, enforcement of labor laws fell as budgets for inspections were slashed and labor courts became less pro-labor than previously. Finally, union bargaining power also fell dramatically, partly due to changes in labor legislation, the decentralization of the collective bargaining process, as well as the increased use of temporary contracts.¹¹

On the other hand, non-wage labor costs increased sharply beginning in the early 1990s (see figure 4.17), due

FIGURE 4.17

Labor legislation and related variables, Peru

Sources: Saavedra and Maruyama 2000, Peruvian Central Bank (BCRP), and National Tax Office (SUNAT).

Note: Non-wage costs are presented as a percentage of wages in December of each year. The index of income declaration denotes the number of contributors (workers and firms) presenting tax statements to national authorities. Note that, actually, a proportion of these contributors may not be necessarily paying any taxes (December of each year).

mainly to the elimination of caps on certain types of contributions. In addition, changes in the regulation of the unemployment compensation individual account (*Compensación por Tiempo de Servicios*) now forced accumulated contributions to be deposited in commercial bank accounts rather than simply counted as a contingent liability and the resources retained as working capital. Further, the 1991 tax reform induced a steady increase in the capabilities of tax authorities to reduce evasion. Saavedra and Torero (2004) show that these movements in labor costs, in fact, have the predicted impact on formal sector labor demand. The doubling of the non-wage costs led to a rise in overall labor costs of 10–11 percent which, at their estimated wage elasticity, could lead to a 3–4 percent decrease in formal salaried employment.

Regulatory explanations for the rise of the informal sector in Argentina are both many and relatively inconclusive, and are reviewed in World Bank 2007a. For example, Bour and Susmel (2000) argue that the low inflation under the Argentine convertibility plan made formal sector wages

sticky and, hence, effectively eliminated downward flexibility and made operating informally more attractive (see World Bank 2007a). The authors put together a year-by-year series of tax rates and labor regulations and compare them to the changes that have taken place in rates of informality (for salaried employees). They show that the increase in rates of informal salaried employment during 1980–88 took place during a period of rising taxes on labor and increased regulation. However, the increase in informality in the period 1995–99 was during a period of declining taxes. Arguably, the impact may have been offset by increased segmentation coming from sliding formal sector productivity and downwardly rigid wages.

Changes in labor law in 1991 and 1995, which provided for short-term labor contracts, have received particular attention as possible triggers for the growth in informality.¹² The reforms introduced fixed-term contracts and provided for special short-term contract regimes for small firms and young workers. Hopenhayn (2004) finds that the 1995 reform substantially increased job turnover rates. Olmedo and Murray (2002) credit the changes with promoting an acceptance of lower job protections and hence, greater informality. They write, “The new labor laws undermine the idea of permanent work with a modicum of economic welfare, social protection (especially in the fields of medical care, old age security and educational opportunities), and job security” (p. 430). Rates of informality among salaried workers rose when short-term contracts were legalized in 1995 and then persisted after such contracts were abolished in 1999. This pattern is consistent with the Olmedo and Murray (2002) claim that informality was made more socially acceptable by labor law reforms and supported by anecdotal evidence that suggests that, in contrast to Peru, enforcement of the labor law was relaxed. This argument, however, cannot explain the large increases in informality that took place during 1980–88, when labor regulation expanded. An alternative may, again, be that as Perry and Servén (2005) note, overvaluation was severe across this 1995–99 period due to the devaluations of neighboring currencies; and, much like the slowdown in the U.S. economy for Mexico, this may have constituted a negative shock to the tradables sector.

Increased incentives to exit

A critical policy conclusion of the analysis in the first section above—that there is a high degree of mobility among formal and informal sectors—is that we need to be concerned

with the incentives generated for working in the two sectors. A nascent literature is emerging precisely around the implications for social policy (see Levy 2006; Maloney 2004) and this will be dealt with in greater degree in the next chapter. The central point is that any wedge that is driven between the contributions a worker makes explicitly in terms of labor taxes or implicitly in terms of lower equilibrium wages, on the one hand, and the perceived benefits, on the other hand, makes formality less attractive. Such wedges may arise for several reasons.

In Brazil, the recent establishment of a minimum pension for all, regardless of contribution, and the lowering of the eligibility age compound the effect discussed by Levy by increasing the net gain from being informal. Further, the constitutional change toward creating universal health insurance may have also become an incentive to informality, although research to date is thin on the point (see Bosch, Goñi, and Maloney 2006) and will be dealt with in greater detail in chapter 5. Carneiro and Henley (2001) suggest that uncovered employment may have risen because employees and employers collude to avoid costly contributions to a social protection system that is perceived to be inappropriate, inefficient, and a poor value for the money.¹³ In principle, then, a universalization of health care delinked from the labor market may have changed the cost–benefit analysis of being enrolled in, and hence contributing to, formal sector benefits programs. In the end, they conclude that this is unlikely, not only because public health services continued to be thought of as substantially worse than the formal sector product,¹⁴ but also because the effective supply of these services was available even for non-contributors several years before the reforms took place, and little progress had been made on implementing the measures contemplated in the 1991 social security reform. However, the fall in participation of Brazilian males across the 1990s suggests some reasonably potent supply side effects.

It has sometimes been suggested that Argentina's high rates of informal employment are a consequence of low public confidence in public services and the social security system, which has resulted in part from repeated economic crises. Argentines do have a relatively low level of confidence in government institutions.¹⁵ In 2005, only 18 percent, 26 percent, and 26 percent of Argentines reported having confidence in political parties, Congress, and the judiciary, respectively. However, there is no clear evidence of a deterioration in confidence in public institutions

over time, other than during the crisis. Hence, it is an open question whether the systematically low level of confidence is related to the time trend toward higher levels of informal work.

Gasparini, Haimovich, and Olivieri (2006) attempt to measure whether the large cash transfers to unemployed household heads implied a disincentive for the program participants to search for a formal job. They find for 2003–04 that the effects were significant and large, reducing the share of workers moving into formal jobs by 5 percentage points. However, the effect vanishes later as the gap between program benefits and the rising wage shrank and would seem an unlikely candidate for explaining the effects since the 1980s.

More generally, it is difficult to know how much the pure tax—that is, income surrendered without corresponding benefits—affects the decision to be informal. Regrettably, research is in its infancy here, and we can offer only two approaches that give a fairly high level of variance. The first approach is to exploit the now well-established finding that self-employment, which we can use as a proxy for informality, decreases with level of development. If, in fact, as argued above, this reflects the rising opportunity cost of independent work discussed earlier, then we can see the imposition of the tax as, effectively, a reduction in this opportunity cost. The second approach uses the Krebs-Maloney (1999) model of efficiency wages that models the firm decision to invest and hire in response to the rate at which formal sector workers quit to enter self-employment, to calibrate the impact of the tax.

The estimated semi-elasticity of self-employment with respect to a change in relative formal sector earnings is 0.03 (Maloney 1998) to 0.05 (Loayza and Rigolini 2006) in the former case compared with the simulated value of 0.3 in the latter. Just as an idea of the broad orders of magnitude involved, if we assume that 10 percent of formal sector earnings is absorbed in unvalued benefits, and perhaps that another 10 percent of the value of earnings is transferred to informal workers in untied social protection programs, leading to a 20 percent decrease in the relative attractiveness of formal labor, then, in the first case, the size of the informal self-employed sector is increased by 0.6 to 1.2 percent and, in the second, by 6.0 percent of the share of the workforce. As an alternative approach, Fernandes, Gremaud, and Narita (2006) simulate the impact of replacing the labor tax on first minimum wage, thereby effectively providing formal sector benefits for free, while maintaining revenue

neutrality. If the value-added tax on capital goods is deductible, then there is a decline in informality of 1.5 percent. If not, however, the decline in physical accumulation decreases the demand for formal workers and leaves informality relatively unchanged. The first and third approaches suggest modest although nontrivial impacts, and the second suggests substantial impacts of these “supply side” concerns on the size of the informal sector.

Reforms along the *de Soto* margin

Finally, probably little responsibility can be placed at the door of policies that seek to make it easier for microfirms to register their workers. As the chapter on firm formality decisions shows, these had relatively small effects in the Sistema de Apertura Rápida de Empresas program in Mexico, the Sistema Integrado de Pagamento de Impostos e Contribuições as Microempresas e Empresas de Pequeno Porte program in Brazil, and the Monotributo in Argentina. Further, in all cases, the results are suggestive that, in fact, the small effects were in the direction of lowering informality.

Conclusion

The evidence from the analysis of gross worker flows suggests that, in Mexico and, to a lesser extent, Brazil, the formal and informal sectors behave across the business cycle as if they were dynamically integrated. In both countries, the flows support self-employment as being a desirable sector. The picture for informal salaried workers is mixed. In Mexico, pro-cyclical patterns of rematching across sectors suggest a sector in line with the previous chapter’s findings about the Dominican Republic: informal salaried jobs are not particularly inferior to jobs that comparable workers have in the formal sector. However, the Argentine and Brazilian evidence suggests a case with a higher proportion of salaried workers queuing for formal sector jobs, consistent with the tremendous decline in formality across the 1990s.

A unified explanation for the expansion of informality across the region over the last decades more generally remains elusive, although it appears to result from a combination of forces—some medium-term macroeconomic in nature, others related to longer-term structural changes. Some of the increase in the early part of the 1990s was due to nontradable/informal booms driven by liberalization of the exchange rate and subsequent inability to adjust formal sector wages downward as difficulties arose surrounding the exchange rate pegs. Trade liberalization probably had little impact overall. There is substantial evidence, at this

point, that changes in labor market regulation and social security taxes had an important impact in Brazil and Colombia. Supply side impacts could plausibly have been important, although the evidence is speculative. Evidence from changes in the difficulty of movement across the *de Soto* margin—of small firms becoming informal—suggests minimal effects working in the other direction.

Policy implications

The previous three chapters offer findings that suggest that labor market policies are potentially important determinants of informality and, hence, there is an agenda for reform. Since experience shows that specific policy recommendations have to be sensitive to country context, this section outlines some general principles without attempting to be comprehensive.

Labor policies work on informality through three channels. First, excessive labor costs, whether arising through labor legislation—exaggerated minimum wages, severance costs, labor taxes, or unrealistic union demands—depress the number of jobs in the formal sector. Other recent reports from the Bank, “Minimum Wages and Social Policies: Lessons from Developing Countries” (Cunningham 2007) and “Job Creation in Latin America and the Caribbean” (World Bank 2007c), as well as the Inter-American Development Bank publication *Good Jobs Wanted; Labor Markets in Latin America* (IDB 2006), investigate in detail the trade-offs between offering strong protections to some workers at the possible expense of excluding others. This chapter has offered evidence that, in several countries, exaggerated labor costs appear to have created the classic segmented labor market, especially during downturns. Further, the experience in the OECD suggests that such regulations have an especially heavy exclusionary impact on young people attempting to find jobs and who, in Latin America, are overrepresented among the involuntary informal salaried.

The evidence for a very high degree of integration of the formal and informal sectors found in several countries and subsectors does not necessarily imply satisfactory labor codes. Severance payments, which, in Latin America, Heckman and Pages (2004) find to be among the highest in the world, for example, may substantially reduce the job creation arising from growth without necessarily segmenting the market.

Second, legislation can create incentives to voluntary informality, which the last chapters have shown to compose a substantial fraction of the sector. This implies that the design

of social safety nets and labor legislation needs to take a more integrated view of the labor market, taking into account the cost–benefit analysis workers and firms make about whether to interact with formal institutions. This analysis incorporates the benefits of such interactions, the availability of informal substitutes, and the weighing of any greater risk exposure against the benefits of being independent and entrepreneurial. Hence, high labor taxes that do not correspond to benefits that workers value cause workers to opt out of the formal labor market. The difficulties of managing work and children also lead women to opt into informal independent work where there is more flexibility. As chapter 7 will show, the provision of substitutes through social security policies untied to the labor contract may offer substitutes to formal protections that further encourage opting out.

Third, labor market institutions can have an important impact on productivity growth. Theory and anecdotal evidence suggest that excessive restrictions on job reallocation or destruction for just cause, or other related state- or union-induced inflexibilities, may have a disincentive effect on the adoption of new technologies and production processes. Productivity gains arising from such innovations account for half of the differences in levels of economic development, the most important determinants of informality, not to mention long-run worker productivity and welfare more generally. Hence, the labor market and its institutions must be viewed as a key part of the national innovation system. These issues take on a heightened importance in the context of more open economies. While there appear to be few long-run adverse impacts of trade liberalization on informality, reaping the benefits of greater integration will require greater attention than interactions between trade and labor regulation.

Overall, the present constellation of often well-intended but heavy-handed labor regulations serves workers and firms poorly, and both could benefit from substantial reform. In particular, rigorous enforcement of a redesigned labor code that combined strengthened safety nets, well-designed worker protections, and worker representation with the flexibility firms need to adapt in a global economy, has the potential to expand formal employment and reduce opting out.

The issues of social protection design are discussed in more detail in chapter 7, but here it is worth mentioning only a few pending reforms of labor protections more specifically. Reducing severance payment to bring existing negative incentives for hiring closer to other regions, while

allowing workers to collect partial indemnities in some cases of voluntary resignation; introducing individual savings accounts to protect against the risk of unemployment, and/or, in countries with adequate administrative capacities, unemployment insurance schemes; establishing more flexible rules concerning dismissals for economic reasons and seasonal work; ensuring fair, transparent, and timely out-of-court mechanisms to resolve labor disputes, possibly including neutral arbitration or mediation councils—all of these actions would facilitate job creation and productivity growth. Minimum wages in most countries are not extremely binding, but, in some and particularly in Colombia, they seem to be a brutally segmenting force that begs for moderation. The problem of lack of flexibility to support child raising poses special challenges since gender-specific labor legislation (for example, restrictions of work hours and indiscriminate maternity benefits) may actually create incentives to discriminate against women, rationing them out of the formal market.

Simply tightening the enforcement of existing laws regarding, particularly, the largely informal microfirm sector may just eliminate jobs, many of which these chapters have shown to be of good quality measured by overall welfare of the worker. At the other extreme, attempting to reduce the weight of labor legislation by creating special classes of less protective contracts can be problematic. On the one hand, their strong incentive effects combined with the weak enforcement capacities of labor ministries can, in essence, create a parallel, unregulated market for jobs that has resulted in adverse effects, such as destruction/substitution away from regulated contracts, higher job turnover, and diminished incentives for training. Further, it may be the case that they contribute to the overall culture of informality (see chapter 8). On the other hand, the limited use of term contracts tied to special provisions (minimum required training investments) to make high-rotation hiring less attractive can help young people, particularly, enter the market. Further, provisions to accommodate different non-wage costs for smaller firms, and flexibility to allow for flexible benefits plans (such as simplified health care/pension plans) under mutual agreement between employers and employees may allow an extension of the overall rubric of labor protections without prejudicing the viability of these firms.

In addition, however, the substandard education and training systems in our region both impede the growth necessary to generate jobs in the more modern sector of the

economy and reduce workers' attachment to it. Informality sharply decreases with education, partly because the opportunity cost of being independent rises. Further, the poor signaling of education quality due to lack of uniform certification or accreditation impedes the entry of young workers into formal jobs. Remedying these failures, perhaps with an expansion of intermediation services, may reduce the information asymmetries that young workers face. Further discussion of necessary reforms in this area is offered in the World Bank regional report "Improving Educational Quality in Latin America and the Caribbean" (2007).

Ongoing upgrading of the workforce through *training*, particularly in rapidly evolving industries, is a central element of the national innovation system, and critical to developing skills used in the modern sector of the economy and to productivity growth. Training, however, is not a substitute for good schools. Firms overwhelmingly choose to train skilled (secondary education or above) workers, and impacts of training programs on older, unskilled workers are very low. Incentives for worker training of youth and older workers without prior experience or sufficient skills—incentives such as tax treatment of investment in human capital similar to that of capital investments, co-investments of workers, and post-training job matching of workers and firms—can ease their entry into the formal workforce. To capitalize on the efficacy of training, care should be taken to ensure homogeneous quality and effective delivery modalities. Optimal government intervention in training systems may require separation of finance and provision, with an unregulated private market and a public subsidy, and a possible new role for public training commissions/regulators to certify private providers of training; may necessitate the building of so-called lifelong learning/competency systems; and may demand subsidizing technical assistance to small and medium-size enterprises for developing human resource strategies.

There is also a role for promoting labor mobility and equal employment opportunities by making greater use of *labor market intermediation services*. Particularly, for women, youth, and minorities, employment centers can provide job search and employment placement assistance that could greatly enhance labor mobility, job matching, and labor market equity, while better linking regional labor markets. With the use of automated systems, these mechanisms have become easier to implement and administer; and while several countries in the region are increasingly making use of them, they remain limited in scope and are underfunded.

Finally, institutional strengthening (staffing, training, technical assistance) of the labor ministries and coordination of relevant public agencies (social security administration, enterprise development agencies, competitiveness councils) are needed so that they can assume their important and increasingly more complex role of facilitating labor productivity growth.

Notes

1. The self-employed-to-formal (SE-F) and formal-to-self-employed (F-SE) transition rates and informal-to-formal (I-F) and formal-to-informal (F-I) transition rates show correlations of 0.85 and 0.26, respectively. The same is emphatically not true in any case between any of the sectors and unemployment that behave as they do in the United States.

2. Hall (2005) calls such inefficient separations "Keynesian."

3. Finally, it can be argued that the pool of unemployed changes its composition during downturns. This has been called the heterogeneity hypothesis. Indeed, crises firing rates increase relative to quits. If fired workers have a greater propensity to work in the informal sector, this could generate the patterns found in the job-finding rates. This possibility is explored by Bosch and Maloney (2006) and the data strongly reject it. Workers with different educational levels, ages, or reasons to be unemployed show similar cyclical patterns of job-finding rates.

4. Another mechanism is also possible. Firms may post vacancies and decide whether to formalize the relationship, depending on the quality of the match. A bad shock reduces the overall posting of vacancies, but it also alters the optimal hiring level of specifically formal workers. Lower overall gain from jobs increases the relative cost of signing formal contracts and, therefore, firms shift to a "cheaper" way of producing. These two combined effects lower the job-finding rate in the formal sector substantially. However, the job-finding rate in the informal sector is determined by two opposing effects. On the one hand, the rate is reduced due to the decrease in the number of vacancies created; on the other hand, it is positively affected by change in the formality levels within the firm.

5. The mechanism may differ. Blau (1987), documenting this tendency across time in the United States, argued that this and similar trends in the OECD were due to the increased opportunity cost of being self-employed. Maloney argues that given that surveys in the United Kingdom, United States, and Germany (Blanchflower and Oswald 1998) suggest that roughly half of salaried workers would prefer to be self-employed and not have a boss, this lower labor productivity in developing countries may simply make this preference affordable. And, in fact, roughly 30 percent of entrants into informal self-employment in Mexico suggest that greater independence was their motivation. Alternatively, Loayza and Rigolini (2006) argue that income is proxy for lower levels of education, rudimentary infrastructure, and laggard technology, although the precise mechanism is not elaborated. Finally, Banerji and Jain (2006) argue that it may be preferences for higher-quality goods across the development process that reduce the market for informally produced goods.

6. Following Goldberg and Pavcnik (2003), the first stage applies a linear probability model to regress the formal dummy indicator on a vector of worker characteristics (gender, age, age², and education indicators) and on a set of industry indicators representing the workers' industry affiliations. The coefficients on industry indicators can be considered "industry formality differentials" stripped of worker characteristics that, in the second stage, are pooled over time and regressed on trade-related industry characteristics (tariffs and imports penetration), either in a fixed-effect or first differences specification (Goñi and Maloney 2007).

7. Because of the nature of calculating the contribution of each factor from the regressions, the sum may not equal 100.

8. A comprehensive survey of the literature studying the size and evolution of the Brazilian informal sector in the labor market can be found in Ulysea (2005), and a summary of stylized facts of this sector covering the 1980s and 1990s is detailed in Ramos (2002), Ramos and Brito (2003), Ramos and Ferreira (2005a, 2005b), Ramos and Reis (1995), and Veras Soares (2004).

9. On the other hand, Cardenas and Gutierrez (1997) find much smaller implicit elasticities of demand, and argue that the larger effects were due to the fall in output and the overvaluation of the peso that drove the rise in unemployment.

10. Firing costs were reduced in two dimensions. First, the tenure-related severance payments schedule was made less steep. Second, and more important, was the elimination of the obligation to reinstate a worker in his or her job if a labor court determined that the dismissal had no just cause.

11. By law, temporary workers may belong to a union, but firms implicitly threatened workers who registered in a union by not renewing the contract.

12. This and the next paragraph are nearly complete plagiarism from the Argentina Poverty Assessment (2007).

13. Carneiro and Henley (2001) estimate that the earnings premium needed in the marketplace to compensate covered workers for having to make social security contributions varies between 7.5 and 12.2 percent of the mean uncovered hourly wage.

14. The public system acts as a floor, available to all but used primarily by the lower classes (Jack 2000). Although evaluation of standards for minimum quality in infrastructure; human resources; and ethical, technical, and scientific procedures in hospitals have been implemented, these practices are far from being universal in the services network (PAHO 2005).

15. This paragraph is taken from the Argentina Poverty Assessment (World Bank 2007a).

References

- Acemoglu, D. 2003. "Patterns of Skill Premia." *Review of Economic Studies* 70: 199–230.
- Albrecht, J., and S. Vroman. 2002. "A Matching Model with Endogenous Skill Requirements." *International Economic Review* 43: 283–305.
- Albrecht, J., L. Navarro, and S. Vroman. 2006. "The Effects of Labor Market Policies in an Economy with an Informal Sector." IZA Discussion Paper 2141, Institute for the Study of Labor, Bonn, Germany.
- Aleman-Castilla, B. 2006. "The Effect of Trade Liberalization on Informality, Wages, and Productivity: Evidence from Mexico." Photocopy. Department of Economics, London School of Economics and Political Science, UK.
- Almeida, R., and P. Carneiro. 2005. "Enforcement of Regulation, Informal Labor and Firm Performance." IZA Discussion Paper 1759, Institute for the Study of Labor, Bonn, Germany.
- Arango, C., and A. Pachón. 2004. "Minimum Wage in Colombia: Holding the Middle with a Bite on the Poor." *Borradores de Economía*, Banco de la República, 280.
- Banerji, A., and S. Jain. 2006. "Quality Dualism." Photocopy. Department of Economics, University of Virginia, Charlottesville.
- Beccaria, L., J. Carpio, and A. Orsatti. 1999. "Argentina: informalidad laboral en el nuevo modelo económico." In *Informalidad y Exclusión Social*, ed. J. Carpio, E. Kleinm, and I. Novacovsky, 143–50. Buenos Aires, Argentina: Fondo de Cultura Económica.
- Besley, T., and R. Burgess. 2004. "Can Labor Regulation Hinder Economic Performance? Evidence from India." *Quarterly Journal of Economics* 119 (1): 91–134.
- Bewley, T. 1999. *Why Wages Don't Fall during a Recession*. Cambridge, MA: Harvard University Press.
- Blanchflower, D. G., and A. J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16 (1): 26–60.
- Blau, D. 1987. "A Time-Series Analysis of Self-Employment in the United States." *Journal of Political Economy* 95 (3): 445–67.
- Bosch, M. 2006. "Job Creation and Job Destruction in the Presence of Informal Labor Markets." Photocopy. London School of Economics, UK.
- Bosch, M., E. Goñi, and W. F. Maloney. 2006. "The Determinants of Rising Informality in Brazil: Evidence from Gross Worker Flows." Photocopy. World Bank, Washington, DC.
- Bosch, M., and W. F. Maloney. 2006. "Gross Worker Flows in the Presence of Informal Labor Markets: The Mexican Experience 1987–2002." Photocopy. World Bank, Washington, DC.
- Botero, J., S. Djankov, R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2003. "The Regulation of Labor." Working Paper 9756, National Bureau of Economic Research, Cambridge, MA.
- Bour, J., and N. Susmel. 2000. "Los Determinantes de la Informalidad Laboral." In *La Economía Oculta en la Argentina*, 225–66. Buenos Aires, Argentina: FIEL.
- Cardenas, M., and R. Bernal. 2003. "Determinants of Labor Demand in Colombia: 1976–1996." Working Paper 10077, National Bureau of Economic Research, Cambridge, MA.
- Cardenas, M., and C. Gutierrez. 1997. "Demanda de Trabajo, Impuestos a la Nomina y Desempleo en Colombia." Photocopy. Fedesarrollo, Bogotá, Colombia.
- Carneiro, F. G., and A. Henley. 2001. "Modelling Formal and Informal Employment and Earnings: Microeconomic Evidence for Brazil." Proceedings of the 29th Meeting of the Brazilian Economic Society, ANPEC (Associação Nacional dos Centros de Pós-Graduação em Economia).
- Contreras, D., E. Puentes, and C. Sanhueza. 2006. "Informality in Chilean Labor Markets." Photocopy. World Bank, Washington, DC.

- Cunningham, W. 2001a. "Breadwinner versus Caregiver: Labor Force Participation and Sectoral Choice over the Mexican Business Cycle." In *The Economics of Gender in Mexico: Family, Work, State, and Market*, ed. E. G. Katz and M. C. Correia, 85–132. Washington, DC: World Bank.
- . 2001b. "Household Risk, Self-Insurance, and Coping Strategies in Urban Argentina." Report 22426-AR, World Bank, Washington, DC.
- . 2007. "Minimum Wages and Social Policies: Lessons from Developing Countries." Photocopy. World Bank, Washington, DC.
- Davis, S., R. J. Faberman, and J. Haltiwanger. 2005. "The Flow Approach to Labor Markets: New Data Sources, Micro-Macro Links and the Recent Downturn." Photocopy. University of Chicago.
- Davis, S., and J. C. Haltiwanger. 1992. "Gross Job Creation, Gross Job Destruction and Employment Reallocation." *Quarterly Journal of Economics* 107 (3): 819–63.
- . 1999. "On the Driving Forces behind Cyclical Movements in Employment and Job Reallocation." *American Economic Review* 89 (5): 1234–58.
- de Ferranti, D., G. Perry, D. Lederman, and W. Maloney. 2001. *From Natural Resources to the Knowledge Economy*. Washington, DC: World Bank.
- de Soto, H. 1989. *The Other Path: The Invisible Revolution in the Third World*. New York: Harper and Row.
- Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1): 1–37.
- Dolado, J. J., M. Jansen, and J. F. Jimeno. 2002. "On-the-Job Search in a Matching Model with Heterogenous Jobs and Workers." IZA Discussion Paper 886, Institute for the Study of Labor, Bonn, Germany.
- Fernandes, R., A. P. Gremaud, and R. D. T. Narita. 2006. "Tax Reform and Formalization of the Economy: Simulating Different Alternatives for Brazil." Photocopy. Department of Economics, University of São Paulo, Brazil.
- Fiess, N., M. Fugazza, and W. Maloney. 2002. "Exchange Rate Appreciations, Labor Market Rigidities, and Informality." Policy Research Working Paper 2771, World Bank, Washington, DC.
- . 2006. "Informal Labor Markets and Macroeconomic Fluctuations." Photocopy. World Bank, Washington, DC.
- Friedman, E., S. Johnson, D. Kaufmann, and P. Zoido-Lobaton. 2000. "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics* 76: 459–93.
- Galiani, S., and P. Sanguinetti. 2003. "The Impact of Trade Liberalization on Wage Inequality: Evidence from Argentina." *Journal of Development Economics* 72: 497–513.
- Galiani, S., and F. Weinschelbaum. 2006. "Modeling Informality Formally: Households and Firms." Photocopy. World Bank, Washington, DC.
- García-Verdú, R. 2007. "Measurement of the Shadow Economy or Shadowy Measurement?" Photocopy. World Bank, Washington, DC.
- Gasparini, L. 2002. "Microeconomic Decompositions of Aggregate Variables: An Application to Labour Informality in Argentina." *Applied Economics* 34 (18): 2257–66.
- Gasparini, L., F. Haimovich, and S. Olivieri. 2006. "Labor Informality Effects of the *Programa Jefes de Hogar*." Photocopy. Center for Distribution, Labor and Social Studies, National University of la Plata, Buenos Aires, Argentina.
- Gasparini, L., and L. Tornarolli. 2006. "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata." Photocopy. World Bank, Washington, DC.
- Goldberg, P., and N. Pavcnik. 2003. "The Response of the Informal Sector to Trade Liberalization." *Journal of Development Economics* 72: 463–96.
- Gollin, D. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110 (2): 458–74.
- Goñi, E., and W. Maloney. 2007. "On the Effects of Trade on Informality. A Note for the Argentinean Case." Photocopy. World Bank, Washington, DC.
- Hall, R. 2005. "Employment Efficiency and Sticky Wages: Evidence Flows in the Labor Market." Working Paper 11183, National Bureau of Economic Research, Cambridge, MA.
- Hanson, G. H., and R. Robertson. 2006. "China and the Recent Evolution of Mexico's Manufacturing Exports." Photocopy. University of California, San Diego.
- Harris, J. R., and M. P. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review* 60 (1): 126–42.
- Heckman, J., and C. Pages. 2004. *Law and Employment: Lessons from Latin America and the Caribbean*. Chicago: University of Chicago Press.
- Holmes, T. 1998. "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders." *Journal of Political Economy* 106: 667–705.
- Hopenhayn, H. 1988. "A Competitive Stochastic Model of Entry and Exit to an Industry." Photocopy. University of California, Los Angeles.
- . 2004. "Labor Market Policies and Employment Duration: The Effects of Labor Market Reform in Argentina." In *Law and Employment: Lessons from Latin America and the Caribbean*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- IDB (Inter-American Development Bank). 2006. *Good Jobs Wanted: Labor Markets in Latin America*. Washington, DC.
- Jack, W. 2000. "Health Insurance Reform in Four Latin American Countries. Theory and Practice." Policy Research Working Paper 2492, World Bank, Washington, DC.
- Jimeno, J., and D. Rodríguez Palenzuela. 2003. "Youth Unemployment in the OECD: Demographic Shifts, Labour Market Institutions and Macroeconomic Shocks." Working Paper 19, European Network of Economic Policy Research Institutes, Brussels.
- Jovanovic, B. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5): 972–90.
- . 1982. "Selection and Evolution of Industry." *Econometrica* 50 (3): 649–70.
- Kennan, J. 2005. "Private Information, Wage Bargaining and Employment Fluctuations." University of Wisconsin, Madison.
- Krebs, T., and W. Maloney. 1999. "Quitting and Labor Turnover: Microeconomic Evidence and Macroeconomic Consequences."

- Policy Research Working Paper 2068, World Bank, Washington, DC.
- Kristensen, N., and W. Cunningham. 2006. "Do Minimum Wages in Latin America and the Caribbean Matter? Evidence from 19 Countries." Policy Research Working Paper 3870, World Bank, Washington, DC.
- Kugler, A. 2000. "The Incidence of Job Security Regulations on Labor Market Flexibility and Compliance in Colombia: Evidence from the 1990 Reform." Working Paper R-393, Inter-American Development Bank, Washington, DC.
- Kugler, A., and M. Kugler. 2003. "The Labor Market Effects of Payroll Taxes in a Middle-Income Country: Evidence from Colombia" IZA Discussion Paper 852, Institute for the Study of Labor, Bonn, Germany.
- Lederman, D., M. Olarreaga, and I. Soloaga. 2006. "The Growth of China and India in World Markets: Opportunity or Threat for Latin American Exporters?" Background paper for the regional study, *Latin America and the Caribbean Respond to the Growth of China and India*. World Bank, Washington, DC.
- Levy, S. 2006. "Productividad, Crecimiento y Pobreza en México: ¿Qué sigue después de Progres-Oportunidades?" Photocopy. World Bank, Washington, DC.
- Loayza, N. V., A. M. Oviedo, and L. Servén. 2006. "The Impact of Regulation on Growth and Informality: Cross-country Evidence." World Institute for Development Economics Research, United Nations University, Helsinki.
- Loayza, N., and J. Rigolini. 2006. "Informality Trends and Cycles." Photocopy. World Bank, Washington, DC.
- Maloney, W. 1998. "Are Labor Markets in Developing Countries Dualistic?" Policy Research Working Paper 1941, World Bank, Washington, DC.
- . 2001. "Self-employment and Labor Turnover in Developing Countries: Cross-country Evidence." In *World Bank Economists' Forum*, ed. S. Devarajan, F. H. Rogers, and L. Squire. Washington, DC: World Bank.
- . 2004. "Informality Revisited." *World Development* 32 (7): 1159–78.
- Maloney, W. F., and J. Nuñez Mendez. 2003. "Measuring the Impact of Minimum Wages: Evidence from Latin America." Working Paper 9800, National Bureau of Economic Research, Cambridge, MA.
- Menezes-Filho, N. A. 1997. "Unions and Profitability over the 1980s: Some Evidence on Union-Firm Bargaining in the United Kingdom." *Economic Journal* 107 (442): 657–70.
- Menzio, G. 2005. "High-Frequency Wage Rigidity." Photocopy. Northwestern University, Chicago.
- Mondino, G., and S. Montoya. 2002. "The Effects of Labor Market Regulations on Employment Decisions by Firms: Empirical Evidence for Argentina." In *Law and Employment: Lessons from Latin America and the Caribbean*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- Mortensen, D., and E. Nagypal. 2005. "More on Unemployment and Vacancy Fluctuations." Photocopy. Northwestern University, Chicago.
- Mortensen, D., and C. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61: 397–415.
- Nickell, S., and R. Layard. 2000. "Labor Market Institutions and Economic Performance." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card. Amsterdam: North-Holland.
- Obstfeld, M., and K. Rogoff. 1996. "Foundations of International Macroeconomics." Photocopy. Massachusetts Institute of Technology, Cambridge.
- Olarreaga, M., L. Castro, and D. Saslavsky. 2006. "The Impact of Trade with China and India on Argentina's Manufacturing Employment." Photocopy. World Bank, Washington, DC.
- Olmedo, C., and M. Murray. 2002. "The Formalization of Informal/Precarious Labor in Contemporary Argentina." *International Sociology* 17 (3): 421–43.
- Paes de Barros, R., and C. H. Corseuil. 2001. "The Impact of Regulations on Brazilian Labor Market Performance." Research Paper R-427, Inter-American Development Bank, Washington, DC.
- . 2004. "The Impact of Regulations on Brazilian Labor Market Performance." In *Law and Employment: Lessons from Latin America and the Caribbean*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- PAHO (Pan American Health Organization). 2005. "Hospital Accreditation Manual." Washington, DC.
- Perry, G., and L. Servén. 2005. "Argentina's Macroeconomic Collapse: Causes and Lessons." In *Managing Economic Volatility and Crises: A Practitioner's Guide*, ed. J. Aizenmann and B. Pinto. New York: Cambridge University Press.
- Pissarides, C. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *American Economic Review* 75: 676–90.
- . 2000. *Equilibrium Unemployment Theory*. 2nd ed. Cambridge, MA: Massachusetts Institute of Technology Press.
- Porto, G., and S. Galiani. 2006. "Trends in Tariff Reforms and Trends in Wage Inequality." Policy Research Working Paper 3905, World Bank, Washington, DC.
- Pries, M., and R. Rogerson. 2005. "Hiring Policies, Labor Market Institutions, and Labor Market Flows." *Journal of Political Economy* 113 (41): 811–39.
- Ramos, L. 2002. "A evolução da informalidade no Brasil metropolitano: 1991-2001." Discussion Paper 914, Institute of Applied Economic Research, Brasília.
- Ramos, L., and M. Brito. 2003. "O Funcionamento do mercado de trabalho metropolitano brasileiro no período 1991-2002: tendências, fatos estilizados e mudanças estruturais." *Boletim Mercado de Trabalho, Conjuntura e Análise* 22: 31–47.
- Ramos, L., and V. Ferreira. 2005a. "Spatial and Sector Patterns of the Evolution of Informality in Brazil: 1991–2003." Discussion Paper 1099, Institute of Applied Economic Research, Brasília.
- . 2005b. "Spatial Pattern of Formal Employment Evolution—1995–2003" Discussion Paper 1102, Institute of Applied Economic Research, Brasília.

- Ramos, L., and J. Reis. 1995. "Salário mínimo, distribuição de renda e pobreza no Brasil." *Pesquisa e Planejamento Econômico* 25 (1): 99–114.
- Rauch, J. E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35: 33–47.
- Saavedra, J., and A. Chong. 1999. "Structural Reform, Institutions and Earnings: Evidence from the Formal and Informal Sectors in Urban Peru." *Journal of Development Studies* 35 (4): 95–116.
- Saavedra, J., and M. Torero. 2004. "Labor Market Reforms and Their Impacts over Formal Labor Demand and Job Market Turnover: The Case of Peru." In *Law and Employment: Lessons from Latin America and the Caribbean*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- Santamaria, M. 2000. "External Trade, Skill, Technology and the Recent Increase of Income Inequality in Colombia." Dissertation. Available at <http://econpapers.repec.org/>.
- Schneider, F., and D. H. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38 (1): 77–114.
- Shimer, R. 2005a. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review* 95 (1): 25–49.
- . 2005b. "The Cyclical Behavior of Hires, Separations, and Job-to-Job Transitions." *Federal Reserve of St. Louis Review* 87 (4): 493–507.
- . 2005c. "Reassessing the Ins and Outs of Unemployment." Photocopy. University of Chicago.
- Ulyssea, G. 2006. "Brazilian Labor Market Informality: A Review of the Literature." *Revista de Economia Política* Oct./Dec. 2006, 26 (4): 596–618.
- . 2006. "Informalidade no mercado de trabalho brasileiro: uma resenha da literatura." *Revista de Economia Política* 26 (3).
- Veras Soares, F. 2004. "Some Stylized Facts of the Informal Sector in Brazil in the Last Two Decades." Anais do XXXII Encontro Nacional de Economia, 142, ANPEC.
- World Bank. 2007a. *Argentina Poverty Assessment. Informal Employment in Argentina: Towards Understanding Its Causes and Consequences*. Washington, DC.
- . 2007b. "Improving Educational Quality in Latin America and the Caribbean." Washington, DC.
- . 2007c. "Job Creation in Latin America and the Caribbean." Photocopy. Washington, DC.

CHAPTER 5

Microfirm Dynamics and Informality

SUMMARY: This chapter describes the patterns of entry, survival, and growth in the self-employment sector of selected Latin American countries. The chapter shows that those patterns match well the predictions of mainstream models of firm dynamics. Entrants into self-employment tend to be found more frequently among workers who have accumulated human and physical capital while working as salaried workers. Higher-productivity firms are more likely to stay in business and create new jobs while becoming increasingly formal in the process. Few firms, however, tend to evolve along this ideal path, with most of them remaining informal and owner-only. While to some extent this can be explained by the low opportunity costs for entry into the sector, policy-induced barriers to formalization may also play a role by impeding microfirm access to technologies and markets.

AS ARGUED EARLIER IN THIS REPORT, THE owners and employees of very small, mostly informal firms constitute a major share of Latin America's labor force. Those workers thus play a very important role in the functioning of the region's labor markets, an issue that has been discussed at length in previous chapters. However, a full understanding of the dynamics of the informal sector also requires a complementary perspective that focuses on the firm rather than on the worker as a unit of analysis. Indeed, to the extent that entry and exit flows in the microenterprise sector are at least partially driven by voluntary choices regarding forms of participation in the labor market, they must also involve calculations related to the expected profitability of running one's own business. Thus, for example, to the extent that exit, not exclusion, is behind the decision of the many formal salaried workers who open new microfirms every year in Latin America, understanding their behavior may require treating them as profit-maximizing entrepreneurs rather than disguised unemployed workers queuing for formal jobs. In that respect, the size and evolution of the sector depend, at least to some extent, on the factors that affect the dynamics of

microenterprises, including the determinants of entry, exit, and growth of small firms.

After briefly reviewing some of the industrial country literature on firm behavior, this chapter presents evidence on the patterns of entry, exit, and growth of microenterprises in Latin America. We then explore the personal and enterprise characteristics associated with microfirm formality. Insofar as our results regarding microfirm dynamics are considerably similar to those obtained for industrial countries, we argue that they support the idea that the informal microenterprise sector is, to some extent, made up of voluntary entrants who choose to be there, taking into account their business prospects in the sector as well as the attributes of the jobs available in other segments of the labor market. We interpret this as evidence that the mainstream view may be well suited to understanding the developing country's microenterprise sector. This approach is consistent with the early work on Kenya by Hart (1972) and the mounting evidence of entrepreneurial dynamism and relative job satisfaction of self-employed individuals—see Bhattacharaya (2002), de Soto (1989), and Maloney (1999, 2004).

Furthermore, we find evidence that formality increases rapidly with firm size and productivity, and it is higher

among those who voluntarily enter self-employment. Thus, some of the same attributes associated with entry and success in self-employment are also correlated with compliance with government regulations. This suggests that formality operates as a normal input into the production process, with a minimal degree of participation in some institutions being a necessary input to growth, and participation increasing with the success of the business. In most cases, however, the large majority of microfirms remain too small to make the benefits of formality overcome its various costs. This could be due to the presence of policy-induced barriers to formalization—for example, high registration costs, stringent tax and labor market regulations, and other factors reviewed in the next chapter—that would in turn limit the growth of informal microfirms and contribute to a vicious continuing circle of low growth and high informality.

This chapter, however, makes the case for a complementary interpretation of the evidence. In particular, given the similarity between microfirm dynamics in developing and industrial countries, we argue that the high rates of microfirm failure and the low rates of job creation and growth, which keep Latin American firms below the threshold over which formality becomes a necessity, could also be the result of the low levels of human capital and managerial ability found among many of the region's microentrepreneurs. This, in turn, could be due to relatively low opportunity costs of entering self-employment, driven in turn by low productivity and salaries in the formal sector. The practical implication of this interpretation is that policy makers interested in altering the incentives that drive the large majority of Latin American microfirms into informality should focus not only on the direct costs and benefits of formality, but also on the factors that, by keeping formal sector productivity low, indirectly increase the number and reduce the average "quality"—in terms of their managerial ability and inherent business prospects—of new entrants into self-employment. In other words, the agenda for reducing microfirm informality should also cover the factors that drive productivity in the formal sector, encompassing improvements in the investment climate and policies aimed at increasing the quantity and quality of human capital.

Conceptual framework: Firm dynamics and institutional development

In the industrial countries, the last two decades have seen the emergence of a set of stylized facts about the personal and firm characteristics associated with entry, survival, and

growth in the microenterprise sector, as well as theoretical frameworks to explain them. However, notwithstanding the increasing importance given to the promotion of micro and small enterprises in development policy circles, there have been few systematic attempts to see how their dynamics approximate those of the mainstream literature.¹ This represents a loss on two fronts. First, if it seems that behavioral differences are not so great, then development policy makers have a wealth of analytical frameworks at their disposal. Second, a finding of kinship with their advanced country counterparts would provide additional evidence on the debate on how we should conceive of the role of the informal microfirm in the developing world.

In the Organisation for Economic Co-operation and Development (OECD), opening a business and being one's boss is often celebrated as a desirable alternative to salaried work; however, in the developing world, the very large unregulated (informal) microenterprise sectors are frequently seen as the disadvantaged segment of a dual labor market in which workers queue for good jobs. The two views have different implications for entry and firm dynamics. In particular, rather than the patterns of entry/exit and growth associated with entrepreneurship dynamics, the traditional least developed countries (LDC) literature on informality would predict patterns consistent with the dynamics of disguised unemployment.

The mainstream view of firm dynamics

The dominant view of the role of self-employment in industrial countries stresses the risk-taking, entrepreneurial nature of the sector, with the celebrated Silicon Valley high tech start-up at its apex. In the classic framework proposed by Lucas (1978), individuals are endowed with a given—and known—level of entrepreneurial or managerial ability, which determines the returns from self-employment. Lucas argued that there is a distribution of entrepreneurial ability in the population: those with a sufficiently high level of proficiency become entrepreneurs, while the rest become wage workers. Among those individuals whose entrepreneurial ability leads them to choose self-employment, the more proficient have firms that are larger and/or more successful. This means that the existence of many small firms does not necessarily imply failure of either labor or credit markets; it may also reflect the inherent distribution of managerial ability in the population.

This mainstream view of microentrepreneurship is supported by evidence showing that, other things equal, some

individuals may derive a larger utility from entrepreneurship than from wage work, thus reducing the net opportunity cost of entering self-employment. Evidence of this has been provided by Blanchflower and Oswald (1998b), who stress that 63 percent, 48 percent, and 49 percent of salaried workers in the United States, the United Kingdom, and Germany, respectively, report that they would prefer to be self-employed, a fraction remarkably similar to the over 60 percent of Mexican entrants to self-employment who report doing so voluntarily (see Maloney 1999). Furthermore, Blanchflower (2004) reiterates the robust finding that those who enter self-employment report higher levels of job satisfaction than employees. Moreover, Hamilton (2000) finds that nonpecuniary benefits—such as being one's own boss—explain the lower conditional earnings generally found in the sector.

Jovanovic (1982) added dynamics to Lucas's view by further assuming that managerial ability is uncertain, and that individuals can gradually learn about their true cost structure only by opening and operating a business. Thus, entrepreneurs cannot know how good their location is, or how good their managerial ability is, until they actually start the business. Entry into self-employment involves a fixed cost, which only those with high expected ability and profits may be willing to pay. After entry, entrepreneurs incorporate the information from their actual profits, revise their ability estimates, and adjust the level of profit-maximizing output accordingly. Firms with consistently lower-than-expected profits tend to contract and eventually go out of business, while unexpectedly high profits cause upward revisions of ability estimates and lead to firm expansion.

The framework proposed by Jovanovic (1982) helps explain why small firms are generally less "productive" or "efficient." Firms that receive favorable cost information tend to revise upward their optimal size estimates, and thus to grow faster. Firms that remain very small are those that have received negative signals concerning their cost parameters. They are less efficient and have relatively high production costs, which lead them to stay small. This framework also helps explain why small firms tend to die more often (see Ericson and Pakes 1995; and Lippman and Rumelt 1982).

Indeed, there is a threshold level of efficiency below which it is not profitable to stay in business. Small young firms are more likely to cross that threshold when negative shocks suggest that in reality their efficiency is below what they had expected. On the other hand, firms that have been in business for a long time will have an accurate estimate of

their costs, so it is very unlikely that new information will be unfavorable enough to induce them to exit. In addition, larger firms are those that found high levels of profits when they started, and hence grew. Thus, their efficiency level is further away from the exit-inducing threshold and they are less likely to fail.²

Empirical evidence favoring the prediction of a negative link between time in business and size, on one hand, and survival and growth on the other, was obtained for the United States by Dunne, Roberts, and Samuelson (1988, 1989) and Evans (1987a, 1987b), for Germany by Wagner (1994), for the United Kingdom by Geroski (1991), for 10 OECD countries by Bartelsman, Scarpetta, and Schivardi (2003), and for 24 industrial and developing countries by Bartelsman, Haltiwanger, and Scarpetta (2004). In a developing-country context, however, one could argue that market and government failures associated with, for instance, credit rationing or inefficient contract enforcement mechanisms could limit the growth of small and newly created firms, which then would not exhibit the relatively higher expansion rates of their industrial-country counterparts. In this scenario, surviving small firms would maintain abnormally high profits given their size and efficiency, a prediction that is consistent with the high rates of return estimated for small enterprises in the developing world (see McKenzie and Woodruff 2006, and the references therein). The next section presents direct evidence on the subject, aimed at exploring the similarities and differences in microfirms' patterns of survival and growth in industrial and developing-country contexts.

Determinants of entry into self-employment

A model of industry dynamics that generates implications that are broadly similar to Jovanovic's, but is especially useful to analyze issues related to entry into self-employment, is the one proposed by Hopenhayn (1992). This model is especially notable for its analysis of the effect of the cost of entry, which can be interpreted as the outside opportunity costs for some resources (for example, managerial ability) used by the firm. Higher costs of entry lead to a lower turnover rate, because more ex ante selection occurs. This could be particularly relevant in developing-country contexts characterized by higher levels of informality and lower productivity in the salaried sector, which in Hopenhayn's model would lead to a lower entrepreneurial ability threshold for entering self-employment and thus to higher entry—and exit—rates.

In regard to the personal characteristics likely to be associated with entry into self-employment, Johnson (1978) and Jovanovic (1979) postulate that since young people are less risk-averse, they would be overrepresented among entrants. A similar prediction would emanate from the standard queuing view of LDC labor markets, with young people being more likely to be rationed out. Empirical evidence for the United States, however, suggests the reverse pattern, with entry *increasing* with age. As an explanation, Evans and Jovanovic (1989) offer a variation of Lucas's model where binding liquidity constraints may lead individuals to delay or forgo profitable business opportunities, reducing entry rates and increasing exit rates among those with low personal assets, including the young. Furthermore, they argue that since credit-constrained individuals are more likely to start small businesses with a suboptimal amount of capital, returns to capital will be higher and smaller firms (with lower capital stocks) will grow faster than firms that entered closer to their steady state.

Evans and Jovanovic's model of credit constraints to voluntary entry is largely inconsistent with the dualistic view of self-employment as an easy entry holding pattern in several ways. Indeed, both the dualistic view of informal labor markets and the advanced country sociological literature that sees the numerous self-employed among certain ethnic minorities as recruited from "misfits"—individuals who lack access to salaried employment due to, for instance, language barriers, a history of unemployment, or limited labor market experience (see Carrasco 1999; Evans and Leighton 1989; and the references therein)—would predict lower entry rates into self-employment coming from salaried work than from unemployment or from out of the labor market. However, the opposite prediction could be derived from the Evans and Jovanovic's view if individuals acquire more capital and knowledge of business opportunities—and to some extent their own managerial ability—while working than while unemployed or out of the labor market. In particular, if formal schooling is relatively poor and most relative human capital is accumulated on the job, then we may find that salaried employment is a logical stepping stone to self-employment.

Moreover, the impact of the level of remuneration earned in these previous jobs also offers some insights that may help distinguish between mainstream and dualistic models. All things equal, we might expect that those earning higher wages in salaried work would be less likely to be "misfits," or unsuited to formal work, and hence less

prone to move into self-employment. However, in the presence of credit constraints, workers earning higher wages in the salaried sector may also be able to accumulate capital faster and hence be more likely to enter self-employment. Moreover, there may be a correlation between previous productivity in the salaried sector, remuneration, and entrepreneurial ability—or at least competence in the chosen field of entrepreneurship. To the degree that this would imply a higher probability of success in self-employment, we may expect those with conditionally higher earnings in the formal sector to enter self-employment while those less skilled may choose not to take the risks.

In sum, the impact of personal characteristics of existing and would-be entrepreneurs could be potentially different in developing-country contexts, at least in the hypothesis that self-employment in LDCs is driven by the presence of dualistic labor markets. Indeed, in this case unemployed individuals, those out of the labor force, young workers, and those with lesser schooling and lower wages should all be more likely to be self-employed, as they would be in a worse position for finding formal salaried jobs. In contrast, the mainstream literature suggests that older, better-educated, and well-paid workers with experience in the salaried sector should have a higher probability of entering, staying, and growing in the self-employment sector as they should be more likely to have accumulated the assets required to start a business, and better positioned to find, assess, and take advantage of good business opportunities.

Finally, two additional covariates appear in the mainstream literature but so far yield ambiguous predictions (see box 5.1 for a summary of empirical findings in industrial country contexts). First, we might imagine that workers with more schooling would find better matches as salaried workers in larger firms that could better utilize their specialized skills. On the other hand, Cressy (1996) and Rees and Shah (1986) argue that more educated individuals can have lower costs of assessing business opportunities, and that human capital may be a complement to managerial ability. This is suggested by Bates (1990), who shows that in the United States, the probability of survival of small businesses is positively related to the level of education of their owners. Second, Carrasco (1999) argues that men who are married could be less willing to take risks. On the other hand, the Mexican sociologist González de la Rocha (1994) suggests that the possibility of combining the self-employment earnings of the household head or

BOX 5.1

Patterns of entry and exit in industrialized countries*Self-employment and age*

Evans and Leighton (1989) find that, for U.S. men, entry into self-employment increases only slightly after the late 20s, exit falls at a decreasing rate with age, and the self-employment rate reaches a plateau when individuals are in their 40s. For Spain, Carrasco (1999) finds that entry rates are highest for middle-age people, especially for those aged 35–45. In the United States, Bates (1990) finds that rates of exit from self-employment are lowest for owners of small businesses aged 45–54. Holtz-Eakin, Joulfaian, and Rosen (1994a) find that exit rates reach their minimum at the age of 47. In addition, for each of the 18 industrial countries considered by Blanchflower and Oswald (1998b), a robust positive relationship emerges between age and the probability of being self-employed.

Self-employment and education

Bates (1990) finds that in the United States the probability of survival of small businesses rises with education. In Spain, entry into the sector also rises with education, as found by Carrasco (1999), particularly for those who become self-employed with employees. In other industrial countries, Blanchflower (2000) finds that self-employment rates are usually highest for individuals with a small number of years of schooling—the United Kingdom being the only exception, with the reverse being the case—followed by those with college education.

Profitability and exit from self-employment

Holtz-Eakin, Joulfaian, and Rosen (1994a) use net income as a proxy for entrepreneurial ability and other relevant factors, such as education, time in business, and gender, not available in their data set. They find that net income has a negative effect on the probability of exiting self-employment.

Previous wages and entry into self-employment

Evans and Leighton (1989) find that higher previous wages lower the likelihood of entering self-employment. However, the statistical significance of that variable is reduced considerably when individuals' labor market characteristics are controlled for.

Unemployment and entry into self-employment

Carrasco (1999) finds that in Spain, for given personal characteristics, the predicted probability of entering self-employment is lower for wage workers than for unemployed individuals. However, the latter become less likely to switch either when they receive unemployment benefits or when aggregate business conditions are unfavorable. In the United States, Evans and Leighton (1989) find that U.S. men who changed jobs frequently, and who experienced relatively frequent and long spells of unemployment, have a higher probability of entering self-employment.

Self-employment and liquidity constraints

Evans and Leighton (1989), Evans and Jovanovic (1989), and Carrasco (1999) find that higher net worth increases the probability of entering self-employment. Evans and Jovanovic (1989) also find that initial assets before entering self-employment are positively related to earnings, at least during the initial years in the sector—as firms grow, the importance of the initial liquidity constraint diminishes. Holtz-Eakin, Joulfaian, and Rosen (1994a, 1994b) find that liquid assets and inheritances increase the likelihood of entry into self-employment, reduce the probability of exit, and increase the earnings of surviving businesses. Blanchflower and Oswald (1998b) provide similar evidence for the United Kingdom; by relaxing liquidity constraints, previous inheritances increase the probability of being self-employed. As additional evidence they report survey data that reveal that most small businesses are started not with bank loans but with one's own or family money, and that shortages of capital are the most common reason for not entering self-employment. Bates (1990) finds that small businesses that were able to secure loans at the time of getting started are more likely to survive.

Both exit rates and employment growth rates of manufacturing firms decrease with size and time in business. Evans (1987a, 1987b) and Dunne, Roberts, and Samuelson (1988, 1989) find that size and time in business are both related negatively to the exit and growth rate of U.S. manufacturing plants.

his/her spouse with the salaries of other family members could reduce overall household income risk.

Determinants for small firm informality

The high levels of informality found among microenterprises in developing economies can be interpreted as resulting from profit-maximizing decisions by private firms subject to environmental constraints that increase the costs or reduce the benefits of formality—for example, high costs of firm or property registration, or market and government failures in credit markets, and contract enforcement mechanisms. A vicious cycle would be at work, by which barriers to small firms' participation in formal market support institutions would limit firm growth, which in turn would help maintain those institutions out of the reach of most small firms. However, a somewhat nuanced alternative view is that small firms in developing countries have an inherently limited demand for those formal services, due to their low productivity levels and limited growth prospects. In this context, a firm's choices regarding the use of formal or informal arrangements would depend on the stage of the firm's life cycle.

This view is related to the burgeoning literature on social capital and informal networks, which has highlighted that absence of participation in state-organized or formal market institutions does not imply that agents do not acquire similar services through less formal means and, to a lesser degree, that informal alternatives are necessarily inferior to formal ones. In this report, we argue that the calculations involved in participation decisions apply to all institutions, both informal and formal. We thus recast the question of formality as the firm's decision of how much to participate in the numerous institutions of civil society: civic organizations, trade organizations, federal and local treasuries, governmental programs such as social security (including pensions and health care), the legal system, the banking system, insurance institutions, health inspections, firm censuses, and so on. We argue that a minimal degree of participation in some institutions is a necessary input to growth for many firms, and that participation increases with the success of the business; it is a normal input into the production process (see Levenson and Maloney 1998).

Firms derive multiple benefits from formality. We can view being "formal" as lying at the end of a continuum of possibilities of participation beginning at the household and extending through communities and networks to the formal institutions of the state. The benefits of formality,

although often overlooked, are numerous and become more important as firms get larger. Consider a select few (and see de Soto 1989; and chapter 6 of this volume):

- *Enforceable/impersonal contracts and credible signaling.* All entrepreneurs have access to social relationships to enforce implicit contracts among their friends and family, who form a small number of their potential customers and employees. Participation in the legal system is needlessly expensive for transactions with these individuals. But this mode of operation is constrained by the ability of the entrepreneur to maintain personal relations with all involved parties, a task increasingly unmanageable as firms expand. Legally recognized, enforceable contracts lend credibility to arrangements, permit entry into long-term commitments, diminish risk, and can reduce monitoring costs. Larger investments require that property rights be secured through the legal system.
- *Access to capital.* Informal capital markets (Besley 1995) may be sufficient to fulfill the firm's external financing needs at low levels of production. However, the small scale and undiversified nature of informal capital markets make them unsuitable for satisfying the firm's financing needs at larger scales of operation; growing firms will turn to formal financial intermediaries such as banks. This, of course, implies costs in terms of better record keeping, and transactions costs more generally.
- *Access to public risk pooling mechanisms.* Demand for risk pooling mechanisms can be expected to increase with firm size. Arnott and Stiglitz (1991) show that where "peer monitoring" generates more information than that available to market insurers, it is an important mechanism in reducing moral hazard and makes non-market (informal) insurance provided by friends or communities an important complement to market insurance. Small firms with relatively little capital may insure against small losses within the family or community rather than pay the premium associated with monitoring pooling across larger groups. As a firm grows, we may imagine that both the magnitude of potential losses and the decreasing ability to peer monitor may lead to increased demand for formal market insurance.
- *Access to business information.* As a firm grows beyond local customers and suppliers, it may also require

more information and business contacts to continue growing. We may, therefore, expect more participation in business associations, among other changes. We can also consider inscription in training programs and the use of other business development services as formal channels for information gathering.

Even though formality also has multiple costs—for example, reporting requirements and fiscal obligations—as firms grow, the relative benefits of formality may eventually exceed the corresponding costs. De Soto claims that Peruvian sidewalk vendors sought, not to avoid, but to pay taxes as a way to establish property rights over their precarious business locations. In reality, although the direct private benefit from paying taxes may be zero (again, assuming no enforcement penalties), there may be ancillary benefits that make compliance worthwhile.

Implications for microfirm dynamics and informality

The combination of the above very stylized concept of participation in formal institutions and the insights from mainstream models of firm dynamics yields a number of implications. First, it is reasonable to expect considerable *heterogeneity in the degree of formality*. The benefits and costs of participation undoubtedly vary across societal institutions, and vary for firms of a different size and age. While there are potential complementarities between different societal institutions, a large number of firms will choose to participate in only a subset of institutions at any point in time. For example, the legal system and bank financing are complements, but a firm may have to register legally before seeking external financing. Thus, informality is not an all-or-nothing state, and the degree of formality varies by firm.

A second implication is that *small firms are disproportionately informal*. Small firms benefit least from participation because of the small scope of their dealings with the public and hired employees (relative to the total volume of transactions undertaken by the firm). Indeed, firms of different sizes may have different degrees of interaction with the public. Because implicit contracts over product quality are cheaper and feasible to enforce with friends and family, the entrepreneur may find it most cost-effective to primarily serve such customers when faced with small sales volumes. At larger volumes (later in the firm's life cycle), friends and family cannot necessarily buy all the firm's output, so sales to the

general public and other firms should increase, leading to greater benefits from formality.

A natural corollary is that “inefficient” firms are disproportionately informal. However, in contrast to other formulations, in this case the causality is not necessarily from informality to inefficiency: high-cost firms choose less formality because they benefit less from it than more efficient firms that produce at higher volumes for longer lengths of time. Similarly, *young firms are disproportionately informal*. This is partly because young firms are more likely to be small. Moreover, conditional on size, the population of young firms contains a disproportionate number that have not received enough signals to figure out whether paying the costs of formality are worthwhile; many eventually will go out of business.

Another implication is that the relatively high mortality rates found among informal firms are not necessarily evidence of the inferiority of informal employment, and could rather reflect the fact that those firms are predominantly “young” and small which, even in an undistorted industrial-country context, would lead to observed high turnover rates. As more efficient firms grow to their equilibrium size, however, both mortality and informality rates would naturally decline.

In sum, combining the insights from the firm dynamics literature with those derived from the literature on institutional development provides a new possible interpretation of the stylized facts on microenterprises and informality. Because they have not assembled enough information on their true efficiency, young firms tend to be small and to have high failure rates. On the other hand, formality is not an all-or-nothing decision, but rather an option in a continuum of participation possibilities. For young and small firms, the benefits outweigh the costs of participation in formal societal institutions. But as more efficient firms survive and grow, their needs for enforceable contracts, formal credit markets, and access to public risk-pooling mechanisms increase, and so does their degree of formality or depth of participation in societal institutions.

Young and small firms do tend to have higher costs and high failure rates and, at the same time, they are more likely to be informal. Although this description corresponds exactly to the standard picture of the stagnant, precarious, unprotected informal self-employment sector familiar in the literature, it is, in fact, the opposite. It emerges naturally from workers trying their luck at entrepreneurship (risk taking), often failing, and not engaging in the formal institutions until they grow. In sum, there may be nothing

TABLE 5.1

Entry probabilities into Mexico's self-employment sector (percent)

Initial/final status	Own account	1–4 workers	5–9 workers	10 or more workers	Total entrepreneurs
Salaried	4.2	1.7	0.2	0.2	6.2
Contract	9.0	3.1	0.1	0.1	12.4
Salaried + contract	4.7	1.8	0.2	0.2	6.9
Unemployment	9.3	1.5	0.1	0.1	11.0
Other work	4.6	3.4	0.2	0.1	8.3
Out of labor force	2.6	0.8	0	0	3.5

Source: Author's calculations, using data on males from the *Encuesta Nacional de Empleo Urbano* 1998–2001.

pathological about the informal microfirm sector, and its existence may be largely unrelated to questions of labor market dualism or even credit market distortions.

Microfirm dynamics in Latin America

Collecting evidence on the entry, exit, and growth dynamics of Latin American and the Caribbean microfirms is of interest for at least three reasons. First, we would like to know if microfirm dynamics in the Latin America and Caribbean (LAC) region are similar to those in the industrial countries. Second, the very high mortality rates among LDC firms contribute to the perception of the sector as “precarious.” If in fact the failure rates are similar to those found elsewhere, it is difficult to make the case that we are dealing with unusually high rates of risk. Third, the demographic characteristics associated with entry and exit provide useful information about the role the sector is playing and the constraints facing it.

Studying the patterns of entry, survival, and growth in the microenterprise sector requires the use of panel data that allow following the same individuals and firms over time. This creates a severe constraint, as most longitudinal industrial surveys exclude small and informal firms. For that reason, this section focuses on only a few countries for which we were able to obtain appropriate data. Our main focus is on Mexico, using combined data from employment and microenterprise surveys.³ Some of the analysis is replicated using a longitudinal household survey for Nicaragua. Finally, we present evidence on the job creation potential of firms of different sizes using the Colombian industrial survey, for a period during which its coverage included micro- and small firms.⁴ In what follows, we define the self-employed category as including all individuals whose main job consists in working in their own businesses—for example, those individuals who report that in their main job

they were either “employers” (“*patrones*”) or “own-account workers” (“*trabajador por su cuenta*”). This is similar to the standard definition used in the U.S. self-employment literature. Evans and Leighton (1989), for instance, define as self-employed all sole proprietors, partners, and sole owners of incorporated businesses.

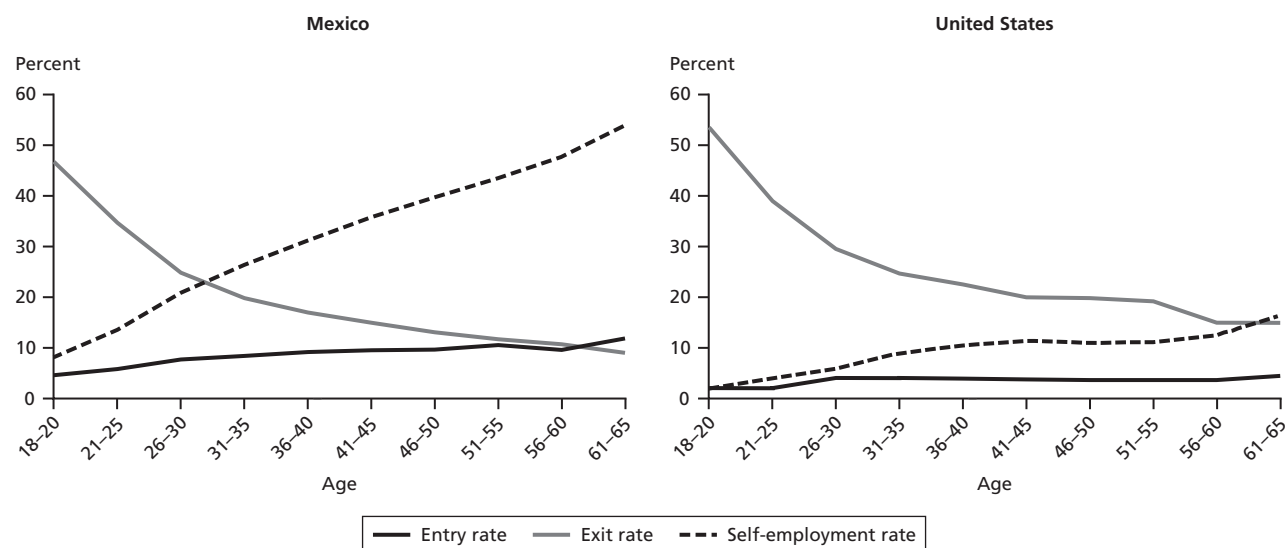
Patterns of entry

Taking the United States as a benchmark, the rates of entry into self-employment in Mexico are of very similar orders of magnitude, with a slightly higher fraction of Mexican wage workers becoming entrepreneurs: 6.2 percent compared to 4 percent per year found by Evans and Leighton (1989).⁵ Table 5.1 presents the probabilities of becoming an enterprise owner for Mexican males previously employed in other segments of the labor market or located outside of the labor force. Rows represent individuals' initial labor market positions, and columns represent labor market statuses a year later. Each cell shows the percentage of individuals who start in a given row-category and end in the corresponding column-category.

New entrants into self-employment are more likely to start their businesses without any employees. Table 5.1 shows that about two-thirds of the salaried workers who transit into self-employment each year—4.2 out of 6.2 percent—do so without hiring any workers, which is consistent with the mainstream firm dynamics literature view of new entrepreneurs “testing the waters” before making significant investment decisions. Table 5.1 also shows that individuals who are outside the labor force have a lower probability of entering self-employment—3.5 percent—and that those coming from contract work or unemployment have higher probabilities of entering the sector—respectively, 12.4 and 11 percent. In all these cases, however, three out of four new entrants start without hiring any employees.

FIGURE 5.1

Entry and exit into self-employment among men



Sources: Evans and Leighton 1989; author's calculations, using *Encuesta Nacional de Empleo Urbano*.

There are also striking similarities in age patterns of self-employment between Mexico and Nicaragua, on one hand, and the United States on the other. The left and right panels of figure 5.1 show the fraction of the male labor force in self-employment by age group, respectively, for Mexico and the United States. The figure also exhibits the rates of entry into that sector from salaried work and the rates of exit from self-employment, also by age group. The self-employment rate (calculated as a fraction of the labor force) is higher in Mexico for every age group by at least a factor of two. Note, however, that the overall age patterns of self-employment rates as well as those of entry and exit are strikingly similar in the two countries.⁶

Both in Mexico and the United States, exit rates decrease while entry rates slowly increase with age. As a result, the percentage of individuals who are self-employed increases with age at a decreasing rate, peaking at the late 40s. As a foreshadowing of the more detailed analyses below, the common upward and downward sloping relationships of age with the rates of, respectively, entry and exit are consistent with the view that older entrepreneurs get a more precise view of their underlying entrepreneurial capacity and hence are more likely to enter and less likely to fail than younger individuals.

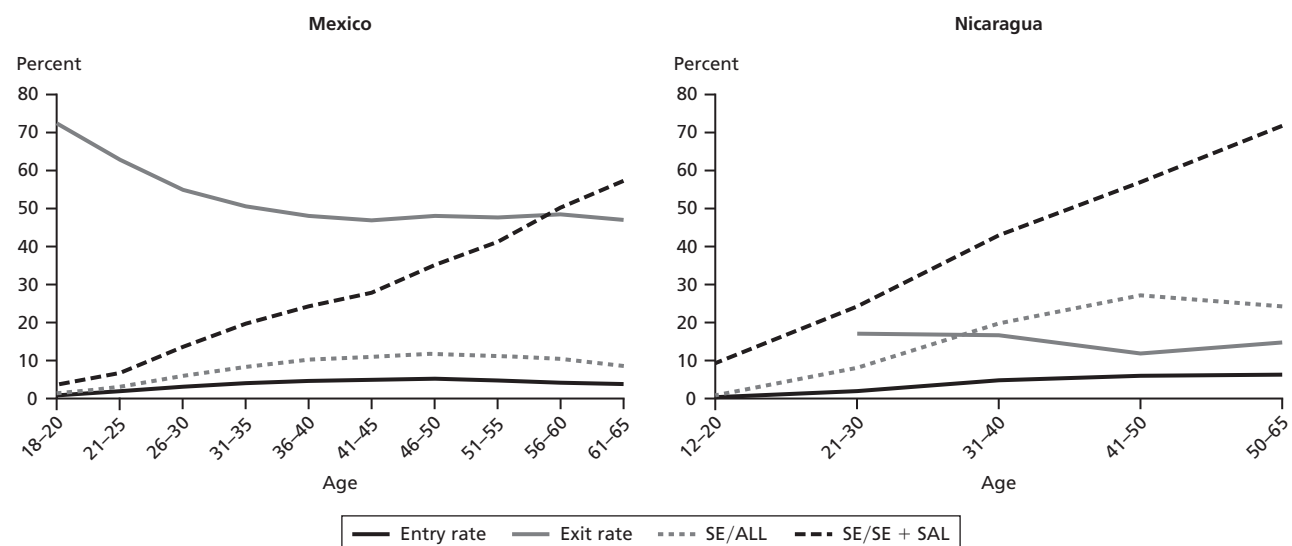
Women exhibit similar patterns of entry and exit into self-employment, which is illustrated in figure 5.2 for the

cases of Mexico and Nicaragua.⁷ Like men, Mexican women's entry rates rise until the early 40s and begin to decline after the age of 50. Their exit rates are much higher and decline at a slower rate until the late 30s, stabilizing at around 50 percent (for yearly transitions). The overall rate of self-employment is thus much smaller than that of men, although it also peaks at the age of 50. However, among working individuals, the share of self-employed women increases monotonically with age at an even steeper pace than that of men; it evolves from less than 5 percent in the early 20s (10 percent in Nicaragua) to more than 50 percent after the age of 55 (70 percent in Nicaragua).

Econometric examination of the effect of individual demographic and labor market characteristics on the probability of entering self-employment among Mexican males confirms the age patterns illustrated in the figures above.⁸ Thus, keeping other personal characteristics constant, and comparing with individuals aged 15 to 20 (for which the rate of entry into self-employment is 2.4 percent), the probability of entering self-employment is 5.7 percentage points higher for those aged 21 to 35, and 9 percentage points higher for those in the 36–50 age bracket. These results are not consistent with the view of the sector as a point of entry into the Mexican labor market, but they are very consistent with the U.S. data. They also provide support to Evans and Jovanovic's (1989) liquidity constraints

FIGURE 5.2

Entry and exit into self-employment among women



Source: Author's calculations, using *Encuesta Nacional de Empleo Urbano* and Nicaragua's Living Standards Measurement Study.

hypothesis, and the view that older workers have a more precise measure of their underlying entrepreneurial capacity.

Educational attainment, on the other hand, has a negative albeit quantitatively small effect on the probability of entering self-employment. However, when we focus only on workers entering self-employment from salaried work and/or when we estimate the effect of education conditioning on individuals' changing labor market status, the negative relationship between education and entry into self-employment levels off and breaks up at higher levels of schooling, suggesting that college graduates (that is, doctors, consultants, and so on) may find the sector attractive. Moreover, when we restrict the definition of self-employment to include only those business owners who employ at least one worker, we find that both secondary schooling and tertiary schooling are positively linked to entry into self-employment. This result is broadly consistent with Carrasco (1999), Evans and Leighton (1989), and Rees and Shah (1986) who found that entry rises monotonically with education.

Also consistent with Rees and Shah (1986) and Carrasco (1999) is the finding that being married is positively associated with entry, which may reflect either that the sector is not riskier than other alternatives available in the labor market (that is, salaried employment) or that, in fact, being married helps diversify risk. Quantitatively, unmarried individuals have a 4.5 percent probability of entry into

self-employment, and that probability increases by about 25 percent for their married counterparts. Moreover, when the estimation is performed conditioning on individuals changing their labor market status, the effect of marriage is about four times larger.

The importance of conditioning on changes in labor market positions becomes clearest in looking at the impact of initial employment and labor force participation status. The unconditional results suggest that those out of the labor force and not studying and those unemployed are more likely to enter self-employment, consistent with traditional dualistic views of the sector as disguised unemployment. However, all of the above segments of the labor market are characterized by very high rates of turnover, so the corresponding workers may well be more likely to move to all sectors at disproportionately high rates. When we condition on changing labor market status, we find that those in formal salaried employment are more likely than most other groups to enter informal self-employment.

Among the unemployed, however, those workers with longer unemployment spells—especially those who have been unemployed for between four and six months—have higher conditional and unconditional probabilities of entering self-employment through opening owner-only businesses. This suggests that even if the sector does not function predominantly as a holding pattern for misfits or

those rationed out of salaried work, it does offer income opportunities for the long-term unemployed.

As for the links between entry into self-employment and the previous job characteristics reported by salaried workers, we find that those employed in firms with at most 10 employees have a much higher probability of entry into self-employment than those coming from larger firms: 12.8 percent compared to about 9 percent for those in firms with 11 to 250 workers, and 5 percent for those coming from firms with more than 250 employees. This may be due to the higher nonpecuniary benefits (such as social security) and job stability offered by larger firms.

For a given firm size, on the other hand, higher prior wages also increase the probability of entry into self-employment. While we cannot tell whether this is the effect of a relaxation of liquidity constraints or if it reflects entrepreneurial ability, the fact that overperformers in the salaried sector are more likely to enter self-employment is at odds with the predictions of both the dualistic approach and the sociological view of that sector as a preferred destination for “misfits.”

As for the effect of the sectors of economic activity in which individuals have been previously engaged, we find that at least in Mexico, entry into self-employment is most likely for those who hold jobs in the construction sector and it is least likely for those employed in manufacturing. Thus, while a salaried worker in manufacturing has a 4.3 percent probability of entering self-employment, the corresponding probability for a salaried construction worker with similar personal characteristics is 11.8 percent. The higher rate of entry into self-employment coming from salaried work in the construction sector persists even when one controls for the fact that this sector also has higher turnover rates—workers there are more likely to change labor market status. One possible

explanation is that the sector offers more opportunities for very small firms, as it comprises mostly individual contractors operating without any employees; in the U.S. construction sector, for instance, owner-only businesses represented 75 percent of all firms, while those with 20-plus workers were just 2 percent, compared to, respectively, 49 and 13 percent in the manufacturing sector. As argued by Tybout (2000) and further discussed in the next chapter, small firms tend to locate in industries where they have smaller cost disadvantages with respect to larger incumbents, and which are characterized by lower levels of capital and skilled-labor intensity.

Patterns of survival and exit

As shown in table 5.2, during any given year about 15 percent of the Mexican self-employed move to salaried work—and the same number of individuals migrate to other labor market positions, namely, to contract work (6.2 percent) and out of the labor force (5.7 percent). The probability of leaving self-employment is higher for those who do not have employees: 35 percent compared to slightly less than 25 percent for those who own larger firms. Moreover, among those leaving self-employment, the probability of switching to salaried work is lower for those who own smaller firms, when compared with owners of larger firms who are less likely to switch to contract work or out of the labor force.

Econometric estimates of the determinants of firm survival using data on Mexican microfirms confirm the age patterns depicted in figures 5.1 and 5.2, with the probability of staying in business increasing with age until the 36–50 age bracket. Thus, while the average self-employed individual aged 15–20 has only a 33.3 percent probability of staying in self-employment, individuals with similar

TABLE 5.2

Exit probabilities from Mexico's self-employment sector (percent)

Initial/final status	Remain entrepreneur	Salaried	Contract	Salaried + contract	Unemployed	Other work	Out of labor force	Total
Own account	65.0	17.0	7.4	24.5	1.5	1.9	7.2	100
1–4 workers	76.6	11.7	5.0	16.7	0.8	2.0	3.9	100
4–9 workers	78.6	13.5	3.2	16.8	0.6	1.3	2.7	100
10 or more workers	78.3	15.2	2.3	17.5	0.3	1.2	2.7	100
Total entrepreneurs	69.9	15.0	6.3	21.3	1.2	1.9	5.7	100

Source: Author's calculations, using data on males from the *Encuesta Nacional de Empleo Urbano* 1987–2001.

personal characteristics in the 21–35 and 36–50 age brackets have probabilities of survival in self-employment of, respectively, 52 and 60.5 percent.

Education has a negative impact on the probability of staying in self-employment, suggesting that the “pull” effect of better employment alternatives for more educated individuals tends to dominate over the “push” effect associated with lower probabilities of business failure observed in the United States among more educated entrepreneurs. However, the effect is quantitatively small—a reduction of about 2 percentage points in the survival probability for individuals with secondary or tertiary schooling—and it becomes positive and significant when the sample is restricted to firms with at least one worker. In other words, when the focus is on larger firms, the evidence for Mexico confirms that for the United States, with schooling and firm survival exhibiting a positive correlation.

Consistent with the various mainstream models of firm dynamics reviewed above, we find that higher conditional wages of business owners—which we interpret as reflecting higher profitability due to unobserved characteristics—tend to increase the probability of staying in business. Thus, a doubling of the net earnings from self-employment is associated with a 4.2 percentage point increase in the probability of staying in that sector. The effect is quantitatively larger when the sample is restricted to businesses with at least one employee besides the owner, for which a doubling of self-employment earnings is related to an 8.8 percentage point increase in the probability of maintaining that initial status.

We also find that survival in self-employment is between 5 and 8 percentage points higher for married individuals—possibly because they can count on unpaid family workers—and about 6 percentage points lower for individuals with a second job. Moreover, workers in the construction sector have a probability of staying in self-employment that is between 10 and 15 percentage points lower than for those engaged in agriculture, manufacturing, and services, while those engaged in commercial activities have a 5.5 percentage point higher survival probability than their peers in the above three sectors. The fact that once again the construction sector emerges as a more likely source of workers who switch sectors could be due to its inherent high rate of firm turnover.

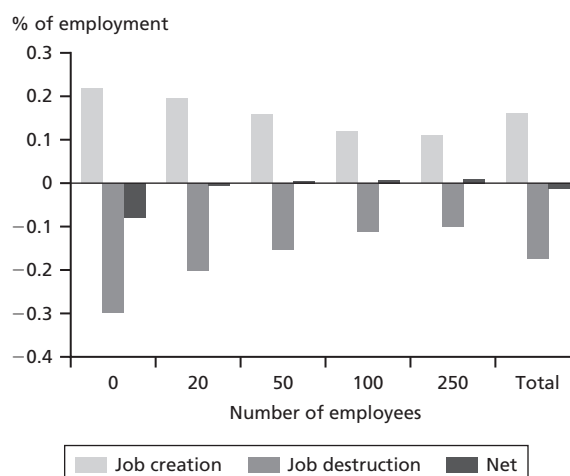
As for the effects of firm size and age on survival in self-employment, we find that they are significant although quantitatively small. In particular, a doubling of the time that a microfirm has been in business is associated with

a 4.3 percentage point increase in its likelihood of survival, while a tripling of its capital stock is related to a 1.4 percentage point increase in that probability. Despite the small magnitude of the effects, these results are consistent both with Jovanovic’s “noisy selection” view and with the bulk of the mainstream empirical evidence that firms get a more precise estimate of their cost structures with experience, and past measures of success are informative for the future evolution of the firm.

Do microfirms grow?

All the available evidence points to the fact that microfirms create relatively more jobs, but destroy even more. This is illustrated by figure 5.3, which shows rates of job creation and destruction among Colombian manufacturing plants calculated for different employment size categories—the rates are calculated as a percentage of employment in the respective size category. Firms with fewer than 20 employees have the highest rates of job creation, but their job destruction rates are proportionally even higher, causing net job creation rates to be the lowest of any size category. As firm size increases, both job creation and destruction tend to diminish, but the latter at a larger rate, causing net job creation to be higher among larger firms. This is consistent with the findings for the United States, where smaller manufacturing firms and plants exhibit very high *gross* job creation rates, but not higher *net* job creation rates (Davis, Haltiwanger, and Schuh 1996).

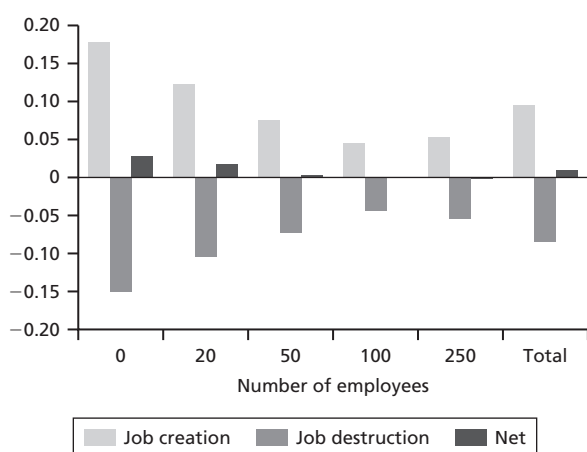
FIGURE 5.3
Job creation and destruction in Colombia, 1977–82



Source: Author's calculations, using the Colombian Manufacturing Survey.

FIGURE 5.4

Job creation and destruction by entering and exiting versus continuing firms in Colombia, 1977–82



Source: Author's calculations, using the Colombian Manufacturing Survey.

Gross and net job creation rates by entering Colombian firms are higher than for continuing firms (figure 5.4). As a share of the size category, roughly four times as many jobs are created by entering firms as by those already existing. Job destruction by exiting firms is roughly equal to that by continuing firms, so that net job creation by entering and exiting firms is positive, while it is negative among continuing firms. Thus, although entering firms more than compensate for the jobs destroyed by the plants that fail, the Colombian data suggest that the low rate of expansion of smaller continuing firms leads them to provide a negative net contribution to employment growth.

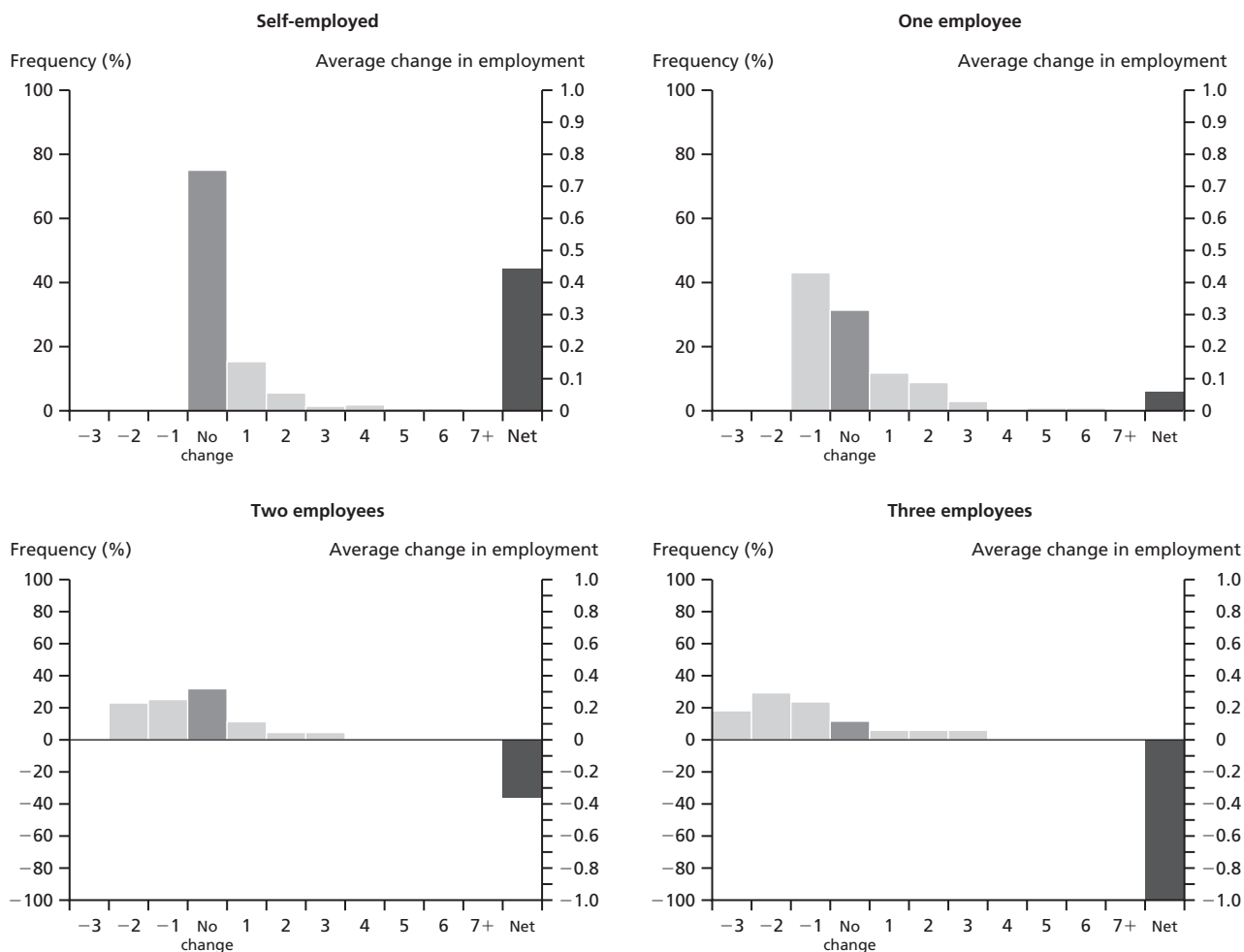
Most of the net job creation among microfirms is from zero to one employee. In the case of Nicaragua, for instance, looking at firms with at most three employees (plus owner) in 1998, we find that for those that were still in business by 2001, employment grew by an average of about 20 percent. For those with two and three employees, the losses actually exceeded the gains, so all the above growth came from firms with at most one employee (figure 5.5). In fact, most was due to those with no employees at all, which on average expanded by 40 percent, despite the fact that 73 percent did not grow at all. Firms with one employee at the beginning of the three-year period also had a positive contribution to employment growth; although 40 percent of those lost their only worker, this was offset by hires by other firms in the category.

Data on the evolution of Mexican microfirms over a 12-month period suggest a broadly similar profile (table 5.3). Over 12 percent of owner-only firms expand to one to four employees across one-year periods. However, 22.1 percent of those between one and four employees contract again to be owner-only. Thus, in absolute numbers, there are slightly less firms that contract from one to four employees than firms that expand into that range. Moreover, over a one-year period, only 5.4 percent of the firms with one to four employees do actually expand; 49.2 percent stay in the same size range and 45.5 percent contract to owner-only or go out of business, with the employer moving into wage work or unemployment, or leaving the labor force.

Given the low rates of employment growth observed among microfirms, it is not surprising that the vast majority of them will reach their steady state at a very small size. Thus, among Mexican and Nicaraguan microfirms with at least three years of time in business and at most 15 workers, 64 percent have no employees, 21 percent have only one, and between 8 and 9 percent have just two workers (figure 5.6). In other words, in those two cases, 93 percent of microfirms do not grow beyond two employees. In fact, the large majority of microenterprises are owner-only. In Brazil and Mexico, for example, respectively, 87 and 80 percent of all microenterprises with fewer than five workers have no paid employees. Moreover, as illustrated in table 5.4, in Argentina, Brazil, and Mexico owner-only firms represent roughly 58–68 percent of all firms. These proportions are smaller than those found in the United States and the United Kingdom, where, respectively, 71.6 and 80 percent of all firms have no employees. It is worth noting that the lower fraction of firms without employees found for LAC is not inconsistent with the fact that, as reported in previous chapters, those countries exhibit higher rates of self-employment. Indeed, the latter are calculated as a proportion of the labor force, while the figures reported in table 5.4 are based on the size distribution of the number of firms—not employment.

Among firms with at least one employee, however, the fraction of microenterprises with between one and four workers is much larger in the LAC countries for which data are available: about 90 percent in Argentina and Mexico and almost 80 percent in Brazil, compared to, respectively, 55 and 65 percent in the United States and the United Kingdom (table 5.5). This is consistent with the evidence presented by Bartelsman, Haltiwanger, and Scarpetta (2004) on the fact that in all countries for which data

FIGURE 5.5

Employment growth of microenterprises in Nicaragua

Source: Author's calculations, using Nicaragua's Living Standards Measurement Study.

TABLE 5.3

Growth and contraction probabilities of Mexican microfirms (percent)

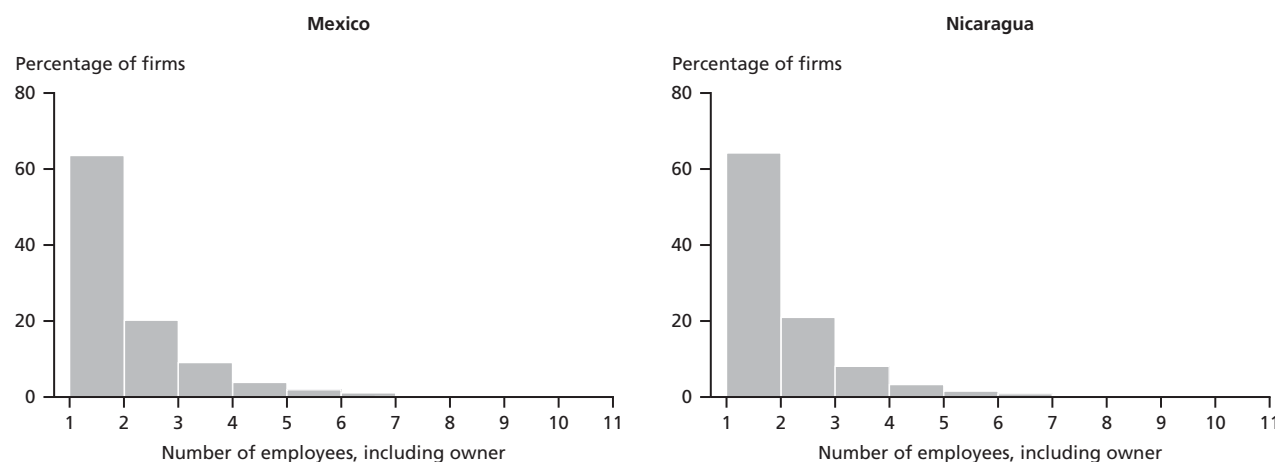
Initial/final status	Own account	1–4 workers	5–9 workers	10 or more workers	Exit	Total
Own account	51.9	12.4	0.5	0.2	35.00	100
1–4 workers	22.1	49.2	3.9	1.5	23.40	100
4–9 workers	7.8	35.1	22.6	13.1	21.40	100
10 or more workers	4.1	15.2	14.4	44.6	21.70	100
Total entrepreneurs	38.4	26.0	2.9	2.6	30.10	100

Source: Author's calculations, using data on males from the *Encuesta Nacional de Empleo Urbano* 1987–2001.

are available—they consider 19 countries, including 9 industrial ones and, from LAC, Argentina, Brazil, and Mexico—small firms with fewer than 20 employees account for between 80 and 96 percent of the total firm population. However, it appears that within the group of firms with fewer than 20 workers, microenterprises are

much more prevalent in developing than in industrial countries. Not surprisingly, Bartelsman et al. (2004) find that the average size of firms is much higher in the United States than in the above three LAC countries, which they show is not related to differences in sector composition—for example, some countries specializing

FIGURE 5.6

Size distribution of microenterprises in Mexico and Nicaragua

Source: Author's calculations, using *Encuesta Nacional de Empleo Urbano* and Nicaragua's Living Standards Measurement Study.

Note: All firms were at least three years old at the time of measurement. Firms pooled 1992–98.

in industries with a smaller efficient scale—but rather to within sector differences in average firm size.

As argued above, a high prevalence of small firms does not necessarily imply the presence of high external barriers to firm growth. As noted by Lucas (1978, pp. 508–23), the size distribution of firms may be determined by the underlying distribution of entrepreneurial ability or the implicit costs of operation at different sizes. That is, urban microenterprises could be optimally small, given the owners' entrepreneurial abilities, and the costs of conforming to the requirements of formal contracting arrangements, participation in credit markets, and such. Thus, the prevalence of small firms could simply reflect small opportunity costs of entry into self-employment. Whether this mainstream view of the sector is in fact relevant in the case of Latin America

is something that we investigate next by exploring the personal and firm traits that are associated with firm growth.

To examine what characteristics seem correlated with microfirm employment growth, we use data on the transition between size-brackets of Mexican microfirms over a one-year period considering only surviving firms.⁹ Consistent with mainstream models, entrepreneurs with conditionally higher earnings (better performance) seem to show a higher propensity to grow. Similar results are obtained for Nicaragua, for both male and female entrepreneurs; those with higher profits relative to their peers with similar levels of human and physical capital are more likely to add workers.

Firm size also appears with the negative sign predicted by the mainstream literature: bigger firms are more likely to have achieved their optimal size. Time in business also has a

TABLE 5.4

Employment size distribution of firms, selected countries (percent)

Number of employees	US (2004)	UK (2005)	Mexico (2004) ^a	Brazil (2003) ^b	Argentina (2003) ^a
With no employees	80.0	71.6	64.4	57.7	68.4
1–4	10.9	18.6	33.0	33.3	28.8
5–19	6.6	7.6	2.0	4.9	2.0
20 or more	2.5	2.1	0.7	4.1	0.8

Source: Author's calculations, using data from U.S. Small Business Administration, U.K. Small Business Service Analytical Unit, Argentina's *Encuesta Permanente de Hogares*, Brazil's *Pesquisa Nacional por Amostragem de Domicílios*, and Mexico's *Encuesta Nacional de Empleo Urbano*.

a. Third and fourth size ranges: 5–14 workers and 15+ workers, respectively.

b. Third and fourth size ranges: 5–10 workers and 11+ workers, respectively.

TABLE 5.5

Employment size distribution of firms with employees, selected countries (percent)

Number of employees	US (2004)	UK (2005)	Mexico (2004) ^a	Brazil (2003) ^b	Argentina (2003) ^a
1–4	54.5	65.7	92.4	78.7	91.1
5–19	33.0	26.9	5.6	11.6	6.3
20 or more	12.5	7.4	2.0	9.7	2.5

Source: Author's calculations, using data from U.S. Small Business Administration, U.K. Small Business Service Analytical Unit, Argentina's *Encuesta Permanente de Hogares*, Brazil's *Pesquisa Nacional por Amostra de Domicílios*, and Mexico's *Encuesta Nacional de Empleo Urbano*.

a. Third and fourth size ranges: 5–14 workers and 15+ workers, respectively.

b. Third and fourth size ranges: 5–10 workers and 11+ workers, respectively.

negative effect, as expected in the context of the literature on developed countries, but its coefficient is not statistically significant. Similarly, consistent with Ericson and Pakes (1992b, pp. 53–82)—although less so with Evans and Jovanovic (1989, pp. 808–27)—we find that higher capital stocks are positively correlated with employment growth.

As for the relationships between firm growth and the age and marital status of entrepreneurs, we find that they are broadly similar to those obtained in the survival analysis: business owners who are married and aged 36–50 are most likely to expand their firms, exhibiting employment growth rates that are, respectively, 11.1 and 9.4 percentage points higher than those of their peers.¹⁰ Interestingly, employment growth for female entrepreneurs increases with their number of children, suggesting that, as has been noted in the literature, most employees of microfirms are related to the owner. This is also consistent with the finding that a key determinant of female children entering the labor force is whether their mother had a microenterprise (Cunningham and Maloney 2001). It may also explain why the rising probability of hiring peaks for both genders around the period when their children are old enough to contribute.

Informality among microfirms

In the literature that associates informal microenterprises with the disadvantaged sector of labor markets segmented by government- or union-induced rigidities, workers enter self-employment involuntarily whereas they queue up for salaried jobs. Informality, in this approach, is not a choice, but rather the option of last resort for otherwise-unemployed workers. In contrast, the parallel tradition that views self-employment through the lens of the *firm* stresses the entrepreneurial dynamism of the sector and the fact that

many microentrepreneurs *choose* to be there. In practice, one is likely to find some evidence for both views, suggesting that the self-employment sector is quite heterogeneous, with a relatively well-off entrepreneurial group coexisting with those involuntarily informal, respectively, the “upper tiers” and “lower tiers” of the sector, in Gary Fields's (1990) formulation. However, as chapter 2 suggests, at least in the case of Latin America, the existing evidence suggests that microentrepreneurs who entered the sector voluntarily are most likely the majority.

Not surprisingly, formality rates are higher among better-performing microentrepreneurs, and for those who entered self-employment voluntarily—who arguably constitute the upper tier of the sector. Thus, as seen in the left panel of table 5.6, among formal Brazilian microfirm owners the fraction that was unemployed prior to entering the sector is 21 percent, compared to 32 percent for nonlicensed firm owners. Similarly, licensed businesses are more likely to show signs of success, such as having plans to expand—46 percent among licensed firms compared to 37 percent for nonlicensed businesses—and their owners are less likely to plan on going back to salaried jobs—6.5 versus 13.4 percent, respectively. It must be noted, however, that the positive correlation between good firm performance and formality could be driven by causal effects in either direction—better-performing firms from the upper tier of the microenterprise sector being more likely to be formal, or formality causing improvements in performance. We will examine this issue in more detail in the next chapter.

That informality is higher among microentrepreneurs in the so-called upper tier of the sector is also illustrated by the fact that, as seen in table 5.7, older business owners—at least until age 47 in Mexico and 45 in Brazil—are less likely to be informal, and so are those with higher levels of

TABLE 5.6

Reasons for starting up, firm prospects, and firm licensing in Brazil

Main reason to start a microfirm	Firms with license (%)	Firms without license (%)	Plans for future	Firms with license (%)	Firms without license (%)
Didn't find a job	20.9	32.2	Expand	45.5	36.6
Profitable business	2.2	1.2	Same level	31.2	31.2
Flexible hours	1.6	2.3	Change activity, remain independent	9.2	9.5
Be independent	27.8	17.1	Find a salaried job	6.5	13.4
Family tradition	11.0	8.1	Don't know	7.6	9.3
To help family income	12.2	20.8	Difficulties to regularize when starting up (2003)	% firms (with licenses)	% firms (with licenses)
Accumulated experience	10.7	8.7	Yes	18	5.1
Make good deal	10.7	7.6	No	57.4	10.4
As a secondary job	2.5	2.1	Didn't try	24.7	84.5

Sources: Pooled *Pesquisa Economia Informal Urbana* 1997 and 2003 (except for "Difficulties to Regularize" from 2003).

Note: Sample restricted to entrepreneurs aged at least 20.

TABLE 5.7

Informality among Mexican and Brazilian microenterprises by age, education, and previous activity of the owner

Variable	Mexico			Brazil		
	(1) No license	(2) No taxes	(3) No social security	(4) No license	(5) No taxes	(6) No social security
Owner's age	-0.019*** (0.000)	-0.014*** (0.000)	-0.012*** (0.008)	-0.009*** (0.000)	-0.001** (0.034)	0.0001 (0.979)
Owner's age squared	0.0002*** (0.000)	0.0001*** (0.000)	0.0001 (0.103)	0.0001*** (0.000)	0.0000** (0.042)	-0.000 (0.637)
Primary schooling	-0.186*** (0.000)	-0.102*** (0.000)	-0.089** (0.045)	-0.048*** (0.000)	-0.008*** (0.001)	-0.0415*** (0.005)
Secondary schooling	-0.365*** (0.000)	-0.218*** (0.000)	-0.205*** (0.000)	-0.107*** (0.000)	-0.022*** (0.000)	-0.170*** (0.000)
Tertiary schooling	-0.523*** (0.000)	-0.357*** (0.000)	-0.390*** (0.000)	-0.247*** (0.000)	-0.039*** (0.000)	-0.3379*** (0.000)
Female	0.075*** (0.000)	0.062*** (0.000)	-0.008 (0.647)	0.095*** (0.000)	0.009*** (0.000)	0.011 (0.225)
Previously unemployed	0.035*** (0.001)	0.034*** (0.000)	0.068*** (0.005)	0.039*** (0.000)	0.014*** (0.000)	0.074*** (0.000)
Observations (n)	28,525	28,525	4,599	45,509	45,509	8,841
Pseudo-R ²	0.189	0.159	0.215	0.220	0.159	0.220

Source: Author's calculations, using *Pesquisa Economia Informal Urbana* and *Encuesta Nacional de Micronegocios*.

Note: Robust *p*-values appear in parentheses. Firm size, time in business, fixed location, sales to large firms, sector, and state are also controlled. Samples of business owners aged 21 and older, without complete tertiary education, operating agricultural activities, and with at most five workers.

p* < .05; *p* < .01.

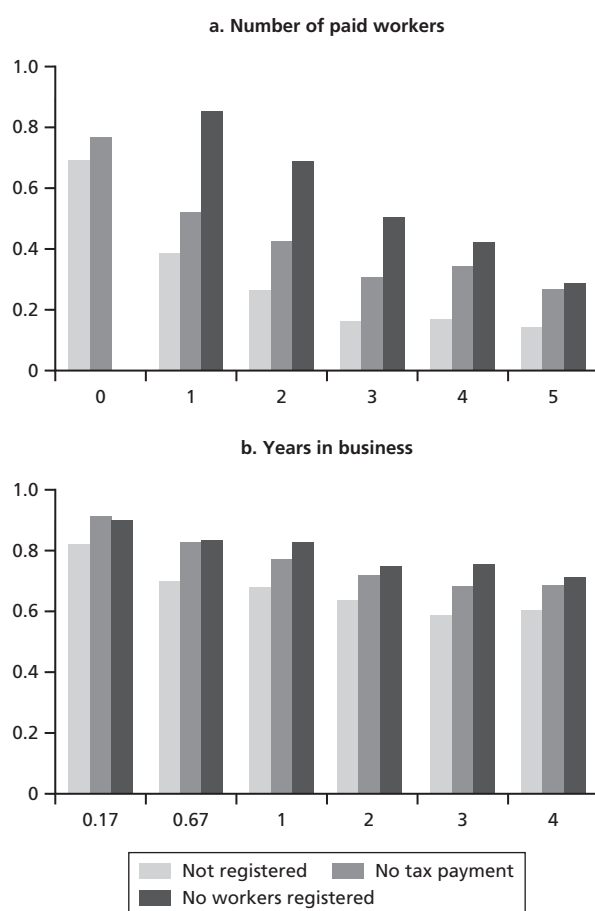
schooling. Thus, for instance, the probability of a microenterprise not being registered is 37 percentage points lower for Mexican businesses owned by individuals with at least some secondary schooling, compared to those who have no schooling at all. The likelihood of not paying taxes or social security contributions also diminishes significantly with the education of the business owner, a result that is also obtained for Brazil. Moreover, after controlling for the effect of age and education, female entrepreneurs and individuals who opened their businesses because they “could not find a job” have a higher probability of being informal. As shown in table 5.7, this is found both for Brazil and Mexico, and for different indicators of informality. This result is consistent with both types of microentrepreneurs—females and previously unemployed workers—viewing their business activity as more temporary than males and voluntary entrants into the sector, respectively.

While most microenterprises are informal, formality rates increase quite rapidly when firms incorporate paid employees. In Mexico, 63 percent of all urban microenterprises are not registered with the federal treasury, and 72 percent pay no taxes. Nevertheless, among those microfirms with at least one paid employee, only 31 percent are unregistered and less than half (46 percent) report paying no taxes. As illustrated in figure 5.7, by the time firms have hired five paid workers the fraction that remains unregistered has fallen to 13 percent and only 28 percent declare paying no taxes. Moreover, while 86 percent of the firms with only one paid employee do not pay social security contributions for that worker, among those with five paid workers, 71 percent report paying social security for at least some of their employees.

Similar evidence emerges from data on Brazilian microenterprises: while 76 percent do not have an operating license and 94 percent do not pay taxes, those rates fall to 51 and 75 percent, respectively, among firms with at least one paid employee. Moreover, among those that employ five paid workers, 67 percent have operating licenses and 23 percent report paying taxes. As shown in table 5.8, the positive effect of firm size on formality persists even when one controls, in a regression framework, for a number of other firm and business owner characteristics. In Colombia, for instance, the probability of a microenterprise with one paid worker not being registered or not paying taxes falls by, respectively, 20 and 15 percentage points in comparison with owner-only enterprises, and further reductions in informality probabilities are found among those with more paid workers.¹¹

FIGURE 5.7

Informality among Mexican microenterprises by number of paid workers and time in business



Source: Author's calculations, using *Encuesta Nacional de Micronegocios*.

Informality also tends to diminish, although less rapidly, with time in business. As suggested by the results in table 5.8, the effect of time in business is stronger for younger firms, but it remains negative—older firms being less likely to be informal—until firms have been in business for 18 years (22 in the case of Brazil). This is illustrated for the case of Mexico in figure 5.7b. Depending on which definition is employed—lack of registration, tax payments, and social security contributions for workers—informality rates drop from between 70 and 90 percent among firms with under a year of existence, to between 60 and 71 percent for those that have been in business for at least four years. Similarly, while only 12 percent of Brazilian microenterprises with at most three years of existence have operating licenses, that fraction is twice as large among older firms. It is worth noting, however, that the

TABLE 5.8

Informality among Mexican, Brazilian, and Colombian microenterprises by paid employment and time in business

Variable	Mexico			Brazil			Colombia		
	(1) No license	(2) No taxes	(3) No social security	(4) No license	(5) No taxes	(6) No social security	(7) No license	(8) No taxes	(9) No social security
One paid worker	−0.322*** (0.000)	−0.255*** (0.000)		−0.122*** (0.000)	−0.031*** (0.000)		−0.202 (0.000)***	−0.146 (0.000)***	−0.937 (0.000)***
Two paid workers	−0.406*** (0.000)	−0.311*** (0.000)	−0.162*** (0.000)	−0.195*** (0.000)	−0.064*** (0.000)	−0.123*** (0.000)	−0.323 (0.000)***	−0.239 (0.000)***	−0.969 (0.000)***
Three paid workers	−0.506*** (0.000)	−0.426*** (0.000)	−0.347*** (0.000)	−0.253*** (0.000)	−0.080*** (0.000)	−0.187*** (0.000)	−0.344 (0.000)***	−0.280 (0.000)***	−0.981 (0.000)***
Four paid workers	−0.528*** (0.000)	−0.434*** (0.000)	−0.469*** (0.000)	−0.273*** (0.000)	−0.108*** (0.000)	−0.260*** (0.000)	−0.349 (0.000)***	−0.283 (0.000)***	−0.981 (0.000)***
Five paid workers	−0.506*** (0.000)	−0.447*** (0.000)	−0.574*** (0.000)	−0.272*** (0.000)	−0.086*** (0.000)	−0.307*** (0.000)	−0.354 (0.000)***	−0.268 (0.000)***	−0.979 (0.000)***
Time in business	−0.011*** (0.000)	−0.013*** (0.000)	−0.009*** (0.000)	−0.009*** (0.000)	−0.001*** (0.000)	−0.008*** (0.000)	−0.297 (0.000)***	−0.266 (0.000)***	−0.017 (0.000)***
Time in business squared	0.0003*** (0.000)	0.0003*** (0.000)	0.0002*** (0.002)	0.0002*** (0.000)	0.0000*** (0.000)	0.000*** (0.000)	0.041 (0.000)***	0.032 (0.000)***	0.002 (0.000)***
Observations (n)	28,525	28,525	4,599	45,509	45,509	8,841	27,177	27,691	27,690
Pseudo- R^2	0.189	0.159	0.215	0.220	0.159	0.220	0.174	0.095	0.665

Sources: Author's calculations using *Encuesta Nacional de Micronegocios* and *Pesquisa Economia Informal Urbana*; for Colombia, calculations performed by Carolina Mejia and Mauricio Cardenas.

Note: Robust p -values appear in parentheses. In the cases of Mexico and Brazil, age, education, and previous unemployment of the owner, as well as fixed location, sales to large firms, sector, and state, are also controlled. Samples of business owners aged 21 and older, without complete tertiary education, operating in nonagricultural activities, and with at most five workers. For Colombia, other controls include having a fixed location and sector. The time-in-business variable is expressed in years for Brazil and Mexico, and is categorical for Colombia.

*** $p < .01$.

decision to operate formally or informally is often made at the time of starting up. As seen in the bottom-right row of table 5.6, almost 85 percent of nonlicensed Brazilian microfirms did not even try to regularize their businesses when they began operating, compared to 75 percent of formal businesses that at least attempted to do so.

With regard to the *sectors of economic activity* where informality is more prevalent, at least in the case of microenterprises, we find that in both Mexico and Brazil informal firms are most frequently found in the construction sector, possibly due to lower firm survival rates in this sector (see previous section), and they are least common in retail trade (table 5.9). In both countries, other

services and manufacturing appear in an intermediary position. Other firm characteristics that are strongly associated with lower informality rates are the fact of operating out of a *fixed location*, and that of selling mainly to large companies. For given firm sizes, having a fixed location is associated with registration probabilities that are up to 25 percentage points higher in the three countries for which we have data on microenterprise formality (table 5.9). Higher probabilities of paying taxes and social security are also found in those three countries among firms operating out of fixed locations. One interpretation of this result is that firms that, because their activities do not need a fixed locale for their operations, could also be

TABLE 5.9

Informality among Mexican, Brazilian, and Colombian microenterprises by age, education, and previous activity of the owner

Variable	Mexico			Brazil			Colombia		
	(1) No license	(2) No taxes	(3) No social security	(4) No license	(5) No taxes	(6) No social security	(7) No license	(8) No taxes	(9) No social security
Fixed location	−0.170*** (0.000)	−0.111*** (0.000)	−0.076*** (0.000)	−0.202*** (0.000)	−0.033*** (0.000)	−0.136*** (0.000)	−0.254 (0.000)***	−0.144 (0.000)***	−0.024 (0.000)***
Large main client	−0.175*** (0.000)	−0.114*** (0.000)	−0.168*** (0.000)	−0.017* (0.082)	−0.036*** (0.000)	−0.048** (0.019)			
Services	0.188*** (0.000)	0.168*** (0.000)	0.054*** (0.000)	−0.021*** (0.000)	0.007*** (0.000)	0.033*** (0.001)	0.196 (0.000)***	0.158 (0.000)***	0.018 (0.000)***
Manufacturing	0.226*** (0.000)	0.176*** (0.000)	0.040** (0.035)	0.107*** (0.000)	0.012*** (0.000)	0.058*** (0.000)	0.170 (0.000)***	0.140 (0.000)***	0.023 (0.000)***
Construction	0.368*** (0.000)	0.271*** (0.000)	0.178*** (0.000)	0.137*** (0.000)	0.013*** (0.000)	0.055*** (0.002)			
Observations (n)	28,525	28,525	4,599	45,509	45,509	8,841	27,177	27,691	27,690
Pseudo- R^2	0.189	0.159	0.215	0.220	0.159	0.220	0.174	0.095	0.665

Sources: Author's calculations using *Encuesta Nacional de Micronegocios* and *Pesquisa Economia Informal Urbana* for, respectively, Mexico and Brazil; for Colombia, calculations performed by Carolina Mejia and Mauricio Cardenas, using data from Colombia National Administrative Department of Statistics.

Note: Robust p -values appear in parentheses. Firm size, time in business, as well as age, education, and previous unemployment of the owner, sector, and state, are also controlled. Samples of business owners aged 21 and older, without complete tertiary education, operating in nonagricultural activities, and with at most five workers. For Colombia, other controls include size and time in business.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

less motivated to formalize themselves because of the lower probability of being caught. Moreover, for firms that could operate either with or without a fixed locale, the desire to remain informal—without being caught—could be the deciding factor for choosing the second option. Note, however, that this would not be without costs, as informal firms that operate on an ambulant basis are also, as shown in the next chapter, less likely to accumulate assets and increase sales and employment.

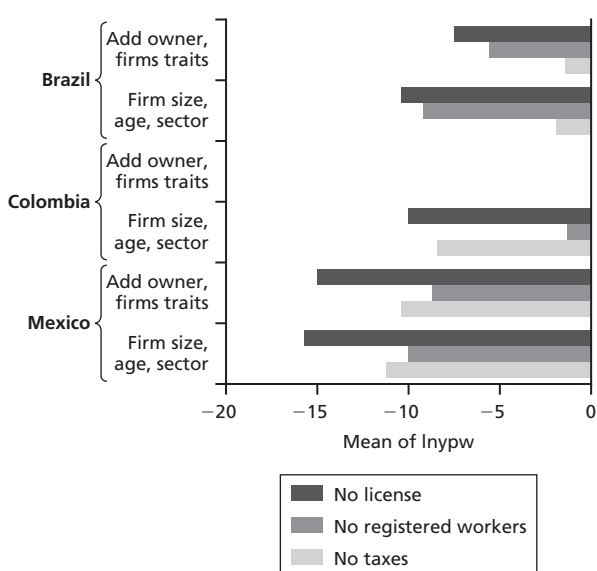
As for firms selling to large companies, being formally registered and paying taxes are likely to be requirements imposed by their clients, which would explain the higher level of formality among microenterprises that cater to large companies. The increase in the probability of being formal for firms having *large companies as clients* is between 11 and 18 percentage points in Mexico. The increase is smaller, while still significant, in Brazil—between 2 and 5 additional percentage points in the probability of being

formal.¹² These results are consistent with the findings of de Paula and Scheinkman (2006), who show that since purchases from informal firms are ineligible for tax credits, value-added tax systems create incentives for the formalization of suppliers located upstream along productive chains—then again, informality also tends to be contagious, as firms that buy from (or sell to) informal firms also have an incentive to operate informally.

Even after controlling for the presence of large clients and a fixed location, as well as for the owners' human capital and other personal characteristics, firms with higher levels of productivity are less likely to be informal. This is illustrated in figure 5.8, which reports the effect of a doubling of labor productivity—equivalent to about one standard deviation in net revenues per worker—on the probability of Mexican and Brazilian microenterprises being informal. The effects are larger in the case of Mexico, where a 100 percent increase in labor productivity is associated

FIGURE 5.8

Effect of doubling labor productivity on probability of being informal (%)



Sources: Author's calculations, using *Encuesta Nacional de Micronegocios* and *Pesquisa Economia Informal Urbana*; for Colombia, calculations performed by Carolina Mejia and Mauricio Cardenas, using data from Colombia National Administrative Department of Statistics.

Note: Coefficients on labor productivity in regressions where other regressors and samples are as described in tables 5.7 and 5.9, respectively, (see notes to those tables).

with a drop of about 15 percentage points in the probability of not being registered, and about 10 and 9 fewer percentage points in the probabilities of not paying taxes and social security contributions, respectively. In Brazil, increases in productivity have a small effect on the probability of paying taxes—between 1 and 2 percentage points higher for a doubling of labor productivity—but they reduce the probability of not having an operating license and not paying social security contributions by as much as 9 and 10 percentage points. As for the evidence on Colombia, we find that a doubling of labor productivity is associated with increases of up to 10 percentage points in formality probabilities.

The apparent negative link between productivity and informality can be interpreted, as mentioned above, as evidence that firms in the upper tier of the microenterprise sector are more likely to become formal. However, as will be discussed in the next chapter, it is also possible that at least to some extent causality also runs in the opposite direction, with access to formality leading to further increases in firm productivity.

Conclusions

This chapter has shown that the patterns of entry, survival, and growth in the self-employment sector in Latin America match well the predictions of mainstream models of firm dynamics developed for industrial-country contexts, suggesting that the behavior of many microfirm owners is driven by *exit* choices and not just by *exclusion* factors. Thus, entrants into self-employment tend to be found more frequently among workers who have accumulated human and physical capital while working as salaried workers. Among the latter, those with relatively higher wages given their human capital—the overachievers—are more likely to become firm owners, and those with more schooling are more likely to open businesses with at least some paid employees. Once in the sector, the owners of more profitable firms are more likely to stay in business and hire more workers. Moreover, as firms grow older and larger, their growth rates tend to diminish and their degree of informality tends to gradually decrease.

Few firms, however, tend to evolve along this ideal path, with the large majority of microfirms remaining owner-only, and a large fraction of them failing. Thus, at least in the short term, the sector is not likely to be an important source of net job creation—if any—for Latin American economies. Two complementary explanations can be proposed to explain the high failure rates and limited growth and job creation of informal firms. On one hand, informality and the factors that may be behind it—to be discussed in more detail in the next chapter—could be to blame, with policy-induced barriers to formalization impeding microfirm access to technologies and markets, which in turn would keep them small and unproductive, and in many cases lead them to exit altogether, thus perpetuating a vicious circle of low growth and high informality. On the other hand, the evidence presented in this chapter is also consistent with an alternative explanation for the type of microfirm dynamics observed in Latin America. In particular, the presence of low opportunity costs for entry into the sector would also lead to a predominance of low-productivity businesses with low growth prospects and high failure rates. In this context, in order to reduce informality, policy makers should focus not only on altering the direct costs and benefits of formality but also on the drivers of formal sector productivity, including measures to improve the investment climate and policies aimed at increasing human capital accumulation.

Notes

1. This section draws heavily on Fajnzylber, Maloney, and Rojas (2006b). Only 6 of the 53 papers mentioned by Blanchflower (2004) in his self-employment literature review concern developing countries, and they focus mostly on the determinants of earnings. To our knowledge, the only previous evidence on the determinants of entry, exit, and growth of microenterprises in developing countries are the papers on Africa by Goedhuys and Sleuwaegen (2000), Liedholm (2002), Liedholm and Mead (1999), McPherson (1995, 1996), and Mead and Liedholm (1998). Other recent studies on firm dynamics in developing countries have focused mostly on larger and/or formal firms, including Aw, Chung, and Roberts (2003), Bartelsman, Haltiwanger, and Scarpetta (2004), and Roberts and Tybout (1997).

2. Other dynamic models also generate these patterns, although with different analytical structures. Ericson and Pakes (1992a, 1995) propose a model of active exploration—as opposed to the passive learning assumption in Jovanovic's model—that incorporates firm-specific sources of uncertainty derived from stochastic outcomes of investments made by firms in order to improve their profitability. Favorable outcomes from the firms' own investments—including those that lead to entry into the industry—tend to move them toward “better” states, while good outcomes of direct competitors move them to less profitable states. As in Jovanovic's model, entry, exit, and investment decisions are made to maximize the expected discounted value of future net cash flows conditional on the current information set.

3. The evidence on Mexico is drawn mostly from Fajnzylber, Maloney, and Rojas (2006b).

4. The above-mentioned surveys are *Encuesta Nacional de Empleo Urbano/Encuesta Nacional de Microgenocios* (ENEU/ENAMIN) for Mexico, the Nicaragua Living Standards Measurement Study (LSMS), and the Colombia Manufacturing Survey.

5. While this could suggest, as noted by Hopenhayn, that at the margin the process of selection of entrepreneurs is poorer in Mexico, one must also bear in mind that the higher rate of entry from salaried work found in Mexico is partly a result of the relatively smaller size of the wage sector in this country.

6. For comparison purposes with Evans and Leighton's figures for the United States, the rates of entry into self-employment reported for Mexico in figure 5.1 are calculated as the fraction of the number of salaried workers who enter self-employment in the course of one year. Similarly, the self-employment rate is calculated by dividing the number of self-employed in a given age range by the sum of salaried and self-employed individuals of that age. Finally, exit rates are the fraction of self-employed workers who move to wage work.

7. Note that in figure 5.2 entry rates are calculated for individuals coming either from salaried work or from out of the labor force, which is a much more common position for women than for men. Moreover, self-employment rates are also calculated as a proportion of the whole population in the same age range. Finally, exit rates in figure 5.2 cover all self-employed women who move either out of the labor force or to other labor market positions.

8. See Fajnzylber, Maloney, and Rojas (2006b) for details on the sample and estimation methodology.

9. Firm size is divided into owner-only firms, 2–5 employees, 6–10 employees, 11–15 employees, 16–50 employees, 51–100 employees, 101–250 employees, and 250 and more employees. The dependent variable is the imputed percentage difference between the mean value of employment in each firm size bracket. By definition, firms that remained in the same size interval have a value of zero.

10. While the positive sign encountered in the marital status variable may well reflect the use of nonpaid family workers, the results are virtually unchanged when the firms with a majority of nonpaid workers are kept in the sample.

11. The estimations with Colombian data were generously shared by Mauricio Cardenas and Carolina Mejia.

12. The quantitatively smaller effect obtained for Brazil could reflect better tax enforcement among smaller firms in Brazil, which would then not be significantly affected by their links to larger ones. Alternatively, if large companies also exhibit high tax evasion, then they would not necessarily impose strict tax compliance requirements to their smaller providers. The second hypothesis is favored by the high level of tax evasion revealed by the 2003 Brazil Investment Climate Assessment survey, in which manufacturing firms reported that one-third of the sales of the average enterprise goes unreported for tax purposes, compared to 23 percent for Colombia and 30 percent for Mexico—the averages for LAC and the OECD are, respectively, 23 and 6.5 percent.

References

- Arnett, R., and J. E. Stiglitz. 1991. “Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?” *American Economic Review* 81: 179–90.
- Aw, B. Y., S. Chung, and M. J. Roberts. 2003. “Productivity, Output, and Failure: A Comparison of Taiwanese and Korean Manufacturers.” *Economic Journal* 113 (491): F485–F510.
- Bartelsman, E., J. Haltiwanger, and S. Scarpetta. 2004. “Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries.” Policy Research Working Paper 3464, World Bank, Washington, DC.
- Bartelsman, E., S. Scarpetta, and F. Schivardi. 2003. “Comparative Analysis of Firm Demographics and Survival: Micro Level Evidence for the OECD Countries.” Economics Department Working Paper 348, Organisation for Economic Co-operation and Development, Paris.
- Bates, T. 1990. “Entrepreneur Human Capital Inputs and Small Business Longevity.” *Review of Economics and Statistics* 72 (4): 551–59.
- Besley, T. 1995. “Savings, Credit, and Insurance.” In *Handbook of Development Economics, Volume II*, ed. J. Behrman and T. N. Srinivasan. Amsterdam: Elsevier.
- Bhattacharaya, P. C. 2002. “Rural-to-Urban Migration in LDCs: A Test of Two Rival Models.” *Journal of International Development* 14: 951–72.
- Blanchflower, D. G. 2000. “Self-employment in OECD Countries.” *Labour Economics* 7: 471–505.
- . 2004. “Self-employment: More May Not Be Better.” *Swedish Economic Policy Review* 11 (2): 15–74.

- Blanchflower, D. G., and A. J. Oswald. 1998a. "Entrepreneurship and the Youth Labour Market Problem: A Report for the OECD." Organisation for Economic Co-operation and Development, Paris.
- . 1998b. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16 (1): 26–60.
- Carrasco, R. 1999. "Transitions to and from Self-employment in Spain: An Empirical Analysis." *Oxford Bulletin of Economics and Statistics* 61: 315–41.
- Cressy, R. 1996. "Are Business Startups Debt-Rationed?" *Economic Journal* 106: 1253–70.
- Cunningham, W. V., and W. F. Maloney. 2001. "Heterogeneity among Mexico's Microenterprises: An Application of Factor and Cluster Analysis." *Economic Development and Cultural Change* 50 (1): 131–56.
- Davis, S. J., J. Haltiwanger, and S. Schuh. 1996. *Job Creation and Destruction*. Cambridge, MA: MIT Press.
- de Paula, A., and J. Scheinkman. 2006. "The Informal Sector." Photocopy. University of Pennsylvania, Philadelphia.
- de Soto, H. 1989. *The Other Path: The Invisible Revolution in the Third World*. New York: Basic Books.
- Dunne, T., M. J. Roberts, and L. Samuelson. 1988. "Patterns of Firms Entry and Exit in U.S. Manufacturing Industries." *RAND Journal of Economics* 19 (4): 495–515.
- . 1989. "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics* 104 (4): 671–98.
- Ericson, R., and A. Pakes. 1992a. "An Alternative Theory of Firm and Industry Dynamics." Cowles Foundation Paper 1041, Yale University, New Haven, CT.
- . 1992b. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62 (1): 53–82.
- . 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62 (1): 53–82.
- Evans, D. S. 1987a. "The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries." *Journal of Industrial Economics* 35 (4): 567–81.
- . 1987b. "Tests of Alternative Theories of Firm Growth." *Journal of Political Economy* 95 (4): 657–74.
- Evans, D. S., and B. Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97 (4): 808–27.
- Evans, D. S., and L. S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79 (3): 519–35.
- Fajnzylber, P., W. F. Maloney, and G. V. Montes Rojas. 2006a. "Does Formality Improve Microfirm Performance? Quasi-experimental Evidence from the Brazilian SIMPLES Program." Photocopy. World Bank, Washington, DC.
- . 2006b. "Microenterprise Dynamics in Developing Countries: How Similar Are They to Those in the Industrialized World? Evidence from Mexico." *World Bank Economic Review* 20 (3): 389–419.
- . 2006c. "Releasing Constraints to Growth or Pushing on a String? The Impact of Credit, Training, Business Associations and Taxes on the Performance of Mexican Microfirms." Policy Research Working Paper 3807, World Bank, Washington, DC.
- Fields, G. S. 1990. "Labor Market Modelling and the Urban Informal Sector: Theory and Evidence." In *The Informal Sector Revisited*, ed. D. Turnham, B. Salome, and A. Schwartz. Paris: Organisation for Economic Co-operation and Development.
- Geroski, P. A. 1991. *Market Dynamics and Entry*. Oxford, UK: Blackwell.
- Goedhuys, M., and L. Sleuwaegen. 2000. "Entrepreneurship and Growth of Entrepreneurial Firms in Côte d'Ivoire." *Journal of Development Studies* 36 (3): 123–46.
- González de la Rocha, M. 1994. *The Resources of Poverty: Women and Survival in a Mexican City*. Oxford, UK: Blackwell.
- Hamilton, B. H. 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-employment." *Journal of Political Economy* 108 (3): 604–31.
- Hart, K. 1972. *Employment, Income, and Inequality: A Strategy for Increasing Productive Employment in Kenya*. Geneva: International Labour Organization.
- . 1973. "Informal Income Opportunities and Urban Employment in Ghana." *Journal of Modern African Studies* 11: 61–89.
- Holtz-Eakin, D., D. Joulfaian, and H. S. Rosen. 1994a. "Entrepreneurial Decisions and Liquidity Constraints." *RAND Journal of Economics* vol. 25 (2): 334–47, summer.
- . 1994b. "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints." *Journal of Political Economy* 102 (1): 53–75.
- Hopenhayn, H. A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5): 1127–50.
- Johnson, W. R. 1978. "A Theory of Job Shopping." *Quarterly Journal of Economics* 92: 261–78.
- Jovanovic, B. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87: 972–90.
- . 1982. "Selection and Evolution of Industry." *Econometrica* 50 (3): 649–70.
- Levenson, A., and W. Maloney. 1998. "The Informal Sector, Firm Dynamics, and Institutional Participation." Policy Research Working Paper 1988, World Bank, Washington, DC.
- Liedholm, C. 2002. "Small Firm Dynamics: Evidence from Africa and Latin America." *Small Business Economics* 8 (1–3): 227–42.
- Liedholm, C., and D. C. Mead. 1999. *Small Enterprises and Economic Development: The Dynamics of Micro and Small Enterprises*. London: Routledge Press.
- Lippman, S. A., and R. P. Rumelt. 1982. "Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency under Competition." *Bell Journal of Economics* 13 (2): 418–38.
- Loayza, N. 1996. "The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129–62.
- Lucas, R. E., Jr. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (2): 508–23.
- Maloney, W. F. 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico." *World Bank Economic Review* 13: 275–302.
- . 2001. "Self-employment and Labor Turnover in Developing Countries: Cross-country Evidence." In *World Bank Economists'*

- Forum*, ed. S. Devarajan, F. H. Rogers, and L. Squire. Washington, DC: World Bank.
- . 2004. "Informality Revisited." *World Development* 32 (7): 1159–78.
- McKenzie, D. J., and C. Woodruff. 2006. "Do Entry Costs Provide an Empirical Basis for Poverty Traps?" *Economic Development and Cultural Change* 55 (1): 3–42.
- McPherson, M. A. 1995. "The Hazards of Small Firms in Southern Africa." *Journal of Development Studies* 32 (1): 31–54.
- . 1996. "Growth of Micro and Small Enterprises in Southern Africa." *Journal of Development Economics* 48 (2): 253–77.
- Mead, D. M., and C. Liedholm. 1998. "The Dynamics of Micro and Small Enterprises in Developing Countries." *World Development* 26 (1): 61–74.
- Pakes, A., and R. Ericson. 1998. "Empirical Implications of Alternative Models of Firm Dynamics." *Journal of Economic Theory* 79 (1): 1–45.
- Rees, H., and A. Shah. 1986. "An Empirical Analysis of Self-employment in the UK." *Journal of Applied Econometrics* 1: 95–108.
- Roberts, M. J., and J. R. Tybout, eds. 1997. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Washington, DC: World Bank.
- Tybout, J. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature* 38 (1): 11–44.
- Wagner, J. 1994. "The Post-entry Performance of New Small Firms in German Manufacturing Industries." *Journal of Industrial Economics* 42 (2): 141–54.

CHAPTER 6

Informality, Productivity, and the Firm

SUMMARY: This chapter focuses on both the determinants and the consequences of informality from the perspective of private firms. The chapter begins by arguing that firm informality encompasses not only the large number of unregistered microfirms described in chapter 5, but also those medium-size and large firms that fail to comply with at least some government regulations. This is illustrated by recent estimates of a large incidence of tax and social security evasion among Latin American firms of all sizes. The chapter then reviews the evidence on the factors that may lead private firms to exhibit different degrees of informality, and analyzes the channels through which regulatory noncompliance can affect individual and aggregate productivity. Finally, the chapter lays out the possible approaches that policy makers could adopt to tackle the issue of firm informality.

WHILE INFORMALITY IS OFTEN associated with small, unregistered microfirms, tax and social security evasion is commonplace among larger Latin American firms. This chapter reviews the main firm characteristics associated with informality among those firms, using data from recently concluded enterprise surveys in a number of Latin American countries. Not surprisingly, we find that the incidence of tax and social security evasion varies considerably across and within countries, and it is generally higher for small, low-productivity firms that started their operations without a formal registration.

With regard to the factors that may help explain this variation, we find that informality tends to increase with the quantity of labor and product markets regulations, and to decrease with the quality of governance—for example, the prevalence of the rule of law and the level of democratic accountability. Moreover, firm informality is positively related to the incidence of corruption, but it can be curtailed by the improvement of market support institutions—the courts, financial markets, and the development of links

between small and large firms. Finally, we show that recent initiatives to reduce red tape and introduce simplified tax and registration systems for microfirms have led to statistically significant increases in the number of new formally registered firms, although further research is required to establish whether the corresponding effects are permanent or temporary, and to resolve existing controversies on their magnitude. In particular, some recent estimates suggest that the simplification of entry regulation leads more former high-ability wage workers to open formal businesses, but it has small or no effects on the formalization of unregistered businesses.

We argue that there are reasons to expect important overall productivity gains if a larger fraction of firms would formalize and if resources could be shifted away from low-productivity informal firms toward the formal sector. This could be the case in a context where increasing returns to scale are relevant, at least among very small firms, and informality is accompanied by a preponderance of small firms. Moreover, unfair competition from informal firms could slow down the process of creative destruction by which inefficient firms are replaced by their more efficient

competitors, and negatively affect the incentives of formal firms to innovate and adopt new technologies. Finally, growth could increase as previously informal firms gain increased access to markets and services. These positive links between formality and growth are supported by firm-level evidence presented in this chapter. Indeed, we find that firms that started operations without formally registering—at least initially—and those located in regions or sectors where informality is more prevalent exhibit, on average, much lower productivity than their peers. Moreover, exogenous increases in formality levels caused by changes in the costs of microfirm formalization are also found to lead to improved firm performance.

In practice, however, policy makers may need to act on several fronts at the same time in order to tilt the cost–benefit analysis of firms toward formality by combining both positive and negative incentives—respectively, “carrots” and “sticks.” Thus, the impact of interventions aimed at reducing the costs of being formal through the removal of regulatory constraints may be larger when accompanied by measures to enhance evenhanded enforcement of regulations and increase the potential benefits of joining the formal sector—for example, through improvements in access to credit, contract enforcement mechanisms, and technical assistance. In particular, recent evidence from randomized experiments suggests relatively high returns to capital among very small Mexican microenterprises, which implies that considerable increases in income could be obtained through incentives to the formalization of small businesses and expanded access to microcredit. These efforts, however, may have larger effects on the subset of higher-productivity informal firms (or high-ability wage earners considering entry into the microfirm sector) that may be closer to the margin separating the formal from the informal sector—for example, because they have more to win from formalizing or because the opportunity costs of operating informally are higher for them. Moreover, for very-low-productivity informal firms it is possible that, as argued in the previous chapter, a larger impact on overall formality levels could be achieved in the medium-to-long term through actions aimed at increasing job opportunities in the formal sector.

Informality among registered firms

The informal sector is often associated with the large number of unregistered small businesses found in most urban centers of the developing world. Thus, what people often have in mind when thinking about the informal sector are small,

unregulated firms that avoid most taxes and labor regulations, and do not comply with most government regulations. Informality, in this approach, is seen as affecting mostly the very low end of the firm-size spectrum.

In a more general approach, as argued in chapter 1 of this volume, it is reasonable to define the informal sector as encompassing all the firms that, at least to some extent, choose to operate outside of the scope of existing regulations. Thus, medium-size and large firms can be considered informal even if they are appropriately registered, provided that, for instance, they underreport their sales for tax purposes; do not register all their workers with the social security administration; or do not comply with some government regulations regarding mandatory operating licenses or permits, as well as product quality and safety regulations.

For analytical purposes, however, it is useful to divide the informal sector into two different subsectors. The first is the informal microenterprise sector, which we described in the previous chapter and is composed of mostly informal, unregistered microenterprises. The second informal subsector, which Djankov et al. (2003) denominate the “unofficial economy,” is made up of firms that are only partially informal, in the sense that they are formally registered but keep a fraction of their workers and sales hidden from government regulators, and/or fail to comply with at least some government regulations related, for instance, to mandatory permits and licenses. While unofficial firms tend to be mostly small, in some countries and regions this subsector can include medium-size and even large companies.¹

Data recently collected through surveys of registered firms in seven Latin American and Caribbean countries confirm that many small, medium-size, and even large firms also exhibit some degrees of informality. This is illustrated in figure 6.1, which shows that sales and employment underreporting (for tax purposes) is commonplace among registered medium-size and large firms in selected Latin American and Caribbean countries of different sizes and levels of income.²

The high variation in levels of tax and social security evasion suggested by this figure, across countries with relatively similar levels of income, implies that cross-country differences in informality are not driven just by levels of economic development. In Brazil and Panama, for instance, firms reportedly evade as much as 30 to 40 percent of their taxes, compared to less than 20 percent in Uruguay and Peru, and less than 5 percent in Chile.³ As discussed in the next section, possible explanations may be associated with the size of tax and social security burdens, labor legislation,

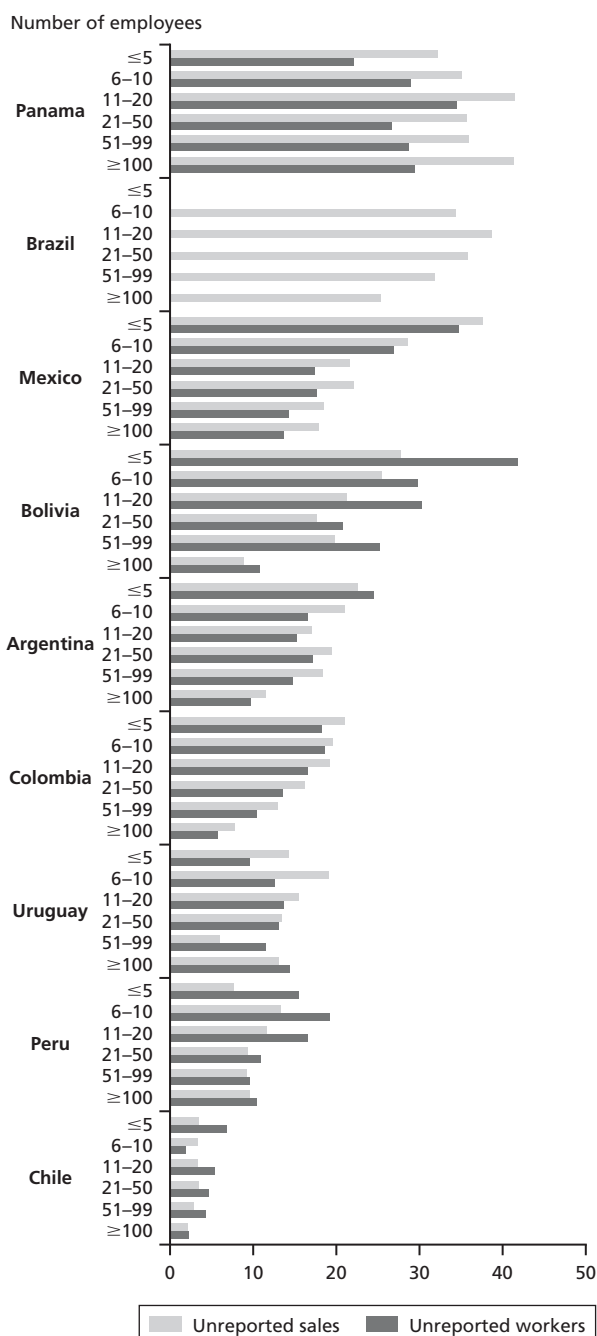
governance issues related to red tape and corruption, levels of regulatory enforcement, and the value attributed by firms to market- and government-provided services.

A second finding derived from figure 6.1 is that, for most countries, the data suggest, as expected, that tax and social

security evasion is more prevalent among smaller firms. In Bolivia and Mexico, for instance, about 35 percent of sales go underreported among microenterprises, compared to around 10 to 15 percent among firms with 100 workers or more. However, in Panama, Peru, and Uruguay, the data do not suggest a clear pattern linking underreporting rates to firm size. Thus, large firms appear to evade taxes and social security contributions at rates that are comparable to those of smaller enterprises—between 10 and 15 percent in Peru and Uruguay, and above 30 percent in Panama.

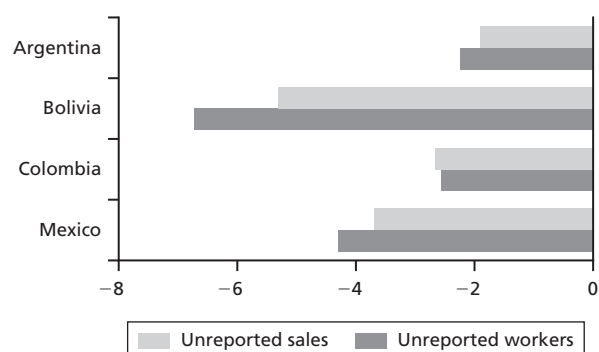
The negative relationship between underreporting rates and employment size in the case of Argentina, Bolivia, Colombia, and Mexico is maintained when the firms' time in business, location, and sector are controlled for in a regression framework (figure 6.2). Thus, hypothetically doubling the size of a firm leads to a reduction in underreporting rates of more than 5 percentage points in Bolivia and 4 in Mexico, with roughly one-half of those effects obtained for the cases of Colombia and Argentina, respectively. A small effect is also found for Uruguay, where the underreporting of sales for tax purposes diminishes by 1.5 percentage points as firms hypothetically double in size. The time that a firm has been in business, however, is not found to be significantly related to informality levels, with the only exceptions being Colombia (where for each additional decade since starting up firms appear to reduce their sales underreporting by about 1 percentage point, and Mexico (where the opposite effect is obtained—increasing underreporting with time in business, at a rate of 1 percentage point for each additional decade since starting up).

FIGURE 6.1

Unreported sales and workers, by firm size (%)

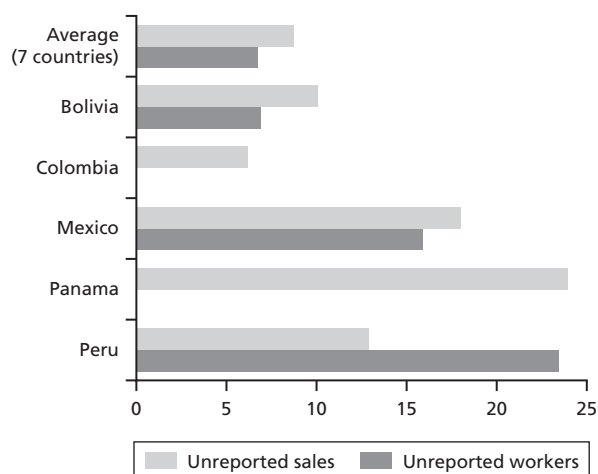
Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

FIGURE 6.2

Effect of doubling employment size on underreporting rates (%)

Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

FIGURE 6.3

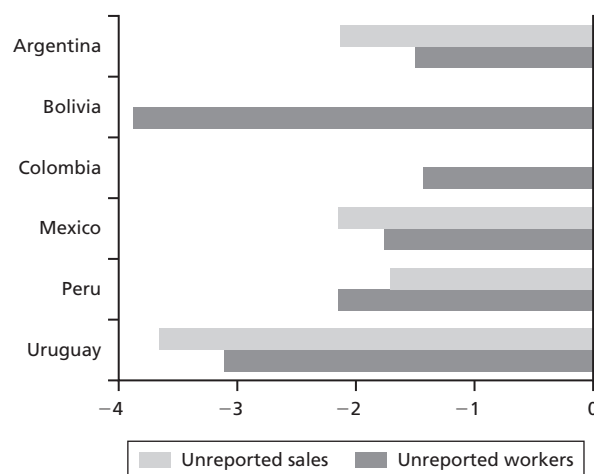
Effect of having started informally on underreporting rates (%)

Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

Another relevant finding common to at least five countries is that firms that start operating in the informal sector exhibit higher rates of sales and employment underreporting than other firms of the same size, sector, and location, even many years after having registered, and after controlling for sector and location characteristics.⁴ Indeed, even after controlling for the time that firms have been in business—on average, more than 20 years, and about 8 years after having registered—starting without a formal registration is associated with rates of tax and social security evasion that are between 6 and 25 percentage points higher than for firms that registered at the time of starting up (figure 6.3). This suggests that efforts to facilitate early registration of new firms have the potential to reduce informality both through increasing the number of tax-paying firms and through reducing tax evasion levels among registered firms.

Within given sectors, however, and for given firm size and time in business, we find that firms with higher levels of *labor productivity* exhibit, in general, lower rates of tax and social security evasion. This is illustrated in figure 6.4, where we report the estimated effects on the rates of sales and employment underreporting of a hypothetical doubling in output per worker (controlling for firm size, time in business, location, and sector of activity). The effects are not significant for Panama, nor for sales underreporting in Bolivia and Colombia. However, for all the other cases, we find that doubling firms' labor productivity—a change of

FIGURE 6.4

Effect of doubling labor productivity on underreporting rates (%)

Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

about one standard deviation in that variable—is associated with underreporting rates that are, on average, 2 percentage points lower.

As suggested above, the negative correlation between productivity and informality is subject to different interpretations. On one hand, more productive firms may arguably have more to lose from operating irregularly (a topic that we discuss in the next section while summarizing the main firm-level determinants of informality). On the other hand, however, it is also possible that productivity is affected by whether firms operate formally or informally, as well as by the general level of informality prevailing in their region and sector (a topic that we cover in the third section of this chapter).

Firm-level determinants of informality

Why do some firms comply with government regulations while others opt for going underground? It is reasonable to assume that private firms voluntarily chose to operate in the formal or the informal sector based on rational profit-maximizing calculations, not unlike those underlying investment and production decisions. In particular, the extent to which firms comply with government regulations is likely to depend on their weighing of the various costs and benefits associated with operating formally or informally. Some of the main factors that firms are likely to take into account are the nature of the regulatory framework,

the extent to which regulations are enforced, and the various opportunity costs associated with operating in the underground economy.

Benefits and costs of informality for private firms

Among the main advantages that firms may consider when opting for informality are the possibility of reducing or eliminating tax payments and social security contributions, and the possibility of avoiding costly and burdensome government regulations, including but not restricted to those affecting labor markets.⁵ Some of the main “benefits” from informality are thus directly linked to the value of taxes and social security contributions that irregular firms are able to avoid.

Other indirect “benefits” are related to cost savings derived from avoiding the often complex administrative procedures associated with tax and regulatory compliance, and to the added flexibility enjoyed by informal firms in their employment and production decisions. Thus, for instance, informal firms arguably enjoy lower hiring and firing costs, and they have more degrees of freedom when setting wages and work hours. Moreover, informal firms may be able to reduce their costs—and potentially increase their sales—as a result of not having to comply with government-imposed standards for products and production processes. And, last but not least, informal firms may be able to circumvent the red tape associated with obtaining government permits and licenses. As documented by the World Bank’s Doing Business reports, many of those procedures are often costly and time-consuming, which may lead some firms to opt for operating informally in order to avoid them.

Potentially countervailing the above cost savings, informal firms need to deal with the risk of being caught and closed down. Since better enforcement of regulations increases the expected value of the fines and other losses derived from being detected by regulators, it reduces the incentives for operating informally. However, since it is neither feasible nor efficient for governments to supervise all individual firms, enforcement tends to be concentrated on larger firms. As a result, informality has the effect of limiting firm growth, both because smaller informal firms are less likely to be caught by government inspectors and because the uncertainty associated with informality discourages investments in illiquid assets. Moreover, in the particular case of informality with respect to labor regulations, irregular firms are likely to have a harder time attracting more educated workers and engaging them in a

longer-term relationship—and that in turn affects their incentives to invest in training and capital goods.

High levels of corruption may also play a role in shaping firms’ incentives for operating informally. Indeed, when caught by government inspectors, irregular firms may have the option of bribing those inspectors to evade fines and other hassles, which may reduce the risk associated with operating informally. Second, in countries where formal firms face a high risk of being extorted by corrupt officials, entrepreneurs may decide to operate informally exactly to reduce their vulnerability to extortion. In fact, there is evidence that this is the case in several so-called transition economies where one of the main motivations for firms’ going underground is to “dodge the grabbing hand” (see Friedman et al. 2000; Johnson et al. 2000; and Johnson, Kaufmann, and Shleifer 1997).

Besides the risk of being caught, informality entails other private costs. Thus, operating in the underground economy eliminates—or at least greatly reduces—access to the courts and other formal contract enforcement mechanisms. This may increase the vulnerability of informal firms in their transactions with other private parties as well as with government. As a result, they may be forced to restrict their transactions to the potentially limited set of trading partners that are deemed trustworthy. This has negative implications not only in terms of social welfare, as it leads to forgoing potential gains from increased trade, but also in terms of reinforcing the above-mentioned factors that tend to limit the expansion of informal firms.

Another important cost associated with informality is given by the narrower set of formal financing mechanisms available to informal firms. Indeed, bank and other formal financial institutions are generally not willing to grant credit to companies that lack proper documentation, including that related to government registration and licensing, tax compliance certificates, and audited financial statements, all of which are generally lacking in the case of informal firms. Moreover, if to evade taxes companies do not register all assets as belonging to the company, their ability to use them as collateral for bank loans may also be limited. Similarly, in the case of firms that hide a fraction of their revenues to elude taxes or other regulations, financial statements may misrepresent their financial soundness and economic prospects, thus reducing their attractiveness to prospective lenders. The same applies to prospective investors, which reduces informal firms’ ability to raise equity capital.

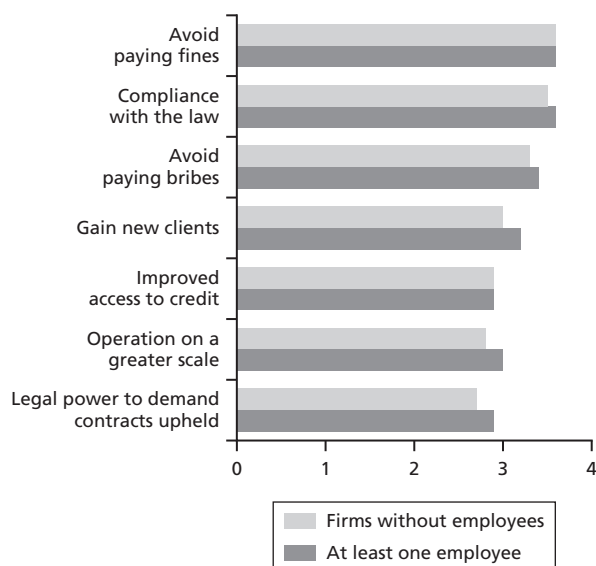
Informality can also forbid firms from benefiting from government support programs targeted at small and medium enterprises (SMEs), as those programs are often restricted to registered tax-paying firms. This can be a serious obstacle to the growth of informal firms, at least if one assumes that SME support programs effectively compensate for market failures that limit the ability of those firms to access formal credit markets and acquire the technologies and human capital needed for their expansion. However, if firms do not place a large value on participating in SME-supporting government programs, or if market support institutions in general are not well developed—for example, credit markets do not function well or contract enforcement mechanisms are slow and costly—then the cost of being excluded from those institutions as a result of operating informally is lower, and a larger share of output is likely to be found in the underground economy.

Evidence on the relative importance of the various costs and benefits of informality—or, equivalently, the main advantages and disadvantages from formalization—is presented in figures 6.5 and 6.6, based on information provided by firms surveyed by the International Finance Corporation in 65 municipalities in Bolivia, Brazil, Honduras, Nicaragua and Peru.⁶ The figures report the average degree of importance of various factors on a scale from 0 to 4,

separating firms with and without employees (beyond the owner). Among the main advantages of formality, the surveyed firms mention the avoidance of fines and bribes, followed by the possibility of gaining new clients and expanding operations—to which, interestingly, firms with at least some employees give more importance than owner-only firms. As for the main disadvantages associated with formalization, the factors to which the enterprises interviewed attribute more importance are the need to renew their licenses every year and the tax obligations resulting from formality—the second one being more important for firms with employees.

Also cited as important advantages of formality, although less frequently than the above-mentioned factors, are the possibility of improving access to credit and using contract enforcement mechanisms—the second one being more important for larger firms. However, when asked about the most important advantage of formality, only 14 percent of the surveyed firms say that they are motivated by the desire to expand or seek new clients, 8 percent mention access to credit as the top reason for formalizing, and just 1.5 percent mention access to the courts. In contrast, 47 percent of the firms say that the top reason underlying their decision to register their enterprise is “to comply with the law” and 29 percent want to avoid fines or bribes. Thus, it appears that the risk of being caught prevails over positive incentives associated with access to markets and services, a finding that, in general terms, applies both to

FIGURE 6.5

Advantages of formalization reported by IFC-surveyed firms

Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

FIGURE 6.6

Disadvantages of formalization reported by IFC-surveyed firms

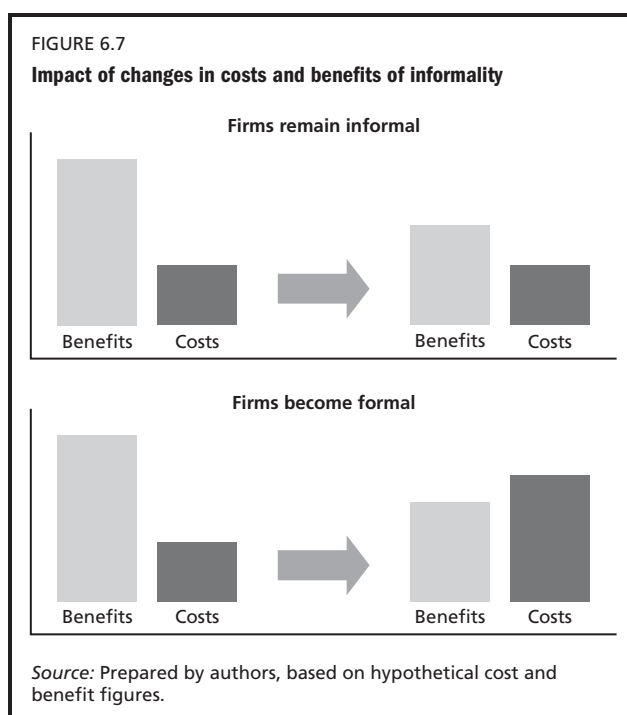
Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

owner-only and to larger firms—the former being slightly more afraid of fines, the latter being more concerned about bribes, and both giving limited importance to issues related to access to credits and the courts.

The relative importance of the various private costs and benefits associated with informality is likely to vary with a number of firm characteristics, including size, time in business, and productivity. Thus, for instance, compared to larger firms, microenterprises may face a smaller risk of being caught by government inspectors when operating irregularly. However, they are likely to have a harder time amortizing the fixed costs associated with regulatory compliance—for example, the costs of firm registration, permits, and licenses—that may be a non-issue for large firms. Moreover, the red tape and monetary costs associated with formalization probably have a larger effect on recently created firms, which may choose to avoid them until they have accumulated sufficient evidence regarding their actual profitability and likelihood of staying in business.

Similarly, firms with inherently low productivity and/or growth prospects are likely to have a lower demand for credit and business development services, as well as for contract enforcement mechanisms, thus being less affected in their informality decisions by the level of development of market support institutions—the courts, financial markets, and the like. This would explain, for instance, why firms belonging to the upper tier of the microenterprise sector are more likely to operate formally, or why registered firms with relatively low levels of productivity are more likely to report higher rates of tax and social security evasion. As formalized by Rauch (1991), to the extent that taxes and regulations are enforced mainly among large, formal sector firms, entrepreneurs with higher managerial ability are endogenously allocated to big firms where they are able to compensate for the corresponding higher regulatory costs. In contrast, smaller and informal businesses are more likely to be run by less-talented entrepreneurs, which would be compensated by the lower costs of informal operations.

It is worth noting that, from a policy point of view, programs aimed at reducing informality by means of affecting its private benefits—for instance, by reducing barriers to entry into the formal sector—may prove ineffective if the costs of operating informally are too low—for instance, because regulatory enforcement and the odds of getting caught are low or because firms place little value on market- or government-provided services available to formal firms. This point is illustrated in the upper panel of figure 6.7,



where a reduction in the benefits from informality proves insufficient to induce firms to formalize because the costs of informality—or, equivalently, the advantages of operating formally—are too low. In those cases, as illustrated in the lower panel of figure 6.7, for policy reforms to have binding effects on firm behavior, efforts may need to be directed both at reducing the benefits and at increasing the costs of operating informally.⁷

As an example, establishing the existence of very high registration costs does not, in itself, imply either that this is why firms don't register or that not registering is a fundamental determinant of average small firm performance. De Soto's (1989) telling anecdotes—for example, the sidewalk vendor who wishes to pay his/her taxes as a way of securing quasi-property rights to his/her pitch—do suggest that the high costs of formalization may impede informal firms from enforcing their property rights and accessing public services, and may limit their access to markets, thus negatively affecting their performance. However, one must bear in mind that registration costs are only one of the factors that informal firms are likely to consider when assessing whether to enter the formal sector. And, depending on the importance of other costs and benefits associated with formality, registration costs may not be the binding constraint for most small informal firms, perhaps because enforcement is limited in any case.

This view is supported by recent evidence on Mexico reported by McKenzie and Woodruff (2006). Using a survey of informal microfirms, they show that the vast majority of them give as the principal reason for not being registered not that registration is too expensive or time-consuming (respectively, 2 and 8 percent of surveyed firms), nor that the costs of operating as registered businesses are too high (4 percent of firms), but that they are too small to make it worth their while (75 percent).

A somewhat different picture emerges, however, from household survey data from Argentina, where most unregistered microfirm owners report that they have remained informal because registration is too expensive (57 percent) or complicated (4 percent), compared to 39 percent who state that they do not register because it would be useless. Data from the Dominican Republic suggest a breakdown similar to Mexico's: only 8 percent of informal microfirm owners report that they remain informal to avoid taxes and social security costs, while 20 percent want to save the time and money involved in formalization. In contrast, almost two-thirds indicate that they are too small to make formality worthwhile, they don't need to be formal, or they don't register because no business like theirs does. Interestingly, 92 percent state that their businesses have not suffered as a result of being informal, and 38 percent indicate that they have actually benefited from their informal status—mainly through lower taxes and the avoidance of government regulations. Similarly, a survey of Guatemalan informal entrepreneurs, performed by the National Economic Research Center, indicates that the majority of them do not perceive any concrete benefit from complying with government regulations (CIEN 2006).

The above evidence leaves open the hypothesis that, depending on each specific country context, registration costs may or may not be binding for most small firms and, in some cases, they could at best be a marginal contributor to informality. If this is indeed the case, further incentives may be needed to entice small firms to enter the formal sector, including both positive and negative incentives—the carrots and sticks mentioned above. While the latter involves increased government enforcement of regulations and potentially higher penalties for evaders, positive incentives range from tax reductions to changes in labor market regulations and improvements in private and public services available to formal firms (for example, credit, contract enforcement, technical assistance, and so forth).

It is worth noting, however, that “pushing” all firms into the formal sector may not necessarily be feasible or good social and economic policy. Indeed, if, as argued by Levenson and Maloney (1998), formality operates as a normal input in the production function of firms, it is possible that the intrinsic cost structure of many informal microenterprises may never, in fact, dictate that they grow large enough to need most of the formal institutions of civil society. Thus, for instance, given their very restricted markets, many of those microfirms may find it more efficient to use informal contract enforcement mechanisms and to operate on the basis of internal sources of finance. As a result, forcing them to formalize or trying to bring them into formal credit or capital markets would amount to “pushing on a string,” and it could lead large numbers of self-employed workers into open unemployment, while pushing formal sector wages downward.

Potentially countering these effects, however, one could argue that overall productivity could increase in the corresponding economies, as surviving microenterprises—the “upper tier” of the sector—become more efficient, thanks to formality, and as the goods and services previously produced by “lower-tier” microfirms are offered by larger and more productive firms. Which effect prevails depends, however, on whether formalization does indeed increase the productivity of some “upper-tier” microfirms, and on whether informality does generate considerable negative externalities on the rest of the economy (topics that we cover in the next section of this chapter).

Cross-country evidence on the determinants of informality

Despite convincing cross-country evidence confirming the relevance of several of the above-cited potential costs and benefits of informality, the data suggest that, to affect the size of the underground economy, policy makers may need to act on several fronts at the same time. The effects of regulations, for instance, appear to depend on the quality of governance. Moreover, as argued above, small changes in only some of the private costs or benefits of informality may not have a binding effect on firms' decisions regarding regulatory compliance.

The presence of a positive relationship between the regulation of firm entry and labor markets, on one hand, and the size of the informal sector, on the other, has been illustrated by Botero et al. (2003) and Djankov et al. (2002). These

authors were the first to construct large cross-country databases covering, respectively, the legal requirements for registering new firms (together with the time and costs involved in the corresponding procedures); and data on employment, collective bargaining, and social security laws.⁸ Djankov and his coauthors show that, for given levels of per capita income, the informal sector tends to be larger in countries where registering a new firm involves a larger number of procedures or where employment and industrial relations laws are more rigid. These results suggest that, at least on average, entry and labor regulations are not driven mainly by public interests, nor are they means for increasing the efficiency with which society operates.

Using a smaller sample but new estimates of the size of the informal sector for 14 Latin American and Caribbean countries, Loayza (1996) shows that informality is positively associated with levels of taxation and labor market regulations, and negatively correlated to the strength and efficiency of government institutions. Loayza and Rigolini (2006) confirm these results in a dynamic framework, showing that, in the long run, informality is negatively and robustly related to the flexibility of business regulations, the value of public services associated with law and order, and the capacity of governments to monitor and enforce formal taxes and regulations.⁹

One caveat to the above results is that the long-run links between regulations and informality may apply differently in countries characterized by strong or weak institutions, with good or bad governance systems. This is exemplified by the finding of Friedman et al. (2000), in a sample of 69 countries, that higher tax rates are not correlated with a larger unofficial economy, and may, in fact, be linked to a *smaller* informal sector. They interpret this result by suggesting that, across the countries in their sample, the incentive to evade high tax rates is outweighed by the larger benefits of formality in countries where higher tax revenues help finance productivity-enhancing public goods and a strong legal environment. Indeed, they find that most of the available indicators of bad governance—including corruption, overregulation, and weak legal environments—are positively and robustly related to the size of the informal sector. Thus, high tax rates can coexist with small unofficial economies, provided that rules and regulations are not enforced in a discretionary way and that levels of corruption are kept under control. In other words, as argued elsewhere in this report,

where tax regulations and enforcement are perceived as being fair—thus increasing “tax morale”—low levels of tax evasion and informality can be achieved without necessarily reducing tax burdens, thereby allowing for an adequate provision of productivity-increasing public goods.

The above findings imply that both the quantity and the quality of regulation matter for explaining cross-country differences in the size of the informal sector. In particular, reducing the quantity of regulations may be a good way of diminishing informality in countries characterized by bad governance, but it may have a much smaller—or even a null or negative—impact where the quality of institutions is high. This is illustrated by the findings of Loayza, Oviedo, and Servén (2005) that labor and product markets¹⁰ regulations are positively related to the size of the informal sector only for countries with low governance quality, below a threshold that corresponds roughly to the levels of Greece, Japan, and Spain. To measure the quality of governance, Loayza, Oviedo, and Servén use indicators of the absence of corruption in the political system, prevalence of the rule of law, and level of democratic accountability. They argue, and their results seem to confirm, that in countries with better quality of governance, regulations are more likely to be driven by valid social goals, as opposed to the interests of particular groups, and their enforcement is probably more transparent and less discretionary. In contrast, where corruption is high, and democracy or the rule of law is weak, increasing the quantity of regulations is likely to stimulate informality. Consistent with the findings of Friedman et al. (2000), Loayza, Oviedo, and Servén (2005) also find that higher levels of fiscal regulations are associated with smaller informal sectors in countries with good governance, but fiscal regulations are unrelated to the extent of informality in countries where the quality of governance is sufficiently low—the threshold corresponding to the levels of Colombia and Pakistan.¹¹

Firm-level evidence on the determinants of informality

Using survey-based, firm-level data for five Eastern European countries, Johnson et al. (2000) confirm some of the above cross-country results. They find, for instance, that among Russian and Ukrainian manufacturing firms, respectively, an average 41 and 29 percent of sales go unreported for tax purposes, compared to between 5 and 7 percent of sales

in Poland, Romania, and Slovakia. This is not surprising, Johnson and coauthors argue, given that managers in Russia and Ukraine face much higher taxes, report much higher levels of bureaucratic corruption and Mafia extortion, and exhibit a lower trust in their legal and court systems. However, firm-level regressions on the determinants of informality using the same data for three of the five countries cited above suggest that only the prevalence of corruption—measured through firms’ reporting of extralegal payments for services or government licenses—has a significant relationship with the percentage of sales unreported for tax purposes, with no effects found for taxation or court efficiency.

Enterprise survey data for Latin America and the Caribbean also suggest that corruption is positively and significantly related to informality. As seen in table 6.1, in five of the seven countries for which data are available (the exceptions being Panama and Peru), companies reporting that bribing of government officials to “get things done” is a common practice in their line of business exhibit rates of revenue and worker underreporting that are between 4 and 8 percentage points larger than those of other firms. As suggested by Johnson et al., (2000), this result could be due to firms’ underreporting some of their activities (sales and workers) in order to hide them from corrupt officials. Alternatively, if causality runs in the opposite direction, bribes could be a condition for remaining partially informal. Moreover, a complementary explanation is that firms that view the government as corrupt may also place a lower value on the public goods that it provides, and thus have lower incentives for contributing to its financing.

To evaluate the effect of labor regulations on informality, we construct a dummy variable for firms stating that those regulations significantly affected their hiring and firing decisions during the previous year. Both for the pooled sample and for three individual countries—Argentina, Colombia, and Mexico—we find that firms constrained by labor regulations evade a higher fraction of taxes and/or social security contributions. In most cases, the cost of severance payments is the aspect of labor market regulations firms most frequently report as the biggest obstacle to hiring more workers (figure 6.8). The only exception, among the eight countries for which data are available, is Mexico, where severance costs are surpassed (as an obstacle) by the costs of health insurance contributions. Health costs are second in importance to severance payments in Colombia,

Panama, Paraguay, and Peru. The relative importance of other labor issues varies across countries, with regulations on temporary work being mentioned more frequently in Argentina and Bolivia (after severance costs), as well as in Colombia and Panama, and retirement benefits being the issue of second-most importance in Uruguay.

In the case of Panama, we also find some evidence of a link between informality and the enforcement of tax regulations. Indeed, in this country we find that informality decreases with the percentage of firms in the corresponding city and sector that have been visited or inspected by tax officials during the previous year. For each percentage point increase in the probability that a firm is visited by tax inspectors, the fractions of unreported sales and workers are reduced by between 1 and 2 percentage points. This result is consistent with evidence on the higher prevalence of informal salaried workers among Brazilian firms located in areas where labor regulations are less tightly enforced. Indeed, Almeida and Carneiro (2005) show that, for each 10 percent increase in the number of fines per 1,000 firms issued due to irregularities associated with unregistered workers, the fraction of informal employees falls by 1.2 percent as a proportion of total employment.

Interestingly, Almeida and Carneiro also find that the lower level of informality resulting from a tighter enforcement of labor regulations is associated with a decrease in labor productivity—a reduction of 3.6 percent for each 1 percentage point decrease in the proportion of informal employment—and lower investments in capital and technology. This result is consistent with the findings of Scarpetta and Tresselt (2004) that higher labor adjustment costs resulting from stricter employment protection legislation may lead, at least in some industries and countries, to lower levels of total factor productivity. In other words, it appears that, in some specific contexts, the added flexibility resulting from informality—which is one of the above-mentioned private benefits from regulatory noncompliance—can facilitate the introduction of new technologies and enable firms to operate more efficiently. These benefits, however, may not necessarily prevail over the various above-mentioned costs of informality, which, as argued below, may lead to a negative relationship between informality and overall productivity.

The results in table 6.1 also provide some support to the hypothesis that, where market support mechanisms and links with large firms are better developed, informality

TABLE 6.1

Firm-level correlates of sales and employment underreporting

Variable	Argentina		Bolivia		Colombia		Mexico	
	Underreported sales (1)	Underreported employment (2)	Underreported sales (3)	Underreported employment (4)	Underreported sales (5)	Underreported employment (6)	Underreported sales (7)	Underreported employment (8)
Corruption dummy	6.436 (0.000)***	5.655 (0.000)***	4.462 (0.088)*	6.603 (0.033)**	3.119 (0.184)	4.784 (0.043)**	7.222 (0.001)***	4.850 (0.021)**
Labor regulations dummy	1.957 (0.267)	4.793 (0.004)***	0.791 (0.799)	4.659 (0.204)	6.771 (0.005)***	7.607 (0.002)***	8.358 (0.006)***	-0.731 (0.803)
Tax inspections (average n)	0.014 (0.898)	-0.054 (0.615)	0.089 (0.701)	-0.370 (0.176)	-0.209 (0.196)	-0.094 (0.564)	0.040 (0.644)	0.056 (0.502)
Contract enforcement (average %)	-0.086 (0.510)	-0.116 (0.352)	-0.170 (0.348)	0.100 (0.640)	-0.019 (0.890)	0.128 (0.356)	-0.288 (0.033)**	-0.227 (0.081)*
Large clients dummy	-5.775 (0.001)***	-2.954 (0.087)*	1.271 (0.673)	-1.378 (0.698)	-3.927 (0.053)*	-3.719 (0.069)*	-2.360 (0.364)	-0.125 (0.960)
Bank loans (% of firms)	-7.292 (0.461)	-10.611 (0.259)	-34.587 (0.232)	1.148 (0.973)	1.679 (0.916)	5.503 (0.732)	17.277 (0.156)	5.986 (0.611)
Observations (n)	744	744	432	432	833	833	1,009	1,009
Correlation of residuals (<i>p</i> -value of independence test)	0.50 (0.00)		0.43 (0.00)		0.64 (0.00)		0.67 (0.00)	

(Continued)

TABLE 6.1

Firm-level correlates of sales and employment underreporting (*Continued*)

Variable	Panama		Peru		Uruguay		Pooled sample	
	Underreported sales (9)	Underreported employment (10)	Underreported sales (11)	Underreported employment (12)	Underreported sales (13)	Underreported employment (14)	Underreported sales (15)	Underreported employment (16)
Corruption dummy	−2.246 (0.680)	0.469 (0.932)	2.281 (0.234)	2.948 (0.171)	7.300 (0.017)**	3.085 (0.311)	4.832 (0.000)***	4.835 (0.000)***
Labor regulations dummy	−4.139 (0.438)	2.513 (0.641)	−2.015 (0.288)	1.141 (0.593)	1.209 (0.624)	2.385 (0.332)	1.937 (0.062)*	3.196 (0.002)***
Tax Inspections (average n)	−1.840 (0.007)***	−1.270 (0.065)*	−0.024 (0.800)	−0.022 (0.834)	0.439 (0.132)	0.337 (0.246)	0.050 (0.237)	0.022 (0.608)
Contract enforcement (average %)	−0.861 (0.191)	−0.366 (0.582)	−0.052 (0.612)	−0.073 (0.528)	−0.568 (0.195)	−0.698 (0.111)	−0.119 (0.009)***	−0.072 (0.122)
Large clients dummy	−2.075 (0.635)	−4.501 (0.308)	−2.107 (0.280)	−4.267 (0.052)*	0.191 (0.949)	0.055 (0.985)	−2.814 (0.004)***	−2.709 (0.006)***
Bank loans (% of firms)	68.822 (0.240)	−15.790 (0.789)	−14.453 (0.423)	1.006 (0.960)	11.987 (0.572)	11.923 (0.573)	−0.594 (0.904)	−4.351 (0.383)
Observations (n)	417	417	575	575	268	268	4,278	4,278
Correlation of residuals (p-value of independence test)	0.55 (0.00)		0.52 (0.00)		0.60 (0.00)		0.58 (0.00)	

Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

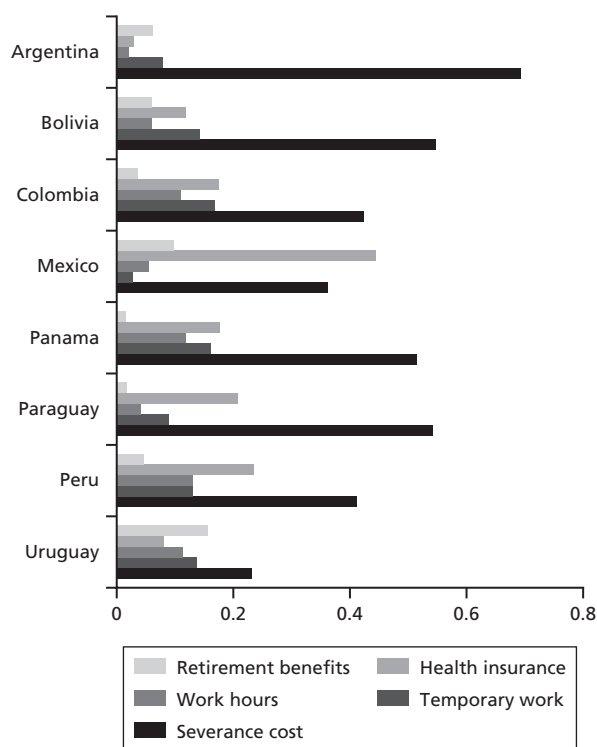
Note: Firm size, time in business, labor productivity, formality status at time of starting up, sector, and location are also controlled for. Estimation was performed for each country, using seemingly unrelated regressions. Probability values are in parentheses.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

FIGURE 6.8

Labor regulations described by firms as biggest obstacle to hiring workers

Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

tends to be lower. In particular, firms that operate in sectors and regions where the use of third-party contract enforcement mechanisms is more frequent have lower tax and social security evasion rates. On average, for the pool of the seven countries for which we have data, an increase of 10 percentage points in the fraction of firms that use third parties to solve commercial disputes—a change of about one standard deviation in that variable—is associated with a reduction of about 1 percentage point in the fraction of unreported sales. Moreover, while we do not find evidence for any impact of access to financial services on informality, the results in table 6.1 suggest that companies that sell mainly to large firms (that is, companies with more than 100 employees) exhibit rates of underreporting that are about 3 percentage points lower.

As for differences in tax evasion rates across sectors, we find in the pooled sample that the highest rates of sales underreporting are in the construction and transport sector, followed by manufacturing, and then by commerce and services.¹² In the case of employment underreporting, the

only significant difference is obtained for manufacturing, which shows higher social security evasion than all other sectors. At the country level, somewhat different patterns emerge in some cases: in Argentina, tax evasion is highest in manufacturing; in Bolivia, employment underreporting is highest in services; and in Mexico, tax evasion is lowest in services. The above cross-sector differences are obtained after controlling for firm characteristics (for example, firm size, productivity, and location, among others), as well as for such sector-specific factors as the incidence of corruption and the strength of regulatory enforcement (see regressors in table 6.1). Thus, there must be other unobserved characteristics of firms operating in those sectors that make them more likely to evade government regulations. While we can only speculate about such unobserved factors, they could include, for instance, a greater ability to avoid government enforcement (as in the case of firms engaged in local urban transport, short-term construction contracts, or locally distributed manufacturing goods), and competitiveness challenges faced by previously protected industries in the context of trade liberalization (as in the case of manufacturing firms using tax evasion as a way to compete with low-cost foreign manufacturers).

The results discussed above also suggest that the factors driving the underreporting of sales and workers are very similar. Moreover, the firms' decisions to evade taxes and social security appear to be quite interrelated, as revealed by the fact that the residuals of the corresponding regressions are highly correlated (table 6.1). These findings lend credence to the hypothesis that firms behave strategically when evading taxes (income taxes, value-added taxes, and so forth) and hiring workers off the books, and that they understandably seek to be consistent in the information they report to the tax and social security authorities to protect themselves from possible audits. Moreover, firms probably decide simultaneously on their levels of tax and social security evasion, taking into account both tax and labor market regulations and enforcement.

A note of caution is in order, however, with regard to the regressions reported in table 6.1, as their explanatory power is relatively small, with *R*-squared statistics ranging from 5 to 20 percent. Thus, while those results illustrate the relevance of a number of factors underlying the large size of the informal sector in several countries of the region, we are unable to explain as much as 80 percent of the tax and social security evasion behavior of the corresponding firms.

Impact of simplified registration and tax systems

While the existing statistical evidence suggests that reducing the time and cost required for firm registration can contribute to increases in the number of formally registered firms, the magnitude of the corresponding effects is still subject to some controversy. In particular, both Bruhn (2006) and Kaplan, Piedra, and Seira (2006) have analyzed the effect of a Mexican program that allows firm registration procedures to be completed within three days—the so-called Rapid Business Opening System (Sistema de Apertura Rápida de Empresas [SARE]) program implemented in about 30 Mexican cities. Kaplan, Piedra, and Seira find that SARE has led to statistically significant, albeit quantitatively small, effects in the *flow* of new registered firms, which would increase by between 4 and 8 percent as a result of SARE, implying about two to five new firms registered and 12 to 19 new formal jobs created per municipality per month. They show that the effects of SARE have been concentrated in the first 10 months after its implementation, which, they argue, suggests that the impact of the program is limited to the formalization of a small fraction of the existing stock of informal firms.

Bruhn (2006), nevertheless, reaches somewhat different conclusions on the impact of SARE, possibly as a result of the use of a different data source (employment surveys as opposed to official administrative records) and a different estimation technique (one based on the different timings of implementation of the program across Mexican cities). Bruhn focuses on the effect of SARE on the fraction of registered businesses, as captured by Mexico's national employment survey. She finds that the program had a much larger effect than the one reported by Kaplan, Piedra, and Seira—namely, a 5.6 percent increase in the *stock* of registered businesses. This implies that SARE can be credited for about 1,000 new registered firms per county, on average. Moreover, Bruhn shows that past informal business owners are not more likely to register their businesses after SARE, but former wage earners with conditionally high wages do become more likely to open a formal business. The fact that Bruhn's estimates imply effects that are much larger than those obtained by Kaplan, Piedra, and Seira can also be attributed to the different types of businesses that are covered in the databases used in each of those reports. Indeed, while Kaplan, Piedra, and Seira look at firms registered with the Mexican Social Security Institute (IMSS), Bruhn's employment data also cover the large majority of

microenterprises that do not have salaried employees, and for whom registration with IMSS is not mandatory. One possible, albeit somewhat surprising, interpretation of the conflicting conclusions reached by Kaplan, Piedra, and Seira and by Bruhn is that SARE could have had a sizable effect on the creation of new owner-only formal businesses, but a much smaller impact on the formalization of existing informal microfirms. In any case, it appears that further research is needed to evaluate the impact on informality of simplified firm registration programs, such as SARE.

Complementary evidence concerning the impact on informality of red tape reduction programs has been obtained from the analysis of the Brazilian Integrated System for Tax and Social Security Payments for Micro and Small Firms (Sistema Integrado de Pagamento de Impostos e Contribuicoes as Microempresas e Empresas de Pequeno Porte [SIMPLES]) program. In a manner different from SARE, however, this program combines simplified firm registration with lower taxes and social security contributions for micro- and small enterprises, allowing for an 8 percent reduction in the overall tax burden faced by eligible firms. Both Fajnzylber, Maloney, and Rojas (2006a) and Monteiro and Assunção (2006) find that, at least during the year following the implementation of SIMPLES, the program led to statistically significant increases in formal registration rates of between 6 and 13 percentage points, depending on the sample and methodology. While further research may be needed to establish whether the effects of SIMPLES were permanent or temporary, the above evidence suggests that there is a potential for increasing microfirm formalization by combining red tape reduction with tax relief measures for microfirms.

Programs to simplify and reduce tax burdens for small contributors, including individuals and small firms, have been implemented in recent years in a number of Latin American and Caribbean countries. These countries include Argentina (the Monotributo program), Bolivia (the simplified tax regime for small firms in selected activities [RTS]), Colombia (simplified value-added regime for small contributors), Costa Rica (RTS), Chile (simplified income tax), Ecuador (simplified value-added tax), Mexico (simplified tax regime for small contributors), Nicaragua (single payment system for value-added and income taxes), Honduras (simplified sales tax), Paraguay (single tax for owner-only firms), Peru (simplified unique tax and special income tax regimes), and Uruguay (small-enterprise tax system).¹³ Given the evidence from the Brazilian case, it appears that most countries

in the region could benefit from a revision of their simplified tax systems for micro- and small enterprises in order to better suit them for the objective of increasing formality—for instance, by combining them with simplified firm registration systems, such as Mexico's SARE.

Impact of informality on firm productivity and economic growth

As argued above, many lower-tier microfirms may choose to operate informally as a result of their limited levels of productivity and growth potential—levels that translate into a small demand for market- and government-provided services that have formality as a precondition. Not surprisingly, when deciding to formalize, they are often more motivated by sticks—avoiding fines and bribes—than by carrots like access to credit or formal contract enforcement. This, however, does not mean that increasing formality does not in itself have the potential for increasing overall productivity, through both static and dynamic channels—a possibility that we review next, both conceptually and from an empirical point of view.

Static versus dynamic effects

To the extent that informality is associated with a preponderance of small firms, there is a concern that it could lead to considerable efficiency losses. This prediction, however, depends on whether returns to scale are constant or increasing. As reviewed by Tybout (2000), the literature on the subject is divided between simulation studies, which often assume decreasing average costs, and survey-based estimates, which generally suggest that the benefits from increasing plant size are relatively small. Thus, while one-person establishments are usually found to be less efficient than firms with at least some employees, returns to scale among the latter firms are very close to unity and, at most, mildly increasing.

This is not to say that increasing returns to scale are not the norm in some specific industries, notably the most capital-intensive ones. However, because of their limited access to capital and skilled labor, and to avoid cost disadvantages, micro- and small firms tend to locate in industries where efficiency losses associated with low scale production are limited. This is facilitated by the fact that demand for such products is negatively correlated with countries' per capita incomes, as Engel effects direct consumer demand toward simpler products that can be efficiently produced with labor-intensive technologies. In other words, in countries where lower levels

of overall productivity drive a large number of firms into informality, consumer demand is also likely to be directed toward products and services whose production does not exhibit increasing returns to scale, so one should not expect large static losses from informality driven by low firm size alone.

Besides potential static inefficiencies associated with the nonexploitation of economies of scale, a parallel concern is that unproductive firms are able to compete with their lower-cost peers by means of avoiding taxes and regulations. Thus, informal firms may be able to stay in business despite having higher operating costs—driven, for instance, by lower levels of entrepreneurial ability. This could slow down the creative destruction process by which innovative, high-productivity firms expand to the detriment of less-productive ones. In other words, to the extent that increases in regulatory enforcement drive out of business a large number of firms that self-selected into informality because of having lower productivity than firms of the same size operating in the same sectors, one could expect potentially large negative effects on aggregate productivity. One caveat to this argument, however, is that, as shown by Almeida and Carneiro (2005) for the case of labor regulations, informality may allow firms greater flexibility in their employment and production decisions, which, in turn, could lead them to operate more efficiently. Whether this effect dominates other factors that could lead to a negative link between informality and firm productivity—for example, the self-selection of unproductive firms into informality, the incentives to operate at a small scale to avoid detection, and the inability to gain access to factor and product markets—is a question that only empirical evidence can help resolve.

A parallel concern is that high levels of informality could also have negative consequences on the incentives of formal firms to innovate and adopt new technologies, which also could reduce overall productivity growth. Several studies have emphasized these potentially negative dynamic implications of informality. Thus, Capp, Elstrodt, and Jones (2005), Elstrodt, Lenero, and Urdapilleta (2002), Farrell (2004), Kenyon and Kapaz (2005), and Palmade (2005) see informality as one of the main causes for the gap in productivity levels between developed and developing countries, inducing distortions in investment decisions and limiting the growth potential of the corresponding economies. Studies performed by the McKinsey Global Institute suggest that informality accounts for around 50 percent of the productivity gap

between countries like Turkey, Portugal, and the United States; and for 30 percent of the productivity gap between Brazil and the United States (Farrell 2004).

The main argument proposed by these studies is that firms that join the informal sector tend to become “trapped in a self-reinforcing dynamic that confines them to subscale, inefficient, low-productivity work” (Farrell 2004, p. 30). Moreover, informality also has negative effects on the investment decisions of formal firms, as it reduces their market share and profitability. The relative cost advantages enjoyed by informal firms as a result of not paying taxes and not incurring the costs of regulatory compliance, however, allow them to stay in business despite their low productivity, which, as mentioned above, could distort competition and limit the process of creative destruction. Moreover, as more productive firms also have fewer incentives to invest in innovation and technology adoption, the McKinsey studies (reviewed in Farrell [2004]) suggest that informality leads to an overall reduction in economic growth.

One caveat to these arguments is that, from a theoretical point of view, technology adoption and innovation could either decrease or increase as a result of unfair competition by informal firms. In particular, as argued by Cunha (2006), if technological change takes the form of the discovery—or introduction into the country—of improved qualities for intermediary goods, informality may have the

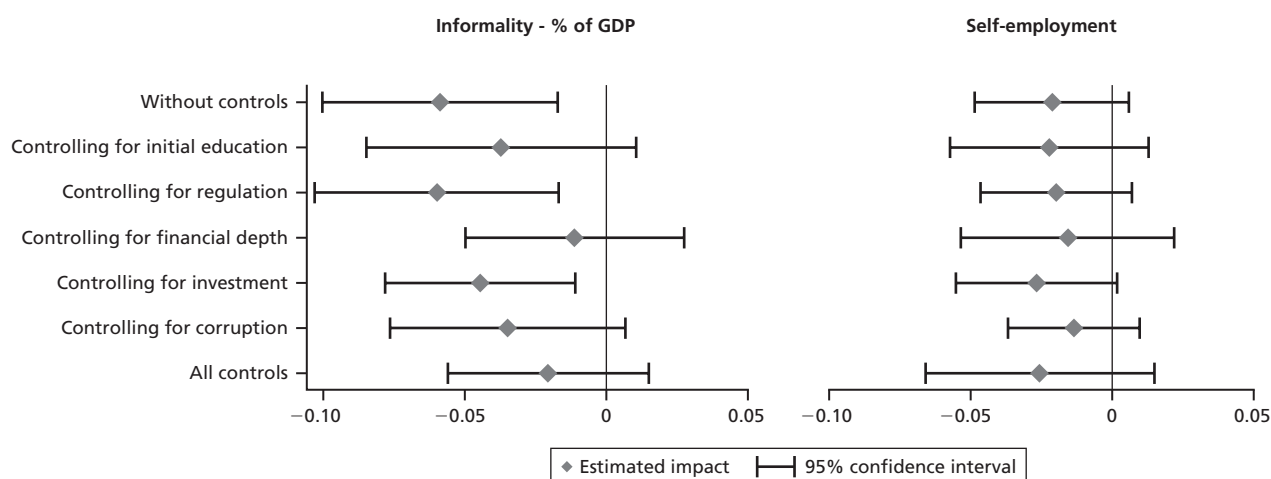
effect of impeding the ability of the frontier quality producer to set a price that would force all other producers—of lower qualities—to leave the market. The impact on research and development investments—thereby including the expenditures involved in adapting foreign technologies to local conditions—is, however, ambiguous. Indeed, while informality decreases the market power and profit levels of frontier producers, it also increases the life span of frontier products by augmenting the quality improvements that are needed to debunk current market leaders. As a result, investments in research and development—and growth—could increase or decrease, depending on which effect dominates.

Aggregate growth effects

Despite the widespread belief that a large informal economy hurts economic growth, cross-country comparisons do not find a robust association between informality and growth. Figure 6.9 shows the estimated impact of informality on growth, using two diverse informality indicators: self-employment and the Schneider (2005) estimates of the proportion of GDP produced by the underground economy in the period 1999–2000. In accordance with common wisdom, under both indicators informality appears to have a negative impact on growth: on average, the regressions suggest that decreasing informal economic activity by 10 percent of GDP is associated with 0.6 percent higher economic growth.

FIGURE 6.9

Estimated impact of informality on growth



Sources: Authors' calculations from Barro and Lee 2000; Fraser Institute 2005; ILO 2005; Schneider 2005; and World Bank 2005.

Note: The dependent variable is average growth of GDP per capita during the period 1990–2003. All regressions include a constant, regional dummies, and initial log of GDP per capita. Education represents the average years of education of the population over age 25. Regulation is an index of regulatory efficiency from the Fraser Institute. Financial depth represents the log of private domestic credit over GDP in 1990.

Nonetheless, the estimated coefficients tend not to be robust. In none of the presented regressions is the coefficient associated with self-employment significantly different from zero at the 5 percent level. To be sure, when we use the macroeconomic estimates of informality, the basic regression without added explanatory variables shows a negative and statistically significant association between informality and growth. However, the relationship loses its significance when we control for education, financial depth, or corruption.¹⁴ Thus, while previous studies have found a negative relationship between informality and growth, they have relied on a very narrow pool of observations (Loayza 1996) or have not controlled for relevant correlates of growth, such as regulation, human capital, and initial GDP per capita (Schneider and Klinglmair 2004). It is worth noting, however, that the fact that we do not observe large effects of informality on growth after controlling for other standard growth determinants may just reflect our inability to empirically distinguish between the direct effects of those variables and indirect effects through informality. In other words, one of the channels through which some of those standard drivers of growth operate could be increasing informality—for example, low human capital reducing the opportunity cost of self-employment, or corruption diminishing the incentives to comply with regulations—but this would not be apparent in the results reported in figure 6.9.

Empirical evidence on creative destruction in Latin America and the Caribbean region

The arguments on the negative dynamic effects of informality offered above should lead us to expect a less vibrant process of industrial evolution in developing countries than in industrial countries—a process that should be reflected in higher productivity dispersion and lower firm turnover. The evidence, however, suggests that the average distance to “frontier” production technologies is similar in studies of developing and OECD countries, with average technical efficiency levels equivalent to about 60–70 percent of the corresponding best practices (Tybout 2000). These measures, it must be noted, are usually based on data that exclude microenterprises and low-productivity, owner-only firms, so they probably underestimate productivity dispersion.

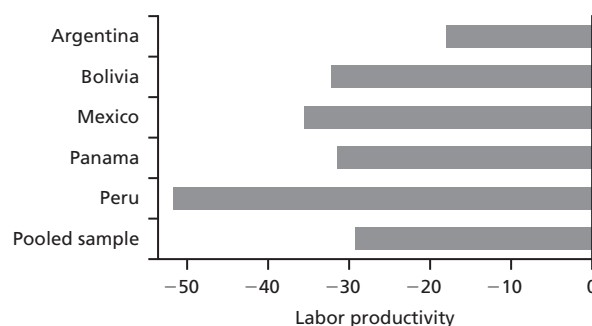
As for the evidence on the rates of firm and job creation and destruction, the evidence on whether it is lower or similar in least-developed countries is ambiguous. Roberts and Tybout (1997) show that rates of firm turnover and market shares of recent entrants into the formal sector are

surprisingly large in developing countries, such as Chile and Colombia. These results, however, do not necessarily constitute evidence that the creative destruction process, initially described by Joseph Schumpeter, is alive and well in those countries, as the observed high turnover rates could be the result of high macroeconomic instability and of the preponderance of small firms—which have inherently higher rates of failure. Thus, when corrected by volatility, turnover rates are not larger in Latin America and the Caribbean than in industrial countries. Moreover, while there is evidence that entrants in Chile and Colombia are slightly more efficient than incumbents, and that new low-productivity firms tend to go out of business rapidly, the impact of this process on overall productivity is found to be relatively small.

Firm-level effects of informality

As evidenced in figure 6.10, firms that report having started operations without formally registering—at least initially—exhibit, on average, much lower levels of output per worker, even after controlling for firm size, time in business, sector, and region. In other words, those that start up informally are clearly at the bottom end of the productivity distribution of the corresponding industries and regions. The difference in labor productivity between those firms and the ones that have always operated formally is 29 percent, on average, for the seven Latin American and Caribbean countries analyzed here. The effects are largest

FIGURE 6.10
Estimated impact of informality on labor productivity (%)



Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

Note: The figure reports coefficients on formality indicators in regressions of the log of output per worker on a dummy for not having formally registered at the time of starting up, controlling for employment size, time in business, and sector and region dummies.

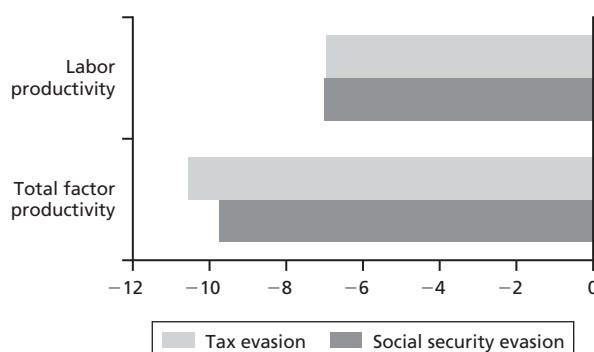
in Peru (50 percent), and, while lower in the remaining countries, they are statistically significant in four out of seven countries, the exceptions being Uruguay, Panama, and Colombia (not shown).¹⁵

Registered firms that report having started informally are only 6.6 percent of all formal firms (in the pooled sample), and respond for just 4.3 percent of total sales and 3.6 percent of employment. Thus, the increase in aggregate productivity derived from hypothetically excluding them from their sectors—for example, by a stricter enforcement of entry regulations—would be relatively small: about 0.8 percent, on average.¹⁶ However, to the extent that labor productivity is higher in registered firms that started up informally but eventually registered, than in similar firms that started informally but never registered, the potential impact on productivity of fully enforcing entry regulations could be much larger. Thus, for instance, in a hypothetical scenario where all nonregistered informal Mexican microfirms with no more than five employees (which represent about 21 percent of total employment in that country) were to go out of business due to stricter enforcement of entry regulations, and assuming that employment would shift to firms with a 35.5 percent productivity advantage (based on the estimate for Mexico reported in figure 6.10, which we consider a lower bound for the productivity differentials between formal and informal firms), the resulting impact on aggregate productivity could be as large as 6 percent.¹⁷ Of course, these are very imprecise, back-of-the-envelope calculations that are aimed only at illustrating a much more general point—namely, the possibility that informality is associated with lower levels of aggregate productivity, which should be considered in any analysis of its social costs and benefits.

Informality, however, also takes the form of tax and social security evasion among registered firms. In this respect, it would be useful to know what would be the impact of marginal reductions in tax evasion on firm and aggregate productivity. With that purpose, we estimate the effect of average tax and social security evasion in a given sector and region on the level of productivity of individual firms operating in the corresponding areas. The results are presented in figure 6.11 for our pooled sample of seven countries. We find that each 10 percent increase in average evasion rates is associated with reductions in labor and total factor productivity of 7 and 10 percent, respectively. These effects are not subject to the criticism of a possible reverse causality from low firm productivity to higher firm informality. Indeed,

FIGURE 6.11

Effects of a 10 percent increase in tax and social security evasion at the industry/region level on individual firm productivity (pooled sample)



Source: Authors' calculations, using the World Bank's Enterprise Survey Database.

Note: The figure reports coefficients on formality indicators in regressions that control for employment size, time in business, fixed capital stocks (in TFP regressions), and sector and region dummies, using a pooled sample of seven countries.

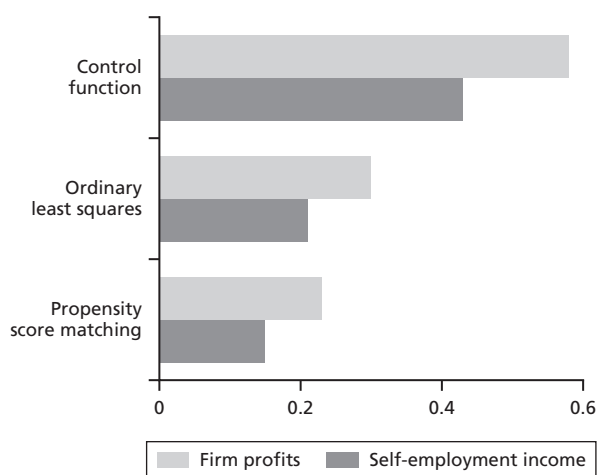
our informality measure is aggregated at the industry and region levels so it could hardly be affected by individual firm productivity. The results suggest that there are significant potential productivity gains to be obtained from increased enforcement of tax and social security regulations.

Further evidence on the presence of a causal link between informality and firm performance has been obtained using microenterprise survey data. In the case of Mexico, Fajnzylber, Maloney, and Rojas (2006c) show that microfirms that report paying taxes exhibit higher levels of profit, even after controlling for employment size and capital stocks. This result is robust to the use of estimation techniques that control for the impact of unobserved personal and firm characteristics—for example, managerial ability—that could affect both the decision to formalize the firm and its performance. Quantitatively, we find that firms that pay taxes exhibit between 15 and 60 percent higher productivity levels, depending on the estimation method and the performance variable used—either firm profits as reported in detailed microfirm surveys or self-employment income as reported in the Mexican Employment Survey (figure 6.12). In addition, there is evidence that owners of formal firms are less likely to go out of business.

One of the channels through which formality could increase firm performance is by facilitating access to factor and product markets. Cull, McKenzie, and Woodruff (2007) show that access to credit among small Mexican retail firms

FIGURE 6.12

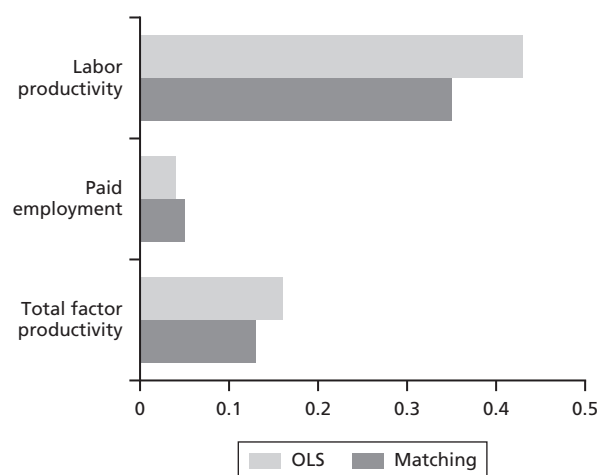
Productivity difference between tax-paying Mexican microfirms and informal firms with similar characteristics



Source: Fajnzylber, Maloney, and Rojas 2006c.

FIGURE 6.13

Productivity and paid employment effects of exogenous increase in formality driven by Brazil's SIMPLES program



Source: Fajnzylber, Maloney, and Rojas 2006a.

is much higher among formally registered firms. While this correlation could reflect the presence of unobserved personal and firm traits linked to both formality and access to finance, it is also consistent with informality reducing microfirm access to formal loans. Thus, for instance, if microlenders make formal registration a requisite for granting loans, formalization could allow credit-constrained microfirms to exploit the sizable returns to investment that have been estimated by Cull and coauthors—between 20 and 33 percent per month for investments of about \$140, using data from a randomized experiment.

Evidence consistent with some of the findings for Mexico has also been obtained for Brazil on the basis of exogenous increases in formality rates associated with the introduction of a simplified tax and registration system for micro- and small firms (the above-mentioned SIMPLES). Indeed, econometric estimates that take advantage of changes in the incentives to formalize, introduced by this program, show that it significantly increased access to credit among eligible firms, and altered the amount and composition of investment toward larger and longer-term projects (Monteiro and Assunção 2006). Moreover, Fajnzylber, Maloney, and Rojas (2006a) show that increases in formality driven by SIMPLES are associated with a higher use of paid labor, higher levels of capital intensity, and increased labor and total factor productivity. In particular, increases in the rates of microfirm registration that can be attributed to SIMPLES are

estimated to be associated with a 5 percent increase in paid employment, a 15 percent boost in total factor productivity, and a 35 percent increase in labor productivity (figure 6.13). At least in the Brazilian context, however, only a small fraction of the revenue-increasing effects associated with formality can be attributed to increased access to credit markets and government-provided technical assistance. In contrast, the greater willingness of formal firms to operate out of a fixed locale is responsible for as much as 50 percent of the increase in revenues among formal microfirms and for a third of the corresponding total factor productivity increase.

Conclusions

The empirical evidence on aggregate negative growth effects of informality is not conclusive, as informality tends to lose significance when other standard growth determinants are controlled for. This, however, could be due to the fact that many of the standard drivers of growth are also likely to affect informality—for example, low levels of human capital or institutional quality leading to both lower growth and higher informality—and it is difficult to separate their direct growth effects from those that operate through larger informal sectors.

The microeconomic empirical evidence, on the other hand, is still quite limited due to the econometric difficulties associated with distinguishing the effects of low

productivity on informality (see chapter 5) from the reverse effects operating from informality to productivity. However, the available evidence suggests that considerable efficiency gains could be derived from the transfer of production from low-productivity informal firms to their more productive formal peers. Similarly, the evidence indicates that the concerns associated with possible negative externalities generated by high levels of tax and social security evasion could be well justified, as firms operating in industries and regions characterized by high levels of sales and employment underreporting exhibit lower levels of labor and total factor productivity. Moreover, there is evidence indicating that exogenous increases in formality are associated with better firm performance, which should, in principle, translate into higher rates of economic growth.

In this context, the interest of policy makers and development practitioners in designing policies and programs to facilitate the formalization of small businesses and increase regulatory compliance by larger firms appears well justified. In particular, the fact that efforts to decrease regulatory burdens have recently become very popular does not come as a surprise. Burdensome regulations and costly bureaucratic requirements are indeed an important determinant of informality that may create barriers for increasing formal entrepreneurial activity. One challenge that governments face in this respect is that of assessing their existing and new regulations to determine the extent to which they are justified by public interests, associated, for instance, with the protection of public safety or the environment. Dealing with this challenge may require comprehensive regulatory assessments aimed at distinguishing relevant from anachronistic regulations, as well as at identifying those regulations that reflect private rather than public interests, and that could represent important barriers to formalization. Examples of such initiatives include national regulatory reviews, such as those implemented in leading transition economies.¹⁸

Similarly, many developing countries are now engaged in reducing the time and cost needed by businesses to obtain various government-issued permits and licenses.¹⁹ Indeed, even well-designed and legitimate regulations may create barriers to formalization if they are badly enforced or administered, creating excessive costs and uncertainty for private businesses. Thus, a complementary approach to reduce firm informality—to be pursued in parallel to regulatory reforms—is the implementation of administrative simplification programs aimed at reducing the transaction

costs associated with operating legitimate businesses. Internet-based technologies and one-stop shops can be effective tools to implement such programs, although their effectiveness can be greatly increased if they are used in conjunction with comprehensive reviews and revisions of existing administrative processes.

We argue, however, that although eliminating unnecessary regulations and reducing excessive red tape could contribute to reducing the size of the informal sector, those actions should not be the exclusive focus of policy makers engaged in attaining that objective. Indeed, the costs of regulatory compliance are only one among other factors that may affect formality decisions. In particular, to attract more businesses into the formal economy, it is crucial to increase the potential benefits of regulatory compliance. This implies facilitating the ability of micro- and small enterprises to tap into formal credit markets and improving the provision of business development and training services available to formal firms. Moreover, it is important to facilitate access to product markets through public procurement opportunities and supplier development programs aimed at increasing links with larger private firms. Other ways of increasing the benefits of formality include improving the quality of legal services available to small businesses and creating mechanisms to provide information to entrepreneurs wishing to formalize their businesses, thereby encompassing advisory services on taxes and regulations, as well as information about financial and nonfinancial services available to them.

Overall, a wider and integrated approach appears to be necessary to switch the incentives of a large fraction of informal firms in the direction of formality. Such an approach would likely have to combine both carrots (for example, lower costs of formalization, better and more efficient government services, and higher access to market- and government-provided services for formal firms) with sticks (such as increasing government enforcement of regulations and the expected cost of being caught). Moreover, as argued later in this volume, it is crucial that both the enactment and the enforcement of regulations are perceived to be fair, as this is vital for maintaining “tax morale” and increasing regulatory compliance.

The correct mix of policies, however, is likely to vary across countries and over time, depending on the relative importance of the various determinants of informality. Moreover, other aspects of public policy should be taken into account, including those related to the social consequences

of drastically reducing the size of the informal sector, which, at least in Latin America, is currently responsible for a large fraction of employment and income-generating opportunities for poor households. In other words, policies aimed at reducing firm informality should be considered in conjunction with the labor market and social protection issues associated with the possibility that large contingents of previously informal workers would have to shift to other segments of the labor market.

Notes

1. As in the study by Djankov et al. (2003), we leave outside the scope of this chapter those “underground enterprises” that are devoted to criminal activities.
2. Since firms are understandably reluctant to reveal information regarding tax and social security evasion, as well as informal payments to corrupt officials, the corresponding survey questions are phrased in terms of the practices of “typical firms in this establishment’s line of business.” This is a standard approach taken to measure the prevalence of corruption used, for instance, by Johnson et al. (2000). Note also that figure 6.1 reports simple averages across firms.
3. The high levels of tax evasion for Brazil are, to some extent, puzzling as tax revenues in that country have increased considerably in recent years, reaching 34 percent of gross domestic product (GDP) in 2002, driven partly by improvements in tax administration efficiency. However, respectively, 40 and 25 percent of federal tax revenues come from indirect cascading taxes and from social security contributions and other payroll taxes that firms, as suggested by the survey data, appear to be quite successful in evading.
4. The controls are those described in table 6.1.
5. De Soto (1989) is the seminal reference on the links between government regulations and informality. In their survey on the topic, Schneider and Enste (2000) list taxes, social security contributions, and the intensity of regulations (including those of labor markets) among the top causes for growth in the “shadow economy.” Loayza (1996) offers a theoretical model illustrating those effects and provides supporting empirical evidence for Latin America. Additional cross-country evidence on the links between tax and regulatory burdens and the size of the informal economy is provided by Botero et al. (2003), Djankov et al. (2002), and Loayza, Oviedo, and Servén (2005).
6. The surveys were done in the context of municipal administrative simplification projects supported by the International Finance Corporation.
7. In the theoretical model proposed by Sarte (2000), for instance, changes in the fixed cost of entering the formal sector do not affect the level of informality—their effect is nonbinding—when the cost of operating informally is relatively low.
8. These databases have been expanded and updated annually through the World Bank’s Doing Business project.
9. Loayza and Rigolini (2006) use the following empirical measures for the above-mentioned variables: the Fraser Institute’s index of credit, labor, and regulatory flexibility; the International Country

Risk Guide’s index of law and order; and the ratio of government expenditures to GDP.

10. The index of product market regulations is a composite of indexes of regulations in the areas of firm entry, trade barriers, financial markets, contract enforcement, and bankruptcy.

11. A related point made by Schneider and Enste (2000) is that complex tax systems can make legal tax avoidance in the official economy more profitable, and thus create disincentives for informality. As a result, fiscal reforms that combine lower tax rates and simpler tax systems could not necessarily lead to smaller informal sectors; the Austrian 1989 reform is mentioned as an example.

12. Within manufacturing, the food and beverage industry exhibits the highest levels of evasion, even after controlling for firm and sector characteristics.

13. See González (2006) for additional details.

14. The relationship remains weak when changing the reference period from 1990–2003 to 1999–2003 or when using five-year averages for the period 1980–2004. As the macroeconomic estimates of informality are cross-sectional, regressions with five-year averages can only be performed with self-employment.

15. It is worth noting that the estimated effects for most countries are based on a sizable number of initially informal firms: between 5 and 14 percent of all firms, or 355 in the pooled sample, 72 in Argentina, 72 in Bolivia, 66 in Mexico, and 104 in Colombia. Our estimates are arguably weaker in Panama, Peru, and Uruguay, where there are, respectively, 9, 20, and 12 of those firms (between 2 and 3 percent of all firms).

16. This is assuming that the employment share of the initially informal firms is taken over by competitors, which, per our estimations, have a 29 percent productivity advantage (figure 6.9).

17. This assumption is subject to the criticism that, as argued in the beginning of this section, informal microfirms tend to locate in industries where efficiency losses associated with low scale production are minimized. Moreover, note that we are assuming that all those working in informal microfirms would find employment in formal firms. We base the estimated employment share of informal microfirms on figures on registration rates from the *Encuesta Nacional de Micronegocios*—27.5 percent for owner-only firms and 62.7 percent for firms with two to five workers—and on *Encuesta Nacional de Empleo Urbano*—based estimates of the employment shares of firms in those size ranges—respectively, 15.3 and 27.5 percent.

18. See Djankov et al. (2003) for a description of the 1995–98 Hungarian regulatory review.

19. See the World Bank’s Doing Business reports for global reviews of country-level reforms and benchmarking exercises.

References

- Almeida, R., and P. Carneiro. 2005. “Enforcement of Labor Regulation, Informal Labor, and Firm Performance.” Policy Research Working Paper 3756, World Bank, Washington, DC.
- Barro, R., and J.-W. Lee. 2000. “International Data on Educational Attainment: Updates and Implications.” Working Paper 42, Center for International Development, Harvard University, Cambridge, MA.

- Botero, J., S. Djankov, R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2003. "The Regulation of Labor." *Quarterly Journal of Economics* 119 (4): 1339–82.
- Bruhn, M. 2006. "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico." Photocopy. Massachusetts Institute of Technology, Cambridge, MA.
- Capp, J., H. Elstrodt, and W. Jones, Jr. 2005. "Reining In Brazil's Informal Economy." *McKinsey Quarterly*. Available at <http://www.mckinseyquarterly.com>.
- CIEN (Centro de Investigaciones Económicas Nacionales). 2006. *Economía Informal: Superando las barreras de un Estado excluyente*. Guatemala City, Guatemala.
- Cull, R., D. McKenzie, and C. Woodruff. 2007. "Experimental Evidence on Returns to Capital and Access to Finance in Mexico." Photocopy. World Bank, Washington, DC.
- Cunha, B. 2006. "Informality, Productivity and Growth." Photocopy. World Bank, Washington, DC.
- de Soto, H. 1989. *The Other Path: The Invisible Revolution in the Third World*. New York: Basic Books.
- Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1): 1–37.
- Djankov, S., I. Lieberman, J. Mukherjee, and T. Nenova. 2003. "Going Informal: Benefits and Costs." In *The Informal Economy in the EU Accession Countries*, ed. B. Belev. Sofia, Bulgaria: Center for the Study of Democracy.
- Elstrodt, H., P. Lenero, and E. Urdapilleta. 2002. "Micro Lessons from Argentina." *McKinsey Quarterly* 2: 1–8.
- Fajnzylber, P., W. F. Maloney, and G. V. Montes Rojas. 2006a. "Does Formality Improve Microfirm Performance? Quasi-experimental Evidence from the Brazilian SIMPLES Program." Photocopy. World Bank, Washington, DC.
- . 2006b. "Microenterprise Dynamics in Developing Countries: How Similar Are They to Those in the Industrialized World? Evidence from Mexico." *World Bank Economic Review* 20 (3): 389–419.
- . 2006c. "Releasing Constraints to Growth or Pushing on a String? The Impact of Credit, Training, Business Associations and Taxes on the Performance of Mexican Microfirms." Policy Research Working Paper 3807, World Bank, Washington, DC.
- Farrell, D. 2004. "The Hidden Dangers of Informal Economy." *McKinsey Quarterly* 3: 27–37.
- Fraser Institute. 2005. "Economic Freedom of the World: Annual Report." Available at <http://www.fraserinstitute.ca>.
- Friedman, E., S. Johnson, D. Kaufmann, and P. Zoido-Lobaton. 2000. "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics* 76: 459–93.
- González, D. 2006. "Regímenes especiales de tributación para pequeños contribuyentes en América Latina." Photocopy. Inter-American Development Bank, Washington, DC.
- ILO (International Labour Organization). 2005. Labor Statistics. Available at <http://laborsta.ilo.org/>.
- Johnson, S., D. Kaufmann, J. McMillan, and C. Woodruff. 2000. "Why Do Firms Hide? Bribes and Unofficial Activity after Communism." *Journal of Public Economics* 76: 495–520.
- Johnson, S., D. Kaufmann, and A. Shleifer. 1997. "The Unofficial Economy in Transition." *Brookings Papers on Economic Activity* 2: 159–239.
- Kaplan, D., E. Piedra, and E. Seira. 2006. "Are Burdensome Registration Procedures an Important Barrier on Firm Creation? Evidence from Mexico." Photocopy. Instituto Tecnológico Autónomo de México, Mexico City.
- Kenyon, T., and E. Kapaz. 2005. "Rising Informality: Reversing the Tide." Viewpoint Series Note 301, Private Sector Development, World Bank, Washington, DC.
- Levenson, A., and W. Maloney. 1998. "The Informal Sector, Firm Dynamics and Institutional Participation." Policy Research Working Paper 1988, World Bank, Washington, DC.
- Loayza, N. 1996. "The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129–62.
- Loayza, N., A. M. Oviedo, and L. Servén. 2005. "The Impact of Regulation on Growth and Informality—Cross-country Evidence." Policy Research Working Paper 3263, World Bank, Washington, DC.
- Loayza, N., and J. Rigolini. 2006. "Informality Trends and Cycles." Policy Research Working Paper 4078, World Bank, Washington, DC.
- McKenzie, D. J., and C. Woodruff. 2006. "Do Entry Costs Provide an Empirical Basis for Poverty Traps?" *Economic Development and Cultural Change* 55 (1): 3–42.
- Monteiro, J. C. M., and J. J. Assunção. 2006. "Outgoing the Shadows: Estimating the Impact of Bureaucracy Simplification and Tax Cut on Formality and Investment." Photocopy. Department of Economics, Pontifícia Universidade Católica do Rio de Janeiro, Brazil.
- Palmade, V. 2005. "Rising Informality." Viewpoint Series Note 298, Private Sector Development, World Bank, Washington, DC.
- Rauch, J. E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35: 33–47.
- Roberts, M. J., and J. R. Tybout, eds. 1997. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Washington, DC: World Bank.
- Sarte, P.-D. G. 2000. "Informality and Rent-seeking Bureaucracies in a Model of Long-run Growth." *Journal of Monetary Economics* 46 (1): 173–97.
- Scarpetta, S., and T. Tresselt. 2004. "Boosting Productivity via Innovation and Adoptions of New Technologies: Any Role for Labor Market Institutions?" Policy Research Working Paper 3273, World Bank, Washington, DC.
- Schneider, F. 2005. "Shadow Economies around the World: What Do We Really Know?" *European Journal of Political Economy* 21 (3): 598–642.
- Schneider, F., and D. H. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38 (1): 77–114.
- Schneider, F. G., and R. Klinglmaier. 2004. "Shadow Economies around the World: What Do We Really Know?" Working Paper 1167, Center for Economic Studies and the Ifo Institute, Munich.
- Tybout, J. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature* 38 (1): 11–44.
- World Bank. 2005. *World Development Indicators*. Available at <http://www.worldbank.org>.

CHAPTER 7

Informality, Social Protection, and Antipoverty Policies

SUMMARY: While formal sector workers often have access to generous social security packages, informal sector workers generally have more limited access to risk management instruments. In response to concerns about large gaps in coverage, several countries have launched or expanded noncontributory assistance programs to help breach the coverage gap, especially among the poor. A clear rationale for public intervention exists; however, weaknesses in the design and implementation of many programs actually exacerbate people's lack of access to protection and create further incentives for informality. Providing effective social protection for the citizens of the region will thus require a reengineering of many countries' existing programs. In the long run, this should include providing universal "essential cover" in health, de-linked from the labor contract and financed by general taxation; it should also include "poverty prevention" pensions aimed at the elderly poor combined, in an incentive compatible way, with individual savings as the mainstay of life-cycle consumption smoothing. Successful implementation of this agenda will, in most countries, require a number of short-to-medium-term measures that increasingly improve the efficiency and design of countries' social protection systems.

THE PREVIOUS CHAPTERS HAVE FOCUSED on characterizing the informal sector of the labor market in Latin America and the Caribbean, as well as those people who work in it. In characterizing informality, one core definition relates to the absence of worker coverage by traditional social security programs—most notably health insurance and pensions, but often also other benefits available to workers by virtue of their labor contract. This chapter focuses on the “flip side” of the informality coin from the perspective of the region’s workers, that is, access to social protection. Specifically, the chapter examines the state of workers’ access to social protection in the region; the role of public social protection—or social risk management—mechanisms; and how deficiencies in the design of social protection programs contribute to the persistence of informal employment, as well to as a larger failure to protect the most vulnerable members of society. It then discusses how the region’s social protection systems may be reengineered

to ensure that programs are incentive-compatible with productive employment, as well as with workers’ and their families’ need for protection against health shocks, poverty in old age, and other debilitating risks.

Informality and social protection: Why policy makers should care

Social protection has been defined in a number of different ways in the literature on social policy and poverty reduction. Most commonly, however, social protection is defined as a range of measures adopted by governments to help people manage risk more effectively—whether in the form of promoting basic income security, protecting people from unanticipated shocks (such as family health shocks or economic downturns), developing and protecting the human capital of society’s poorest members to strengthen their ability to prosper in the labor market, and/or ensuring basic service access to those outside the reach of traditional government (or private sector) programs. Social protection programs

typically focus on helping people prevent, mitigate, or cope with a range of risks. In doing so, programs that are well designed and successfully implemented can contribute not only to household welfare, but also to long-term economic productivity, growth, and development. Social protection is typically delivered via a range of social insurance and/or social assistance programs (box 7.1).

Nearly all Latin American countries are characterized by what have been called “truncated welfare systems.” These are social protection systems in which, historically, formal sector

actors (public and private sectors, employers and employees) have contributed to social security programs and have, in return, been covered by relatively generous, multidimensional benefit packages—often including health insurance, old-age pensions, disability and workers’ insurance, and, in some cases, housing, child care, and sports and recreation benefits. At the same time, those outside the formal sector, both in urban and rural areas, have had much more limited access to formal risk management instruments or other government benefits.

BOX 7.1

Social protection—strengthening people’s abilities to manage risk and promoting long-term productivity, growth, and development

People in developing countries face a range of risks. Some risks, such as economic recessions, harvest losses, natural disasters, and wars, affect whole societies or large groups. Others, such as the illness of family members, loss of household breadwinners’ jobs, and crime, may affect only individual households. *Social protection* is defined as a range of public interventions that support society’s poorest and most vulnerable members and help individuals, families, and communities better manage risks by helping them prevent, mitigate, or cope better with adverse events. Public social protection can comprise a range of mechanisms, including regulation, government financing, direct provision of services, or provision of conditional or unconditional transfer payments. Intended to augment, not replace, family, community, and market-based risk management mechanisms, such interventions complement national economic policies and support strategies for poverty reduction and human development.

Public social protection measures in Latin America and the Caribbean often are categorized into two main groups: *social insurance* and *social assistance*. Social insurance includes a range of contributory programs, including old-age pensions, health insurance, disability, and professional risk insurance, that are intended to help cushion the impact of shocks affecting income/earnings, health, and employment, and thereby prevent families from falling into poverty. These programs generally fall under the heading of social security. Social assistance includes a variety of noncontributory safety net programs, such as workfare, assistance to the disabled and indigent, and cash transfers, all of which help individuals

and families deal with temporary or chronic poverty at different stages of their life cycle, and/or strengthen their capacities to achieve higher standards of living.

Well-designed public social protection programs can compensate for missing insurance markets or other private risk-mitigation instruments, and thus they create opportunities among the poor for more productive investments and higher incomes. Moreover, certain types of safety net programs—such as “conditional transfers” in which payments are made contingent upon family investments in children’s health or schooling—both provide short-term income support and strengthen longer-term investments in children’s education, health, well-being, and productivity.

Other policies and institutions also play important roles in social risk management. Labor market policies and institutions play a critical role—by influencing the nature and extent of the risks workers face; by providing the framework in which certain programs, such as pensions, health insurance, or workfare are accessed; and by providing opportunities for skills development or technical training so that people can find more remunerative employment. Public health systems in developing countries often perform a safety net function, providing subsidized health services to the poor and others who lack health insurance. And where financial institutions function well, access to financial services can also contribute in important ways to households’ ability to manage risk, as a complement to public social protection programs.

Sources: de Ferranti et al. 2000; World Bank 2001, 2003b.

While some workers, on having done an implicit or explicit cost–benefit analysis, voluntarily opt out of the system of social security, others may not have that choice. For example, by virtue of residing in less favorable geographic locations (such as poor or remote rural areas), or lacking opportunities to work in the formal sector (their employers may not offer benefits), many do not have access to formal risk management instruments. Still others—those who through the course of their working lives frequently move between formal and informal employment—may find they are not eligible for some benefits during key stages in their life cycles due to program design features or eligibility requirements. Historically, there has been a strong correlation between income and social security coverage and, concomitantly, between poverty status and *lack of access* to formal risk management instruments.

There is ample evidence that individuals and families employ a number of strategies to manage risk—with varying degrees of success. If so, is lack of access to formal social protection programs really a problem? The evidence, which will be elaborated upon in this chapter, indicates “yes,” on at least three fronts. First, from the perspective of *household welfare*:

- While even poor, informal sector workers engage in private risk management strategies, the evidence indicates that households are only partially successful in protecting themselves against the impoverishing effects of shocks, whether caused by illness, disability, unemployment, or loss of income in old age.
- Poor and near-poor households generally have fewer assets and/or risk management instruments at their disposal, making them particularly vulnerable to shocks. Indeed, inadequate access to risk management instruments can lead to families engaging in harmful “coping” activities—for example, removing their children from school—that can make it more difficult for them to escape poverty, and serving to perpetuate poverty across generations.

Second, from the *societal perspective*:

- A growing body of evidence indicates that, in addition to adverse welfare effects, too much uninsured risk can have negative productivity and income effects at the microeconomic level (that is, house-

holds and firms), and even at the level of aggregate output (see, for example, Ravallion 2003).

- Failure to cover losses due to health or income shocks can impose external costs on society. Indeed, the existence of externalities, along with other market failures (for example, information failures), provides a powerful rationale for public intervention to improve societal outcomes.

Finally, from the perspective of the *social protection system*, the truncated welfare systems of Latin America and the Caribbean have tended to result in small, inefficient risk pools and forgone savings, due to the absence of scale economies.

Partly in response to concerns about the truncated welfare state and partly due to political pressures associated with democratization, several countries in the region have launched or expanded noncontributory assistance and/or targeted poverty reduction programs over the last decade to help breach the coverage gap. These programs have often contributed in important ways to the welfare of the poor. At the same time, some of these programs have inadvertently brought with them their own set of challenges from the perspective of the labor market. As this chapter will discuss in further detail, problems in the design of social protection programs—and of the social protection system as a whole—can serve to generate adverse incentives in the labor market, stimulating informality and leaving people without adequate protection from key risks. They can also have adverse effects on economic productivity and growth.¹

Thus, the design of social protection—both the traditional “Bismarckian” social security system linked to the labor contract and subsequent noncontributory assistance programs—appears to have generated a number of critical challenges to ensuring access of country populations to adequate risk management instruments, particularly among the poor. In this context, this chapter

- reviews the state and recent evolution of social protection in Latin America and the Caribbean;
- examines the role and limitations of households’ private risk management strategies and outlines the rationale for public social protection in the region;
- analyzes the key challenges policy makers face in making adequate risk management instruments available to the regions’ citizens working in both the formal and informal sectors; and

- outlines key policy directions for ensuring that the populations of Latin America and the Caribbean are adequately protected from key risks.

In developing directions for policy, the chapter builds upon the lessons from recent World Bank studies on social protection policy, based on the concept of social risk management and on the economics of insurance.² Individuals and societies can respond in a variety of ways to the prospect of economic losses associated with such shocks as illness, disability, or loss of income due to old age, and the

social risk management concept organizes people's responses to shocks into three broad categories: prevention (ex ante), mitigation, and coping (ex post) (see Holzmann and Jorgensen [2000] and World Bank [2001]). The *comprehensive insurance framework*³ offers a tool for determining which mitigation instruments and preventive measures will be most effective, given the expected size, frequency, and extent of externality of a range of possible financial losses; the framework can also be used to identify when coping is an efficient, ineffective, or even damaging course of action (box 7.2).

BOX 7.2

The comprehensive insurance framework—providing guidance on social protection policy making

In classical theory, individuals facing the likelihood of financial loss from an adverse event can either insure against such a loss or take steps to lower the likelihood that the loss will occur. The challenge that people face, therefore, is to choose the optimal mix of market insurance, self-insurance, and self-protection. Both market and self-insurance transfer income from the “good” states to the “bad” states of the world. Market insurance pools risks across individuals, compensating for differing risks among them. Where market insurance is available, it can be purchased at a price—the insurance premium—which reflects the size of the prospective loss and the probability of the bad state coming about. Self-insurance—essentially individual or family saving—does not involve risk pooling or compensation for risk differentials. Although it has no explicit price, its cost can be imputed from the expense people incur to save, for instance, in forgone consumption. Self-protection refers to measures that individuals or households take, ex ante, to prevent an adverse event from occurring. Self-protection reduces the probability that losses will occur, and may reduce the size of a loss, should one occur. For simplicity, the literature often refers to market insurance as *risk pooling*, to self-insurance as *saving*, and to self-protection as *prevention*. When individuals and households do not insure through risk pooling or saving, or engage in prevention, they are often forced to cope with the costs of losses in the wake of a shock.

The comprehensive insurance framework provides guidance on which mitigation instruments—risk pooling or saving—and preventive measures will be most effective, given the expected *size*, *frequency*, and *extent of externality* of

a range of possible financial losses; the framework can also be used to identify when coping is an efficient, ineffective, or even damaging course of action. From the perspective of financial protection, when prospective losses are small and infrequent, it is more efficient for individuals to *cope* with the loss after the fact than to *insure*. In that sense, full insurance is not efficient. But as prospective losses become more frequent, it becomes relatively more efficient to engage in *prevention*—to lower the probability of the loss occurring—and *saving*—to cover the costs of the loss. When a prospective loss becomes less frequent but increases in size, it becomes more efficient to engage in *risk pooling*. For rare but large prospective losses, households also have incentives to engage in *prevention*, again to lower the likelihood that the adverse event will occur. For losses that are both frequently occurring and catastrophic in size, there is little that individuals, households, or even markets can do on their own; specific measures to create a larger risk pool are required.

When markets fail or are missing, policy intervention can be justified to strengthen people's ability to manage risk. While insurance markets may fail due to the nature of risk (prospective losses are both frequent and large), they also commonly fail due to information problems. On one hand, individuals may lack critical information on the nature and extent of the risks that they face, leading to a lack of demand for appropriate risk management instruments; on the other hand, problems of “adverse selection” and “moral hazard,” well-known in the insurance literature, can raise the price of risk mitigation instruments, pricing the poor out of the market, or cause

markets not to form at all. Finally, the failure of individuals or households to manage risks appropriately can often impose negative externalities on society. For example, in countries where a significant number of people fail to insure, governments face a dilemma in which politicians cannot credibly refuse to provide aid when a large number of people suffer losses, and the burden of these losses is transferred onto current and future taxpayers. Similarly, in the case of health, communicable diseases that go untreated due to lack of insurance can have important negative spillover effects on large swaths of society. Indeed, the degree of externality posed by health risks and the public-good nature of many health treatments and interventions present a powerful additional justification for policy intervention.

The comprehensive insurance framework provides guidance for public intervention regardless of the prospective loss under consideration (for example, due to illness, disability, unemployment, or income loss in old age):

- *Government should provide (or help provide) the instruments that the market cannot (or will not) provide.* Risk pooling to cover certain losses (for example, lost earnings from becoming unemployed, the risk of poverty, disasters, and certain frequently occurring or preexisting health events with catastrophic costs) does not exist in many contexts due to information problems. Government can step in to correct market failures by providing risk-pooling instruments.
- *Government should provide (or help markets provide) superior instruments where only inferior instruments are available.* For risks best covered with individual savings, private agents may turn to “bad” saving instruments (for example, property or other nonliquid assets for precautionary saving) because “good” instruments (such as diversified financial assets; safe, reliable, and competitively priced forms of liquid savings; or credit) are not available. Moreover, poorer households simply may not have much margin to save. Government can intervene to foster the development of more efficient instruments for saving through prudential regulation of capital, credit, and insurance markets, as well as provide direct subsidies for households that are too poor to hold savings or debt.
- *Government should help households build and protect their human capital.* Investing in human capital—education, hygiene, and primary and preventive health care—can be an effective and powerful means of prevention and can strengthen families’ abilities to cope effectively. However, where credit is constrained, individuals may under-invest in human capital. To ensure appropriate human capital development, governments can subsidize investments in education and health for the poor. Due to the nature of health shocks and the presence of externalities, government can support work to broaden and strengthen risk-pooling instruments.
- *It is less costly to help households mitigate losses than to cope with them.* The instruments for individuals and families to pool risks and save are not always available. The resources to take preventive measures are often scarce. Where individuals and families are constrained, adverse coping can result. Some prevention, risk pooling, and saving are always desirable. Effective policy should thus place priority on enabling individuals to insure against losses through risk pooling and saving and to lower the probability of losses through prevention, rather than merely coping after a shock.

While market failures and externalities justify state intervention, the comprehensive insurance framework also highlights the risk of “government failure.” This can occur when governments “overreach,” that is, where the state tries to do too much by providing coverage that households and markets can—or should—provide themselves. Examples of this include when governments subsidize pensions for upper- or middle-class groups or subsidize insurance for small and frequent health events (phenomena that have been common in Latin America and the Caribbean). Overreach can lead to unsustainable fiscal burdens, even when coverage rates are low. Government interventions can also be “misaligned,” that is, focused on the wrong mix of instruments given the expected size and frequency of loss, or excluding clear “insurables” (such as low probability, high expected loss events) in mandatory risk-pooling packages.

Source: Adapted from Baeza and Packard 2006; Gill and Ilahi 2000; Gill, Packard, and Yermo 2004; and Packard 2006.

The state of social protection in Latin America and the Caribbean

The truncated welfare state in Latin America and the Caribbean has stemmed largely from what is often called the “Bismarck” model of social security, which has been developed in most countries in the region since the early part of the 20th century. The core of the Bismarck model is that social security coverage is based on the form of the worker’s labor contract. In general, workers acquire rights to a package of benefits—health insurance, disability, pensions, and so on—via employment in the formal sector of the economy rather than by virtue of working per se (or by virtue of a more inclusive concept, such as “citizenship”). Such benefits are typically financed through a combination of payroll taxes on the part of the worker, employer contributions on the part of the firm, and, in many cases, subsidies from the government. In some instances, non-wage costs of employment (payroll taxes plus other benefits) can be very high, affecting both employers’ and workers’ informal-formal sector decisions. In Colombia, for example, total non-wage costs have been estimated at 53 percent of payroll, the highest in the region, followed by Mexico whose

non-wage costs account for over 47 percent of payroll (see Mason et al. forthcoming).

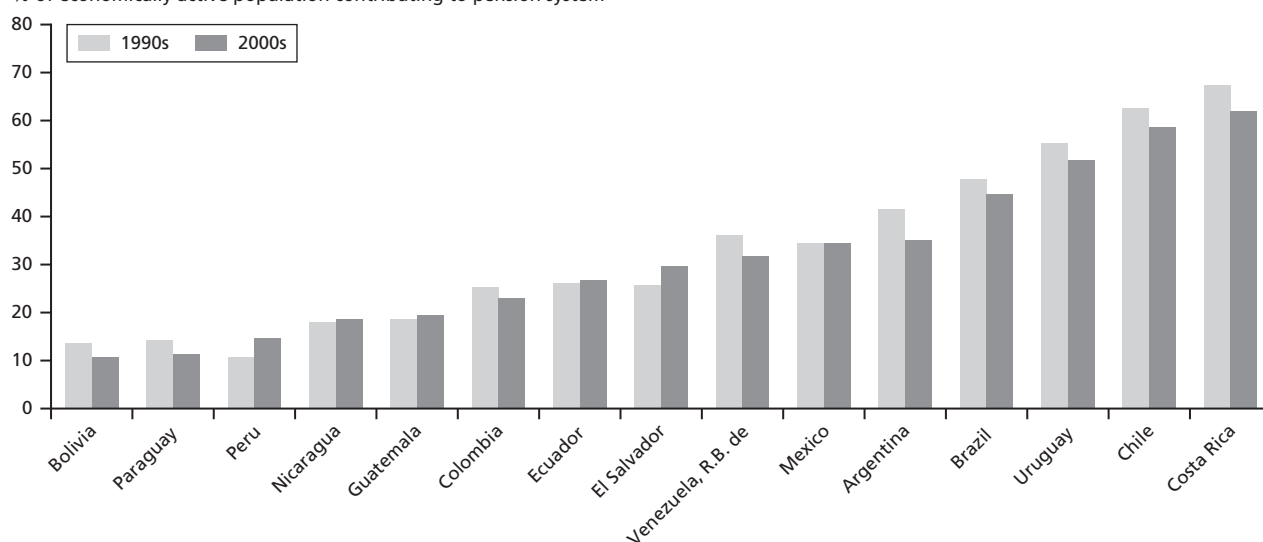
In response to fiscal pressures, as well as concerns about low coverage and persistently large informal economies, a number of reform measure changes have been undertaken in Latin American and Caribbean countries since the 1980s, particularly with respect to pensions. This has improved the fiscal sustainability in some countries, but increases in coverage have remained elusive. As discussed briefly in chapter 1, social security coverage in most Latin American countries remains relatively low, with limited progress recorded over the last decade (figure 7.1). Indeed, in examining changes in pension coverage between the mid-1990s and the early to mid 2000s, Rofman and Lucchetti (2006) find that in 9 out of 15 countries (for which comparable household survey data exist) coverage rates have actually declined over time, and even in the few countries where progress has been made (for example, Peru), advances have been small.

As discussed in chapter 4, a combination of factors—macroeconomic, structural, and demographic—appears to have contributed to observed patterns since the 1990s. This

FIGURE 7.1

Pension coverage rates in Latin America and the Caribbean

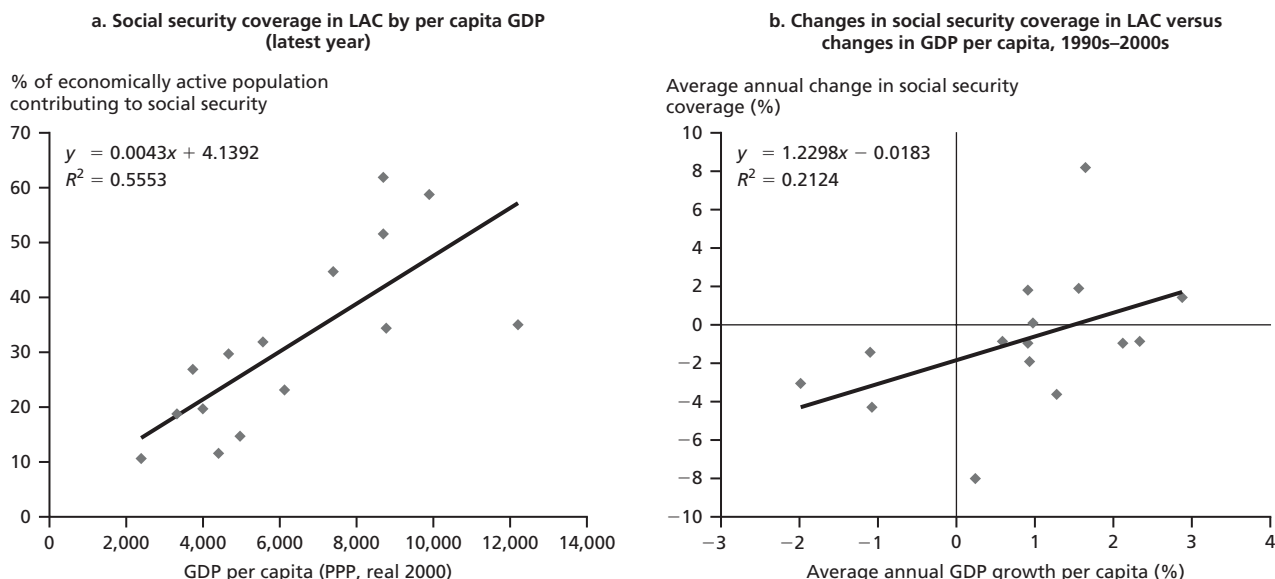
% of economically active population contributing to pension system



Source: Rofman and Lucchetti 2006.

Note: The years for which household survey data are available across countries in the region are not identical. As such, Rofman and Lucchetti (2006) present figures based on available data in each country that comes closest to the 1995–2004 period. The years used are as follows: Argentina 1995–2004; Bolivia 1999–2002; Brazil 1995–2002; Chile 1996–2003; Colombia 1996–99; Costa Rica 1995–2004; Ecuador 2000–04; Guatemala 1998–2000; Mexico 1998–2002; Nicaragua 1998–2001; Paraguay 1999–2004; Peru 1999–2003; El Salvador 1995–2003; Uruguay 1995–2004; Republica Bolivariana de Venezuela 1995–2004.

FIGURE 7.2

Social security coverage and GDP per capita

Sources: Author's calculations. Data on social security coverage from Rofman and Lucchetti (2006); data on GDP per capita purchasing power parity from *World Development Indicators*, various years.

includes increases in labor force participation rates (particularly among females) during the 1990s, growth in nontradable sectors driven by currency appreciations in the early 1990s, and increased labor market rigidities, at least in Colombia and Brazil.⁴ In addition, social security reforms themselves may have lowered the perceived benefits of the affiliation relative to its costs. In the case of pensions, for example, reforms that were designed to increase fiscal sustainability often raised contribution rates, minimum contribution periods, and retirement ages, while lowering replacement rates. These changes in the relative benefits and costs of pension schemes may have more than offset the salutary effects associated with closer links between contributions and payments. As will be discussed further in this chapter, a number of factors related to the design of social security programs and to the incentives created by the broader constellation of social protection programs—both social insurance and social assistance—appear to affect coverage levels.

Data indicate that social security coverage has failed to increase (informality has failed to decline) in the region despite net economic growth over the period. As was highlighted in chapter 1, global data show that informality

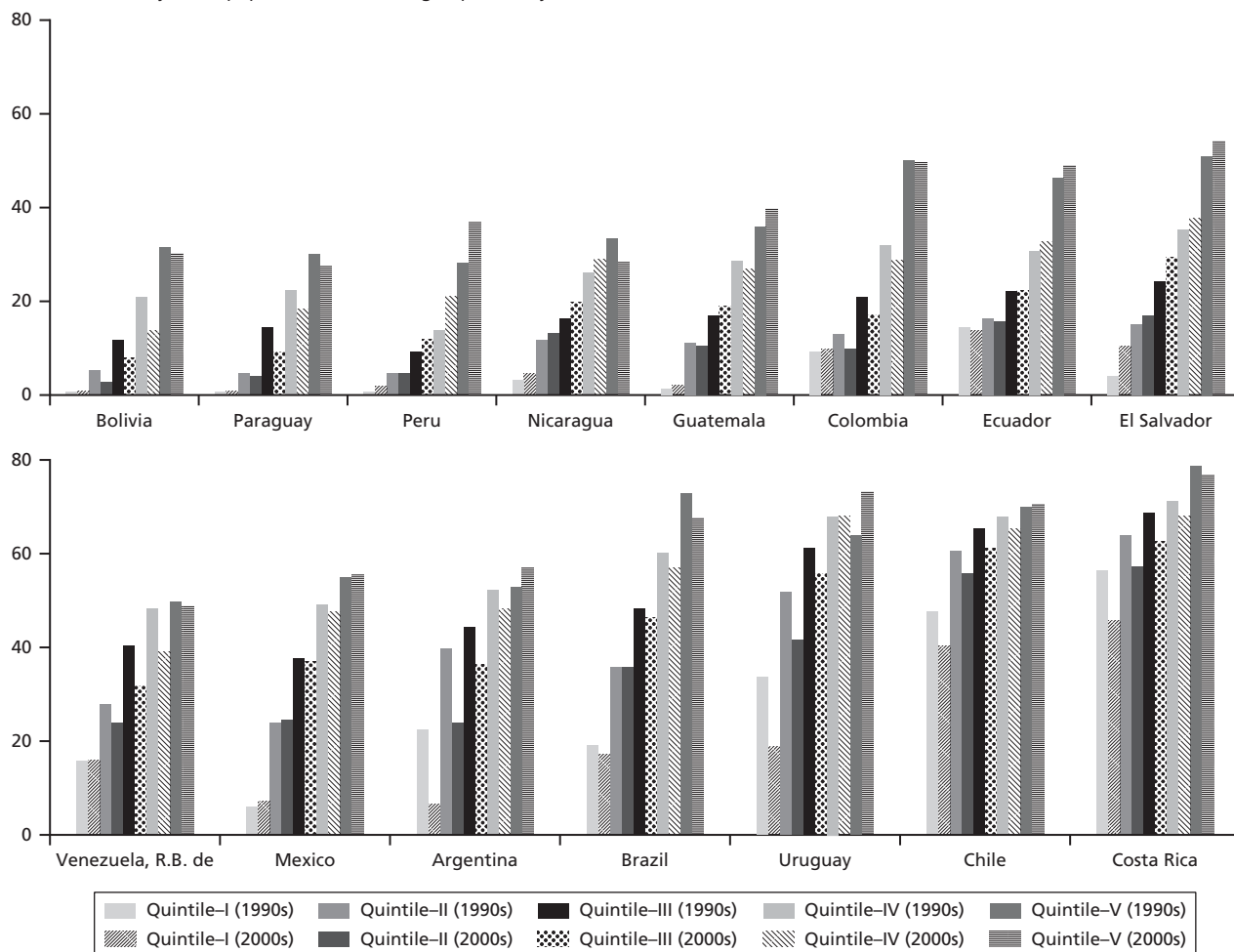
tends to decline (social security coverage tends to rise) as national income rises. Indeed, the same relationship between gross domestic product (GDP) per capita and social security coverage that holds globally is observed when only Latin American and Caribbean countries are analyzed. This is true whether one looks either at GDP levels or GDP growth (figure 7.2a and b). As can be seen in panel b of the figure, however, several countries that experienced positive economic growth did not increase social security coverage between the mid-1990s and the early to mid 2000s; in fact, coverage actually declined in several positive-growth countries. This suggests that the factors leading to increased informality—whether macroeconomic, structural, demographic, or related to social protection design—frequently offset the positive effects of growth.

Consistent with the evidence presented in the earlier chapters, in nearly all countries in the region, coverage rates are significantly lower among low-income than high-income workers (figure 7.3). In most of the region's countries, the poorest are practically excluded from the system. The situation is similar when considering the employed population, showing that pension systems suffer from major inequities in terms of access, even if differential unemployment is

FIGURE 7.3

Coverage rates in Latin America and the Caribbean, mid-1990s to early/mid-2000s, by quintile of per capita income

% of economically active population contributing to pension system



Source: Rofman and Lucchetti 2006.

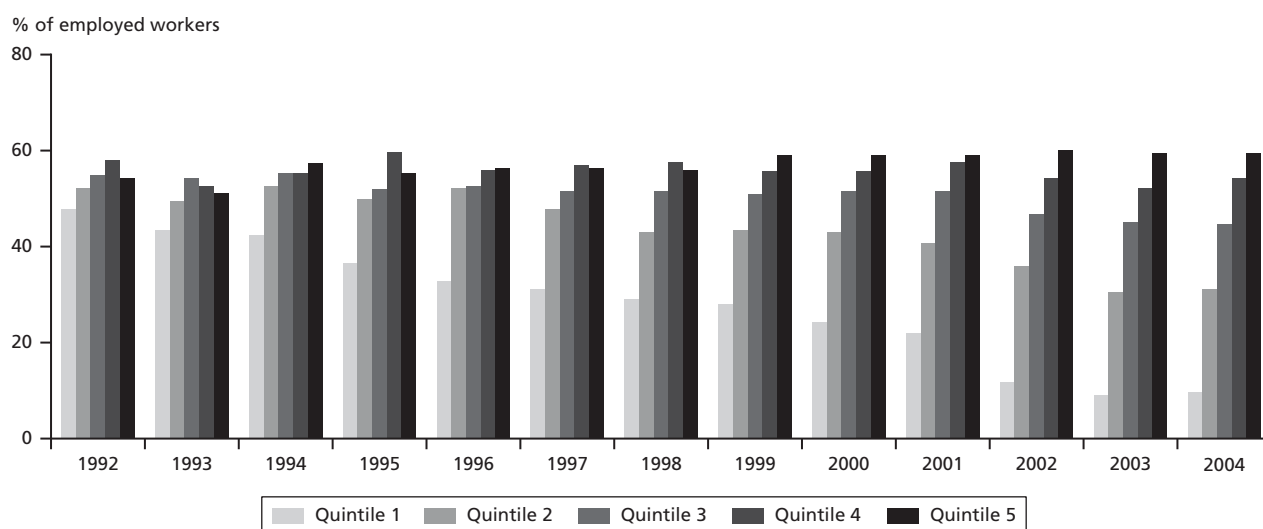
ignored. Most countries have coverage rates that are above 50 percent for the highest quintile, but none has a rate higher than 50 percent among the poorest workers. Analysis of coverage rates by worker education levels shows similar patterns (see Rofman and Lucchetti 2006).

In addition, inequalities in access to social security have tended to increase over time (figure 7.3). In Argentina, Uruguay, Chile, and Costa Rica, increases in inequality occurred, at least in part, because of noticeable declines in coverage among those in the poorest quintile. Perhaps the most dramatic instance of this occurred in Argentina (figure 7.4). While coverage rates in Argentina actually increased slightly among those in the top quintile (from 54

to 59 percent of employed workers), it declined among all other income groups. Declines in coverage were largest among those in the lowest quintile (from 48 to just under 10 percent [Rofman and Lucchetti 2006]).

Beginning in the 1990s—and partly in response to concerns about limited access to formal social security programs by large swaths of the region's population—a number of countries in Latin America and the Caribbean have launched or expanded poverty reduction and/or social assistance programs to help provide support and coverage to the poor or extreme poor (for example, Argentina, Bolivia, Brazil, Chile, Mexico, and Peru, among others). These efforts have taken a number of forms—including

FIGURE 7.4
Pension coverage rates in Argentina, by quintile of per capita income



Source: Rofman and Lucchetti 2006.

subsidized health care and insurance (*Acceso Universal para prestaciones integrales y Garantías Explícitas* [AUGE] in Chile, the *Regimen Subsidiado* in Colombia, and *Seguro Popular* in Mexico), noncontributory pension programs for the poor (*Bono Solidario* [BONOSOL] in Bolivia, the Rural Pension Scheme in Brazil, and *Pension Asistencial* [PASIS] in Chile), workfare programs (*Jefes de Hogar* in Argentina), and conditional cash transfer programs (*Bolsa Escola* in Brazil, *Familias en Acción* in Colombia, and *Oportunidades* in Mexico).

Figure 7.5 presents the coverage of the elderly in Latin America between the mid-1990s and the 2000s. As with participation of the working-age population in formal social security schemes, formal pension coverage among the current elderly population is extremely low in many countries in the region; rates are at 60 percent or higher only in Costa Rica, Argentina, Bolivia, Chile, Uruguay, and Brazil. While formal pension coverage has increased in several countries since the 1990s—for example, in Colombia, República Bolivariana de Venezuela, Costa Rica, and Panama—some countries have actually registered declines in coverage of the elderly. Coverage in Argentina shows a particularly large decline, falling from 77 percent in 1995 to 65 percent in 2004.

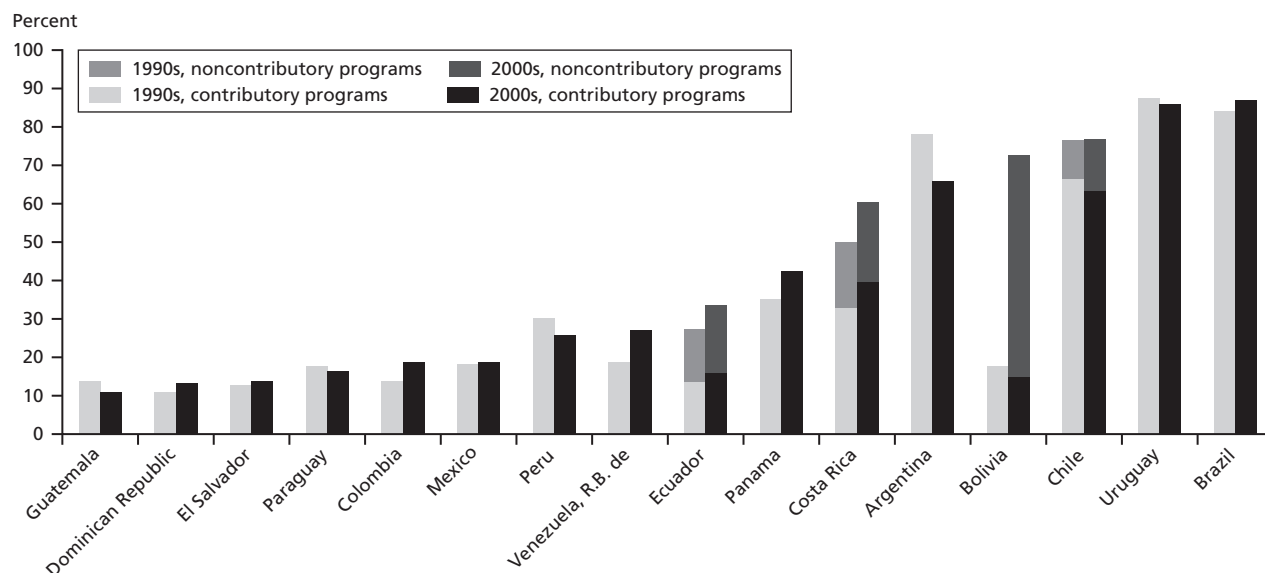
In this context, several countries in the region operate noncontributory pension programs, some of which have had an important impact on coverage of the elderly.

Uruguay and Argentina have small noncontributory benefits that cover some impoverished individuals, age 70 or older, who do not qualify for contributory retirement benefits. In Brazil, there is a large quasi-noncontributory system that covers rural workers.⁵ Other countries, such as Mexico, have some noncontributory schemes at the subnational level, including a program that covers the elderly in Mexico City. Unfortunately, data available from household surveys do not generally distinguish between beneficiaries of contributory and noncontributory programs (so most bars in figure 7.5 only show total coverage, not what proportion of the coverage is due to the contributory, quasi-, or noncontributory programs). There are a few countries—including Bolivia, Ecuador, Chile, and Costa Rica—where data on coverage under contributory and noncontributory programs are collected. As can be seen, where coverage of noncontributory programs is measured, they can often have a substantial impact on coverage of the elderly. Most dramatic is the case of Bolivia's BONOSOL program, which extends pension coverage to 58 percent of the elderly who are not otherwise covered by formal pension programs.⁶

Beyond pensions, several countries have also introduced programs to provide subsidized health insurance (for example, AUGE, Chile; *Regimen Subsidiado*, Colombia; *Seguro*

FIGURE 7.5

Pension coverage among the elderly in Latin America and the Caribbean, contributory and noncontributory programs



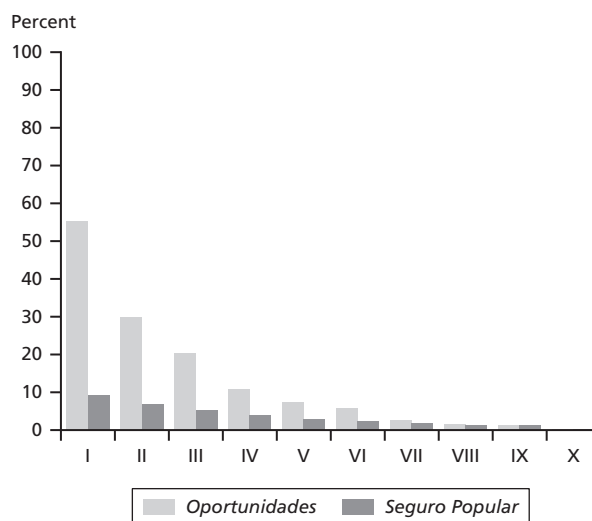
Source: Rofman and Lucchetti 2006.

Popular, Mexico) and/or conditional cash transfer programs focused on strengthening the human capital of the poor (for example, *Bolsa Escola*, Brazil; *Familias en Acción*, Colombia; *Red Solidaria*, El Salvador; *Programa de Asignación Familiar*, Honduras; Programme Advancement Through Health and Education, Jamaica; *Oportunidades*, Mexico; *Red de Protección Social*, Nicaragua). In some cases, these programs have had considerable outreach to the poor. In Mexico, for example, the *Oportunidades* program (formerly known as *Progres*a) has reached roughly 5 million very poor families, about two-thirds of whom live in rural Mexico.

As can be seen in figure 7.6, poverty-targeted programs such as *Oportunidades* and *Seguro Popular* in Mexico have done a much better job of reaching the poor than traditional (contributory) social security programs. While analysis of household survey data (as shown in figure 7.3) indicates that the distribution of formal social security benefits is very regressive, coverage of *Oportunidades* and *Seguro Popular* is highest among the poor. Indeed, around 55 percent of households in the poorest decile were covered by *Oportunidades* in 2004, and coverage declines significantly as household income rises (figure 7.6). Coverage of *Seguro Popular* is similarly progressive, albeit at lower levels of coverage in 2004; nonetheless, coverage of the program has grown considerably since that time.⁷

FIGURE 7.6

Coverage of two poverty-targeted assistance programs in Mexico: *Oportunidades* and *Seguro Popular* (by decile, 2004)



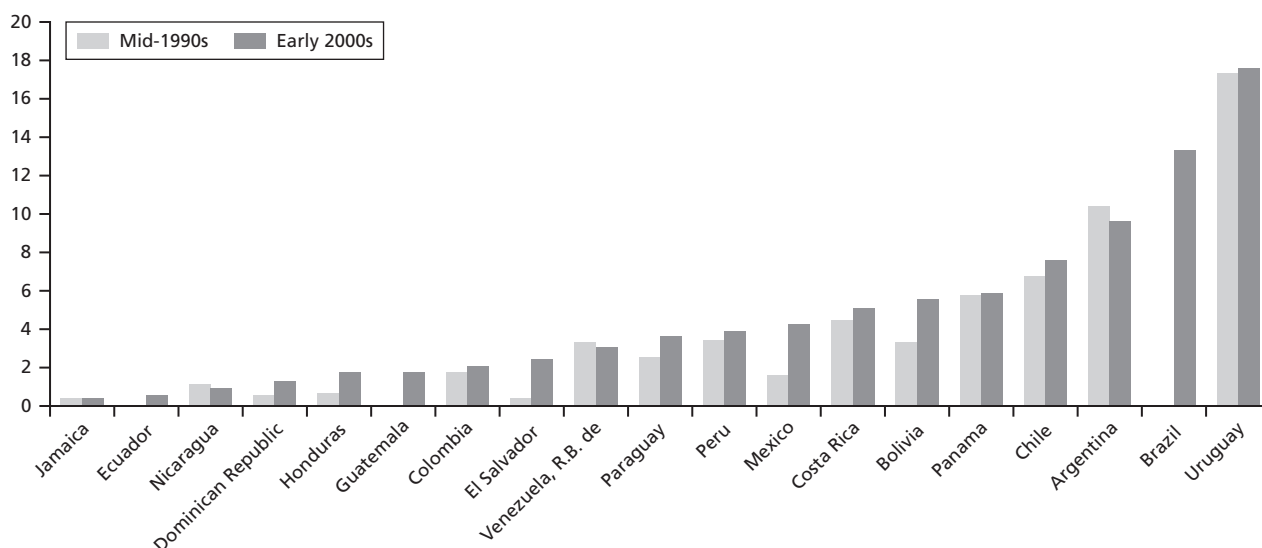
Sources: Adapted from Mason et al. forthcoming, based on data from Mexico's 2004 national income and expenditure survey, *Encuesta Nacional de Ingresos y Gastos de los Hogares*.

Note: Some *Seguro Popular* beneficiaries also participate in *Oportunidades*. Seventy-two percent of *Seguro Popular* beneficiaries participate in *Oportunidades* in the first decile, but the overlap decreases very rapidly with income.

FIGURE 7.7

Public social protection spending in Latin America and the Caribbean

% of GDP



Source: World Bank staff calculations.

Note: No mid-1990s data were available for Brazil, Ecuador, and Guatemala.

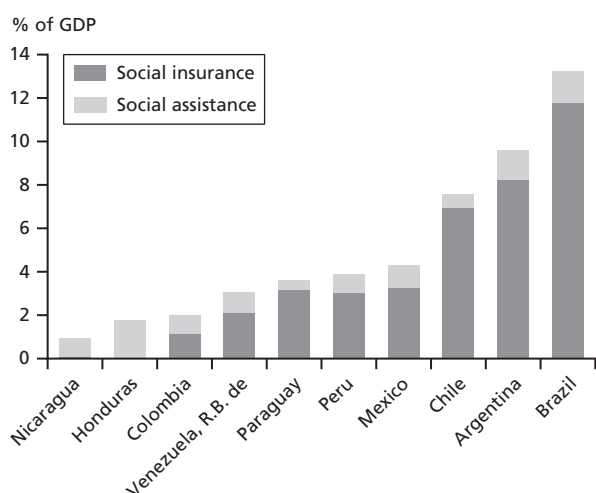
A recent study on income redistribution and public transfers in Latin America and the Caribbean shows that, in general, social assistance programs have done better at reaching the poor than have traditional social security programs, such as health insurance or pensions (Lindert, Skoufias, and Shapiro 2006). It is important to note, however, that the quality of targeting varies across types of programs and across countries, and that not all assistance programs are targeted in a progressive manner. Scholarships, for example, as well as many food-based programs in the region, often lack effective, pro-poor targeting and are therefore more likely to benefit nonpoor families and individuals than the (intended) poor.

Regional efforts to extend social protection programs more effectively toward the poor can be seen, in part, in the figures on public spending on social protection between the 1990s and the 2000s. Average levels of public spending on social protection in Latin America and the Caribbean increased over the period, from around 4.0 to 4.8 percent of GDP (figure 7.7). This is compared to 14.4 percent in the Organisation for Economic Co-operation and Development (OECD), 8.3 percent in the United States, and 16.3 percent in continental Europe.⁸ Clearly, the data show significant variations in the levels of spending across Latin

American and Caribbean countries. For the countries in which disaggregated data exist, it can be seen that spending on social insurance (that is, formal social security) generally make up the lion's share of spending; with a few exceptions (such as Nicaragua and Honduras), spending on social insurance accounts for more than two-thirds of total social protection spending (figure 7.8). At the same time, existing data suggest that the share of spending going to social insurance has declined slightly in several countries since the 1990s, as the prevalence of poverty-oriented assistance programs has increased.

This growth of social assistance programs over the last 10 to 15 years has had at least two important effects on social protection in the region. First, an increasing number of poor people have access to state benefits that play an important social protection or poverty reduction function—whether to help the poor build their human capital, protect them against basic health risks, or provide basic income security in old age. These benefits can often have an important positive impact on the recipients, even though the unit transfers tend to be considerably smaller on a per-person (or per-household) basis than formal social security benefits (Lindert, Skoufias, and Shapiro 2006). Indeed, there is a growing empirical body of literature on the

FIGURE 7.8
Public spending on social insurance and social assistance in Latin America and the Caribbean, early 2000s



Source: World Bank staff calculations.

positive impacts of conditional cash transfer programs, such as *Oportunidades* in Mexico or *Familias en Acción* in Colombia, on human capital outcomes of the poor (see Rawlings and Rubio [2005] for a recent summary of impact evaluation findings). Similarly, Brazil's rural pension program has contributed to observable declines in old-age poverty in that country (World Bank forthcoming).

At the same time, the growth of social assistance programs has led to an increasing fragmentation in social protection systems in the region, in which the poor (and some nonpoor) receive similar—albeit more modest—benefits for free as formal sector workers receive by virtue of their payroll contributions. As will be discussed later in more detail, such well-intended efforts to make social benefits available to those outside the formal sector may—due to lack of incentive compatibility with contributory programs—be creating disincentives to the formalization of the workforce. To the extent that such disincentives exacerbate distortions in the labor market, they may also have adverse productivity effects.⁹

Private risk management and rationale for public social protection

While informal employment means that workers are not directly affiliated with formal social security programs, this is not to say that informal workers are left completely

without risk management instruments. In some cases, workers may have access to health insurance or other formal risk management instruments by virtue of the employment status of a spouse, a parent, or other relative.¹⁰ Moreover, individuals and households in Latin America and the Caribbean typically engage in a range of private risk management strategies to prevent, mitigate, or cope with a variety of risks and shocks. While, in principle, private risk management strategies could include the purchase of private insurance (for example, health, disability, and so on), in practice, private insurance markets remain relatively thin in the region, and access to private insurance instruments is generally confined to the wealthier segments of societies. In this context, a number of private risk management strategies are identified in the literature for Latin America, including income diversification, adjustments of household labor supply, drawing down of household savings, sale of assets, and adjustments in household spending and/or consumption patterns. Such strategies are used both before the fact to reduce the likelihood of an adverse shock (or to mitigate the likely impact of a shock, should one occur), and after the fact to help cope with or deal with the effects of an adverse event.

Evidence from Mexico, for example, indicates that households send additional members into the labor force in response to real or expected employment shocks (Cunningham 2001). Evidence from Guatemala, Nicaragua, El Salvador, and Honduras shows not only that family members increase their hours worked to mitigate the impact of an adverse event, but that they draw down financial savings or other assets, if necessary, to protect their income and consumption levels (World Bank 2003a). Migration and remittances also make up a key element of household risk management in a number of Latin American and Caribbean countries—both ex ante, as preventive measures, and ex post, to soften the impact of a shock (see Arias [2004] and Beneke de Sanfeliú and Shi [2004] for evidence from El Salvador). In Argentina, during the 2002 economic crisis, families altered their consumption patterns, buying fewer luxury goods and spending less on necessities, educational materials, and children's health visits (World Bank 2003a). When households' abilities to cope with shocks are stretched, as they were during the 2002 crisis in Argentina or after hurricane Mitch in Nicaragua, families often rely on wider social networks, including memberships in community, religious, or neighborhood organizations, that can provide an alternative

source of resources—as loans or gifts—in the event of an adverse shock (World Bank 2003a; Klugman, Kruger, and Withers 2003).

It is important to note that while some household risk management strategies, such as developing diversified income-earning portfolios, increasing adult labor supply, or drawing down financial savings, may be seen as appropriate and rational responses to the reality of risk and shocks, other strategies, such as engaging in distress sales of productive assets like land, withdrawing children from school, or deferring utilization of preventive or curative health services, may generate their own risks to long-term family productivity and welfare. During the 2002 crisis in Argentina, spending less on education and on children's medical care raised critical concerns about the long-term effects on children's human capital and long-term family welfare (World Bank 2003a). Moreover, a recent study of Mexico indicates that children who are removed from school in response to a shock are one-third less likely to continue school than those who find ways to continue their education in the face of a shock (Sadoulet et al. 2004). Thus, in the absence of adequate insurance or other formal risk management mechanisms, families can risk long-term losses in their human capital, adversely affecting their economic productivity and increasing the likelihood of intergenerational transmission of poverty.

A number of recent empirical studies—within and outside Latin America—have tried to measure how effectively households smooth their consumption in the face of adverse income shocks. Specifically, these studies measure the extent to which a shock to household per capita income translates into a shock to household per capita consumption. While the specific findings differ from country to country, these studies find that households are partially—but not fully—effective at mitigating the impacts of shocks to their incomes. Indeed, recent evidence from several Latin American countries, including Mexico, Nicaragua, and Peru, suggests that households in these countries are able to protect (or insure) only between 60 and 75 percent of their per capita consumption in the face of an income shock (Glewwe and Hall 1998; Klugman, Kruger, and Withers 2003; Skoufias 2004). In other words, a 10 percent shock to a household's per capita income translates into a 2.5–4 percent decline in its per capita consumption on average.

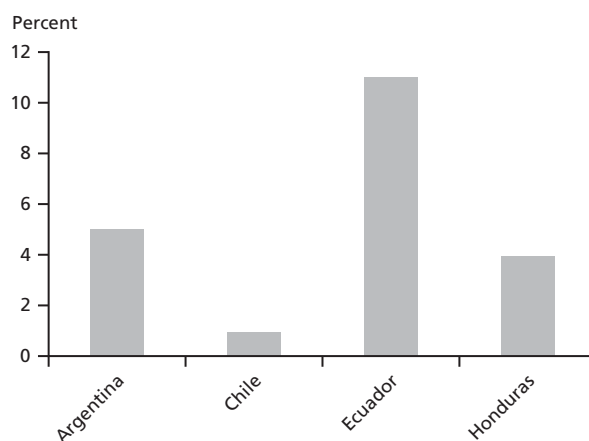
The evidence also indicates that poor households are often less able to smooth their consumption in the face of shocks than nonpoor households. A recent study of income

dynamics in rural El Salvador indicates, for example, that poor households take considerably longer to recover from an income shock than nonpoor households (Rodriguez-Meza and Gonzalez-Vega 2004). Evidence from rural China also indicates that income shocks have deeper impacts on poor households (Jalan and Ravallion 1999). While an income shock of 10 percent had only a 1 percent impact on the consumption of the wealthiest households, it caused a 4 percent decline in consumption among the poorest households.

Indeed, despite the best efforts of workers and their families to protect themselves from risk, households that lack access to formal risk management instruments remain vulnerable to the impoverishing effects of shocks. The evidence suggests that families are particularly vulnerable to the effects of health shocks (Baeza and Packard 2006; Skoufias 2004). In addition to treatment costs, households bear the cost of productive time lost from work, as well as the opportunity costs due to days spent taking care of family members who are ill. Costs associated with illness can drive families into poverty; for the already poor, the costs of a health shock can make escape from poverty more difficult. Evidence from Argentina, Chile, Ecuador, and Honduras indicates that the impoverishing effects of health shocks can be significant. In Argentina, 5 percent of all nonpoor households fell into poverty for at least three months in 1997 due to out-of-pocket spending on health (figure 7.9).

FIGURE 7.9

Nonpoor population who fell below the national poverty line due to out-of-pocket health expenditures



Sources: Baeza and Packard 2006, citing Maceira (2004) for Argentina; Bitran, Giedion, and Muñoz (2004) for Chile; Montenegro (2004) for Ecuador; and Fiedler (2004) for Honduras.

In Ecuador, 11 percent of nonpoor households fell below the national poverty line for at least three months in 2000 due to the direct costs of health care. In Honduras, in 2000, 4 percent of nonpoor households fell into poverty, at least temporarily, due to health-related spending.

Cross-country analysis of public and private spending on health indicates that out-of-pocket spending on health care is relatively high in Latin America and the Caribbean, as compared to other regions (Baeza and Packard 2006).¹¹ In Latin America and the Caribbean, 85 percent of all private spending on health care represents out-of-pocket spending by households. This compares with 72 percent, on average, in Europe and the OECD (which also have lower levels of private health spending, as a share of total health spending, than does Latin America). Moreover, the poor generally pay a higher share of health costs out of pocket than the nonpoor. And while participation in well-designed and well-functioning risk-pooling schemes reduces the likelihood of falling into poverty in the wake of a health shock, relatively few of the poor (or near-poor) in the region participate in effective risk pooling (Baeza and Packard 2006).

Another result of low social security coverage in the region is that poverty among the elderly in Latin America and the Caribbean is often very high. A recent 19-country study on old-age security found that the incidence of poverty among the elderly in the region, measured in terms of household per capita income, is often significantly higher than among the population as a whole (Bourguignon et al. 2004). Indeed, this was found to be the case in 11 out of 19 countries studied. Elderly who fall outside the formal social security system find themselves at particular risk. Simulation analysis carried out by the authors suggests, however, that programs that provide minimum income support to the elderly poor could have an observable, positive impact on old-age poverty. This is consistent with real-life experiences in Brazil and Bolivia. In Brazil, as noted above, the rural pension program is credited with contributing to observable declines in old-age poverty in recent years.¹² A recent study on Bolivia indicates that BONOSOL payments have had a positive and significant impact on food consumption among recipient households, particularly in rural areas (Martinez 2005).

The role of the state

Together, the evidence suggests that although households rely on a number of private mechanisms for risk management, this reliance on informal mechanisms does not afford

many families adequate protection from impoverishing risks and shocks. In this context, what is the role for public social protection policies and programs? From an economic perspective, if private markets for insurance (or other risk-pooling or savings instruments) existed and functioned well, if actors in the market for risk management instruments had adequate information, and if there were no externalities associated with inadequate risk management, then there may not be a compelling reason for the state to intervene. But private insurance markets are often missing—or are extremely thin—in Latin America and the Caribbean.¹³ Information problems abound, as do negative externalities associated with insufficient insurance.

In terms of information, workers and their families commonly lack sufficient knowledge of the nature and extent of the health risks that they face, as well as the likely direct (and indirect) costs associated with illness of the family breadwinner or another family member. They also have less-than-complete information on what their future needs will be—for example, at retirement. This latter issue causes what is commonly referred to as “myopia” in the pensions literature. Information problems—such as adverse selection and moral hazard—well-known in the economics of insurance literature, similarly plague those who would supply insurance or other risk-pooling instruments. Such problems can work to raise the price of risk mitigation mechanisms beyond the reach of low-income groups; information problems are also notorious for causing the private market for risk-pooling to fail—and even preventing these markets from forming in the first place (Packard 2006). Information-related problems are often aggravated in low-income, high-poverty environments.¹⁴

Insufficient risk management instruments can also have important negative externalities that argue for public intervention. An individual’s failure to manage risk can impose costs on others in society in one of several ways. First, in countries where a significant number of people fail to insure, governments face a “Samaritan’s dilemma”; politicians cannot credibly refuse to come to the aid of a large number of people who suffer a loss, and the burden of these losses is transferred onto current and future taxpayers (Packard 2006). But the external costs of individuals who are not adequately insured can extend well beyond the tax burden; in the case of health, in particular, communicable diseases that go untreated due to lack of insurance (or financial protection) can have health, and therefore, basic welfare impacts on large swaths of society.¹⁵

In sum, missing or failed markets for insurance or other forms of risk pooling, information problems, and externalities all serve to establish a role for the state in social protection from an economic perspective. Moreover, well-designed public intervention in social protection—discussed in greater detail later in the chapter—can itself help to strengthen the efficiency of risk-pooling mechanisms, through widening and deepening of the risk pool (which is often too small and fragmented, especially in developing-country contexts). But there are also very valid reasons for public social protection that go beyond an economics rationale. Among the most important include valid distributional concerns in Latin America and the Caribbean's highly unequal societies (see de Ferranti et al. 2004), and concerns for poverty reduction. Such justifications, while not purely economic in nature, speak to the types of societies that Latin American policy makers would like to foster (or as will be discussed in chapter 8, how to redefine the nature of the social contract in the region).

Challenges for social protection in the face of informality

Given the high human costs associated with lack of access to appropriate risk management instruments, such as health insurance and old-age security, and the clear rationale for public intervention to ensure basic access, Latin American and Caribbean policy makers face an important challenge; to ensure that individuals and families have access to suitable risk management instruments in the face of significant levels of informality. At the same time, it is important to note that ill-designed interventions may serve to make things worse, not better. Indeed, as will be shown, there is ample evidence of “government failure” in Latin America and the Caribbean that needs to be addressed as part of any actions to strengthen risk management among the region's citizens. Therefore, an important, related challenge is to ensure that the design and implementation of risk management instruments (and related programs) are consistent with improved risk management for the people of the region, as well as with increased productivity and sustained economic growth.

As has been highlighted in earlier chapters, the informal sector is highly heterogeneous. Some individuals (for example, many self-employed workers) are in that sector as a matter of choice, the result of an explicit or implicit assessment of the costs and benefits of formal versus informal employment, and/or in light of their personal characteris-

tics and endowments. For these workers, part of the explicit or implicit calculations may involve their perception of the relative costs and benefits of contributing to and affiliating with formal social security. Other individuals, such as many informal salaried workers, find the choice to affiliate to social security made for them by employers and firms that have chosen (for whatever reason) to opt out of the system. Chapters 5 and 6 have addressed the issue of firms' relationships to informality. This chapter focuses in more detail on attributes of social protection programs and, more broadly, social protection systems (the agglomeration of social protection programs in a country) that affect how workers value the prospect of working in the formal versus the informal sector. Understanding these factors can help policy makers design a smarter, more effective system of risk management for their citizens.

Costs, benefits, and the design of social security programs

A number of factors affect the locus of costs and benefits that workers see with respect to participation or nonparticipation in formal social security programs, including physical accessibility (proximity) of benefits and services for affiliated workers, the quality of services offered within the scheme, workers' valuations of the sometimes-complex bundle of goods and services that they are compelled to “purchase” via payroll taxes, program rules that make it difficult for workers to qualify for certain benefits, and design peculiarities that “force” affiliated families to pay twice for the same services. These factors are now examined in turn.

Accessibility of benefits

For many workers, particularly those in remote rural locations, the lack of program-authorized health facilities makes the choice of opting into the system untenable, even if they would otherwise be inclined to affiliate with a social security system. The unavailability of the basic amenities associated with social security makes the costs of contributing well above the benefits. Levy (2006a, 2006b) highlights that this is a critical issue facing millions of workers in rural Mexico and in less-developed regions of that country.

Program quality

Similarly, workers will also be inclined to favor informal employment if they perceive that the value of services

provided is low relative to its cost (that is, the cost of their payroll tax contributions). In the case of health, for example, workers may not value health insurance if the service at authorized health facilities is unpredictable, involves long waiting times, and is of low quality. To the extent that low service quality drives some workers—particularly low-risk, high-income workers—to affiliation with health insurance, this not only reduces participation in the system but serves to “fragment” the risk pool, hampering the effectiveness of the insurance model.

Perceived value of the “bundle of benefits”

In Mexico, the benefit package offered by the Mexican Social Security Institute (IMSS) has eight mandatory components: health insurance, retirement pensions, disability insurance, professional risk insurance, life insurance, day care centers for workers, sports and cultural facilities, and housing credits. In Colombia, formal sector workers make mandatory payroll tax contributions toward a number of benefits, including health insurance, pensions, and professional risk insurance, as well as toward worker training via the Colombian Training Institute (SENA), to the Colombian Institute for Family Welfare (ICBF), and to the *Cajas de Compensación Familiar*, which provide services ranging from cash subsidies to supermarkets to recreation for workers. While some workers may value *all* elements of this package, others may value only *some* of them. For example, while single mothers may value child care benefits very highly (indeed, they may be willing to pay more than they are asked to contribute via payroll costs), other workers—such as those without children—may not value it at all. In cases where workers do not value one or more components of the benefits package, those components may be considered by the worker as a “tax” on their earnings.

In some cases, workers may face situations in which they perceive contributions for part or all of the package of benefits provided under social security as a “pure tax.” In some cases, this may be an issue of government credibility. For example, if workers lack confidence in their government’s commitment or ability to actually provide promised benefits—for instance, pension benefits to be paid 20 years in the future—then workers may view making pension contributions as a pure tax. Moreover, certain aspects of social security design may cause workers to perceive social security contributions as a pure tax. In Brazil, for example, the Social Security Law (*Lei Orgânica da Previdência Social*) guarantees all Brazilians over the age of 67 a benefit of one

minimum wage, if they do not have other sources of income. Thus, the value of a pension benefit for a worker who earns close to the minimum wage would be practically the same, whether or not he contributes to social security—although the worker can access the benefit earlier if he or she is a formal sector worker (Fernandes, Gremaud, and Narita 2006). In Mexico, while all workers (and employers) are required to contribute to the housing subsidy, they are unable in practice to access those funds, due to the nature of the Mexican housing market (Levy 2006b; Mason et al. forthcoming).¹⁶ In this context, again, workers will view social security contributions—or at least the relevant component—as a pure tax on their earnings.

Paying twice for the same benefits

Another common feature of formal social security in the region is that members of the same family—for example, husband and wife—are required to contribute individually for health insurance benefits, even though nonworking wives (or husbands) would be covered by their working spouses’ insurance. In other words, members of the same family are essentially required to pay twice for the same set of benefits when both work in the formal sector. In this case, where one formal sector earner is already contributing for health insurance, the contribution of the second earner may be viewed as a pure tax on his or her earnings. Indeed, in Colombia, 44 percent of informal salaried workers and 54 percent of informal independent workers state that the main reason they do not contribute to health insurance is that they are already covered by a relative’s insurance plan (see chapter 2, table 2.13). This particular design issue has become increasingly important over the last few decades as female labor force participation rates have risen in the region. In the case of Argentina, Galiani and Weinschelbaum (2006) argue that significant numbers of female entrants to the workforce chose to be informal because their husbands were already eligible for benefits, and that this phenomenon led to a 13 percent increase in informality between 1974–76 and 1999.

Labor movements in and out of the formal sector

Latin American and Caribbean social security systems are also designed under the implicit assumption that workers will spend their entire working lives in the formal sector. This approach is consistent with a long-standing view across Latin America that labor markets are highly segmented and that flows of workers between the formal and informal

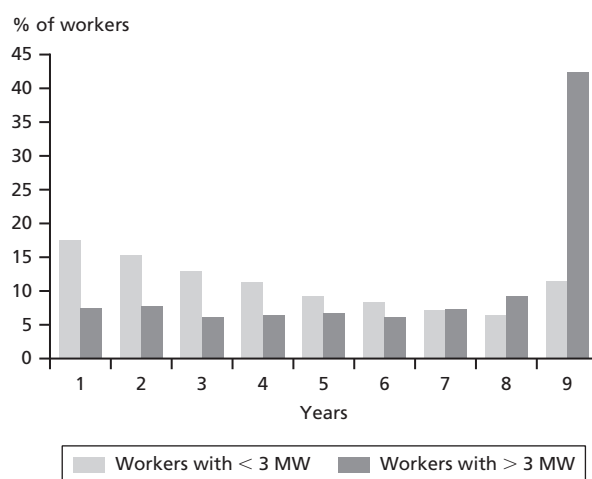
sector jobs—and particularly from formal to informal sector jobs—are small. As was discussed in earlier chapters, there appears to be more mobility and a greater flow of labor in and out of the formal sector than was envisioned in the design of either health insurance or pensions.

Data suggest, for example, that there is considerable movement of individual workers in and out of the formal and informal sectors, even over relatively short periods of time. Analysis of Mexico's 2005 National Urban Employment Survey indicates that about 11 percent of high-wage workers in the formal sector during the first quarter of 2005 (those earning more than three minimum wages) had moved to the informal sector by the end of the year (Levy 2006a). About 16 percent of low-wage workers in the formal sector (those earning less than three minimum wages) had moved into the informal sector from the beginning to the end of the year. Flows moved in the other direction as well; roughly 11 percent of both low- and high-wage workers who were informal in the first quarter had moved into the formal sector by the end of 2005.

Overall, movement of Mexican workers is greater among low-wage workers than among higher-wage workers. This can be seen from data on the duration of roughly 9 million IMSS workers in the formal sector between 1997 and 2005 (figure 7.10); only 11.6 percent of low-wage IMSS workers spent the entire nine-year period in the IMSS system. Moreover, on average, low-wage workers spent just less than half of the period (4.3 years) in the system. In contrast, over 42 percent of the higher-wage workers spent the full nine-year period in the system. These workers also spent a longer amount of time in the IMSS system over the period—6.5 years, on average. It is worth noting, however, that while higher-wage workers spent more than two years longer in the system than low-wage workers during the period, they still averaged roughly 2.5 years out of the system over the period.¹⁷

This mobility of labor has potentially important implications for workers' access to benefits and, thus, the design of social protection. For example, under current regulations, IMSS workers are required to accrue 25 years of work experience "in the system" in order to qualify for the Minimum Pension Guarantee (MPG). But if the years of IMSS affiliation implied in figure 7.10 are representative, then it would take a low-wage worker roughly 50 years of work to qualify for the MPG. For all practical purposes, therefore, the MPG would be unattainable. Moreover, to the extent that the data reflect worker movement in and out of the

FIGURE 7.10
Distribution of Mexican workers' years insured by IMSS,
1997–2005, for low- and high-wage workers in IMSS in 1997



Source: Adapted from Levy (2006a), based on Razu (2006).

Note: Of the almost 9 million workers analyzed, nearly 6 million were "low-wage" workers, defined as earning less than three times minimum wage, while roughly 3 million were "high-wage" workers, defined as earning more than three times minimum wage.

system, this means that, over any given time period, IMSS-affiliated workers—especially low-wage workers—spend a considerable amount of time without access to IMSS health insurance. This suggests that under Mexico's current social protection system, some workers may spend some time periods covered by IMSS, and other periods covered by ministry of health programs (for example, *Seguro Popular*) with access to separate and mutually exclusive health facilities. This represents significant inefficiency in the system.

Similar conclusions can be drawn from data on years of worker registration with the main social security institution in Uruguay, *Banco de Previsión Social* or BPS (Bucheli et al. 2006). First, these data also suggest that there is considerable movement in and out of Uruguay's formal sector during a person's normal working life. This can be seen in figure 7.11. If the data had a bimodal distribution—in which one group of workers never or almost never contributed to BPS and another group always or almost always contributed over their working lives—then one could say that the data show evidence of strong labor market segmentation. But, as can be seen from the figure, the data show a continuum of cases, with peaks around 17, 23, 25, 30, 35, and 38 years of service, suggesting more mobility than segmentation.¹⁸

FIGURE 7.11

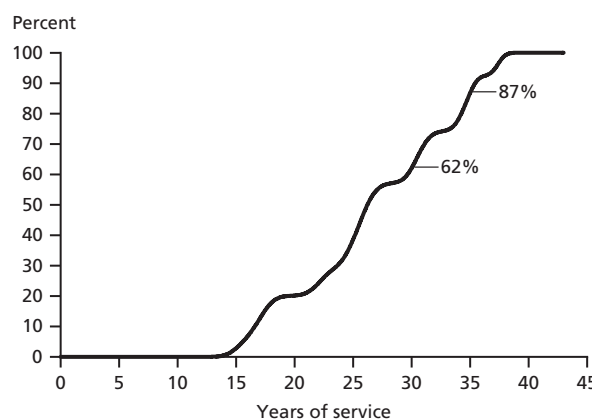
Density function for years of formal sector work among 60-year-olds in Uruguay (accumulated ages 18–60)



Source: Bucheli et al. 2006.

FIGURE 7.12

Cumulative distribution of years of formal sector service among 60-year-olds in Uruguay (accumulated ages 18–60)



Source: Bucheli et al. 2006.

Second, the data suggest that there is a high risk that a very significant share of the workers registered in the work history of the BPS will not be able to comply with the condition of contributing 35 years to access the pension when they reach the usual ages for retirement (figure 7.12) (Bucheli et al. 2006). In contrast to Mexico, 35 years of contributions are required not only to get a minimum pension, but to attain the right to any pension at all (if the worker is below the age of 70).¹⁹ As can be seen from the figure, Bucheli et al. estimate that by age 60, roughly 87 percent of all workers will not have accumulated the 35 years of service necessary under the law to qualify for pensions (figure 7.12). The authors estimate similarly that 72 percent of 65-year-olds will not have accumulated the necessary number of years.

As in Mexico, the Uruguayan problem tends to be more serious among low-income workers than among high-income workers (and among private sector workers as opposed to those in the public sector). Low-income workers in Uruguay tend to be at the bottom of the distribution of years of formal sector work, while high-income workers tend to be at the upper end of the distribution (Bucheli et al. 2006).²⁰ Bucheli et al. estimate that just over 50 percent of 60-year-olds from the highest earnings quintile will achieve 35 years of formal service (and registration in BPS), but that almost no one from the lowest earnings quintile will. As such, low-income workers in Uruguay—as in Mexico—face a particularly serious risk of not acquiring the years of service necessary to qualify for pension benefits.

For those workers—particularly low-income workers—who do not envision achieving the required vesting period, such rules may provide a strong disincentive to entering the formal sector in the first place.²¹

Just how important are these factors? A large body of empirical literature from Latin America and beyond shows that high rates of payroll tax have a negative impact on formal sector employment. For example, Packard (2002) finds that the higher payroll taxes for social security are associated with lower numbers of contributors in the workforce. Fiorito and Padrini (2001) arrive at similar results in an analysis of labor taxes in developed economies.

Several authors have estimated semi-elasticities of self-employment with respect to a change in relative formal sector earnings (Krebs and Maloney 1999; Loayza and Rigolini 2006; Maloney 2001). These authors find a range of estimates, depending on methodology used,²² and this has implications for the possible size of the impact. If, for example, one assumes that 10 percent of the formal sector earnings package is absorbed in unvalued benefits and perhaps that another 10 percent of the value of earnings is tied to inaccessible benefits (leading to a 20 percent decline in the relative attractiveness of formal labor), then the authors' estimates suggest the size of the informal self-employed sector would increase by between 0.6 and 6.0 percent, as a share of the workforce. In the case of Brazil, Fernandes, Gremaud, and Narita (2006) simulate the impact of eliminating payroll taxes for unskilled workers (those earning up to one minimum wage). If combined

with deductions of value-added tax (VAT) paid on capital investment, these authors find that eliminating payroll taxes on unskilled workers would lead to a decline in informality of 1.5 percent.²³

Adverse incentives within the broader social protection system

The fact that there is greater mobility between the formal and informal sectors than has generally been thought creates challenges not only for the design of specific social security programs, but for social protection systems as a whole. As mentioned earlier in the chapter, the recent growth of social assistance programs has also led to increasing fragmentation in social protection systems in the Latin American and Caribbean region. A key benefit of this approach has been that an increasing number of poor people have access to programs that play an important social protection or poverty reduction function, providing free but relatively modest benefits to eligible families—those whose livelihoods exist largely outside the formal sector. The main challenge relates to the fact that expanded social assistance programs provided freely (and sometimes provided conditionally on working in the informal rather than the formal sector) may themselves be creating disincentives to the formalization of the workforce who are taxed via payroll contributions to gain the right to social security benefits.

An example of this can be seen in the case of Mexico. As noted, the benefit package offered by the IMSS has eight mandatory components, including health insurance, retirement pensions, disability insurance, professional risk insurance, life insurance, day care centers for workers, sports/recreational and cultural facilities, and housing credits, paid for in part by workers' payroll taxes. In 1995, the government of Mexico launched a conditional cash transfer program, called *Progresa* (subsequently renamed *Oportunidades*), which provided cash transfers to poor families on the condition that families make specified human capital investments. The design of the program is oriented toward strengthening the human capital of the poor through incentives and investments in school-aged children. Furthermore, the design of the program is such that support will end prior to the time that a student-beneficiary graduates from school and prior to the time that he or she makes a labor market decision (for example, agriculture versus industry versus services, rural versus urban, informal versus formal).

In the last few years, however, several new social assistance programs have been launched or proposed for poor workers working in the informal sector. These include a subsidized package of financial protection in health, called *Seguro Popular*, with an insurance premium that is 100 percent subsidized for extremely poor families; a pension program, MAROP (*Mecanismo de Aborro para el Retiro Oportunidades*), that includes noncontributory transfers for the elderly and savings incentives for those in *Oportunidades* who work in the informal sector; and *Fondo Nacional de Habitaciones Populares* (FONAHPO), a program of housing subsidies for poor informal sector workers.²⁴ More recent proposals are on the drawing board for universal health insurance for all pregnant mothers and children under five years of age, and subsidized child care for working mothers in the informal sector.²⁵

Several features are noteworthy about these recent or emerging social assistance instruments—health insurance, pension, housing, child care—relative to the IMSS package. First, although the benefit packages are not as generous as those offered by IMSS, they are less costly than the IMSS package and, in some cases, essentially free. Second, eligible beneficiaries can choose which of the programs they would like to participate in; there is no mandatory bundle. And, third, beneficiaries must work outside the IMSS (formal sector) system to be eligible for these programs. So, key questions that potential beneficiaries face are: Do I search for a job in the formal sector where I can receive a bundle of benefits that I will pay for via my payroll contributions? or, Do I stay in the informal sector and get similar—although somewhat less generous—benefits that are essentially free? The answers lie, in part, in how highly workers value the *net* benefits of social security versus social assistance benefits. While this emerging system of assistance programs is too young to have enabled measurement of its effects on formal versus informal employment, the creation of a dual system—social insurance versus social assistance—may be creating incentives for greater informal employment (or slower formal employment growth) in Mexico.

Levy (2006b) models the labor market decisions of utility maximizing workers between the formal and informal sectors, given a predetermined and mandatory (take-it-or-leave-it) bundle of social security benefits available to formal sector workers, on one hand, and an unbundled collection of social assistance benefits available to informal sector workers on the other. Consistent with the situation

in Mexico and elsewhere in Latin America and the Caribbean, workers must pay for social security benefits via payroll taxes, while social assistance benefits are available for free to workers outside the formal sector. Workers may differ in their valuation of social security benefits as a result of inherent differences in their personal preferences or due to differences in the quality of program benefits and services. When workers' valuations of the relative benefits differ, they self-select into formal and informal sector jobs, affecting the composition of the labor market. A general conclusion of Levy's model is that when social (protection) policy consists of many programs and policies with distinct rules of access (formal versus informal job status) and forms of financing (payroll taxes versus general revenues), it can generate unanticipated labor outcomes; these resulting outcomes can have adverse effects on productivity and long-term growth (Levy 2006b).

While the Mexico case may be particularly illustrative, Mexico is by no means alone on this issue. Colombia has a subsidized health insurance regime (*Regimen Subsidiado*) that is available to poor informal sector workers; Argentina has an emergency employment program (*Jefes de Hogar*) in which participation is conditional on not being employed elsewhere. In fact, a recent study of the informality impacts of the *Jefes de Hogar* program in Argentina provides some preliminary evidence of how important these incentive effects might be at the specific program level. In the midst of one of the most serious economic crises of its history, Argentina implemented *Jefes de Hogar* (*Jefes*), a large poverty alleviation program. This program combines the features of a workfare and a conditional cash transfer program. *Jefes* is aimed at providing cash transfers to those unemployed household heads with children at school. The belief that poverty is closely related to unemployment led Argentina to include the unemployment requirement as a targeting device.

In principle, conditioning on unemployment implies a full taxation on outside incomes for the program participants; that is, getting a job means losing their program benefits. Under certain circumstances, therefore, *Jefes*'s unemployment requirement may create a disincentive for beneficiaries to search for a formal job—although in practice, monitoring of *Jefes*'s program requirements is not perfect, so the incentive effects may not bear out in reality. The effect of *Jefes* on informality is, thus, an empirical matter. In this context, Gasparini, Haimovich, and Olivieri (2006) assess the impact of *Jefes* on labor informality

between 2002 and 2005, during a period of strong economic growth and formal employment growth in Argentina. The authors investigate whether male *Jefes* participants were less likely to accept formal jobs in this booming economy than their nonparticipant counterparts, taking advantage of a short panel of data from the Argentine household survey.

Gasparini, Haimovich, and Olivieri (2006) find some evidence of an informality-incentive effect due to *Jefes*. Specifically, the authors find that the share of *Jefes* participants who found a formal job was significantly lower than the corresponding share of nonparticipants with similar observable characteristics. In some cases the difference is not only statistically significant but economically large. For instance, when carrying out the propensity score matching with the radius method, during the 2003–04 period, the difference between male *Jefes* participants and *Jefes* nonparticipants in the share of workers moving to a formal job is 5 percentage points. The authors also find that the informality-incentive effect is sensitive to the relative wage offers in *Jefes* and in the formal sector. Specifically, the effect of *Jefes* on informality vanishes during the 2004–05 period when the gap between the *Jefes* transfer (fixed in nominal terms) and wages in the formal sector greatly widened.²⁶

In sum, evidence from the *Jefes* program in Argentina suggests that the design of social assistance programs and the incentives they create do indeed influence workers' decisions to seek formal versus informal employment. Additional analysis in other country and programmatic contexts would be useful in deepening policy makers' understanding of just how strong are these incentive effects, both in absolute terms and relative to other factors. In the meantime, several factors may be thought to affect the relative strength of the incentive effects in a given context. These factors include the following:

- The more generous noncontributory program benefits are relative to contributory program benefits, the stronger the incentive problem. This could include relative wages, as in the case of *Jefes*, or the relative size of non-wage benefits, all other factors (including service quality) being equal.
- Program incentive effects are likely to operate most strongly among workers at the margin of formal and informal employment decisions. In that context, social assistance programs (including

noncontributory health and/or pension benefits) are likely to have a smaller effect on labor market outcomes. When they are well targeted toward the very poor, and when the social security coverage gap is relatively large, formal sector employment is concentrated among those in the upper deciles. The issue of incentives between social security and social assistance is likely to be more important in the relatively more developed countries in the region, with moderate-to-high levels of both social security and social assistance coverage.

- Certain types of social assistance programs may not have strong incentive effects on the workers' decisions to seek formal versus informal employment. Programs such as *Oportunidades* in Mexico (or similar conditional cash transfer programs) that limit their focus to strengthening children's human capital—for example, education, health, and nutrition—but do not try to influence labor market choices, would not be expected to create strong incentives for any specific forms of labor market participation.
- Conversely, those programs that tie participation to a particular labor market state—for example, working outside the formal sector—would be expected to more strongly affect informal-formal sector decisions.

Reengineering social protection to protect all citizens

Poor access to basic risk management instruments, combined with the existence of market failures (information problems, externalities) and pervasive government failure, highlights an urgent need to rethink—indeed, “reengineer”—social protection policy and systems in much of Latin America and the Caribbean. Specifically, given the high and persistent levels of informality in the region, it will be important to rethink the traditional Bismarckian model of social protection in which protection depends on the specific form of the labor contract. The collection of evidence suggests that a broader notion of who has access to basic risk management instruments is needed—one based on assuring the basic protection and welfare of countries' *citizens* rather than of workers, as traditionally and narrowly defined. Chapter 8 focuses in detail on the nature and evolution of the social contract between the Latin American State and its citizens. This section focuses on potential implications of informality, risk,

market failure, and government failure for the design of the region's social protection systems.

Providing essential cover in health

For guidance on *how* governments should intervene to protect their citizens, it is useful to return to the comprehensive insurance framework discussed earlier in the chapter (box 7.2). Baeza and Packard (2006) illustrate the application of this framework in the context of health. Illness comes in a variety of forms. Most frequent illnesses are not serious, nor do they imply large costs or financial losses. In fact, for 80–90 percent of health events that households will experience in their lifetimes, people rarely go to the doctor. Most symptoms—headaches, the common cold, adult diarrhea, and even minor fever—last fewer than three or four days and can be easily treated with rest and nonprescription pharmaceuticals. To mitigate the financial losses from these relatively small, frequent symptoms, most households are better off relying on prevention (for example, good nutrition and good hygiene, preventive and primary medicine, and exercise) and saving (that is, individually assuming the cost of treatment and medication).

However, for less frequently occurring, more serious illnesses, the cost of treatment can increase rapidly. Indeed, for conditions such as complicated flu, pneumonia, bacterial bronchitis, or urinary infections, the cost of diagnosis, treatment, and resolution can be substantial. Medical consultation becomes critical to identify more serious conditions. To cover the potential financial consequences of less-frequent sicknesses that are costly to treat, households are better off relying on risk-pooling arrangements. As the cost of treating health events grows—and because the direct and indirect costs of an illness and its treatment can be impoverishing—it becomes critical for individuals and households to find an effective mechanism to pool risks (Baeza and Packard 2006).

Yet, as noted above, private markets for health insurance are thin or are missing in most Latin American and Caribbean countries, due in part to information failures. The small and fragmented risk pools created by existing health insurance schemes are relatively inefficient. Moreover, in the case of health, losses that go uncovered have the potential to impose significant external costs on others. As such, there is a strong case for public intervention to strengthen risk-pooling mechanisms and to expand the pool. Since public resources for health insurance tend to be heavily constrained in the region, *there is a case for extending*

broad coverage via a package of minimum or “essential” direct cover of catastrophic losses with high external costs (Packard 2006). To ensure the greatest possible coverage, such a package should be de-linked from the labor contract and financed through general taxation. Since even moderate health costs can be catastrophic for the poorest households in the region, the exact composition of the package should reflect this fact.

There are a number of arguments in favor of de-linking the provision of essential cover in health from the labor contract and, instead, financing such provision through general taxation. General taxation is potentially the most efficient and also the most equitable financing mechanism for risk pooling, depending on the progressivity of tax collection instruments and subsequent patterns of public spending (Mossialos et al. 2002; Savedoff 2004, cited in Baeza and Packard 2006). Financing essential cover through general tax revenues also has the benefit of ensuring that health risks are effectively pooled across the widest possible risk pool. In the case of essential cover, Packard (2006) argues that, given externalities, the social costs of individuals failing to cover themselves and their dependents are high enough that “there is a clear risk management rationale to take the ‘choice to cover’ out of the hands of employers and workers” (p. 25) by shifting financing away from payroll contributions and toward general revenues. Moreover, in countries such as Mexico or Colombia where subsidized health regimes have been put in place to fill health coverage gaps left by the formal sector, general-revenue financing alleviates the problem of misaligned labor market incentives associated with a system in which payroll tax–financed social security for the formal sector “competes” with general revenue–financed (and government-subsidized) assistance programs for workers in the informal sector (Levy 2006b).

While national health systems in Latin America and the Caribbean have long sought, in principle, to cover a wide range of health conditions via the public health services, these systems have generally not conformed to sound insurance principles, nor have they been able to deliver on their promise with quality services. Overreach on the part of government has led not only to low-quality service in the public health system, but unsustainable fiscal situations in some countries. In recent years, however, several countries have attempted to define more modest but implementable minimum-benefit packages, such as the *Regimen Subsidiado* in Colombia, *Seguro Popular* in Mexico, and the AUGE

package in Chile (box 7.3). It is important to underscore that the idea of general-revenue financing of health insurance should apply solely to minimum essential cover; both the economics of insurance and a country’s real-life fiscal constraints dictate that additional coverage, if desired, should be available on a contributory basis. In this context, there is a critical role for public policy in fostering more efficient health sectors, including the strengthening of both insurance and provision functions.²⁷

Strengthening old-age security

There is also a case for providing essential cover to the elderly in the form of a poverty prevention pension, focused on the poor, as part of a broader multipillar pension system. As a form of social insurance, risk pooling would be central to the poverty prevention element, and, because of the social costs associated with people falling into poverty at an older age, there is a clear risk management rationale for de-linking access to this poverty prevention pension from the form of the labor contract, and financing it through general revenues (Packard 2006). At the same time, the comprehensive insurance framework highlights the importance of individual saving as a core element of a broader system of old-age security. Indeed, *given the high probability of income loss in old age, saving should be the mainstay for earnings replacement during old age* (Gill, Packard, and Yermo 2004). Lessons from recent experience make clear that savings pillars should closely link benefits to contributions and do so in a similar way for most workers, regardless of the status of the labor contract. In this context, the individual capitalization schemes that have been introduced in Latin America over the last few decades are fully consistent with the saving objective.

In pursuing more effective protection against old-age poverty in the face of informality, Latin American and Caribbean countries face several important challenges regarding the design and operation of the poverty prevention and savings components, as well as the relationship between the two. Recent policy research on pension reform in Latin America and the Caribbean suggests, for example, that from an institutional perspective, the poverty prevention component is best financed and managed separately from the savings component. Among other things, the institutional imperatives associated with managing a minimum poverty reduction payment are distinct from those associated with the management and supervision of investments under individual capitalization schemes (Gill, Packard, and Yermo 2004). Moreover, weak design of one element can undermine

BOX 7.3

Extending health insurance coverage by correctly aligning risk-pooling instruments

The question of how best to protect households from the financial consequences of illness and disability has been the subject of lengthy debate in Latin America. The core of this discussion has focused on the trade-off between the *breadth* of coverage (how many households can count on *some* health protection) and the *depth* of coverage (the package of care they can count on). The constitutions of most Latin American countries guarantee the right of citizens to good health and access to health services. In theory, the public health system covers treatment of all health conditions. However, in practice, health systems are financially constrained, and resort to rationing or lowering the quality of care in order to comply with their budget constraints. At least until the mid-1990s, constitutional mandates existed alongside a large, unsatisfied demand for health care, and the promise of universal health care was not accompanied by effective instruments to achieve that. The casualties of this failure tended to be the poor. Beginning in the mid-1990s, however, several countries have introduced legislation and sector reforms to transform health benefit packages into well-defined and explicit entitlements guaranteed to all citizens. These reforms typically include legal mechanisms for households to demand their entitlements from the state.

Since 2002, Chile has been gradually introducing a new health insurance package (Plan of Universal Access with Explicit Guarantees, or AUGE). Rather than continuing to offer unrealistic promises to cover all health needs, AUGE establishes a guaranteed minimum package of basic health cover for all Chileans, along with guarantees of attention within specified time periods, set ceilings on co-payments, and full subsidies for the poorest households. The new system is designed to eliminate rationing and improve service. AUGE levels out premiums and co-payments and reduces “cream skimming” by insurers. It represents an enforceable “patient’s bill of rights” for affiliates of FONASA (*Fondo Nacional de Salud*), the National Health Service, and is set up to reduce waiting lines and other inefficiencies that can increase out-of-pocket costs, particularly for poorer households. AUGE was first tried out on a pilot basis

with the population covered by FONASA. The pilot experience yielded positive results by increasing access to the services included in the package. Since January 1, 2004, AUGE has been mandatory for all citizens and all public and private providers of health insurance. Both households covered by FONASA and those covered by the private health insurers (called ISAPREs) are entitled to the same minimum package of health services.

While it is too early to assess the effect of the Chilean reform on lowering vulnerability to poverty or easing the burden of health expenses for the poor, the “explicit entitlements approach” embedded in the AUGE is already having both intended and unexpected effects in Chile and elsewhere in the region. For example, it has forced policy makers to revisit the *breadth* and *depth* discussion of health coverage. Before the reforms, all health services were theoretically available to all citizens, but, in practice, nothing was actually guaranteed. The result was a limited and poor-quality package of services, particularly for the poor.

Under the new explicit entitlement approach, a basic package of health coverage becomes legally binding for governments, throwing open the debate of whether to guarantee a limited package to all or an extensive package (often focused on the most costly health care) to a few. Moreover, these reforms have led to a much closer dialogue between ministries of finance and health because the legally binding nature of the package reduces the space for fiscal adjustment in the health sector. Errors in defining the coverage of the package can thus have significant fiscal consequences, putting a premium on efficiency and management in the health sector. Indeed, although largely unintended, most related reforms have provided governments with a powerful instrument to focus on broader efficiency-enhancing reforms in the health sector. Guaranteeing a package requires clarity regarding the quality of delivery, which, in turn, requires complex monitoring systems, provider payment systems, and contracts or quasi-contracts between the public financing agency and health service providers.

Source: Baeza and Packard 2006.

the impact of another. For example, excessively generous or badly designed poverty prevention components serve to reduce incentives for personal retirement savings and/or exacerbate incentives for informal sector employment.

So what does this imply for the design of a poverty reduction component of a pension system? Given institutional and fiscal constraints that exist in the region—and that vary significantly across Latin American and Caribbean countries—it is useful to draw on the discussion of incentives above, and to distinguish between those countries with large coverage gaps, in which incentive issues may not be so important in the short run, and those with moderate-to-high levels of social security and social assistance coverage, where the incentive effects are likely to have greater importance, even in the present. More specifically, these incentives could include the following:

- For countries with large coverage gaps, it makes sense to focus first on the implementation of social assistance pensions that are well targeted to the poor (or extreme poor) and scaled to be fiscally sustainable.²⁸ In this context, an important design challenge will be to set the level of the pension benefit large enough to provide protection against poverty, but not so large that it creates significant adverse labor market or saving incentives in the future when the coverage gap between the social assistance pension and social security closes.²⁹ An important message here is that even in designing a well-targeted social assistance pension, policy makers should consider not only short-term concerns about coverage, but also longer-term concerns about incentive compatibility as the gap between contributory and noncontributory programs closes over time.
- For countries with moderate-to-high levels of social security coverage, where incentive issues are likely to be more important in the short term, the need to focus on the issue of incentive compatibility is more immediate. In such cases, in addition to having a well-targeted (and well-means-tested) poverty prevention pension, some progressive reduction in the benefit level with income would be advisable to help maintain appropriate incentives throughout the broader system. In this regard, the pension reform recently proposed by the government of Chile may hold some valuable lessons for other countries in the region. The Chilean proposal focuses on providing

old-age security for all, including the poor and those who work outside the formal sector. In doing so, it would move Chile from having two different systems—a defined contribution pillar system and a targeted social assistance pension—to a single integrated system. The reform seeks to establish an appropriate mix of pooling and saving instruments, an effective institutional structure for supervision and management of the new system, and, importantly, incentive compatibility of the poverty prevention and saving components of the system (box 7.4).

What about the design of the savings components in the face of pervasive informality? A key issue here is whether (or how) to mandate old-age savings among those who, by definition, operate outside of countries' administrative and enforcement mechanisms. Finding the appropriate balance between the mandatory savings pillar (so-called pillar 2) and the voluntary savings pillar (so-called pillar 3) has been the focus of recent analysis of multipillar pension schemes in Latin America (Gill, Packard, and Yermo 2004). Part of the argument for mandating savings revolves around concerns about information failures and "myopia" that might lead workers and households to under-save for retirement. Yet, working to ensure old-age savings opportunities for those in the informal sector would seem to argue for strengthening voluntary savings instruments as part of a multipillar system. It also argues for providing greater flexibility within the system to help raise the perceived benefits of savings relative to its costs. In this context, there will be gains from ensuring greater portability of savings and benefits as workers move in and out of the formal sector. Providing greater flexibility in savings and investment options, perhaps based on people's stage of the life cycle—for example, and allowing lower levels of mandatory contributions for younger workers, allowing younger workers to select investments with higher risk-return profiles, while supporting lower-risk portfolios for older workers—may also serve to raise people's demand for retirement savings.

Unbundling complex, multidimensional benefit packages

As previously noted, a number of countries' social security systems require contributions to complex, multidimensional benefit packages. As has also been noted, to the extent that workers do not value one or more components of the package, those components represent a tax on

BOX 7.4

Old-age protection in the new millennium: Chile's proposed pension reform

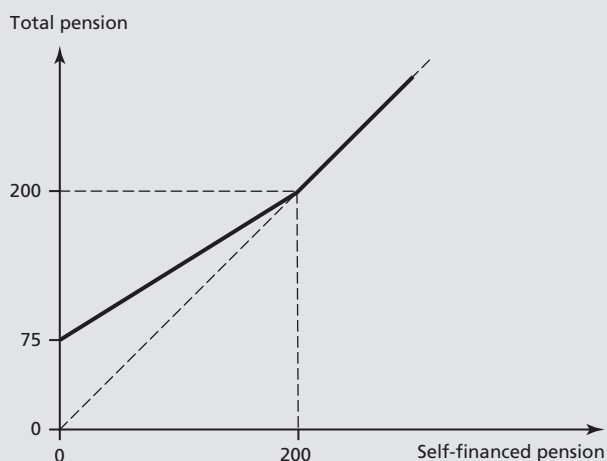
In December 2006, the Chilean government sent to Congress a draft law to reform the pension system. The proposed reform contains a broad set of measures, including the replacement of the existing targeted social assistance pension (PASIS) with a new solidarity pillar, extension of coverage to self-employed workers, norms to promote gender equity, and several norms aimed at improving the performance and supervision of the individual accounts pillar. It is an ambitious and comprehensive proposal that builds on the existing system.

To prevent old-age poverty, the current Chilean social security system includes a minimum pension as part of its contributory pension scheme, as well as a social assistance pension. The minimum pension is available only to individuals who have contributed for at least 20 years to individual accounts, and there is some evidence that many workers will not achieve the required contribution history before retiring. At the same time, the social assistance pension is targeted to the poorest. Thus, there is a concern that a significant share of the population might end up with very low—or no—pensions.

The proposed reform would create a new solidarity pillar that provides benefits to individuals in the lowest 60 percent of income distribution who either have not contributed to the system or have low contribution levels. Those who have not contributed would receive a pension of Ch\$75,000 (approximately US\$142) per month. Those who have low contribution levels would receive a supplementary payment, such that the sum of their self-financed pension and the supplement would be no less than Ch\$75,000 and no more than Ch\$200,000 (approximately US\$380) per month. The size of the supplementary payment declines with the size of the pension that is self-financed so that workers who can self-finance a pension of Ch\$200,000 or more receive no supplement (see box figure). Until that point, however, the total pension increases as the self-financed pension does, so that workers receiving the supplement still have positive incentives to continue contributions.

The participation of self-employed workers in the current pension system is voluntary in Chile, and only about 5 percent of the self-employed have chosen to participate. The proposed reform would extend the same program

Benefit structure under Chile's proposed pension reform
(workers' total pension as a function of their self-financed pension)



Source: Forteza 2007, based on Government of Chile 2006a, 2006b.

benefits to the self-employed as to salaried workers, including the solidarity benefits. Over time, participation of self-employed workers would become mandatory, unless the person explicitly opts out. It is worth noting that even though Chile currently has low coverage of self-employed workers, as in other Latin American countries, it does not have as high levels of informality among the self-employed (since most of them are registered and pay other taxes). As such, the enforcement challenges associated with expanding coverage among the self-employed in Chile look relatively less onerous than in other countries in the region.

The reform effort also seeks to promote reductions in the commissions charged by the pension fund administrators and to increase the return of the investments. Current commissions are considered high and are attributed to the lack of sufficient competition. The Superintendent of Pensions will organize annual auctions of new affiliates based on the level of the commission; the administrator who offers the lowest commission will get the new affiliates. The administrator will be required to keep the commission offered to new affiliates for at least 18 months and extend this rate to all its affiliates. The reform also seeks to raise the return of the funds by providing more flexibility to choose the composition of

(Continued)

BOX 7.4

Old-age protection in the new millennium: Chile's proposed pension reform (Continued)

the portfolio. With more flexible supervision, administrators will have more responsibility for returns on the funds.

The reform would also modify the institutional structure for managing and supervising the new system. Specifically, a new Social Security Institute (*Instituto de Previsión Social*) would be in charge of the system of solidarity pensions while a Superintendent of Pensions (*Superintendencia de Pensiones*) would supervise the whole system, including the Social Security Institute. In undertaking the reform, the government of Chile envisions a five-year transition period, with a progressive roll-out of the solidarity pillar. The fiscal implications of the reform, while not trivial, appear manageable. The presidential advisory commission overseeing reform of the social security system estimates that the new solidarity pillar will cost roughly 1 percent of GDP in 2025. The total fiscal cost of pensions in that year would thus grow from a projected 1.6 percent of GDP in the absence of the reform to 2.5 percent of GDP with reform.

The pension system envisioned in the Chilean reform represents a movement away from what is predominantly a single savings pillar to a more balanced multipillar, or multi-tier, system. The proposal strengthens and better integrates the assistance and redistributive components with the individual savings component of the system. It improves the capacity of the pension program to protect the less fortunate, without neglecting incentives and avoiding strong fiscal impacts. The proposal also represents a valuable contribution in terms of the policy-making process. For example, there were extensive consultations with stakeholders conducted by the commission appointed by the government to develop the preliminary set of proposals. In addition, the government chose to build its reforms on the existing system. The result is a balanced project that improves the existing system without disregarding the achievements made under previous reforms.

Source: Forteza 2007, based on Government of Chile 2006a, 2006b.

workers' earnings, generating disincentives to formal sector employment. Such taxes adversely affect incentives for formal employment.

From the perspective of public social protection, some of the components included in these complex packages have no clear risk-pooling or risk management rationale. Nor is there clear economic justification as part of a social security package on the basis of providing public goods or positive externalities. Similarly, there is no clear justification for payroll-tax financing. Take, for example, the cases of the sports and recreation or housing-related benefits in Mexico. Both of these benefits represent private goods; neither involves risk pooling or social risk management.

In such cases, social security packages would benefit from "unbundling" of what might be called "nonessential" elements. Unbundling would involve shedding—or making voluntary—those elements of the current social security packages with no risk-pooling or risk management rationale, or that finance private as opposed to public goods. The objectives of unbundling are to increase the benefits relative to the costs of social security packages and focus social security on its core social protection functions.

If governments see the need to provide a housing benefit, at a minimum that benefit should be made voluntary or "elective." However, the rationale for using payroll taxes as an instrument to finance such a benefit is unclear. If the objective is to ensure that the poor have access to adequate housing, then a more efficient and appropriate way to provide that benefit would be through direct subsidies targeted to the poor and financed by general revenues.

Some countries, such as Colombia, also provide subsidies to early childhood development financed through payroll taxes. Although investments in early childhood development do have a strong "public goods" dimension, the rationale for financing them through payroll taxes remains unclear. They, too, are better financed through general taxation.

Potential costs of social protection reform: Financing essential cover

Reengineering Latin America's social protection systems and, in particular, ensuring minimum essential cover in health and pensions via *de-linked* social protection programs will have important fiscal implications for most countries in the region. First, de-linking implies eliminating payroll

taxes (at least in these areas) and replacing them with revenue sources that are less distortionary, such as the VAT, or income or property taxes. Second, in most countries in the region, providing essential cover will imply the need for greater public resources to health and pension coverage. The precise amounts of additional resources required and the appropriate revenue source in each country will depend on a number of factors, including but not limited to the size of the coverage gap, the size of the benefit package being offered by the government, and the country's fiscal space.

Calculating the fiscal impacts of providing essential cover is a difficult exercise, as estimates are sensitive to changes in assumptions. In the case of health, for example, estimating the cost of providing essential cover is tricky, as the results are extremely sensitive to the definition of the size and content of the benefit package. Moreover, in a number of countries, there are potentially important fiscal savings associated with micro-efficiency reforms in the health sector that would be important to facilitating essential cover in health. Measuring the extent of those savings and how they offset the additional costs of expanding coverage is also a challenge. Nonetheless, attempts to estimate the potential costs are important, and can be useful in identifying general orders of magnitude of impact.

Recent World Bank estimates in the case of Mexico are illustrative. In terms of providing universal essential cover in health, it is assumed that the unit cost of such provision is equal to the unit cost of *Seguro Popular*, Mexico's subsidized health insurance program for the poor. Launched in late 2004, *Seguro Popular* covers a packet of basic health care, including preventive, outpatient, inpatient, emergency, and surgical interventions.³⁰ This is a much more modest package of coverage than is afforded by the health insurance packages in IMSS and ISSSTE (the Social Security Institute for Mexican civil servants at the federal level). In this context, it is assumed that wealthier households would be free to purchase additional (supplementary) health insurance in the private market, according to their preferences and willingness to pay. Under this scenario, it is estimated that the *additional* fiscal costs of providing essential cover in health (that is, costs above and beyond what the government of Mexico currently spends on health care provision) would be approximately 1.3 percent of GDP.

The potential fiscal costs of providing a poverty prevention pension were also estimated for Mexico—specifically, the costs associated with providing a targeted social assistance pension to all elderly among the extreme poor. As with

health, the expected fiscal impact of essential old-age cover depends critically on several factors, including desired coverage levels, expected demographic changes, the proposed size of the benefit, whether benefits are indexed to prices or wages, and whether savings incentives are offered. For the purpose of this exercise, estimates were made using a proposed benefit of half a minimum wage (roughly equivalent to the extreme poverty line in Mexico), indexed to real wages.

Under this scenario, and accounting for the aging of the Mexican population, it is estimated that the incremental fiscal costs of implementing a targeted social assistance pension would be 0.3 percent of GDP between now and 2010, rising to 0.7 percent of GDP in 2050. Data from other countries in the region suggest that the fiscal impact of extending pension coverage will vary depending on the program's design and the country's institutional characteristics—although the estimates are generally consistent with the findings for Mexico. The government of Chile, for example, estimates that the proposed solidarity pillar will translate into a cost of around 1 percent of GDP by 2025 (box 7.4), while a forthcoming study on pensions in Brazil estimates that spending on the government's extensive rural and social assistance pension schemes is equivalent to about 2 percent of GDP (World Bank forthcoming).

What might such reforms imply from the perspective of the tax system and, in particular, if such programs were to be financed via general taxes rather than through payroll taxes? Analysis was undertaken for Mexico to estimate the potential effects of financing these reforms via the VAT. Specifically, it was assumed that all households would be taxed under an increased VAT. It should be noted that the current VAT in Mexico includes a number of exemptions, such as on food and medicines. Estimates of the impact on VAT rates were undertaken—with the assumption that these exemptions were kept in place as rates were raised.

One concern sometimes voiced about raising the VAT is its costs to the poor. To address potentially adverse impacts on the poor of an increased VAT, the estimates assume that the poorest Mexicans are directly compensated for the additional tax impact.³¹ While direct compensation for higher value-added taxes would not be provided to the nonpoor, it should be noted that those working in the formal sector would simultaneously benefit from the fact that they would no longer have to contribute to the payroll tax. Moreover, given that part of the payroll tax is paid by employers, current formal employees may be expected to receive part of the current employer tax burden back in the form of higher

salaries.³² While these latter benefits to formal workers are noted, they were not included explicitly in the VAT calculations.

Estimates of the cost of VAT financing of essential cover for all Mexicans were undertaken, including provision for the poorest 20 percent of Mexican households to be compensated for their increased tax burden. This is equivalent to the roughly 5 million households currently covered under the *Oportunidades* conditional cash transfer program, and the compensation could in principle be targeted using the *Oportunidades* beneficiary registry. These estimates suggest that to cover additional public spending equivalent to 1.3 percent of GDP—the amount necessary to provide minimum essential cover in health—an increase in the VAT rate of about 6.4 percentage points would be necessary, in the absence of a payroll tax and under the current VAT structure, including exemptions. That would imply an increase from the current VAT rate of 15.0 percent to one of 21.4 percent, a change that is not trivial.

In the case of providing a targeted social assistance pension, resource needs increase over time before leveling out, and this is reflected in the necessary VAT increases. Specifically, financing a well-targeted social assistance pension would require an increase of the VAT rate by approximately 1.3 percentage points at present, but would require a cumulative increase in the VAT rate of roughly 3.4 percentage points by 2050 (as the costs associated with providing a social assistance pension, under current assumptions, rise). It should be noted, that, under both the health and pension scenarios, generalizing the VAT to include taxes on food and medicines—were such a move politically feasible—would have a significant mitigating impact on the need to raise rates, even if compensation of the poor for broader VAT taxation were included.

Relying on the VAT to finance a de-linked social protection system may not be the optimal solution in all countries, however; indeed, other instruments, such as income or property taxes, may be more appropriate, given the specific context. Recent analysis suggests that in Brazil, for example, switching from a payroll tax to a VAT in its current form could have a negative impact on economic growth in the long run (Fernandes, Gremaud, and Narita 2006). This is because the VAT in Brazil does not completely exempt investment goods from the taxation, and this additional taxation on capital goods has a long-term negative impact on growth. Nonetheless, Fernandes, Gremaud, and Narita also find that, in Brazil, if measures were taken to permit

exemptions of the VAT on capital goods, then the adverse effect on productivity and growth would disappear.

In considering a shift from payroll tax to general revenue-financed social protection, it is important to recognize that policy makers in Latin America are not starting with a blank slate, but operate in a well-established policy, institutional, and political environment. That said, most OECD and many developing countries—all with their own history of policies and institutions—have already started to make the transition. Even countries with long traditions of Bismarck-type social insurance systems, such as Spain, are moving toward general taxation to finance minimum levels of risk-pooling coverage (see box 7.5). In fact, countries that have opted for general-revenue financing typically started with fragmented voluntary and then mandatory social insurance systems similar to the types found in Latin America and the Caribbean. Even in Germany, birthplace of the Bismarckian social insurance model, recent proposals to reform segments of the welfare system aim at moving health finance toward general taxation.

The legacy of past social protection policies creates important challenges for policy makers in switching from today's model based on payroll taxes to a new model based on general-revenue financing. However, lessons from recent reform countries, such as Spain, can provide some guidance; some experience with this type of shift also exists in the Latin American and Caribbean region. Chile's pension reform of 1981 included shifting some elements of payroll tax-financed social insurance to general revenues, and was accompanied by a substantial decrease of payroll taxes (Gruber 1995). Thus, technical solutions exist for how to implement the shift to a social protection model that can cover all citizens.³³

Managing the transition from here to there

Latin America and the Caribbean countries exhibit a high degree of heterogeneity in levels of social protection coverage (the “flip side” of informality). They also demonstrate great heterogeneity in social protection spending and the capacity to dedicate additional resources to social protection in the short-to-medium term. Social protection institutions and institutional capacity also differ from place to place. So while some countries may be in a position—and have the political will—to effect significant changes in the structure and financing of social protection in the near term, others will need to focus on more incremental challenges, as they move toward longer-term goals of ensuring universal access

BOX 7.5

De-linking health coverage from employment status: Spain's shift to general-revenue financing of essential social insurance

The health-financing system in Spain underwent radical changes in the 1980s and 1990s, when that country shifted from a social insurance system financed by payroll taxes to a national health service financed by general taxation. Today, almost 100 percent of public expenditures in health are financed from general taxation.

Spain's transition to democracy and the constitution of 1978 gave new impetus to health care reform. A separate organization was established within the social security system for the administration of health care services. Most health care programs and organizations were consolidated under the umbrella of the Ministry of Health, which was established as an independent entity in 1981.

In 1986, the General Health Care Act was approved following almost four years of public and parliamentary discussion. The act provided a unified legal framework for many of the previous piecemeal reforms and called for a tax-based financing system. Publicly managed health services were consolidated in a single national organization and in a small number of regional organizations (Cataluña and the Basque Country), within the framework of the newly decentralized state. The Spanish National Health System was subsequently devolved to the 17 Autonomous Communities that, since 2001, have fully managed their Regional Health Services and together make up the Spanish National Health System.

Consistent with the reforms introduced by the General Health Care Act, the sources of funding for the health care system were drastically modified in 1989. Beginning in that year, new budgets were financed 70 percent from general taxation and only 30 percent from payroll tax contributions.

In the mid-1990s, consistent with agreements signed by political parties and trade unions (known as the *Pactos de Toledo*), 100 percent of financing would come from general taxation and individual contributions were to be progressively phased out by 2000. In 1999, one year ahead of schedule, the entire health care budget in Spain was financed from general taxation. Today, Spanish regions receive funding for health care as part of the general funding from the central government. Funding is proportional to population adjusted for such factors as age distribution, number of temporary residents, and services provided to the national system or to neighboring regions. Supplemental funding is also supplied from fiscal revenues raised in the region.

Currently, only workers' compensation for work-related injuries and diseases is financed from individual contributions of employers and employees.

Source: Fernandez 2004, in Baeza and Packard 2006.

to basic risk management mechanisms in health and old-age security.

Indeed, in pursuing a long-term vision and strategy for minimum essential cover in social protection, most countries in the region will likely need to pursue a series of discrete measures, rather than “big bang” types of reforms. This may be particularly pragmatic given that the needed reforms require the support of a diverse set of institutional and political actors. In this context, it will be important to ensure that any short-term measures countries take are consistent—or at least *not inconsistent*—with the governments' long-term visions and agendas. This will be especially important if governments choose to pursue more integrated social protection systems that align social objectives of better risk management with economic objectives of higher productivity and growth.

To this end, governments in the region would benefit from taking several sets of incremental reforms in the short-to-medium term that would contribute in important ways to the long-term goals of extending social protection in health and old-age security to all its citizens. These reforms are beneficial in their own right. At the same time, they can help pave the way toward a system of essential cover that facilitates greater mobility in the labor market, not based on the informal-formal distinction but on sensible design of social protection and on flexible, incentive-compatible labor market institutions. Key priorities include the following:

- *Unbundling of complex, multidimensional benefit packages.* The objective of unbundling would be to increase the benefits relative to the costs and improve the public

goods content of social security packages. While health and pension benefits would remain at the core of the benefit package, unbundling would involve shedding—or making voluntary—those elements of the present social security packages that currently represent pure or partial taxes on workers or finance “private” as opposed to “public” goods.

- *Improving program/benefit quality in health and pensions.* Again, the objective here would be to increase the relative benefits associated with participation in the social security system. Key areas for attention include:
 - *micro-efficiency reforms in the health sector* of most countries in the region. Recent studies of the health sector in the region (for example, Baeza and Packard 2006; Mason et al. forthcoming) highlight several key areas that are critical to strengthening social (and financial) protection in health, including:
 - * separation of the purchaser/insurer and health care provider functions,
 - * moving from historical budget processes to production-based budgeting and strategic purchasing,
 - * developing minimum benefit packages (essential cover) that are consistent with the fundamental principles of insurance, and
 - * establishing—or strengthening—consumer choice among service providers.
 - *efficiency reforms in the region's pension systems.* Among the priorities are:
 - * lowering costs and administrative fees to affiliates and improving risk management (Gill, Packard, and Yermo 2004); further reductions in commissions, along with efforts to raise net rates of returns, would go far to improve the attractiveness of the funded savings pillars,
 - * strengthening voluntary savings instruments for old age, and
 - * creating greater flexibility in the savings pillars, including greater portability and flexibility in savings and investment options to reflect different stages of workers' life cycles (for example, lower mandatory contributions for younger workers and/or greater flexibility in the investment-risk profile across the life cycle); such flexibility would help raise the implicit benefit–cost ratio among potential affiliates.

- *Strengthening program design to account for and enable greater worker mobility.* This should include:
 - *revising overly burdensome vesting periods*, where they exist, in the region's pension programs to facilitate worker access and mobility. In this context, Latin American and Caribbean countries should consider adopting the types of points programs used in some European countries that enable workers to access pension benefits in proportion to their time in service.
 - *enabling spouses and/or secondary family workers to opt out of specific benefits*, such as health insurance, so as to avoid making families pay twice for the same coverage.
 - *fostering portability of health benefits and retirement savings across jobs, sectors, and occupations.*
- *Establishing consistency and incentive compatibility of program structures and benefits* across different parts of countries' social protection systems (both within social security and across social security and social assistance) to minimize adverse incentives and productivity effects. This includes efforts to “harmonize” rules, eligibility requirements, and benefits levels across programs and institutions—both in health and in old-age security.

While the actions listed here constitute an important short-to-medium-term reform agenda for most countries in the region, these actions are particularly important in countries in which there is a high degree of integration and movement between the formal and informal sectors, as well as in countries where the gaps in coverage between formal social security and emergent social assistance programs are relatively small.

Conclusion

Nearly all countries in Latin America and the Caribbean remain characterized by “truncated welfare systems,” in which those in the formal sector have access to an often generous multidimensional package of social security, while those outside the formal sector—whether in urban or rural areas—have much more limited access to government benefits and/or formal risk management instruments. Recent progress in extending social security to uncovered portions of the population has been disappointing, at best. Overall, social security coverage has failed to increase (informality has failed to decline) despite economic growth over the period; and coverage has actually declined in a number of countries over the last decade. Moreover, in nearly all

countries in the region, coverage rates are significantly lower among low-income than high-income workers. In many countries, the poorest workers and families are practically excluded from the system.

Partly in response to concerns about the truncated welfare state, several countries in the region have launched or expanded noncontributory assistance and/or poverty reduction programs over the last decade to help breach the coverage gap and provide support and coverage to the poor or extreme poor. A number of these efforts (for example, conditional cash transfer programs) have contributed in important ways to the short-term welfare and long-term opportunities of the poor. These efforts are welcomed and are to be encouraged. At the same time, the creation and expansion of certain types of assistance programs, especially those that link eligibility to work in the informal sector (or to being unemployed), may be creating incentives that constrain the growth of formal sector employment as well as long-term economic growth in the region. These incentive effects are particularly serious in countries where workers perceive that the benefits of affiliating with social security are low relative to the costs. In sum, the region faces both new and long-standing challenges in providing appropriate risk management instruments to its people.

Against this background, this chapter has reviewed the state and recent evolution of social protection in Latin America and the Caribbean; examined the role and limitations of households' private risk management strategies and outlined the rationale for public social protection in the region; analyzed the key challenges policy makers face in making adequate risk management instruments available to the regions' citizens, both those working in the formal sector and those in the informal sector; and proposed several key directions for policy to ensure that all citizens of Latin American and Caribbean countries are adequately protected from key risks, especially those associated with health shocks and poverty in old age. Several key messages emerge from the analysis:

- First, despite ample evidence that individuals and their families engage in a number of private risk management strategies, many people in the region—particularly those in low-income households—remain vulnerable to the impoverishing effects of health shocks and to poverty in old age.
- Second, private insurance markets are generally thin or missing in the region and market failures (for

example, information problems and externalities) abound. As such, there is a strong case for public intervention to strengthen households' abilities to manage risk.

- Third, alongside market failures that justify public action, there is evidence of widespread government failure in the design and implementation of social protection. These problems serve to exacerbate people's lack of access to appropriate risk management mechanisms. For example,
 - at the level of *specific programs*, design problems raise the costs of participating in social security relative to its benefits, causing some workers to opt out of the system; design issues also serve to hinder some workers' eligibility for benefits;
 - at the level of social protection *systems*—the constellation of contributory social security and noncontributory social assistance programs—programs often compete, creating adverse labor market incentives and outcomes.
- Finally, the combination of inaccessibility to basic risk management instruments, market failure, and government failure creates an urgent need to rethink and, in fact, “reengineer” social protection in the region—looking beyond the traditional Bismarckian model in which protection is based on one's labor contract to one that ensures protection to people on the basis of *citizenship*.

So what does this imply for the future of social protection in the region? Drawing from the economics of insurance, the chapter has outlined a long-term agenda for social protection reform in the region. In the case of health, because shocks that go “uncovered” can impose significant external costs on society, there is a case for providing a package of minimum essential direct cover, de-linked from the labor contract and financed through general taxation. In the case of old-age security, there is also a case to provide essential cover in the form of a poverty prevention pension focused on the poor, and as part of a broader multipillar pension system that includes provisions for individual retirement savings. Whether countries should focus first on a well-targeted social assistance pension for the poor or on development of a more integrated system of old-age security depends on a number of factors, including the size of the current coverage gap, the country's fiscal space, and local institutional capacity. In either case, it will be important

for policy makers to ensure incentive compatibility between the poverty prevention pension and the savings component, which should be the mainstay of earnings replacement during old age.

For a variety of reasons, including those related to fiscal and institutional capacity, movement to minimum essential cover in health and old-age security—de-linked from the labor market and financed by general taxes—represents a long-term agenda for many countries in the region. In this context, it will be important for countries to orient their short-to-medium-term policy agendas in ways that are consistent—or at least *not inconsistent*—with their longer-term vision. This will be critical if the region's governments are to align more effectively the objectives of better social risk management with those of higher productivity and growth.

To this end, governments in the region will benefit from making incremental reforms that improve the efficiency of existing systems, as well as establishing greater consistency and incentive compatibility across program benefit structures. Several sets of actions will contribute to short-term improvements in social protection while moving countries in the direction of essential cover in the long term. This includes efforts to improve the benefit–cost ratios of programs via unbundling of complex, multidimensional benefit packages; undertaking microefficiency reforms in countries' health and pension systems; and harmonizing rules, eligibility requirements, and benefits levels across programs and institutions. Such measures will enable greater worker mobility and provide the foundation for more effective social protection for all citizens.

Notes

1. Depending on how programs are financed, legitimate efforts to cover increasingly large segments of the population could also have unintended negative effects on a country's long-term growth. For example, Levy (2006b) argues that increases in social protection spending in Mexico have come at the expense of investments in infrastructure that are necessary to keep Mexico's economy competitive in the global economy and, if continued, will have adverse effects on the country's growth performance. Similar concerns have been raised about the costs of high (and inefficient) levels of social protection spending on infrastructure investment and, thus, future growth and competitiveness in Brazil (see World Bank forthcoming).

2. This conceptual framework has been developed and extended progressively in a series of studies on social protection in Latin America and the Caribbean, including Baeza and Packard (2006), de Ferranti et al. (2000), and Gill, Packard, and Yermo (2004).

3. The framework is based on Ehrlich and Becker (1972) and Gill and Ilahi (2000).

4. For a more detailed discussion of the factors that have driven increases in informality in the region, see chapter 4.

5. While the rural pension system in Brazil is formally contributory, contributions are not linked to salaries or individual income, but to rural production. Benefits are not linked in any way to past contributions, but are defined in relation to the minimum wage, and the system is heavily subsidized. For a detailed discussion, see Schwarzer and Querino 2002.

6. Indeed, the BONOSOL program covers nearly 75 percent of the population and in a very egalitarian manner. As can be seen in figure 7.5, in the absence of this scheme (that is, relying only on contributory pensions), Bolivia would be among the lowest-coverage countries in the region (Rofman and Lucchetti 2006).

7. Absolute levels of coverage by *Seguro Popular*, which was launched in 2004, have increased since the data shown in figure 7.6 were collected. Newer data suggest that *Seguro Popular* covered as many as 11.5 million individuals by the first quarter of 2006 (Gakidou et al. 2006). According to the October 2005 census, 21 percent of *Seguro Popular* affiliates belong to the second income (wealth) decile and 19 percent belong to the poorest decile. The fraction declined sharply as income increases.

8. See Lindert, Skoufias, and Shapiro (2006), who present OECD data from 2001. Using a smaller sample of Latin American countries, they find that average public social protection spending in the region was approximately 5.7 percent of GDP around the late 1990s and early 2000s.

9. To the extent that financing of expanded assistance programs comes at the expense of infrastructure or other investments that increase a country's competitiveness, such efforts may also have unintended negative effects on long-term growth (see Levy 2006b; World Bank forthcoming).

10. As chapter 2 shows, obtaining access to health insurance through a relative is an important reason for not affiliating in Colombia (table 2.13).

11. Notable exceptions to this are seen in Colombia and Uruguay (Baeza and Packard 2006).

12. This appears to be the case, notwithstanding important inefficiencies in Brazil's pension system. For more detailed analysis of the strengths and weaknesses of the Brazilian pension system, see World Brazilian Bank (forthcoming).

13. In Mexico, for example, only 2 percent of households have private insurance (Mason et al. forthcoming), while in Colombia, less than 5 percent of formal independent workers and less than 1 percent of informal workers (either independent or salaried) have private health insurance (see ch. 2, table 2.12).

14. Indeed, in Bolivia, 42 percent of informal salaried workers and 55 percent of informal independent workers reported that the main reason for not contributing to social security was lack of knowledge about how the system works (see ch. 2, table 2.12).

15. High rates of immunization play a prevention role not only for those directly immunized, but also for society in general. As such, they have a strong public goods element to them. So, too, do goods

that help prevent the transmission of sexually transmitted diseases (for example, HIV/AIDS), such as condoms, which not only serve to protect the individual users, but also help to slow the spread of such diseases more generally (Packard 2006).

16. Data collected by the IMSS between 1997 and 2005 indicate that roughly two-thirds of the almost 9 million IMSS-affiliated workers these are “low-income workers,” earning less than three minimum wages (Levy 2006a). It is possible that most if not all of these workers are unable to access the IMSS housing subsidy (or their related contributions).

17. These figures almost certainly understate mobility across the formal and informal sectors in the Mexican economy, as the data do not capture rural workers. It is not uncommon for rural workers to migrate for work—and move in and out of formality—on a seasonal basis (Levy 2006a, 2006b).

18. At first glance, these patterns appear to be at odds with other recent findings for Argentina and Uruguay, based on similar types of administrative data. For example, Farall et al. (2003) and Bertranou and Sánchez (2003) for Argentina and Lagomarsino and Lanzilotta (2004) for Uruguay present bimodal densities of pension contributions (although with significant numbers of individuals in intermediate cases, so there still appears to be a continuum of cases in their data). It is important to note, however, that these other analyses do not control for the age of the workers. This is important, as younger workers (aged 18–26) tend to be overrepresented among those with a low density of contributions, while older workers (aged 40–48) tend to be overrepresented among those with a higher density of contributions. Once the age is controlled for, one finds the type of multimodal distribution found in Bucheli et al. (2006) and shown in figure 7.11. The authors thank Alvaro Forteza for clarifying this issue.

19. Only when the worker reaches age 70 is he or she entitled to a special (small) pension—although even that requires a minimum of 15 years of contributions to qualify.

20. As might be expected, Bucheli et al. (2006) also find that achieving the required years of formal sector service is more difficult among private sector than among public sector workers.

21. Another related issue concerns the general lack of portability in the region’s pension systems, whether in and out of the formal sector, across the private and public sectors, or across specialized occupational schemes that exist in some countries. Although there is little or no direct evidence to date on the effects of the lack of portability of pensions on informality, it stands to reason that by curbing job mobility, lack of portability also impedes entry (and exit) from the formal sector.

22. Maloney (2001) and Loayza and Rigolini (2006) use cross-sectional estimation techniques on cross-country data, while Krebs and Maloney (1999) do simulations based on labor transitions in Mexico.

23. In the Brazilian case, if capital is not exempted, the authors find a decline in physical accumulation that decreases the demand for formal workers in the long run and leaves informality relatively unchanged (Fernandes, Gremaud, and Narita 2006).

24. It should be noted that the MAROP program was launched in 2006 by the Fox administration, but as of the writing of this chapter, it has not been implemented.

25. Universal health insurance for pregnant mothers and children under five years of age and subsidized child care for those working in the informal sector were campaign proposals during the presidential campaign of Enrique Calderón and, at the time of this writing, were under development by the new Calderón administration.

26. While Gasparini, Haimovich, and Olivieri’s (2006) results are illustrative, they need to be interpreted with some care. Their findings are not robust to all econometric specifications tested. Moreover, they note that the control group may well differ from the treatment group in some dimensions that are not observed and cannot be controlled for in the data. If such factors affect the probability of finding a formal job, then their econometric results could be biased.

27. See the section titled “Managing the transition from here to there” later in this chapter. For more detail on leverage points for improving health sector efficiency in Latin America and the Caribbean, see Baeza and Packard (2006).

28. Estimates of the likely fiscal costs of a well targeted social assistance pension are presented later in the case of Mexico.

29. How social assistance pension benefits are adjusted for inflation over time can also affect the incentive structure across programs. If benefits are adjusted by the Consumer Price Index rather than by real wages, they will keep their value in terms of protection against poverty, but become an increasingly less attractive alternative to participation in the savings pillars as the economy and people’s real incomes grow.

30. The unit cost of *Seguro Popular* was equal to Mex\$8,500 in the fall of 2004. This included a cost of Mex\$7,500 for the basic package, but an additional Mex\$1,000 allocated for catastrophic health care. It should be noted that although *Seguro Popular* is called “insurance,” at present there are very few premiums charged to affiliates under the program. In principle, premiums will be charged on a graduated scale, based on household income, with the poorest Mexicans receiving a full (100 percent) subsidy, but with progressively wealthier households paying progressively higher premiums (see Mason et al. forthcoming).

31. A study by Levy and Dávila (2003) examines the option of raising value-added taxes as a revenue-enhancing measure and then providing compensating transfers to the poor. They similarly find this approach a viable and efficient option.

32. Estimates of the potential costs of promoting health insurance portability in Chile (Baeza and Copetta 1999) suggest that moving from payroll tax to general revenue-financing of health insurance could have a neutral fiscal effect or even lead to tax savings—at least in the Chilean context—if greater formal sector participation and earnings due to the removal of payroll taxes translate into higher earnings, and, thus, more income tax revenues.

33. For a more detailed discussion of informality, taxation, and the social contract, see chapter 8.

References

- Arias, O. 2004. “Poverty in El Salvador during the 1990s: Evolution and Characteristics.” Photocopy. Poverty Reduction and Economic Management Unit, Latin America and the Caribbean Region, World Bank, Washington, DC.

- Baeza, C., and C. Copetta. 1999. "Análisis conceptual de la necesidad y factibilidad de introducir mecanismos de ajuste de riesgo en el contexto de portabilidad de los subsidios públicos en el sistema de seguros de salud en Chile." CLAISS, Santiago, Chile.
- Baeza, C., and T. Packard. 2006. *Beyond Survival: Protecting Households from Health Shocks in Latin America*. Washington, DC: World Bank.
- Beneke de Sanfeliú, M., and M. Shi. 2004. "Dinámica del ingreso rural en El Salvador." Photocopy. FUSADES, San Salvador, El Salvador.
- Bertranou, F., and A. Sánchez. 2003. "Características y determinantes de la densidad de aportes a la Seguridad Social en la Argentina 1994–2001." International Labour Office, Geneva.
- Bitran, R., U. Giedion, and R. Muñoz. 2004. "Fondos de riesgo, ahorro y prevención: estudio regional de políticas para la protección de los más pobres de los efectos de los choques de salud: Estudio de caso de Chile." Background paper for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Bourguignon, F., M. Cicowiez, J. J. Dethier, L. Gasparini, and P. Pestieau. 2004. "What Impact Would a Minimum Pension Have on Old Age Poverty? Evidence from Latin America." Paper presented at the Keeping the Promise of Old-age Security Conference, Bogotá, Colombia, June 22–23.
- Bucheli, M., N. Ferreira-Coimbra, Á. Cortez, and I. Rossi. 2006. "El acceso a la jubilación o pensión en Uruguay: ¿cuántos y quiénes lo lograrían?" Economic Commission for Latin America and the Caribbean, United Nations, Montevideo, Uruguay.
- Cunningham, W. 2001. "Breadwinner Versus Caregiver: Labor Force Participation and Sectoral Choice over the Mexican Business Cycle." In *The Economics of Gender in Mexico: World, Family, State, and the Market*, ed. E. G. Katz and M. C. Correia, 85–132. Washington, DC: World Bank.
- de Ferranti, D., G. E. Perry, F. H. G. Ferreira, and M. Walton. 2004. *Inequality in Latin America: Breaking with History?* Washington, DC: World Bank.
- de Ferranti, D., G. Perry, I. Gill, and L. Servén. 2000. *Securing Our Future in the Global Economy*. Washington, DC: World Bank.
- Ehrlich, I., and G. Becker. 1972. "Market Insurance, Self-Insurance and Self-Protection." *Journal of Political Economy* 80: 623–48.
- Farall, R., D. Aisenberg, M. Bourquin, E. Caviglia, F. Dabusti, S. Mychaszula, A. Sanmiguel, and D. Zagaglia. 2003. "Estudio sobre la frecuencia de los aportes en el SIJP. Una primera aproximación hacia las historias laborales." In *Historias Laborales en la Seguridad Social*, 17–36. Social Security Secretariat Publication Series, Ministry of Labor, Employment, and Social Security, Argentina.
- Fernandes, R., A. P. Gremaud, and R. D. T. Narita. 2006. "Tax Reform and Formalization of the Economy: Simulating Different Alternatives for Brazil." Photocopy. Department of Economics, University of São Paulo, Brazil.
- Fernandez, J. M. 2004. "Background Note on Spain's Health Reform for the Regional Study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*." World Bank, Washington, DC.
- Fiedler, J. 2004. "Out-of-pocket Health Expenditures, Risk Pooling, Savings, and Prevention: A Honduras Case Study." Background paper for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Fiorito, R., and F. Padriani. 2001. "Distortionary Taxation and Labor Market Performance." *Oxford Bulletin of Economics and Statistics* 63 (2): 173–96.
- Forteza, A. 2007. "Old-age Protection in the New Millennium: Chile's Proposed Pension Reform." Background note for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Gakidou, E., R. Lozano, E. González-Pier, J. Abbott-Klafter, J. T. Barofsky, C. Bryson-Cahn, D. M. Feehan, D. K. Lee, H. Hernández-Llamas, and C. J. L. Murray. 2006. "Assessing the Effect of the 2001–06 Mexican Health Reform: An Interim Report Card." *Lancet* 368: 1920–35.
- Galiani, S., and F. Weinschelbaum. 2006. "Modeling Informality Formally: Households and Firms." Photocopy. World Bank, Washington, DC.
- Gasparini, L., F. Haimovich, and S. Olivieri. 2006. "Labor Informality Effects of the *Programa Jefes de Hogar*." Photocopy. CEDLAS (Centre for Distribution, Labor and Social Studies, National University of La Plata), Buenos Aires, Argentina.
- Gill, I., and N. Ilahi. 2000. "Economic Insecurity, Individual Behavior, and Social Policy." Office of the Chief Economist, Latin America and the Caribbean Region, World Bank, Washington, DC.
- Gill, I., T. Packard, and J. Yermo. 2004. *Keeping the Promise of Social Security in Latin America*. Washington, DC: World Bank.
- Glewwe, P., and G. Hall. 1998. "Are Some Groups More Vulnerable to Macroeconomic Shocks Than Others? Hypothesis Tests Based on Panel Data from Peru." *Journal of Development Economics* 56: 181–206.
- Government of Chile. 2006a. "Informe del Consejo Asesor Presidencial para la Reforma Previsional." Santiago, Chile.
- . 2006b. "Reforma Previsional. Protección para la Vejez en el Nuevo Milenio." Santiago, Chile.
- Gruber, J. 1995. "The Incidence of Payroll Taxation: Evidence from Chile." Working Paper 5053, National Bureau of Economic Research, Cambridge, MA.
- Holzmann, R., and S. Jorgensen. 2000. "Social Protection as Social Risk Management: Conceptual Underpinnings for the Social Protection Sector Strategy Paper." *Journal of International Development* 11: 1005–27.
- Jalan, J., and M. Ravallion. 1999. "Are Poor Less Insured? Evidence on Vulnerability to Income Risk in Rural China." *Journal of Development Economics* 58: 61–81.
- Klugman, J., D. Kruger, and K. Withers. 2003. "Consumption Risk and Smoothing during Disasters: The Case of Hurricane Mitch in Nicaragua." Photocopy. World Bank, Washington, DC.
- Krebs, T., and W. Maloney. 1999. "Quitting and Labor Turnover: Microeconomic Evidence and Macroeconomic Consequences." Policy Research Working Paper 2068, World Bank, Washington, DC.
- Lagomarsino, G., and B. Lanzilotta. 2004. "Densidad de aportes a la Seguridad Social en Uruguay: Análisis de su evolución y

- determinantes a partir de los datos registrales de historia laboral (1997–2003).” Document prepared for the team for workers’ representation at the Banco de Previsión Social, Montevideo, Uruguay.
- Levy, S. 2006a. “Productividad, Crecimiento y Pobreza en México: ¿Qué sigue después de Progres-Oportunidades?” Photocopy. World Bank, Washington, DC.
- . 2006b. “Social Policy, Productivity and Growth.” Background paper for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Levy, S., and E. Dávila. 2003. “Taxing for Equity: A Proposal to Reform Mexico’s Value-added Tax.” In *Latin American Macroeconomic Reforms: The Second Stage*, ed. J. A. Gonzalez, V. Corbo, A. O. Krueger, and A. Tornell. Chicago: University of Chicago Press.
- Lindert, K., E. Skoufias, and J. Shapiro. 2006. “Redistributing Income to the Poor and the Rich: Public Transfers in Latin America and the Caribbean.” Social Protection Discussion Paper 0605, World Bank, Washington, DC.
- Loayza, N., and J. Rigolini. 2006. “Informality Trends and Cycles.” Photocopy. World Bank, Washington, DC.
- Maceira, D. 2004. “Mecanismos de protección social en salud e impacto de shocks financieros: El caso de Argentina.” Background paper for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Maloney, W. 2001. “Self-employment and Labor Turnover in Developing Countries: Cross-country Evidence.” In *World Bank Economists’ Forum*. Washington, DC: World Bank.
- Martinez, S. 2005. “Pensions, Poverty and Household Investments in Bolivia.” Photocopy. Department of Economics, University of California at Berkeley.
- Mason, A. D., J. Saavedra, M. L. Escobar, G. Lopez-Acevedo, and M. Rubio-Sanchez. Forthcoming. “Strengthening Social Protection in Mexico: Recent Progress, Future Challenges.” Mexico Transition Policy Notes, World Bank, Washington, DC.
- Montenegro, F. 2004. “Household Health Expenditures, Financial Protection, and Poverty in Ecuador.” Background paper for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Mossialos, E., A. Dixon, J. Figueras, and J. Kutzin. 2002. *Funding Health Care: Options for Europe*. Berkshire, UK: Open University Press.
- Packard, T. 2002. “Are There Positive Incentives from Privatising Social Security? A Panel Analysis of Pension Reforms in Latin America.” *Journal of Pension Economics and Finance* 1 (2): 89–109.
- . 2006. “On Social Protection Policy and the Informal Economy: An Essay.” Background note for the regional study, *Beyond Survival: Protecting Households from Health Shocks in Latin America*. World Bank, Washington, DC.
- Ravallion, M. 2003. “Targeted Transfers in Poor Countries: Revisiting the Trade-offs and Policy Options.” Social Protection Discussion Paper 0314, World Bank, Washington, DC.
- Rawlings, L., and G. Rubio. 2005. “Evaluating the Impact of Conditional Cash Transfer Programs.” *World Bank Research Observer* 1: 29–55.
- Razú, D. 2006. “Competencia Desleal entre Políticas Públicas en México: El Modelo de Seguridad Social vs. Programas Asistenciales.” Photocopy. Mexico City.
- Rodriguez-Meza, J., and C. Gonzalez-Vega. 2004. “Household Income Dynamics and Poverty Traps in El Salvador.” Photocopy. Ohio State University, Columbus, OH.
- Rofman, R. 2006. “El Sistema de Pensiones.” Bolivia Policy Notes, World Bank, Washington, DC.
- Rofman, R., and L. Lucchetti. 2006. “Pension Systems in Latin America: Concepts and Measurements of Coverage.” Social Protection Discussion Paper 0616, World Bank, Washington, DC.
- Sadoullet, E., F. Finan, A. de Janvry, and R. Vakis. 2004. “Can Conditional Cash Transfer Programs Improve Social Risk Management? Lessons for Education and Child Labor Outcomes.” Social Protection Discussion Paper 0420, World Bank, Washington, DC.
- Savedoff, W. 2004. “Tax-Based Financing for Health Systems: Options and Experience.” Health Financing Issues Paper EIP/ FER/HFP/PIP.04.4, Department of Health Systems Financing, Expenditure, and Resource Allocation, World Health Organization, Geneva.
- Schwarzer, H., and A. C. Querino. 2002. “Beneficios sociales y los pobres en Brasil: Programas de Pensiones no Convencionales.” In *Pensiones No Contributivas y Asistenciales*, ed. F. Bertranou, C. Solorio, and W. van Ginneken, 63–124. Santiago, Chile: International Labour Organization.
- Skoufias, E. 2004. “Household Vulnerability and Its Determinants in Rural and Urban Mexico.” Photocopy. World Bank, Washington, DC.
- World Bank. 2001. *Social Protection Sector Strategy: From Safety Net to Springboard*. Washington, DC.
- . 2003a. “Argentina Crisis and Poverty 2003: A Poverty Assessment.” Report 26127-AR, Poverty Reduction and Economic Management, Latin America and the Caribbean Region, Washington, DC.
- . 2003b. “Volatility, Risk, and Innovation: Social Protection in Latin America and the Caribbean.” *Spectrum*, Washington, DC.
- . 2005. “Shocks and Social Protection: Lessons from the Central American Coffee Crisis.” Report 31-8570-CA, Poverty Reduction and Economic Management and Human Development Sector Management Units, Latin America and the Caribbean Region, Washington, DC.
- . Forthcoming. “Brazil: Towards a Sustainable and Fair Pension System.” Photocopy. Latin America and the Caribbean Region, Washington, DC.
- . Various years. *World Development Indicators*. Washington, DC.

CHAPTER 8

The Informal Sector and the State: Institutions, Inequality, and Social Norms

SUMMARY: This chapter argues that the state–society interactions underlying informality, whether leading to exclusion or to exit, ultimately constitute an indictment of the overall effectiveness and legitimacy of the state. As noted before, burdensome business and labor market regulations, poorly designed social protection systems, and weak enforcement capabilities bias the cost–benefit assessments of firms and workers in favor of informality. Furthermore, a collective perception of ineffectiveness, unfairness, and illegitimacy of the state’s actions, in terms of who it represents and serves, can give rise to a social norm of noncompliance with taxes and regulations (a “culture of informality”), which further undermines the state’s capacity to enforce the law and to provide effective public services. This is partly related to the inability of the state to redress the long-standing high inequality in the region, resolve social tensions, and uphold the rule of law; and it crystallizes in a dysfunctional underlying “social contract.” In Hirschman’s terms, it is a lack of voice and loyalty. Thus, addressing informality in Latin America requires bolstering the competency and legitimacy of the state by delivering the correct set of “carrots” in more equitable public policies and programs to foster a sense of greater inclusion and responsiveness, while carefully considering the incentives created to become formal or informal, and by wielding adequate “sticks” through political resolve and evenhanded enforcement of laws and regulations.

THIS REPORT HAS VIEWED INFORMALITY through the lens of the relationship between economic agents and the state. Several chapters have discussed how specific policies in the areas of labor legislation, business regulation, social protection, and taxation can have a critical impact on microlevel decisions that foster or preclude agents’ participation in the formal economy. These policies define a comprehensive set of incentives to which individuals and firms respond by weighing the costs and benefits of participating in different markets and the enforcement efforts and capabilities of the state. This chapter argues that these economic responses are also influenced by how agents—individually and collectively—perceive and define a relationship with the state. In some situations, there are “social norms” that are influenced by the perception of the effectiveness of the state and the collective projects it represents, and that foster willingness to comply with taxes and

other regulations. For instance, the high tax compliance rates observed in developed countries cannot be explained only by the deterrent effect of the chance of being caught evading and the ensuing penalties.

In Latin American economies, with large informal sectors, the individual cost–benefit analysis leading to *exit* from taxation and other regulations, and from participating in the formal circuit of taxes and transfers, in the Hirschmanian sense, may be influenced by a collective perception that the state’s action is ineffective, inefficient, and unfair. In a social exchange view, willingness to comply with regulations may be affected by the individual perception of the effectiveness of the government in providing services. Moreover, a collective perception—the “sentiment” within a group—about the performance of the state influences the social norm regarding compliance. For instance, if more people operate in the informal sector, it might be easier for each citizen to do so; for a given level of state enforcement

effort, one person will be less likely to be sanctioned if lots of people also evade taxes or do not comply with other regulations; the psychological/ethical costs of evading will be lower if most people in one's peer group do so, and hence one's tax morale and disposition to comply with regulations will be lower. This interdependent behavior might generate a social multiplier that might make a specific social norm more prevalent. This type of mechanism might be behind what has been dubbed a "culture of informality" that prevails in many Latin American countries.

Informality is also a reflection of mechanisms that *exclude* large segments of the citizenry from education, health care, and judiciary services; and from economic opportunities through a segmented labor market and imperfection in other factor markets. This exclusionary process is related to the extremely high and persistent levels of inequality, which are rooted in differences in power, voice, and influence; and which, as reported here, seem empirically highly correlated with informality.

Latin America's low level of trust in the state, its culture of informality, and often the design of its regulations and policies reflect what we might call a "dysfunctional social contract" under which the state is not complying with its designated roles and individuals therefore see little point in playing by its rules. We use the expression *social contract* to refer to some degree of societal consensus over basic aspects of the operation and role of the state relative to the private sector and among citizens. In this usage, *social contract* refers to key aspects of a social equilibrium, including beliefs and actions of citizens, organized groups, and state actors. Among the aspects that enable us to characterize this contract are the structure of taxation and social expenditures (Lledo, Schneider, and Moore 2004), the performance of the state in using citizen's taxes in delivering public goods, and the structure and effectiveness of the social protection systems (Birdsall and Menezes 2005; ECLAC 2006).

Heuristically, this situation may be characterized by a "bad equilibrium" where certain norms are being upheld, implicitly or explicitly. The state might be ineffective, able neither to enforce well its regulations—the sticks—nor to lure firms and individuals to formality by providing quality public goods and services for all—the carrots.

Many Latin American countries seem to be in a situation where the share of informality is high; trust in the state is low; tax morale and regulatory compliance are low and, hence, enforcement of regulations and tax collection are generally low; and public provision of public services is

inequitable and of low quality. Firms, workers, and citizens in general make decisions based on the state's enforcement capacity and on their perceptions about the effectiveness of the state and of prevalent social norms. These decisions, in turn, affect the capacity of the state to enforce regulations and provide high-quality public goods and services for all.

Alternatively, an economy may be in a "good equilibrium" where a large mass of the economy operates formally and public goods and services are provided effectively and where the government is able to enforce the agreed regulations and taxes—which, in turn, is facilitated by a social norm that induces people to comply. Even if economies are, in general, not "stuck" in any of these equilibria, in many cases they have features that resemble these feedback loops that might impede the reduction of informality.

Throughout this report we have described many features that relate to different measures or dimensions of informality and that, indeed, are reflections of a systemic failure. Among those features, we have

- low levels of participation in the social security system
- low coverage of many social insurance schemes, especially among poor people
- a large number of small firms (and larger ones) that partially or completely evade tax, labor, and business regulations
- low-quality regulation that increases red tape
- exclusion in the access to property rights, judiciary services, and other public services
- low-quality public provision of many social services (such as health care or education)
- low levels of trust in the state and in the fairness of dominant arrangements
- low and uneven enforcement
- with exceptions, low levels of tax collection, related to low compliance and low tax bases.

Each of those features is, in itself, a reflection of a dysfunctional individual and collective interaction with the state, intimately linked with the state's inability to perform effectively and equitably its main roles—roles of remedying market failures, coordinating the provision of public goods, and maintaining a level and equitable playing field. Seen from a less state-centered perspective, the features mentioned above may also be a reflection of dysfunctional social equilibrium in the "horizontal" relationships among citizens—how they interact with and the degree to which

they trust each other. In other words, several of the phenomena that may be related to the notion of informality are, in the end, a reflection of the way individuals interact both with the state and with each other.

In this chapter, we explore several characteristics of the social contracts that prevail in the different countries of the region. This chapter also will present suggestive evidence of the kind of feedback loops mentioned above. We dwell on the different manifestations of informality that concern the relationship between the individual and the state, emphasizing how social norms and social interactions also might affect the decision maker. We discuss and present some evidence related to perception and performance of the state and how they correlate with informality. The chapter then focuses on the roles of taxes and how tax compliance is related to the real and perceived role and effectiveness of the state. We also discuss the structure of taxes and transfers and its relation to inequality and informality, as well as other elements that illustrate the exit or opting-out mechanisms that are observed in the region.

Social norms, the state, and informality

Latin America's high level of informality is a manifestation of disconnects between the state and citizens, in part as a result of failures of the state in its various roles. Here we present some evidence of firms' and individuals' perceptions of the performance of the state, and how these perceptions correlate with indicators of informality. To frame the analysis of these correlates, we provide first a framework to discuss different dimensions of the decision to participate in the formal economy.

Dealing—or not—with the state

Throughout the report, we have stressed the cost–benefit analysis that individuals undertake in deciding what sector to work in, whether to register their firm, and whether to pay their taxes or risk detection and punishment. In this section, we further explore how individual perceptions about the performance of the state and how social norms and social interactions might also impinge on these decisions to opt in or out.¹ Consider the case of a worker or a microentrepreneur who has two employment/business opportunities that are somewhat comparable in some other dimensions (such as net earnings), but that differ in the fact that one is formal and requires complying with all regulations, and the other is informal. From the individual point of view, the occupational choice entails having a position

simultaneously in several realms. As an illustration, consider this nontaxonomic list:

1. obtaining the protection of labor laws, paying contributions, and getting social security benefits for the worker and his or her family
2. avoiding the costs of being caught not complying with regulations, given state enforcement technology
3. contributing to the provision of public goods and services to society as a whole, and trusting the state in doing it
4. avoiding the peer-pressure cost of being singled out as a cheater.

To assess costs and benefits, the worker will take into consideration a number of factors that have the state as the main actor. Among others, the informality decision might be affected by the following:

- *The direct costs and benefits of formality:* The assessment of point (1), above, depends on the costs of labor market regulations and a comparison of those costs with the valuation of the benefits provided to the individual (for example, comparing payroll tax payments with health benefits received, as discussed in chapter 7). Note, however, that what the government can provide depends on its fiscal capabilities, which, in turn, depend on the decisions of agents whether to be formal (and pay taxes and contributions). Massive opting out may generate a negative feedback loop that might move the country to a bad equilibrium.
- *State enforcement capacity:* The assessment of point (2) will depend on the perceived capacity of the state to enforce labor, tax, and other regulations. This is a function of the enforcement technology used and the perceived probability of being caught. This is influenced by individual risk aversion and by the collective perception of this probability. In addition, however, the formation of these collective perceptions may be influenced by social multipliers. Another negative feedback loop is possible here because low enforcement capacity implies fewer resources for the state.
- *Individual perceptions about government effectiveness:* The assessment of point (3), above, might depend on how effective and fair the individual perceives government institutions to be in fulfilling its role. The individual might not be amoral and might decide to comply

with regulations not only as a result of the cost–benefit analysis of a utility maximizer, but might also be factoring in the degree of citizen/taxpayer satisfaction with, confidence in, and trust in government. A perception of state incompetence, unfairness, and corruption may affect willingness to comply. The literature finds that, in collective action settings, individuals adopt not a purely materialistic, calculating posture, but a more complex, emotional, and *reciprocal* stand.²

- *Social norms and the social multiplier*: The assessment of point (4) depends on what others believe about the role and performance of the state, that is, the social norm—the pattern of behavior that constitutes a customary rule that coordinates actions among people (Young 2006) and is sustained by social approval—that influences collective behavior toward the state. Deterrence is not enough to explain the observed levels of tax and regulatory compliance, and at least parts of those levels are explained by social interactions. Moreover, if one’s own behavior changes not only because of the influence of an exogenous determinant but also because of the change in the behavior of a reference group, the result is a social multiplier.

The revision of these dimensions of the formality decision illustrates that informality connects to issues pertaining to the reality of and perception about the state. The answer to each of the questions above depends on some characteristics and capabilities of the state and the services it provides, which affect the assessment of the private benefit of engaging with the state. But it also depends on a system of individual and collective beliefs about the effectiveness and fairness of the existing arrangements.

Informality and performance and perceptions about the state

Many Latin American states share the characteristics of what can be called an *exclusionary* state, in itself another manifestation of an imperfect social contract. It is a contract as long as it is an implicit arrangement through which the society has given the state some of the roles mentioned above.

However, in Latin America this contract is failing to define social and economic arrangements that are inclusive and that provide fair rights and responsibilities to all. This generates a perception that the state is not complying fully with its role in the social contract. Manifestation of this is a political equilibrium where certain groups are exempted

from paying directly through taxes, in part because it is too costly for the state to go after them. At the same time, however, they are excluded from the benefits of being protected by the systems regulated by the state or are excluded from receiving public services. Among many others, an example of this exclusion is the inadequate access that the poor have to the judiciary. If, for instance, individuals do not perceive that property rights will be enforced, communities will maintain traditional mechanisms to enforce property rights.³ Another example of exclusion is the existence of large segments of the population that are not covered by any social protection mechanism, the so-called truncated welfare state discussed in chapter 7. And, in a most extreme case of informality generated by an exclusionary process, large segments of the country, often in rural areas, have never been reached by the state. A serious manifestation of this is the problem of unregistered births and undocumented citizens which is still significant in several countries, as the evidence available for Bolivia, Brazil, the Dominican Republic, Nicaragua, and Peru demonstrates (see box 8.1).

A perception that there is an *incompetent state* due to corruption or ineffective governance generates low trust in the state. In addition to directly reducing the benefits of engaging in formal interactions, this might reduce the willingness to comply voluntarily with regulations. Investment climate survey data concerning the efficiency of governments to deliver services show that Latin America appears to fare particularly poorly,⁴ featuring the lowest proportions of firms being satisfied with the efficiency of government services (panel [a] of figure 8.1). Across countries, there is a negative relationship between a proxy of informality and governments’ effectiveness (even after controlling for the level of development). Panel (b) of figure 8.1 shows significant negative correlation between a proxy for government effectiveness⁵ and a proxy for informality. Informality also tends to be higher in countries that are perceived as more corrupt (figure 8.2). As discussed in chapter 6, data for five Latin American countries suggest that the perception of corruption is positively and significantly related to informality.⁶ Companies reporting that bribing government officials to “get things done” is a common practice in their line of business exhibit higher rates of revenue and worker underreporting (see table 6.1). This could be interpreted as firms hiding some of their activities from corrupt officials. But a complementary explanation is that firms that view the government as corrupt are less

BOX 8.1

The extreme of informality and exclusion: Being undocumented

A significant share of the population in several Latin American countries lacks official identity cards, and in this sense is largely invisible to the state. This is an extreme of the continuum of informality by which groups of individuals are absent from the formal circuits of the economy and any interaction with the state. Lack of documentation can have serious consequences, such as making an individual vulnerable to exploitation by employers because of the inability to seek legal redress, preventing access to basic public services and transfer programs, and curtailing the accumulation of human capital through public education. Due to the nature of the phenomenon, regional estimates of the magnitude of this population are hard to come by. However, some representative statistics are available. The Inter-American Development Bank (IDB 2006) reports that demographic and health surveys conducted in Nicaragua in 2001 showed that 17 percent of people over 15 years of age did not have a national identity (ID) card. In the Dominican Republic, a 2004 living conditions survey estimated that about one-fourth of the poorest segment of the population lack birth certificates and identity cards. A survey for a planned project in Argentina indicates that 14 percent of potential beneficiaries lacked national ID cards in one municipality and 17 percent in another municipality. The UK Department for International Development found that, in 2002, between 750,000 and 2 million Bolivian citizens were “functionally undocumented.” Furthermore, it was found that, in some areas of Bolivia, 90 percent of the population lacked a valid form of identification, and the majority was not included in the civil registry at all.

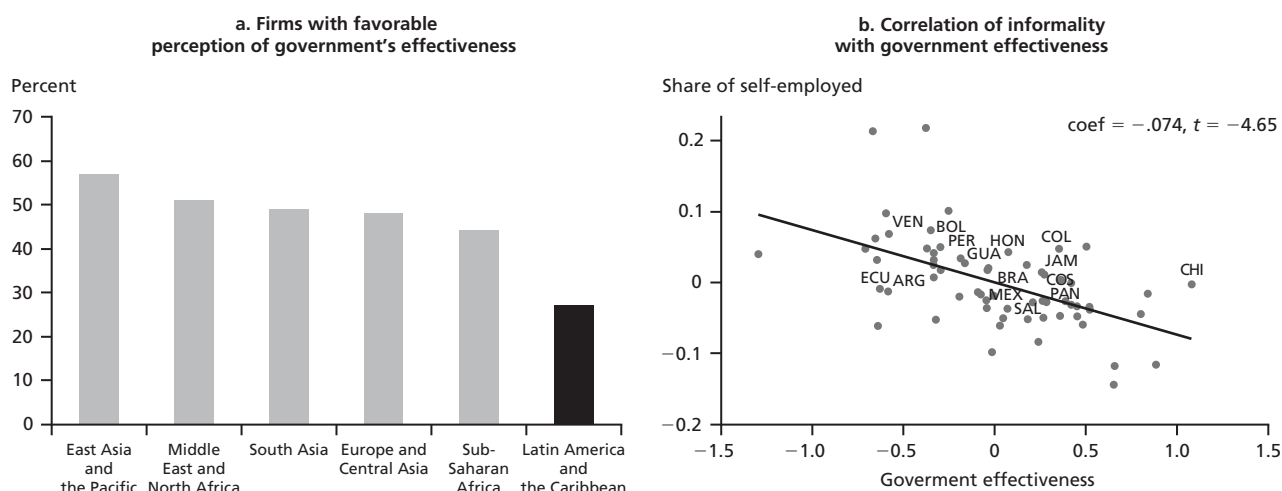
Undocumented adults were, in most cases, once undocumented children who have continued living in a state of official nonexistence. Statistics for birth registrations are more readily available in the region, Duryea, Olgiati, and Stone (2006) find that the lack of birth registration varies from 8.4 percent in Peru to 25.8 percent in the Dominican Republic. They also find that birth underregistration rates are much higher in rural areas than urban areas. In Peru and Bolivia, rural rates are between 25 and 40 percent higher than in urban areas; however, in Nicaragua, rural rates were 200 percent higher than urban rates. Furthermore, the highest birth underregistration rates in Nicaragua are found in areas suffering extreme poverty,

an indication that a key reason for the high rates is the prohibitive cost of traveling to an urban center where there is a government office. Other reasons that are often suggested are a lack of knowledge of the importance of birth certificates and the legal quagmires surrounding registration (for example, if parents are undocumented, it may be very difficult or impossible to register a child).

The effects of not being registered are serious. Using household surveys, Duryea, Olgiati, and Stone (2006) found that for the nearly 300,000 children between 7 and 9 years old who did not matriculate in Brazil, the main reason cited for not enrolling was the lack of documentation. Econometric analysis for the Dominican Republic found that lack of documentation was one of the strongest predictors of school enrollment in higher grades. The International Program on the Elimination of Child Labor—International Labour Organization found that 50 percent of children living in Velleda Morales, a largely indigenous community in Honduras, did not attend school because they lacked a birth certificate. The inability to increase one’s human capital through education has long-term consequences on earnings potential. In contrast, the Inter-American Development Bank (IDB 2006) reports that access to health care services for nondocumented people is much easier in most Latin American countries than access to education—with the exception of Colombia where health services require two forms of identification.

Countries in the region have responded to the registration crisis by combining social targeting programs and registration initiatives. The Inter-American Development Bank (IDB 2006) presents a number of examples, such as Chile where registration has become a part of the nationwide Chile Solidario program that offers, among other things, health care services, counseling, and education services to poor families. Similarly, the Brazilian government has introduced special boats that offer a multitude of services to isolated populations living deep in the Amazon region, including medical services, birth registration, civil marriage, voter registration, and military conscription. Similar initiatives are being developed in the Dominican Republic. Despite these initiatives, additional government efforts and innovations are needed to include these invisible segments of the population in the mainstream of the economy and the social contract.

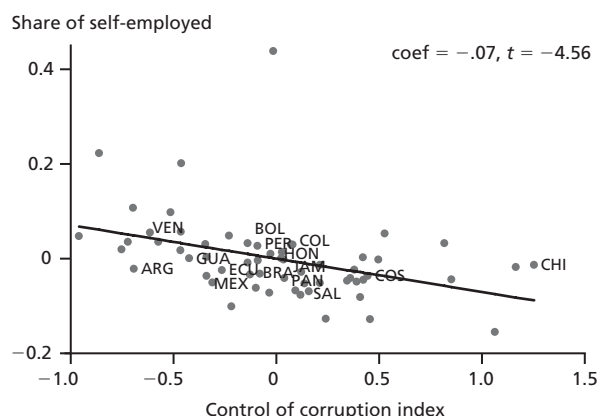
FIGURE 8.1

Informality and government effectiveness

Source: Author's estimates, based on World Bank *World Development Indicators* and *Worldwide Governance Indicators*, (2005); and investment climate surveys.

Note: Figure b shows partial correlations controlling for gross domestic product (GDP) per capita at purchasing power parity (PPP). Government Effectiveness Index measures the quality of public service provision, the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies. Higher values correspond to a more effective government. Firms' perception of government effectiveness is defined as the proportion of firms that report that their government is efficient in delivering services.

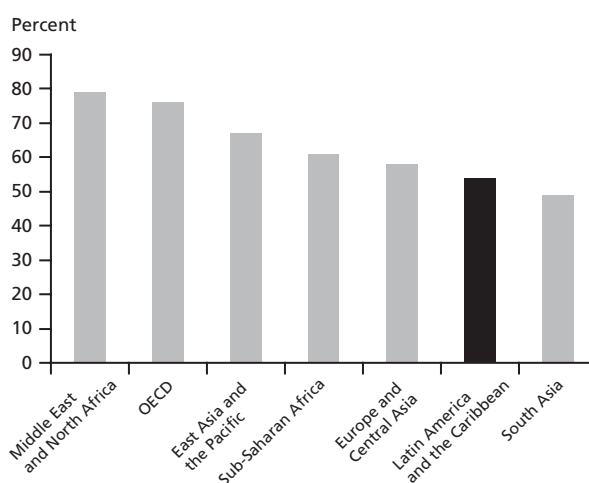
FIGURE 8.2

Informality and corruption

Source: Author's estimates, based on World Bank *World Development Indicators* and *Worldwide Governance Indicators*, (2005).

Note: Figure shows partial correlations controlling for GDP per capita at PPP. The control of corruption index is a measure of perceptions, where corruption is defined as the exercise of public power for private gain, with higher values corresponding to less perception of corruption.

FIGURE 8.3

Share of firms confident that the judiciary will enforce contractual and property rights

Source: Investment climate surveys, (2005), as reported in World Bank (2007).

willing to pay taxes and finance the state. As will be discussed later, this unwillingness relates to the discussion of the so-called tax morale.

Many judicial systems in Latin America are perceived as incompetent, inefficient, and unfair; and this perception

generates very low levels of trust in the judiciary. On average, firms in Latin America are less confident that their judicial system will enforce contractual and property rights disputes than are firms in other regions of the world (figure 8.3). Within the region, there are large differences,

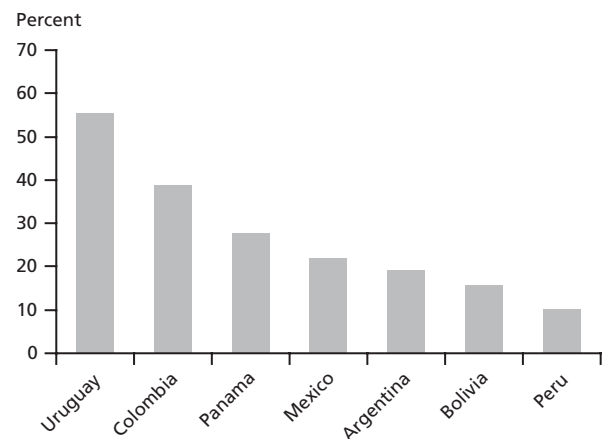
with more than 70 percent of firms being confident in the judicial system in Chile and Costa Rica compared with less than 30 percent in Ecuador and Guatemala. Evidence presented by Biebesheimer and Payne, (2001) shows that although 65 percent of citizens were reasonably confident in their judicial systems in Europe, only 35 percent expressed such confidence in Latin America. In questions asked in recent investment climate surveys, which included a larger sample of small firms, the percentage of firms that perceived courts as fair, impartial, and uncorrupt was less than 25 percent in Panama, Mexico, Argentina, Bolivia, and Peru (figure 8.4).

As discussed by Ronconi (2006), there is ample evidence that people do not trust the judiciary, in part because judges are perceived as corrupt and as benefiting the wealthier side in a lawsuit (Eyzaguirre 1996). Although systematic evidence is lacking, he argues that it is plausible that low levels of trust in the system translate into fewer employees whose labor rights might have been violated bringing their cases to courts because those employees have low expectations of success or because they perceive that they will lack the financial resources to pay for their defense—a perception that often is justified. This perception reduces the probability that employers will be fined and, hence, induces a social norm of noncompliance.

As observed in figure 8.5, impartiality of courts, which the previous figures show is perceived as low in a sample of

Latin American countries, is clearly negatively correlated with informality, even after controlling for GDP. More generally, an indicator of the rule of law from the WGI database, which reflects perceptions of enforcement capacity, of the efficiency of the police and judicial systems, as well as the popular observance of the law, is negatively correlated with informality. Broadly, Latin American countries follow closely the regression line in these graphs.

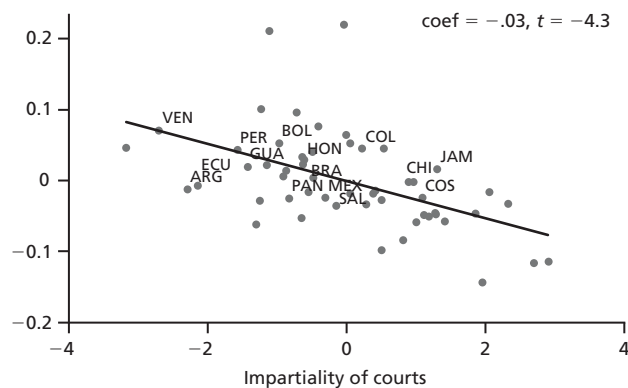
FIGURE 8.4
Share of firms that consider the court system fair, impartial, and uncorrupt



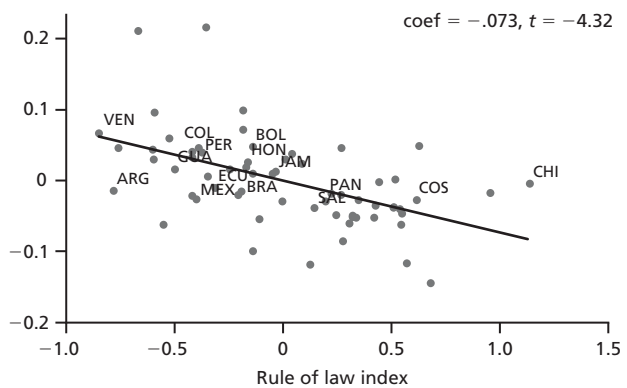
Source: Author's estimates, based on investment climate surveys (2005).

FIGURE 8.5
Informality and state competence indicators

Share of self-employed



Share of self-employed



Source: Author's estimates, based on *World Development Indicators* (2005); Fraser Institute (2004); and *Worldwide Governance Indicators* (2005).

Note: Figure shows partial correlations controlling for GDP per capita at PPP. Impartiality of courts is defined as the degree to which a trusted legal framework exists for private business to challenge the legality of government actions or regulation. The rule of law index measures, in broad terms, the respect of citizens and the state for the institutions that govern their interactions.

A related view of informality posits that it is the consequence of weak Latin American states that assign themselves an unbearable and unmanageable regulatory load, with a scope that goes beyond their enforcement capacity—what Centeno and Portes (2006) call a *frustrated state*.⁷ Chapters 1 and 2 reviewed the evidence supporting the belief that regulatory barriers to entry are positively related to informality. Djankov et al. (2002) and Loayza, Oviedo, and Servén (2005), for example, find that regulatory barriers to entry are positively related to informality.

Loayza and Rigolini (2006) put several of the elements of this literature in a theoretical framework and hypothesize that the size of the informal sector increases with the regulatory burden but decreases with the efficiency in the provision of services and with compliance enforcement. This long-run relationship is tested empirically through a regression in which informality, proxied by the ratio of self-employment to total employment, depends on the flexibility of the business environment, the quality of public services (such as the prevalence of the rule of law, the efficiency of police and judicial systems), and an indicator of the government's ability to monitor taxes and enforce regulations, proxied by government expenditures as a percentage of GDP.⁸ As generally found in the cross-country literature, GDP explains approximately 80 percent of the variation of this proxy of informality

(table 8.1). The first two variables, business flexibility and law and order, that represent the opportunity cost of informality, and government expenditures (which the authors use as a proxy for the capacity to enforce) that represent the direct cost of informality all have the expected signs and are significant when they replace GDP in the specification. When the four variables are included, significance is lost because of the strong correlation with GDP per capita.

Within the same theoretical framework, a simple empirical exercise was conducted to analyze alternative institutional indicators and alternative variables to approximate informality, namely, the share of the informal economy, the ratio of self-employment to total employment, and the ratio of workers without contribution to pensions or social security to the total number of workers.⁹ Even controlling for GDP per capita, in most cases the institutional variables have a significant and negative sign (table 8.2). The dummy for Latin America is significant in specifications for two of the proxies of informality.¹⁰

A final strand of the literature emphasizes the possible existence of a *captured state*, where a political equilibrium is preserved such that elites (business, public sector, or labor) interact explicitly or implicitly with the state to maintain rents, even if that implies the exclusion of certain segments of the population. The typical situation is that of a populist

TABLE 8.1

Long-run informality relationships

Self-employment rate (as ratio to total workers)	OLS		
	(I)	(II)	(III)
GDP per capita (in logs, annual)	−0.0759*** 0.0043		−0.0516*** 0.0060
Business flexibility (index from Fraser Institute, range: 0–10, country average)		−0.0293** 0.0111	−0.0167* 0.0092
Law and order (as % of GDP, country average)		−0.0457*** 0.0072	−0.0191*** 0.0050
Government expenditure (as % of GDP, country average)		−0.0050** 0.0022	−0.0015 0.0015
Constant	0.9065*** 0.0388	0.6954*** 0.0666	0.9030*** 0.0424
Observations (n)	525; 42	525; 42	525; 42
R ²	0.80	0.72	0.85

Source: Loayza and Rigolini 2006.

Note: OLS = ordinary least squares. Robust, country clustered estimations, standard errors.

* significant at 10 percent.

** significant at 5 percent.

*** significant at 1 percent.

TABLE 8.2

Indicators of informality and institutional indicators

Indicators	Definition 1: Share of informal Economy			Definition 2: Share of self-employed			Definition 3: Share of persons without access to pension		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
GDP (in logs)	-5.23 (3.72)***	-3.12 (2.20)**	-3.83 (2.49)**	-7.88 (5.55)***	-7.60 (5.05)***	-6.68 (4.41)***	-22.70 (10.41)***	-21.67 (9.11)***	-23.53 (9.28)***
Regulatory quality index	-3.42 (1.96)*			-3.44 (2.00)**			-5.59 (2.07)**		
Rule of law		-6.19 (3.69)***			-3.60 (2.05)**			-6.36 (2.33)**	
Government effectiveness			-5.03 (2.78)***			-4.82 (2.75)***			-3.79 (1.28)
Latin America	8.89 (3.27)***	6.34 (2.33)**	7.52 (2.74)***	2.45 (1.07)	0.71 (0.28)	0.74 (0.31)	14.06 (3.69)***	11.97 (2.99)***	13.61 (3.38)***
Observations (n)	110	110	110	60	60	60	98	98	98
R ²	0.47	0.52	0.49	0.74	0.74	0.75	0.82	0.82	0.81

Source: Author's estimates, based on *World Development Indicators* and *Worldwide Governance Indicators*, 2005.

Note: Absolute value of *t*-statistic is shown in parentheses. Regulatory Quality Index: higher values correspond to a more effective regulatory policy.

* significant at 10 percent.

** significant at 5 percent.

*** significant at 1 percent.

government, with a political base in the middle and low-middle class, that is unwilling to eliminate subsidies to public services, even if that crowds out expanding the services to poorer segments with less political clout.

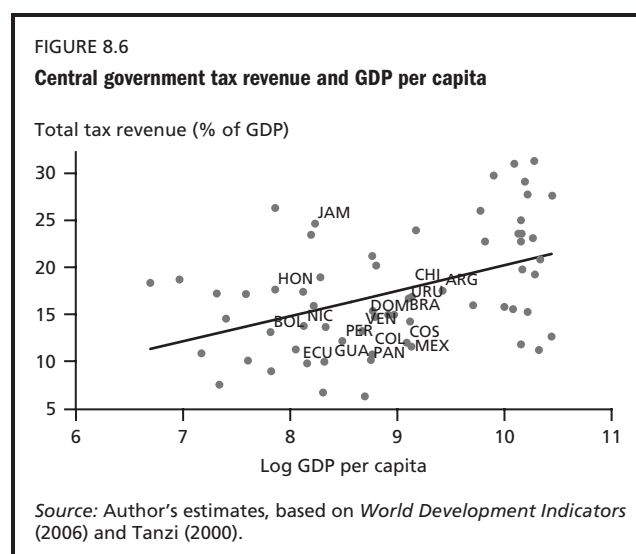
The tax side of the social contract in Latin America

Tax compliance is one of the main components of the decision to engage in the formal economy. In fact, as mentioned in chapter 6, some of the private “benefits” of informality are related to cost savings derived directly from eluding the payment of taxes and avoiding the often complex administrative procedures associated with tax and regulatory compliance. The large informal sector both reflects inefficiencies and inequities of the tax system and defines the magnitude of the challenge to improve tax collection in the region. As long as tax compliance is a social transaction between citizens and the state, it is a key element of the social contract. We review here the main features of the tax system in Latin America and discuss tax compliance behavior and selected policy issues.

Some stylized facts of the tax structure

Tax revenues in Latin America remain below the international norm. Figure 8.6 shows that almost all Latin American countries lie below the regression line relating tax collection to per capita GDP. Lledo, Schneider, and Moore (2004) show that the average tax take has been persistently smaller in Latin America than in OECD countries, while it has been fairly similar to the Asian average (see table 8.3). The median country in the region collects 4 percentage points of GDP less than would be expected, given its level of development (Perry et al. 2006). The average figure for the regions stands at 14.2 percent of GDP in 2000–03 (15.8 percent of GDP including social security contributions). Only Brazil, Nicaragua, Trinidad and Tobago, and Jamaica have tax burdens—without social security contributions—above 20 percent. Guatemala, Paraguay, and the República Bolivariana de Venezuela barely reach 10 percent.

Indirect taxes, in the form of taxes on domestic and internationally traded goods and services, represent the bulk of Latin American tax revenues. They make up about



60 percent of total tax revenues—almost twice as important as taxes on income, profits, and capital gains. By comparison, in developed countries tax revenues from international trade and from domestic goods and services make up 40 percent of total tax revenues (see table 8.4). Undertaxation of income, wealth, and property is a persistent characteristic of Latin American tax systems. In Latin America, direct taxes are about a third of tax collections; in Europe, half; and in North America, that share reaches 80 percent. All the difference is explained by individual income taxes. These are a third of total tax collections in Europe and more than 60 percent in North America. In Latin America, the share is in the single digits—lower than any other region of the world. The share of corporate direct taxes, however, is not small, and it is only below that in Asia. This situation is a reflection of a very low capacity to

observe and monitor incomes; and it is consistent with a concentration of taxes on businesses, presumably medium-size and large businesses that are easier to monitor. The percentage of corporate income tax revenues triples the percentage of personal income tax revenues. Further, when benchmarked by GDP per capita, Latin American collections of personal income taxes appear unduly lower than those in comparable countries. It is interesting to note that in all regions, except Latin America and Africa, direct taxes have increased their share during the last decades.

Tax collection has been rising modestly in most of Latin America since the early 1990s (Lora 2006b). Most of the increase in tax collection since the late 1980s is explained by value-added tax (VAT) collections, which increased from 2 percent of GDP in the mid-1980s to an average of 5 percent in 2003,¹¹ becoming one of the most important sources of revenue for governments in the region. The income tax collection, however, remained flat at around 4 percent of GDP during the same period, while revenues from trade, excise, and other taxes fell sharply (figure 8.7).

Part of the trend observed in VAT collection is related to a small increase in the VAT rate since the mid-1980s, reaching a median rate of 15 percent by 2003 (figure 8.8). There is still great rate disparity in the region, with rates ranging from 5 percent in Panama to 25 percent in Brazil. Tariff tax rates and dispersion fell sharply as a consequence of trade liberalization.

Corporate and personal income tax rates also experienced important changes. Corporate rates decreased from 41 percent in 1985 to 29 percent in 2003. To encourage investment in an environment with higher capital mobility, these rates were set below the maximum rate in developed economies. The personal income tax shows quite

TABLE 8.3
Comparative perspective of tax burdens and structures (percent of GDP)

Economic region	Tax revenue		Direct		Domestic indirect		International trade	
	1975–85	1986–97	1975–85	1986–97	1975–85	1986–97	1975–85	1986–97
OECD	28.3	34.2	9.9	12.4	8.6	10.5	1.0	0.5
Asia (South and Southeast)	14.0	15.8	5.3	5.6	4.2	5.8	4.0	3.4
Africa	17.6	18.8	5.7	6.3	4.3	5.7	6.4	5.6
Latin America	14.9	15.2	3.8	3.4	5.0	5.5	3.1	2.2
Caribbean	23.3	22.4	6.4	0.0	6.7	6.2	6.3	5.4

Source: Lledo, Schneider, and Moore 2004.

Note: Total tax revenue includes social security contributions.

TABLE 8.4

Tax structure by region, selected years, 1975–2002 (percentage of total tax revenue)

Region	Income taxes			Domestic goods and services			International trade
	Total	Individual	Corporate	Total	General consumption	Excises	
<i>North America</i>							
1975–80	78.4	56.9	20.5	15	7.7	6.5	6.6
1986–92	78.8	63.5	14.4	17	9.8	6.3	4.3
1996–2002	83.3	66.3	15.8	14.8	8.8	5.1	1.8
<i>Latin America</i>							
1975–80	32.7	11.1	17.6	40.4	17.1	19.3	26.8
1986–92	31.1	8.5	17.6	47.3	20.9	21	21.5
1996–2002	30.4	6.2	18.5	56.3	34	16.1	13.3
<i>Western Europe</i>							
1975–80	42.7	33.3	8.5	50.6	28.6	16.5	6.7
1986–92	43.4	32.9	9.3	53.4	33.4	14.9	3.2
1996–2002	47.2	32.8	13	52.4	31.8	15	0.3
<i>Asia</i>							
1975–80	38.8	22.9	20.5	37.2	14.3	18.3	24.1
1986–92	39.3	20.8	19.2	39.5	17.4	16.7	21.2
1996–2002	46.9	24.2	21.4	40.2	19.6	15.3	12.9
<i>Africa</i>							
1975–80	32.1	14.6	16.1	29.7	18.4	13.5	38.2
1986–92	27.4	14.6	11.4	31.9	18.3	11.9	40.7
1996–2002	30.7	17.7	11.6	36.2	21.8	11.3	33.2

Source: Bird and Zolt 2005.

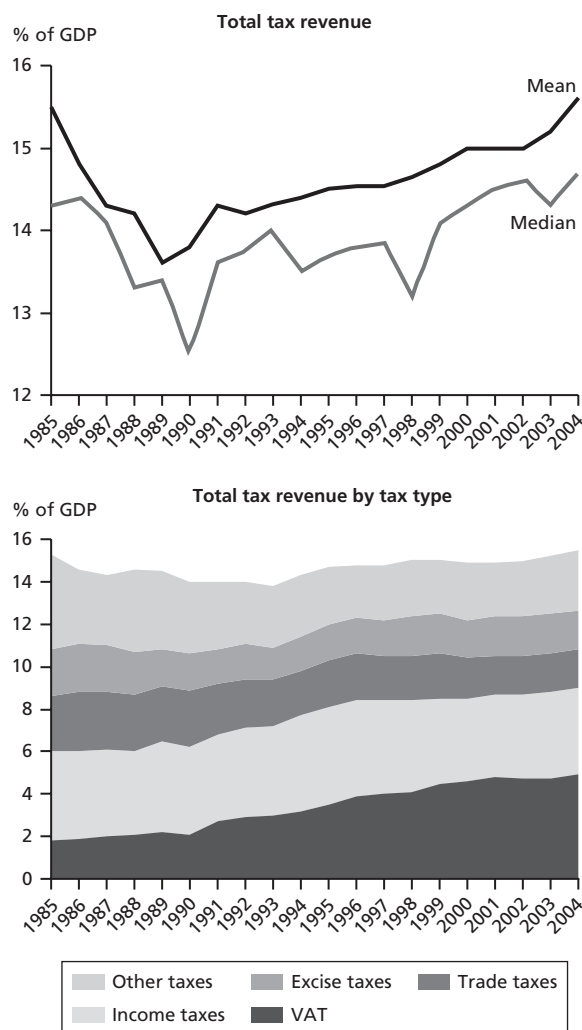
varied approaches in the region. For example, Bolivia has a flat rate tax of 13 percent, whereas Colombia has a number of progressive rates (132 in all), ranging from 0.26 percent to 35 percent; Chile similarly has a very progressive schedule with eight brackets and with minimum and maximum rates of 5 percent and 40 percent, respectively. Personal income maximum rates also decreased sharply in the period. On the other hand, personal exemption levels increased from an average of 60 percent of GDP per capita in 1985 to 230 percent in 2003 (an unusually high level by international standards), and the income levels taxed at maximum rates (the cutoff for the upper-income bracket) were lowered sharply. It is important to note that, in some countries, the personal income tax exemption allows not only for the poor but also for many middle-income and rich workers to be exempted from income taxes.

Overall tax productivity (the ratio between real and potential tax collection, given the basic or maximum rate) has increased during the last 15 years. VAT productivity

rose from 24 percent in 1985 to 34 percent in 2000, while income tax productivity also increased but at a lower pace (figure 8.9). These average figures show still very low levels in the region, although they hide a great deal of heterogeneity. There is great disparity in the region in the overall effectiveness in using the existing tax structure. For example, as shown in figure 8.10, VAT productivity ranges from a high of 0.64 in Chile to a low of 0.17 in Guatemala.

Differences in productivity across countries are probably linked to different VAT structures and to tax expenditures (that is, to multiple tax exemptions and to deductions granted to specific sectors, particular geographic regions, or, in some cases, to specific taxpayers (see table 8.5), as well as government's ability and willingness to administer efficiently the existing taxes. Exemptions are particularly pervasive in the region and are a manifestation of state capture by specific groups. In addition to direct loss of revenues, these loopholes and distortions usually create opportunities for tax elusion. There are also important differences

FIGURE 8.7

Tax revenue of Latin America and Caribbean countries

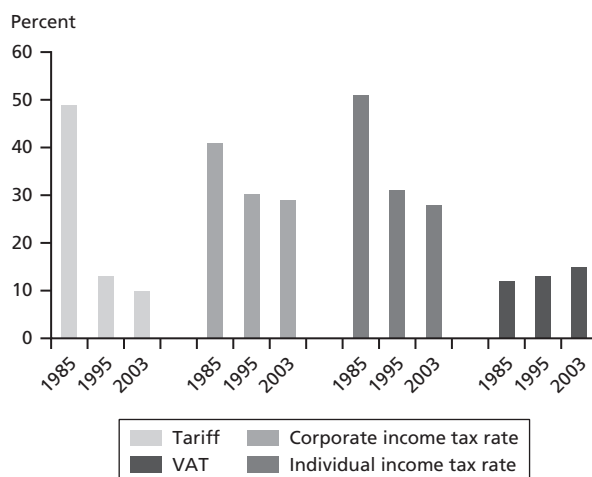
Source: Based on Lora 2006b.

in the administration capacity. Chilean tax administration practices would appear to be over three times more effective than those in Guatemala in raising VAT revenues (Alm and Martinez-Vazquez 2007).

From a different perspective, the productivity of VAT is undercut by informality and difficulty in bringing small businesses and individuals supplying certain services into the system. As observed, there is a clear inverse correlation between VAT productivity and informality (figure 8.10).

So the relatively low level of tax collection in Latin America is a result of tax systems characterized by poor capacity of tax administration, excessive exemptions, and narrow tax bases. As will be discussed in the following section, however, the rich in Latin America pay a much larger share of taxes than the share paid by the rich in developed countries.

FIGURE 8.8

Average tax rate of Latin America and Caribbean countries

Source: Lora 2006b.

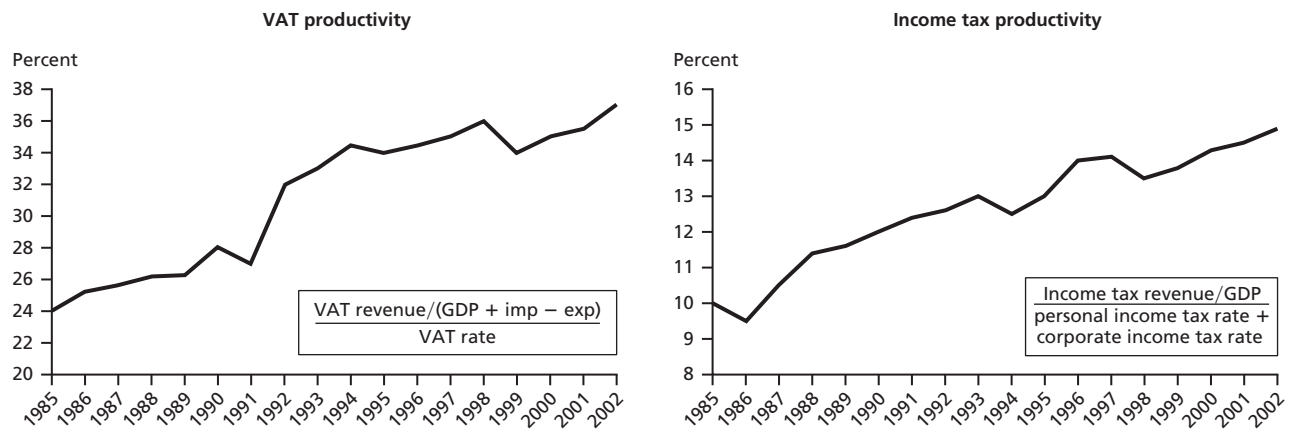
The cross-country literature also supports the view that the higher the tax burden, the larger the informality. In cross-country studies, Johnson, Kaufmann, and Zoido-Lobaton (1998), using indicators of perceptions of the tax burden—including both how high tax rates are and the discretionary power of authorities in administering and operating the system—find a large positive effect on informality. They also show that what matters to explain informality are the administration and operation of the tax system rather than the established rates as a key correlate. Friedman et al. (2000) find that higher tax rates are not correlated with a larger informal sector. In fact, the opposite might be true: countries with high tax rates may be those with high benefits of formality.

Tax compliance behavior

From a utilitarian viewpoint, individuals regard tax obligations casually and display no particular moral aversion to evading if they feel they can safely do so. The taxpayer is an isolated expected utility maximizer who makes rational portfolio decisions under uncertainty, given an informational set. In this view, people pay taxes exclusively depending on their perception of being detected and sanctioned (Alm et al. 1995; Cowell 1990). They do not relate tax payment to perceptions of the quality or fairness of the public services received, to perceptions of a moral obligation, or to any social norm. If enforcement is weak and the possibility of a sanction is low, predicted tax evasion would be high. The policy implication of this approach is that

FIGURE 8.9

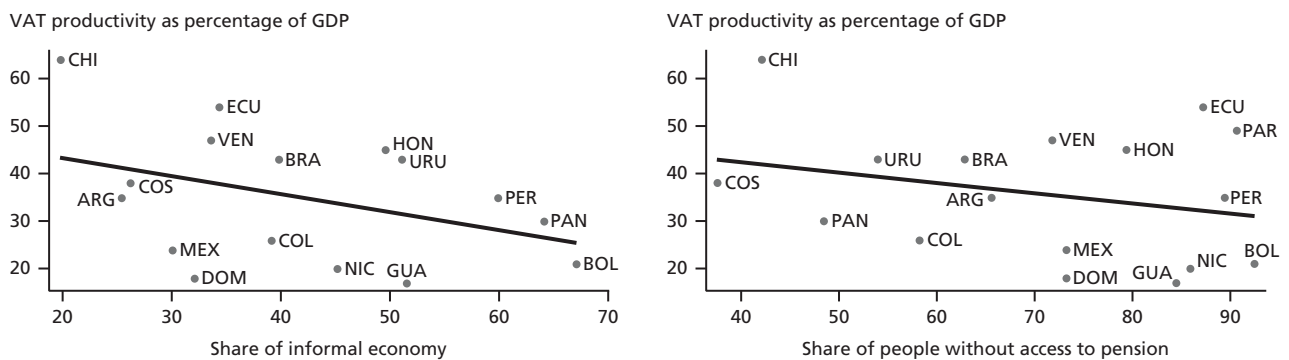
Average VAT and income tax productivity in Latin American countries



Source: Lora, Cardenas, and Mercer-Blackman 2005.

FIGURE 8.10

VAT productivity and informality indicators



Source: Author's estimates, based on Alm and Martinez-Vazquez (2007), and World Development Indicators, 2005.

TABLE 8.5

Tax exemptions in Latin America (percent of GDP)

Country	Total
Argentina	2.4
Brazil ^{a,b}	1.4
Chile	4.2
Colombia	9.2
Ecuador	4.9
Guatemala	7.3
Mexico ^a	6.3
Peru	2.5
Uruguay	5.3

Source: Based on Gómez-Sabaini 2005.

Note: GDP = gross domestic product.

a. Corresponds to federal or central government.

b. Contributions from profits and social security are included.

compliance with taxes—and presumably with most regulations—will depend on the economic consequences of detection and punishment; therefore, greater compliance will depend on deterrence mechanisms.

But tax compliance seems influenced by factors beyond economic ones. For instance, in the United States, the percentage of individual income tax that is audited is quite small, often less than 1 percent; typically, it is even lower in Latin America (Alm and Martinez-Vazquez 2007). And, in any case, very few noncompliance complaints are punished effectively (Tanzi 2000). Given the level of fines and audit rates in most countries, and the available estimates of risk aversion, deterrence models are not able to predict the observed levels of compliance (Alm and Torgler 2006; Feld and Frey 2007), and differences in public attitudes toward

tax laws seem to play a role. In some nations, individuals tend to view paying their taxes as an important civic obligation and, for that reason, are motivated to pay.

An individual's compliance is related to her or his belief that compliance is the social norm. Perception of fairness, trust, and legitimacy in the system might influence a social norm that leads to higher voluntary compliance. Citizens perceive their taxpayer relationship with the state as one of exchange or as a contract. They will avoid taxes if they perceive they are not getting quality government services for the taxes levied on them. Cowell (1990) posits also that the degree of taxpayers' satisfaction with the government affects evasion decisions. If they perceive that this relationship is not in equilibrium, moral costs of evading fall and tax morale is crowded out (Torgler 2005). Using survey and tax return data from a sample of 800 well-off taxpayers in New York, Scholz and Lubell (1998) found that trust in government and in fellow citizens' keeping their side of the social contract significantly influences tax compliance "even after controlling for the influence of any internalized sense of duty and self-interested fear of being caught" (p. 412). Bergman (2002) discusses the contrast between northern Europe, where tax evasion historically has been lower than in Italy, Greece, and Portugal. He also discusses how, on the Iberian peninsula, democratization and the expansion of the welfare state led to improvements in compliance. This strand of the literature views tax compliance as influenced by a "social exchange," a social transaction

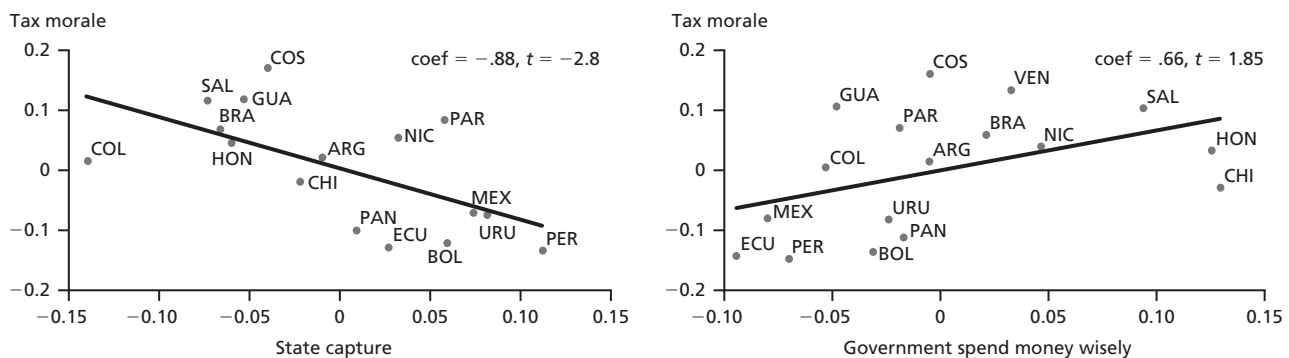
between states and citizens and, as such, views this exchange as the "bedrock of the social contract" (p. 290).

These social norms of tax compliance are measured in the empirical literature by what is termed *tax morale*.¹² For example, Torgler (2005) performs a multivariate analysis of tax morale for Latin America using data from the *Latinobarometro*. The 1998 data come from approximately 15,000 individuals in 10 countries in the region. He finds that Mexico stands out as a country with low tax morale, while South American countries generally have lower tax morale than do Central American and Caribbean countries. Also, he shows that a large majority of individuals perceive tax collection as largely arbitrary and unfair (only 23 percent of those surveyed by *Latinobarometro* in 2003 thought tax collection was "impartial"). Spicer and Becker (1980) provide evidence of a "fairness effect" whereby those who believe they are not being treated fairly by the tax system are more likely to evade.

Cross-country data provide suggestive evidence of the relationship between willingness to comply with tax regulations and perceptions of government's performance. Figure 8.11 shows a clear negative correlation between tax morale and the perception that the government is run according to the interests of a few, a measure of state capture. There is also a positive correlation between tax morale and the perception that the government spends taxpayers' money wisely. In both cases, correlations are significant even after controlling for GDP.

FIGURE 8.11

Tax morale, state capture, and perception that the government spends taxpayers' money wisely



Source: Author's estimates, based on data from *World Development Indicators* (2005), and *Latinobarometro* (2004).

Note: Figure shows partial correlations controlling for GDP per capita at PPP. State capture is proxied by an indicator of the perception about the economy being run according to the interests of a few. To construct the indicator we ask: "In general terms, would you consider that the country is governed according to the interests of a few or is governed for the benefit of the country?"

But perceptions of individuals regarding the effectiveness and fairness of state actions may change through time. For example, Martinez-Vazquez and Torgler (2005) studied changes in tax morale in Spain, using data from the World Values Survey and the European Values Survey, to analyze the evolution at four benchmark years: 1981, 1990, 1995, and 1999/2000. Spain has undergone fundamental changes in the role and effectiveness of the public sector since its

transition to a democratic system after 1975 and after joining the European Union. The country has adopted major tax policy and tax administration reforms, an extensive redirection in public expenditures with the development of a social welfare system, and a significant push for decentralized governance. Martinez-Vazquez and Torgler argue that Spain succeeded in designing general institutional reforms, including tax policy and tax administration reforms, as well

BOX 8.2

Local taxation and social norms

Sokoloff and Zolt (2005) maintain that, in the 19th century, North American and Latin American economies raised national/federal-level taxes in a very similar way, with high reliance on trade and excise taxes. Where the two areas differed was in taxation at the local level. Local governments were far more prominent in North America, and they relied for their revenue primarily on property taxes—a fact that suggests a rather progressive taxation structure. Local governments grew on the basis of these resources and were given by local communities the task of providing public education and investing in roads and other infrastructure. This dynamic suggests closer local control, higher accountability, and the generation of a social norm that induced people to pay taxes. In Latin America, local governments were much smaller, and both local and provincial or national governments did not rely on property taxes; rather, they relied on mechanisms that placed a smaller relative burden on the elites (see table 8B.1).

In the 20th century in North America, the federal government's share of taxation started to increase, and, at both the national and the local levels, the structure of taxation shifted to one that relied on property, income, and sales taxes. In Latin America, however, local governments and local taxation never grew significantly. Most of the increase in taxation there occurred at the national level and was much more timid. As shown in table 8B.1, there was basically no significant increase in the tax take as a percent of gross domestic product during the first six decades of the 20th century, and only between the 1970s and the 1990s was there some increase. This increase in taxation involved greater reliance on consumption taxes in the period, reaching basically half of the tax collection in the 1990s. Property taxes made a very small contribution to revenues collected.

TABLE 8B.1

National-level government tax revenue in Latin America, 1900–2000 (percent of GDP)

Country	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
Argentina	10	7	5	7	8	10	10	8	13	10	14
Bolivia	—	—	—	—	—	—	5	10	5	14	18
Brazil	10	11	9	8	10	7	7	10	10	24	23
Chile	—	—	—	—	9	11	17	16	32	21	24
Colombia	—	—	—	—	4	7	8	10	12	13	14
Costa Rica	—	—	—	—	—	10	12	14	18	23	21
Mexico	5	4	—	6	7	9	8	9	16	16	15
Peru	—	—	—	—	—	11	16	16	17	13	16
Uruguay	—	—	—	—	—	—	—	—	22	24	28
Venezuela, R. B. de	—	—	8	9	12	18	27	19	26	24	20

Source: Sokoloff and Zolt 2005.

Note: — = not available.

as in changing the extent to which citizens identify themselves with the state and the national institutions of the country itself—processes that had a strongly positive effect on tax morale and practically doubled the tax effort for general revenues in the country (from 22 percent of GDP in 1976 to 40 percent of GDP in 2002). Concomitantly, a major increase in formal employment was observed during the 1990s, and tax revenues collected from small firms increased dramatically (Farrell 2006).

The cross-country correlations presented above suggest that willingness to pay taxes is related to perceptions about the performance of the state, among other things. Bergman (2002) goes a step further and tries to test explicitly tax compliance at the country level using micro data surveys in Chile and Argentina. He finds that, in Chile, citizens comply more and are more willing to abide by the rules. Confidence in public institutions generates trust in their ability to use public resources to fund social policies and fight poverty. In Argentina, on the other hand, there is lower satisfaction with less trust in public institutions, and this is behind the observed lower levels of solidarity (see box 8.3).

In a related report, Bird, Martinez-Vazquez, and Torgler (2006) explore the role of governance and, more generally, the quality of the responsiveness of governments to citizens' demands over tax effort. Their basic hypothesis is that although traditional "supply factors" (such as trade openness, sectoral composition of output, GDP per capita, and structure of tax bases) clearly matter in explaining tax effort, there is also a need to account for citizen attitudes in response to government performance as shaped by societal institutions. To account for such "demand factors," or societal institutions, they study the explanatory power of quality of governance indicators (including government effectiveness, rule of law, control of corruption, regulatory quality, and voice and accountability). Using a cross-section of developed and developing countries, and controlling for the above-mentioned supply factors, they show that these "demand" factors matter and that the quality of institutions have a strong positive effect on tax effort.

In addition to this "social exchange" mechanism—that is, I comply because I trust in the state, and the government is responsive, and so I have a moral obligation—tax compliance decisions are also affected by social interactions (Andreoni, Erard, and Feinstein 1998). In a multi-country study of tax compliance, Cowell (1990) reports, "These systematic differences among countries and among groups within one country cannot be dismissed as innate

differences in taste or temperament. Inconvenient though it might appear for neat, individualistic models of economic behavior, people do seem to take into account the 'climate' within the group or groups to which they belong" (p. 102). This relates to the literature of *strong reciprocity*, which is the behavior of so-called emotional and moral reciprocators who condition their contributions to collective goods on the contributions of others, even in fleeting transactions with multiple actors whose behavior they cannot keep track of and whose identities they can never discern (see also Falk and Fischbacher 2005). Schelling (1978) initially posited the idea that interdependent behavior can generate multiple equilibria through a "social multiplier effect" in which one person's behavior influences his or her neighbor's behavior (Glaeser, Sacerdote, and Scheinkman 2003).¹³ An individual's perception of the extent of evasion is a powerful predictor of compliance behavior: the higher an individual believes the rate of tax cheating will be, the more likely he or she is to cheat as well. Individuals prefer to contribute if they believe others are inclined to contribute, but free-ride if they believe that others will do so (see, for instance, Alm, Sánchez, and De Juan 1995; Andreoni, Erard, and Feinstein 1998; Kahan 2005; and Sheffrin and Triest 1992). These behaviors that condition tax compliance and the incentives to informality on "what others do" tend to generate strategic complementarities:¹⁴ if more people operate in the informal sector, it might be easier for anybody to do so; for a given level of state enforcement, an individual will be less likely to be sanctioned if other people also evade taxes. Conversely, given specific changes in tax policy, each person's tax compliance might increase not only because of the policy change, but also because of the change in peers' behavior. Levi (1998) provides a double contingency approach: citizens will pay taxes according to a social norm of reciprocity between taxpayers and the state; at the same time, they will comply based on the perceived fairness of a collective behavior.¹⁵ Using data from *Latinobarometro*, Torgler (2005) finds that knowing about other individuals avoiding taxes has a significant negative effect. He finds that if people trust that others will obey the law and if people trust the president of the country, tax morale is higher.

The interaction of trust among individuals, a perception of fairness in existing institutional arrangements, and a perception that the state fulfills its part of the social contract do not seem to hold in Latin America, and the state

BOX 8.3

Tax compliance, social norms, and trust in the state: The contrasting cases of Chile and Argentina in the late 1990s

Bergman (2002) contributes to the literature on compliance by providing us with two interesting case studies. He uses a multivariate ordinary least squares regression to determine the relationship between trust and perception variables taken from surveys conducted in Chile and Argentina and the “reported willingness of taxpayers to comply with taxes in order to fund social policies” (p. 294).

The surveys Bergman used were conducted in Argentina in 1997 on a sample of 549 individuals, and in Chile in 1998 on a 1,200-person sample. In both countries, 70 percent of the active population paid taxes. In the 1990s, Chile’s level of value-added tax (VAT) evasion averaged 22 percent, compared with a rate of 55 percent in Argentina during the same period. Seventy percent of Chilean taxpayers report VAT evasion as very hard or somewhat difficult. This is in contrast to 67 percent of Argentines who “agree or strongly agree with the statement, ‘People think that in this country evading taxes is easy’” (p. 269). In addition, both countries show a different prioritization of social welfare. In Chile, 76 percent of respondents somewhat agreed or totally agreed with the statement “I am ready to pay more taxes if they are channeled to benefit the poor.” Argentina’s survey shows wide dissatisfaction with social services. Bergman (2002) reports that Argentines “do not challenge the legitimacy of the tax system” (p. 269), but are dissatisfied with the quality of public services.

Using that information Bergman finds that, in Chile, both perceptions of others’ honesty and disapproval of cheating are correlated with citizens’ *willingness to increase their tax burden*, and are statistically significant (see table 8B.3.1). Trust in public agencies (police, public health, courts, customs, and treasury) also has a positive effect on willingness to pay more taxes. The model does not find any effect on the responses caused by the individual’s subjective estimate of the tax authority’s detection capacity.

In the case of Argentina, the dependent variable is actually less stringent because it inquires only about *the willingness to comply with taxes*—not to increase one’s tax burden—to fund social programs. Satisfaction with the performance of public services (health, education, security, and infrastructure) has a positive influence on willingness to comply; tolerance of cheating has a significant

TABLE 8B.3.1

Determinants of reported willingness to increase tax burden, Chile

Independent variable	Coefficient
Collective perception of honesty	0.079**
It is justified to cheat	−0.130***
Trust in and approval of public services	−0.163***
Perception of detection capability	−0.036
$R^2 = 0.12$	

Source: Bergman 2002.

** significant at 5 percent.

*** significant at 1 percent.

TABLE 8B.3.2

OLS—effects of attitudes and tax experience of the willingness to comply, (controlled for age and gender), Argentina

Independent variable	Coefficient
Satisfaction with public services	0.173***
Tolerance of cheating	−0.211***
Feels guilt at evasion	0.289**
Social sanction by peers if taxes are not paid	0.144**
$R^2 = 0.26$	

Source: Bergman 2002.

** significant at 5 percent.

*** significant at 1 percent.

negative coefficient; and, finally, the detection variable again seems to have no effect. Social sanctions and guilt—variables not available for Chile—show, surprisingly, that there is a lack of social norm and individual guilt to pay taxes (see table 8B.3.2).

Bergman concludes that there is a strong link between citizens’ satisfaction with public services and their willingness to comply with tax responsibilities or to increase their tax share in Argentina and Chile.

Chile has a higher level of “social solidarity” (that is, willingness to pay higher taxes) as a result of the populace’s trust in public institutions’ commitment to alleviating poverty. In contrast, Bergman concludes from the Argentine empirical analysis that the lack of satisfaction and trust in public sector institutions is behind the low levels of solidarity. Moreover, legal enforcement there is perceived as weak, as is social enforcement of norms.

Source: Based on Bergman 2002.

might lack the legitimacy for citizens to feel obliged to comply with these regulations. Even if systematic micro evidence is still scarce in the region, the cases of Chile and Argentina suggest a large contrast in the social norm that prevails in those countries. In the same way, attitudes data from *Latinobarometro* suggest that there is high variance of tax morale in the region. Low tax morale leading to tax evasion is related to the Hirschmanian exit option: if the state does not give people public goods that they value, and there is no societal pressure for contributing to or participating in the circuit of taxes and transfers established by the state—and, consequently, there is a social norm of non-compliance—then people will have a strong incentive to go and remain underground. This is illustrated in figure 8.12, which suggests that tax morale is, as a consequence, robustly negatively related to informality in Latin America, using two different indicators of informality. This suggests that countries with high informality tend to be those where the social norm is not conducive to complying with tax regulations. In the case of Latin America, in addition to the problems of perception that the state is weak and does not respond to the interests of the majority, the social norm may be affected by the operation of tax administrative systems. If there is a widespread perception that the government is not willing to fully detect and penalize evaders, then noncompliance is legitimized and the prevailing social norm is reinforced.

Latin American countries currently lack the capacity to enforce tax collection as in OECD countries. In a way, it is a

matter of political will because the technology is available and administrative capacity increases are definitely within the range of options of the much-modernized Latin American tax systems. As an example, in 1947, the United States had a GDP per capita that was lower than the one in Argentina and somewhat above the ones in Chile, Costa Rica, and Mexico in 2000. However, in that year the United States already raised 56.6 percent of its revenues from personal and corporate income taxes (Schmitt 2005). In Latin America, that figure today is 30 percent. In table 8.6, survey data reveal some of the problems faced in Latin America. Levels of bribery and corruption are, in general, higher than in other regions. Underreporting and evasion seem also to be higher, as sales amounts reported to authorities by a typical firm in Latin America are about three-fourths of actual sales, with East Asia and the Pacific the only region where that figure clearly is lower.

There is large scope to increase collection through improvements in tax administration and tax policy. Bird, Martinez-Vazquez, and Torgler (2006), Schmitt (2005), and Tanzi and Zee (2000) discuss several areas where the technology is available, and the challenge is more in the political realm. Among others, Latin American tax authorities have to move aggressively in professionalizing tax administration management and in putting firewalls between tax administration decisions and tax policy decisions. Significant progress has already been made in areas like computerization, control of large taxpayers, and use of withholding systems, together with third-party information. In addition,

FIGURE 8.12

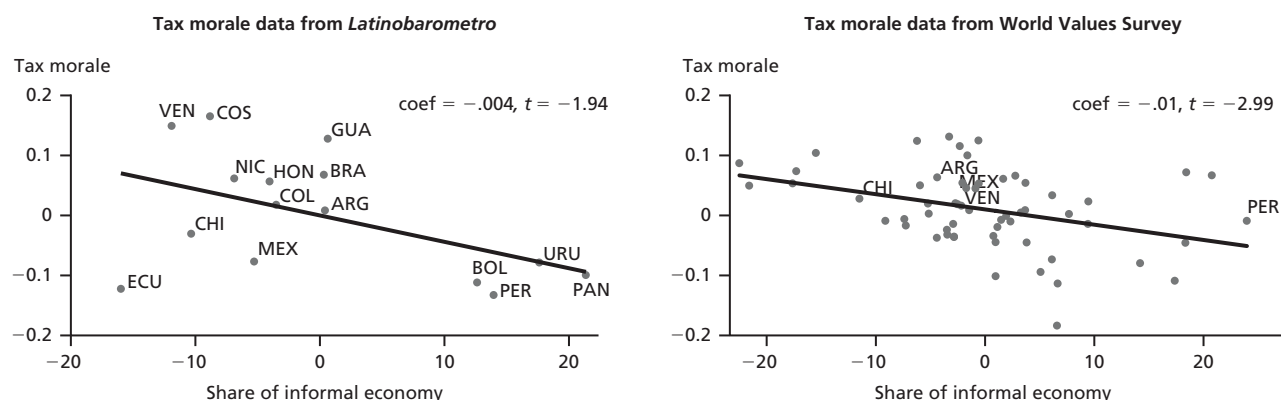
Tax morale and informality indicators

TABLE 8.6

Comparative efficiency and corruption in tax administration: Survey evidence for 2005–06

Indicator	East Asia and Pacific	Europe and Central Asia	Latin America and Caribbean	Middle East and North Africa	OECD	South Asia	Sub-Saharan Africa
<i>Taxes</i>							
Average time firms spent in meetings with tax officials (days)	4.91	2.78	2.89	3.52	1.65	3.37	5.08
Time to prepare and pay taxes (hours)	270.06	437.92	548.80	281.36	197.19	331.71	394.00
<i>Corruption</i>							
Unofficial payments for typical firm to get things done (% of sales)	1.81	0.76	1.48	2.72	0.13	1.28	1.64
Firms expected to give gifts in meetings with tax inspectors (% of total)	33.59	42.84	30.40	40.09	28.26	44.27	18.86
Value of gift expected to secure government contract (% of contract)	1.82	1.36	4.08	1.30	0.55	2.04	3.52
<i>Informality</i>							
Sales amount reported by a typical firm for tax purposes (% of total)	69.30	89.35	76.51	73.55	93.55	93.7	78.39

Source: Alm and Martinez-Vazquez 2007; data from the World Bank Enterprise Surveys.

Note: OECD = Organisation for Economic Co-operation and Development.

in most countries in the region, significant efficiency gains can be obtained by eliminating or lowering exemptions that create loopholes in value-added, income, and property taxes.

But incentives to register with the tax authorities are necessary to increase the tax base—both carrots and sticks. Examples of carrots are programs like the earned income tax credit that provides tax credits to the labor income of families whose annual earnings remained below a certain threshold, often gradually phased out (see box 8.4). These types of programs have had tremendous success in reducing poverty in OECD countries, and have the virtue that transfers are implicitly conditional on the family having someone in regular employment.

Another example of policies oriented to lure the taxpayer is changing the tax administrator's approach to a "service paradigm" so as to enhance its role as a facilitator and a provider of services to taxpayers. This might comprise actions promoting taxpayer education, developing services to assist taxpayers in filing returns, broadcasting advertisements that link taxes with government services, lowering compliance costs, simplifying taxes and their payment, and promoting a taxpayer—and a tax administrator—"code of ethics."

On the side of "sticks," audits and penalties are core instruments of tax enforcement policy. The audit function in most of the region is weak, often underfunded, under-

staffed, and compounded with a legal system too weak to criminalize evasion effectively. Many of the steps that can be taken to improve capacities on this front are technically feasible in almost all countries. Although it is difficult to disentangle their deterrent effects empirically, there is consensus that auditing, fines, and business closures are essential tax enforcement tools. In one of the few studies in Latin America, Bergman (2003) examines the cases of Argentina and Chile, two countries that, as noted before, at the beginning of the 1990s, found themselves with very similar tax policies and macroeconomic conditions but that have since diverged in their tax compliance; today, Argentina has approximately double the VAT evasion of Chile, which is 22 percent. Bergman posits that the key reason for this has been the inability of Argentina's tax system to create a permanent, credible threat to the noncompliant. Argentina has granted a total of 24 major amnesties since 1974, and surveys show that the population believes bribing auditors is common and tax audits have low detection rates. Furthermore, audits in Argentina have been performed almost exclusively on large taxpayers, and smaller firms are aware of that fact. In contrast, Chile has had no major amnesties, and the population perceives it is very hard to bribe auditors. Although it also focuses more on high-level taxpayers, Chile's audit selection system includes in its design a larger number of smaller firms. The efficacy of audits has been very

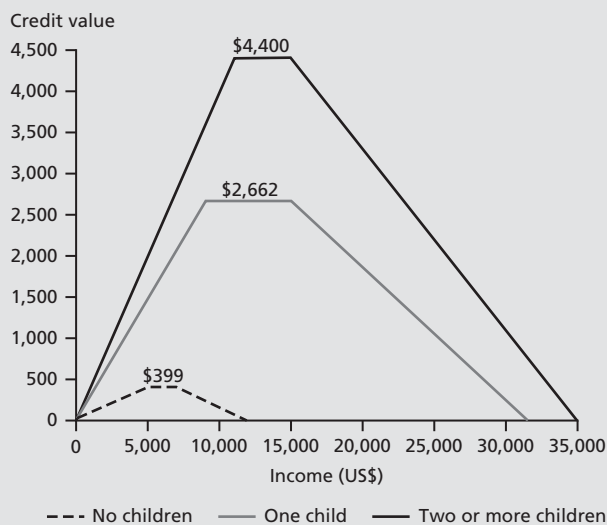
Earned income tax credits: Transfers that encourage formal employment

Fighting poverty and creating incentives to work are challenges faced by many governments. One experience that potentially can offer insights for Latin American countries is the earned income tax credit (EITC). In 1975, the United States introduced an innovative social policy tool that used employment incentives as opposed to public assistance to aid low-income people by providing tax credits to the labor income of families whose annual earnings remained below a certain threshold. More than 30 years later, the EITC has become the largest antipoverty program in the United States: in 2005, 22 million poor working families benefited from the \$34 billion in credit. Other countries, such as Belgium, Canada, France, the Netherlands, New Zealand, and the United Kingdom have since introduced their own versions of earned income tax credits. Using such credits in countries with large informal sectors could induce large segments of the population to register, increasing the ability of the state to observe income for tax or transfer purposes.

The basic design of the credit is this: for each additional earned dollar, workers receive a tax credit, up to an income level considered by the government as the national poverty threshold; after this, the credit plateaus and steadily phases out until an earnings ceiling is reached where eligibility stops. Different thresholds apply to different numbers of children (see figure 8B.4.1). Schedules

also differ among countries (see figure 8B.4.2); indeed, the Dutch have chosen not to phase out at all when the maximum credit level is reached. Their motivation has been to respond to the key theoretical criticism of the labor tax credit, which is that it might provide disincentives to working or earning more in the phase-out range (Hoffman and Seidman 2003; Hotz and Scholz 2003).

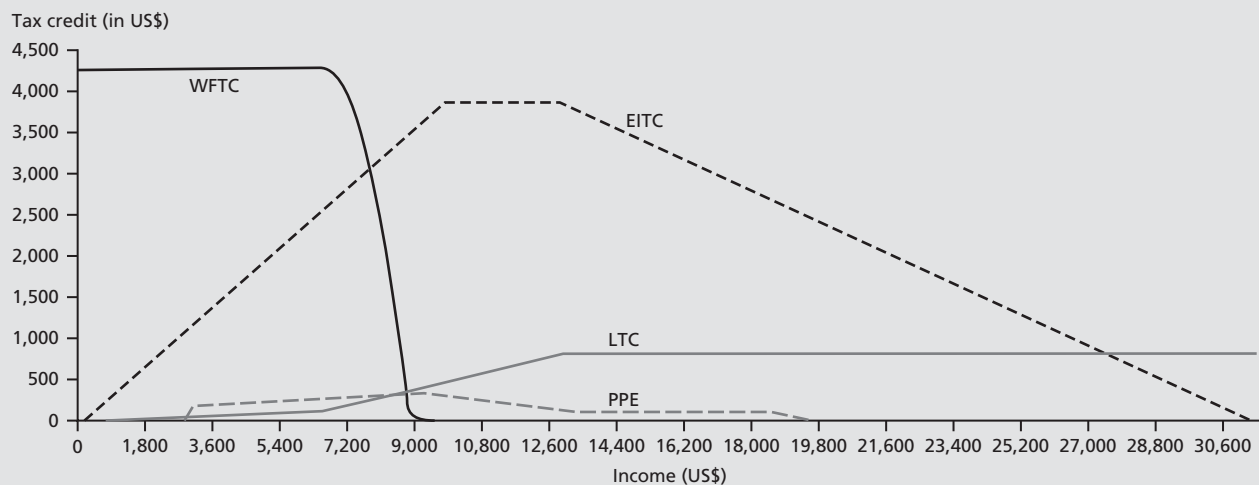
FIGURE 8B.4.1

Value of the EITC, by income, unmarried filers,* 2005

Source: Holt 2006.

Note: Data from Internal Revenue Service. (*) Married couples filing jointly are eligible for slightly higher credit amounts in the "phase-out" range of the EITC.

FIGURE 8B.4.2

Labor tax credits in selected countries, 2001

Source: Detragiache 2001.

Note: WFTC = working families tax credit, UK; EITC = earned income tax credit, USA; LTC = labor tax credit, the Netherlands; PPE = prime pour l'emploi, France.

Other internationally differentiating characteristics are whether the annual credit is refundable or if it applies only to cover one's tax burden, whether families without children are eligible, and whether there is a minimum number of hours one must work to be eligible.

The U.S. EITC has been studied extensively, and the foremost objective of the credit—to reduce poverty—has been a widely documented success. Between 1995 and 1999, the EITC was responsible for an overall reduction in the U.S. poverty rate of 1.5 percentage points (Hoffman and Seidman 2003). Estimates suggest that, without the EITC, the poverty rate among children in the United States would be 25 percent higher (Greenstein 2005). Indeed, the credit has been so successful that, by 2006, 18 individual states in the United States had instituted similar credits for state taxes. The program has also positively affected participation rates among one of the most vulnerable populations in the United States, single mothers. Meyer (2001) finds that 60 percent of increases in single-mother labor participation between 1984 and 1990 are due to the EITC and other tax.

The EITC is not the only tax credit model that seeks "to make work pay." In 1999, the United Kingdom introduced the working families tax credit (WFTC). Brewer et al. (2005) find that the WFTC was responsible for a 5.11 percent increase in the labor supply of single mothers in the United Kingdom between 1999 and 2003. The direct poverty effects of the WFTC have been more ambiguous. A key feature of the WFTC, which has now

been split into the child tax credit and the working tax credit, was that the credit encouraged more hours of work by entitling only individuals working more than 16 hours a week to the credit and providing small additional credit for those working more than 30 hours a week.

The earned income tax credit model's ability to be tailored to fit specific country environments is a characteristic that makes this formula attractive. In the Latin American and Caribbean context, over the medium term, tailored credits could be not only a potent tool for poverty reduction but also an instrument to combat the region's high level of informality. People would have greater incentives to register with tax authorities to receive the credit, providing corresponding incentives for individuals working in informality to exert pressure on their employers to become formal. In addition, disbursement of the credit using official financial intermediaries could be used as a tool to encourage participation in the financial system. In terms of innovative regional design and targeting, we can envision, for example, that poor families involved in the care of the elderly would receive an adjusted credit in the same way that families with children currently do. A primary issue would be the development of good targeting systems in the absence of good income indicators to ensure that the poor are the ones receiving the working credit. Targeting models currently used in conditional cash transfer systems, for example, could provide guidance as to how, in the Latin American context, a working credit initiative could be realized.

different, with Chile focusing on enhancing compliance by matching computerized and third-party information. Argentina relied more on penal sanctions that end up being less effective, given the low expected probability of being detected.

Bergman (2003) provides complementary evidence of the specific impact of audits on tax evasion in Argentina. He finds that, in both Argentina and Chile, audited and sanctioned taxpayers decreased their future compliance in subsequent years, presumably in an effort to recoup the losses prompted by the audit fines, but the decrease was more moderate in Chile. The author attributes this to the fact that, in Chile, the threat posed by the tax authorities is considered credible, as discussed above. The conclusion of this analysis is that tools of deterrence may have unforeseen

effects on tax evasion in a country where tax authorities do not have the ability to produce a credible threat of detection and evenhanded enforcement of sanctions.

Inequality, taxes, and transfers

One of the elements that influence the social norms related to compliance with regulations is the perception of how resources given to the state will be used. On one hand, deficient public services encourage people to opt out of public service systems. On the other hand, a state that is perceived as unfair will lack legitimacy. One element that feeds the perception of unfairness is the real and perceived structure of incidence of taxes and transfers along the income scale. Are taxes excessively concentrated and deemed unfair by some? Are transfers—public expenditures,

in general—excessively concentrated? Are there segment of the population that are excluded from the taxes and transfer mechanism? Are public resources used to equalize opportunities in a way that is consistent with the implicit social consensus? Do they generate an exclusionary mechanism that leaves part of the population out of the social contract and consequently renders the state less legitimate?

We discussed above some stylized facts of the tax side of the equation. As we saw, tax collection in the region is below its expected value, given the level of development. However, social spending does not show a clear pattern. Overall, the ratio of social spending to GDP in Latin America seems to be in line with the region's level of development. There is, however, a large variance within the region (Lindert, Skoufias, and Shapiro 2006). Argentina, Brazil, and Uruguay spend around 20 percent of GDP in social areas, whereas, at the other extreme, the Dominican Republic, Ecuador, Guatemala, and El Salvador show social expenditures below 7 percent of GDP. Mexico stands out as a richer country with relatively low social expenditures as a percentage of GDP. Similar to the small increase in taxes observed since 1985, there is clear evidence of a slightly larger increase in social expenditures. Without exception, social expenditures have increased in all the countries of the region, particularly in countries that started at a low basis in the early 1990s. For the region as a whole, the increase was from 12.8 percent of GDP in 1990/91 to 15.1 percent in 2002/03.

Despite this increase in social spending, patterns of regressivity remain and within-country inequities continue to be very large. In education, spending is mostly progressive, except for tertiary education, and has increased in all countries. But, despite the improvements, quality indicators perform poorly and within-country variance is extremely large. In the case of health care, Latin America has witnessed undeniable progress in the provision of basic services, such as basic health and nutrition services, and indicators like maternal mortality and immunization rates have improved dramatically. But, again, differences in the quality of access by income groups are still extremely large (see de Ferranti et al. 2004; World Bank 2006). Spending on social security, which has expanded the most, is fairly regressive. The pension system is the worst offender, and one of the biggest problems, as explained in chapter 7, is the low coverage. Some social assistance programs are fairly progressive, but, in general, they represent a small fraction of spending. In addition, in Latin America there is abundant evidence of deficiencies and inequities in access to other

services, such as to the judiciary and to police protection. Very much to the point of this report, informal workers are one of the key excluded groups.

In an exercise that aggregates social spending, Breceda, Rigolini, and Saavedra (2007) compare patterns of progressiveness for a sample of Latin American countries with the United States and the United Kingdom. They find that, in addition to significant differences in average levels, there is a strong contrast between Latin America and the United Kingdom in terms of the progressiveness of social spending. In Latin America, social spending is slightly biased in favor of the rich: on average, social spending to the poorest quintile corresponds to 92 percent of social spending to the richest quintile, against 233 percent for the United Kingdom.¹⁶ Latin America seems, therefore, to be closer to the U.S. model, which has a ratio of 107 percent. Although variation in the Latin American sample is quite large (in Honduras—the least progressive state in the sample—the poorest quintile receives 57 percent of what the richest quintile receives, whereas in Colombia the ratio climbs to 108 percent), all countries remain far away from the progressiveness of the European welfare states, and all are less progressive than the United States.

The key element of this discussion for the issue at hand is that there is a large segment of the population for which the state is not providing basic services, for which the state is not a guarantor of minimum opportunities, and for which the provision of public goods is insufficient and, equally important, inequitable.

Inequality and the system taxes and transfers

It is well known that Latin America is the most unequal region in the world, as measured by income distribution (de Ferranti et al. 2004; Perry et al. 2006). Recent research work at the World Bank (2006) has shown that inequality of opportunities is one of the key factors. Differences across individuals in access to assets, public goods, services, and voice (for which individuals could hardly be responsible because the differences arise at the moment of birth) determine differences in their ability to contribute to development and to build their own future. Accumulated differences in opportunities translate eventually, among other ways, to differences in income.

A significant part of the observed inequality is due to the effects of state intervention (or lack thereof) more than to pure market outcomes. As discussed in Perry et al. (2006), if income inequality in Latin America is compared with that

of Europe, there is evidence that a significant part of the difference between the levels of disposable income inequality in the two regions is due to the different impact of taxes and transfers: they reduce market income inequality considerably in Europe, and very little in Latin America.¹⁷ Interestingly, the same can be said even when comparing the Latin American countries with the United States, a country that has the reputation of not being too distributive (Alesina and Glaeser 2004). The inability or unwillingness of Latin American political and economic systems to improve the distribution of income is not a new event, nor has it passed unnoticed by previous observers. For instance, Kuznets (1955) argued about the “the failure of the political and social systems of underdeveloped countries to initiate the governmental or political practices that effectively bolster the weak positions of the lower-income classes” (p. 24; cited in Beramendi and Díaz-Cayeros 2006).

Breceda, Rigolini, and Saavedra (2007) present an incidence analysis of both social spending and taxation for seven Latin American countries, and make the comparison to the United Kingdom and the United States. Consistent with previous studies, they find that, in Latin America, absolute levels of social spending are fairly flat across income quintiles—and in some countries they are even regressive (see figure 8.13). Nonetheless, they find taxation to be quite progressive, particularly income and value-added taxes. On average, the richest income quintile in Latin America pays 22 times more taxes than the poorest quintile. This remains close to the difference in the United States (19 times) and much higher than in the United Kingdom (6 times). Moreover, in Latin America, the share of taxation paid by the richest income quintile averages 61 percent. This remains significantly higher than the share paid by the richest quintile in the United Kingdom (43 percent), and similar to what the richest quintile contributes in the United States (58 percent).

Both features make the structure of social spending and taxation in Latin America closer to that of the United States than to that of the United Kingdom (where social spending is more progressive and taxation is less so). The comparison, therefore, strongly indicates that Latin American countries resemble a minimalist welfare state similar to the one in the United States, more than a Europe-like one. The extremely high inequality levels observed in Latin America make the transition difficult toward a more progressive welfare state. In particular, although the rich in Latin America are taxed equal to or less than the rich in

many European countries (as a proportion of their income), given the high inequality, their contribution to total tax revenues is much larger (despite the elements of state capture mentioned above, including the high level of tax exemptions that, in many cases, favor the rich disproportionately). This may place a strain on the social contract because the richest quintiles have to pay for most of the state provision of public services without seeing sizable benefits in return. These findings give further support to the argument that an increase in the efficiency of the system and an increase in the tax base are the key avenues to increase tax collection in the region.

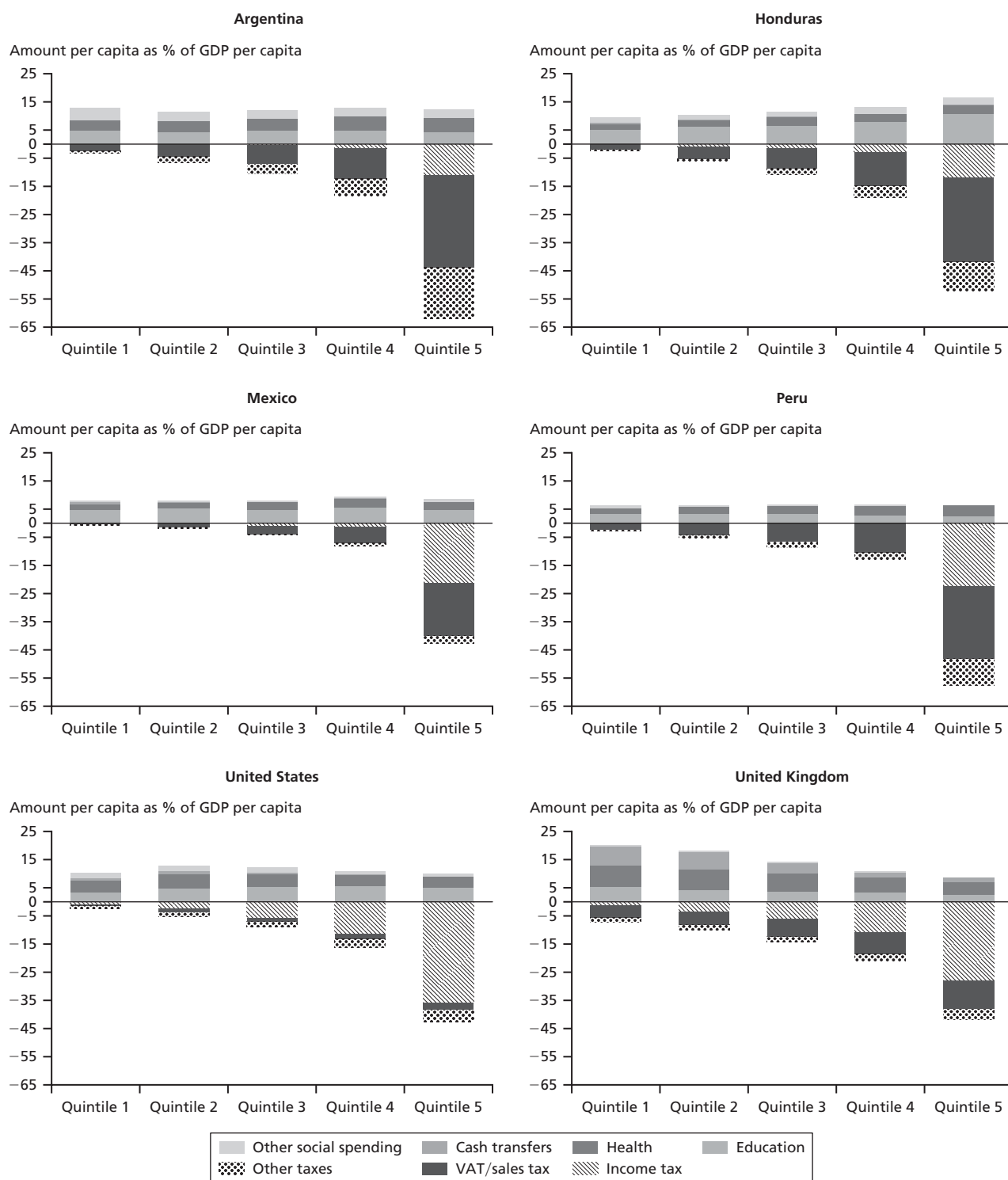
Other studies that look only at taxes find small progressivity and, hence, a small redistributive effect of taxes in Latin America (Chu, Davoodi, and Gupta 2000). Moreover, Engel, Galetovic, and Raddatz (1999) find that, in Chile, which has the highest tax productivity of the region, taxes are slightly regressive. They argue that the more unequal the pretax distribution—as is the case in Chile and most of the region—the less the redistributive effect of progressive taxation. A lot more is achieved through better taxation, fewer loopholes, and higher levels and quality of spending.

An unfortunate characterization of Latin America is one of low-quality, ineffective provision of public services that reflects an unresolved problem of high inequality of opportunities and is correlated with an extremely high level of inequality of current incomes. This is within a context of a taxes and transfers framework that does not redistribute effectively.

In addition to how inequality interacts with the taxes and transfers structure, there are several channels through which inequality might be linked explicitly to informality, although the intuitive relationship is not well studied yet. Ahmed, Rosser, and Rosser (2004) find that the informal sector share is a significant determinant of income inequality in a sample of 52 countries (as cited in Davis [2007]).

Chong and Gradstein (2007), explore the opposite channel, which is of inequality as a mechanism that generates more informality. They model and test that relationship empirically and find a significant positive relationship. Further, they argue that the effect of high inequality may be exacerbated in the context of low institutional quality. The reason they postulate is that, given market imperfections, when institutional quality is low, protection of property rights in the formal sector is weak and resources largely are up for grabs. “Poor individuals whose endowments are

FIGURE 8.13

Social spending and taxation by income quintiles

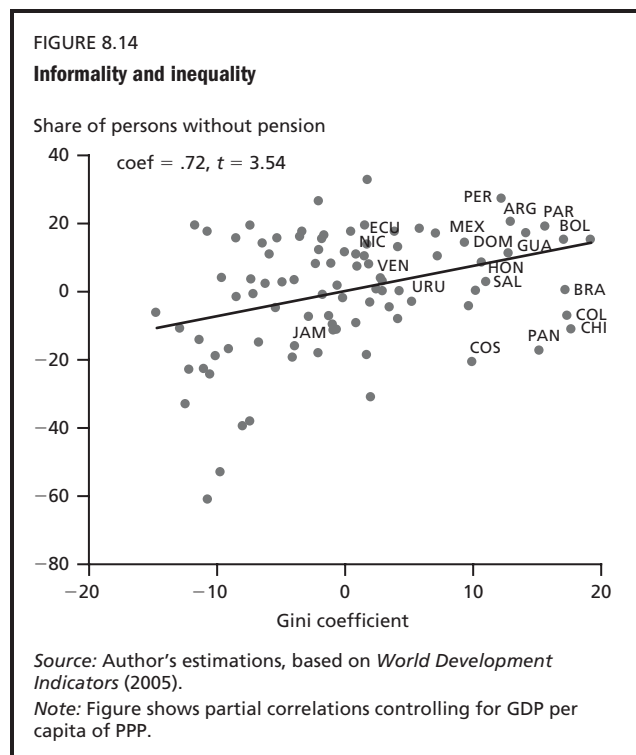
Source: Breceda, Rigolini, and Saavedra 2007.

relatively limited are at a disadvantage in extracting a larger share of the resources, hence, find it beneficial to move into the informal sector, where although less productive, they are able to fully retain their production output. High inequality, exacerbated by low institutional quality, magnifies this effect, implying a positive relationship between inequality and the size of the informal sector” (p. 160).

This result is very relevant in the context of Latin America. Income inequality reflects itself in differences in voice, power, and influence. The lack of influence usually leads to capture and to a perception that the state is run according to the interests of an elite. Among other effects, this leads to lower tax morale and higher informality. Precisely, better institutions might be hampering the possibility that inequality of income may lead to inequality of power.¹⁸ Bird and Zolt (2005) show evidence of a negative effect of inequality on tax effort. They argue that highly unequal distributions of income, typical of Latin America, can lead to low levels of solidarity by the elites toward poorer groups. For example, low tax effort in Central America is often interpreted as the unwillingness of a small elite to finance publicly provided goods because they can opt out and finance their own services.

But an argument that is valid for the elites who might influence policies in their favor is not valid for the relatively rich or the middle class. When there is extreme inequality, it is more difficult to collect revenues in a fair and efficient way. In Latin American countries, despite paying proportionally less than their peers in richer countries, the relatively rich (that is, the small middle class that exists in Latin America) pay a disproportionate share of all revenues. And, given that services provided by the state are of low quality, what they get back from the state is not aligned with what they pay—although they may be getting more service from the state than the poor are getting. This misalignment implies that the opportunity cost of tax compliance is even higher.

We performed several cross-country estimations to test the robustness of the result of a positive relationship between inequality and informality, given the level of development and other institutional characteristics. As shown in figure 8.14, the two variables are positively related after controlling for GDP, and the results are robust to using different informality indicators. We wanted to further analyze whether the effect of inequality is conditioned by institutional characteristics. However, the econometric analysis (not reported) shows that both inequality and GDP are highly correlated with institutional and structural variables that, by them-



selves, can explain a high share of the variation in informality. It is then difficult to disentangle the independent effect of inequality on informality from the effect of institutional variables. The evidence presented by Chong and Gradstein (2007) is suggestive, but more analysis is needed to understand better how the institutional setting may affect the channel through which inequality affects informality.

Informality: A reflection of a broken social contract?

Is Latin America in a “bad” equilibrium?

Economies can land in different equilibria if social norms and social interactions that lead to specific collective behaviors are strong.¹⁹ You make the queue or skip the line. You stop at the red light or you keep moving. As discussed above, tax and regulatory compliance, one of the consummate collective action problems of public policy and a key factor behind informality, may be affected by social interactions. Individuals will be more inclined to pay taxes if they believe the government is complying with its fair share of the social contract and is using public resources effectively. They will be more willing to comply if they believe others do so, prompting still others to comply, and so forth and so on until a highly cooperative state of affairs takes root. Conversely, people will be inclined to evade tax obligations

if they believe others are inclined to do so. Such interdependencies tend to generate patterns of collective behavior with specific reinforcing mechanisms.

Latin America seems to be in an equilibrium in which trust in the state is low; tax collection and compliance with regulations are low in most countries; and public provision of public services is inequitable and of low quality, both directly because of lack of resources and because of low government effectiveness, even given the available resources. This unfortunate characterization reflects an unresolved problem of high inequality of opportunities correlated with an extremely high level of inequality of current incomes. This happens in the context of a structure of taxes and transfers, that does not redistribute effectively.

Low taxes are mostly related to low tax productivity. And, as discussed above, this is mainly related to low tax compliance, narrow tax bases, and excessive exemptions. The empirical analysis presented above suggests that tax compliance is affected by deterrence and by a complex of factual beliefs and emotional dispositions, whereby social norms are not necessarily conducive to complying with taxes and other regulations—a fact that has been proxied by low tax morale. The evidence in Latin America also points to a negative relationship between tax morale and informality. These actually are different sides of the same coin. Low willingness to pay taxes and incentives to operate informally are shown to be empirically correlated with perceptions of incompetence of the state (proxied by indicators of government effectiveness, rule of law, impartiality and fairness of courts, for example). Evidence also suggests the importance of fairness in the use of public resources. In this context, informality is fueled by a structure of regulations that, in a cost–benefit analysis, may provide a rationale for firms and workers to operate underground, as discussed in chapter 5, fostered by a social norm not conducive to complying with regulations and by weak social and administrative sanctions. In this context, there are multiple examples of ways in which both the poor and the rich might feel excluded and find it convenient to opt out, usually partially, from the system.

Opting out . . . but not completely

Telling the story in terms of the beliefs and actions of agents at different levels of the socioeconomic scale, we might say that, in many Latin American countries, the rich, the middle class, and the poor might feel they are not getting a

fair deal—through different mechanisms—by the current arrangements, and that they are justified in avoiding making a contribution to the system whenever possible.

From both ends of the socioeconomic scale we tend to observe a large amount of exit, in Hirschman's terminology. We can use a cost–benefit decision scheme, like the one we presented in the second section of this chapter, now from the point of view of a citizen from the highest income quintile who is considering his or her relationship with the state and with society at large when it comes to deciding whether to evade taxes. In many Latin American countries, such a “rich” citizen is likely to opt out of public services and into the higher-quality private provision of old-age insurance,²⁰ security services, education, and health care. A citizen is likely to feel that such state services are of very little value to him or her. When assessing how valuable such state services are for the population at large, he or she is likely to share the perception that these things that are of little worth to him or her are not useful, effective, of high quality, or fair for citizens in other strata. This negative view of the worth of state services is likely to be compounded by a generalized perception of patronage and corruption in the government generally and in social assistance particularly; and by the social norm externalities that provide an implicit validation of widespread tax evasion.²¹ In some Latin American societies, such as Argentina, we see an increase in socioeconomic segregation, which might bring further difficulties for weaving together a social contract down the road. Poor and rich citizens tend to live in worlds further apart in terms of the schools their children go to, where health services are received, and where their homes are located, with gated communities and shanty towns as clusters, each time placed farther apart.

However, many of the rich, usually the richest of the rich, may also be capturing the state and using its relatively larger influence to maintain tax privileges or lobby to maintain oligopolistic structures. Corporatist groups and certain unions may use political pressure to maintain privileges that perpetuate unequal structures of power and of opportunities. In any of these cases, we are confronting weak and incompetent states—not necessarily small states—on one side, incapable of constraining the influence of powerful groups (Guerrero, López-Calva, and Walton 2006) and, on the other side and related to this, incapable of providing services and public goods in a fair manner.

Poor people feel even more disengaged. The Hirschmanian notion of exit applies also to the lower end of the socioeconomic distribution: the poor do not pay many taxes, but they also do not get much from the state (see box 8.5). They feel an adverse differential access to public goods, to property rights, to protection under the law, and to judiciary services. This fosters the use of informal mechanisms and reduces the incentive to participate in the formal circuit. The poor will organize themselves for self-protection as well, will invade public property, and, in some unfortunate cases, will take justice into their own hands. Poor citizens are more likely than their wealthier neighbors to be informal, to participate in clientelistic networks, and to have a negative view of the state and the extant social arrangement. A difference with respect to the wealthy is that, in many cases, the poor have never been part of the formal system anyway. More than “exiting” the system, the poor have never entered the system. This situation is part of a culture of informality in which the state is basically absent and a social contract is basically broken.

It is easy to see that precisely these attitudes have important reinforcement effects. The low willingness to contribute is part of the explanation of the fiscal limitations that impinge on the low quality and coverage of the services provided by the state and that, in turn, feeds back into low trust in the state and low tax morale, as discussed above. There is no clear evidence that the process of opting out is increasing, but there also is no evidence at all to the contrary. At least it is safe to say that some people opt out (exit) from a more inclusive social contract, while others continue to be excluded from it. It happens not only in the “flat” horizontal sense, but also in a “vertical” sense in which whole groups “collectively” exit through what one might call “local” reciprocity dynamics.

The specific experience regarding all the different aspects of exit and exclusion that are behind the large informal sector in Latin America varies notoriously. Below, we describe briefly three particular cases to illustrate this heterogeneity.

Heterogeneity in Latin America

Most Latin American countries attempted at some point in the 20th century to form a social contract that centered on the labor market, with some strong institutional actors, such as unions. That model, linked to the development model in vogue at the time, provided some progressive

incorporation and looked like the construction of a social contract that allowed for the inclusion of increasing segments of the population. But that welfare system reached full coverage, was strongly stratified, and turned out to be financially unsustainable. Many segments of society, including rural citizens and urban marginal areas, never participated.

As a result of an explicit agenda to improve fiscal sustainability and to insert some “market principles,” that system and the political institutions behind it were reformed. The social security system was reorganized along market principles. It is not clear yet that the new system will imply a fast incorporation of segments that traditionally have been left out. Moreover, the evidence shown in this report and elsewhere points to increasing informality in the 1990s.

But different countries seem to be going in different directions, and societal consensus regarding the potential future path varies. In Argentina, for example, there is evidence of increasing inequities in access to social protection and less clarity about the effectiveness of the current arrangements. One possible reading is that now we have a patchwork of elements of the previous covenant, mixed with the recent add-ons, in the form of social assistance programs that were effective during the time of the crises and efficient in some dimensions, but that have not yet added up to anything coherent, let alone an integrated set of social policies that has garnered societal consensus. Some interactions between increasing informality and a transformation of politics seem to be feeding a negative loop. Countries such as Argentina that, several decades ago, seemed to define crucial political decisions in a centralized bargaining arena defining national policies now are moving to an increasingly territorialized model of construction of political power in which focused social assistance programs, informality, clientelism, and new forms of political participation (such as *piquetes*) seem to reinforce each other.

Chile is a very different case. It is the only economy in the region that has been able to attain very low levels of extreme poverty, with a Chile Solidario program that has more beneficiaries in 2003 than what the *Encuesta de Caracterización Socioeconómica Nacional* survey reported as extreme poor. One of the key—and first—elements of the Chilean process that started with the return to democracy in 1990 was its ability to negotiate a tax reform as one of its first steps. It was able to convince the elites (businesses

BOX 8.5

Expansion of private security services in Managua

Hirschman (1970) posits that when discontent with the performance of an organization to which one belongs, a person has two options: “exit” the organization and find another whose performance is superior, or “voice” one’s dissatisfaction in the hope of prompting the organization to improve. In that framework, the phenomenon of wealthier Latin American citizens seeking alternative health care, education, and security providers is seen as their exercise of the exit option. In a flight to quality, they abandon the state and seek private providers. Such abandonment can have far-reaching effects on the state. Citizens may be less inclined to pay taxes for services they don’t use or value, so the state’s capacity to perform its duties will be weakened. The provision of public security to a nation is one of the oldest and most firmly established responsibilities of the state. Centeno and Portes (2006) aptly describe the state of public security in the region at the dawn of the 21st century in these terms: “In Latin America, the incapacity of the state to protect the citizenry has led to the massive growth of private security services, the withdrawal of the wealthy into fortress-like gated communities. . . . In the absence of credible enforcement of rules, people take things into their own hands . . .” (p. 14).

Rodgers (2004) explores issues of security, criminality, and spatial segregation in Managua, Nicaragua. The city

has seen an annual growth in the rate of crime of approximately 10 percent since 1990. Crimes against people represent the greatest increase—362 percent over 13 years. Rodgers reports that “. . . regionally, the Nicaraguan police force has the lowest number of police personnel per capita and per crime, the lowest budget per crime, the lowest budget per police officer, and the lowest average salaries in Central America. As crime has exploded in the capital so has the private security market” (p. 117). He states that, in 1990, there was one registered private security firm in the capital. This number climbed to 14 in 1996 and was 56 in 2003. In that same year, 9,017 people were registered as private security guards. The increasing trend has not abated, despite the government’s efforts to increase resources for the police force and a 33 percent increase in the number of officers between 2000 and 2005 (Gómez 2005). Recent statistics show that, in 2005, there were 67 private security companies covering 4,153 locations, with 9,329 guards and 6,805 weapons. Gómez delves further into the phenomenon of private security in Nicaragua to reveal a crossover between public and private security provision. He observes that some of the highest-ranking former police officers have become active in providing private security.

The message seems to be that when the state doesn’t deliver, people take things into their own hands.

and high-income earners who would pay about two-thirds of the new tax burden) that, in the midst of uncertainty, it was a small price to pay for the return to socioeconomic peace. The combination of tax reform and explicit social policy objectives made this a step toward building a social contract. The reform was approved by Congress six weeks after the new *Concertación* government started (box 8.6). Chile reformed its pension system and has recognized that the increase in coverage is still too small. It is therefore proposing a package of reforms to the system aimed at achieving universal coverage in a way that would be incentive compatible with keeping individual savings as the mainstay of the system and that is fiscally sustainable (see chapter 7). Trust in institutions is high and tax productivity is the highest in the region. Informality is less than 15 percent. Inequality of income in Chile, however, is still

extremely high and, symptomatically, very high in the policy agenda.

The opposite of Chile might be Guatemala, where trust in institutions is low and the presence of the state in several socioeconomic realms is very limited. Tax collections are among the lowest in the region, and informality is around 80 percent. The peace accord signed in 1996, after 30 years of civil war, included a political pact for the reconstruction of the country; and it established the need for increasing tax collection by at least 50 percent before 2000 (that is, to 12 percent of GDP). The increase in tax collection was earmarked for education, health care, housing, and justice. In 1998, the government rescheduled the agreement. In 1999, a large commission prepared the draft of the new fiscal pact. During 2000, the negotiations among all political parties, the business elites, and other organizations failed,

BOX 8.6

Negotiating tax reform and the start of the social contract, Chile, 1990

In a difficult transition environment, the center-left *Concertación* coalition, which won Chile's first democratic election in 17 years, successfully passed a tax reform bill within the first months of the new party's term, earmarking greater tax revenue for increased social expenditures. Boylan (1996) identifies two main reasons for this unexpected policy success: a willingness to be moderate and an engaged negotiation strategy. The final version of the tax package consisted of four main components. It included an increase in the corporate tax rate from 10 to 15 percent and a change in the progressive income tax categorization system that put more people in the highest bracket. Third, to control tax evasion, the highest contributors in the agriculture, transportation, and mining sectors were no longer taxed on estimated profits but on actual profits. Finally, the VAT increased from 16 to 18 percent.

Boylan (1996) argues that it was crucial that the details of the tax reform were negotiated in extraparlimentary meetings with a targeted group from the major opposition party. The targeted politicians of *Renovación Nacional* (RN), a party whose agrarian base strongly identified with the outgoing regime, were known for their technical competence, consensual approach to politics, and their desire to transform the right into a democratic force. *Concertación* was thus able to develop the tax reform with a small, less hostile faction of opposition, who then sold the reform to the rank and file of their party. This way, by the time the reform was introduced into parliament, it was almost a fait accompli.

In addition to this negotiation strategy, led by Alejandro Foxley (then minister of finance), policy success resulted from *Concertación*'s willingness to be moderate, as evidenced in the final substance of the reform. Initially *Concertación* wanted to raise the percentage of gross national product tax revenue to 3 percent rather than 2 percent. The RN negotiators pushed for the reduction of the corporate tax increase from 20 to 15 percent. From the RN's perspective, the business elite's foresight that the "social debt" of the years of dictatorship had somehow to be paid and its realization that, compared with other taxes in the region, the proposed taxes were not high played a role in ensuring the successful negotiations and passage of the reform. Finally, *Concertación* accepted the RN's proposed increase in the VAT, thus showing a willingness "to tax their own," not only the business elite. In terms of selling the reform to the general public, the earmarking of the additional revenue to explicit social policy gave a sense that the reform was a step toward building a social contract. Furthermore, Boylan (1996) argues that Chilean tax reform was more than a targeted policy success: the moderate and risk-averse strategy followed by the new government "played a crucial role in soldering the fragile rule-making environment at the delicate moment of regime change" (p. 8).

Source: Based mainly on Boylan 1996.

among other reasons, because of the lack of leadership in the executive (ICEF 2005). Some changes implemented in 2001 were harshly resisted by the business elites, which in that year alone presented 31 constitutional appeals against tax increases. The courts lined up with business interests, and the tax take is still around 11 percent of GDP. No need to continue the narrative. Point made.

Mexico is an intermediate case. Filgueira (2005) classifies it among those dual social states with a large social protection system that covers a bit less than half of the population and with rigid labor legislation that hampers formal job creation. Mexico reformed its private pension system, which has made it financially healthier but has

failed to increase coverage. Minimum pension programs have been launched in several states, and the flagship antipoverty program, *Oportunidades*, covers 80 percent of the poor and is considered an international best practice. But further expansion of minimum pension programs or any other initiative aimed at universalizing social insurance clashes with a hard budget constraint. The Fox administration was never able to pass a tax reform, and Mexico is the OECD country with the lowest tax collection—well below what would be expected for its level of development. Trust in institutions is weak, and the social norm regarding tax compliance is probably behind the low VAT productivity of 24 percent. Organized public sector labor is politically

very strong, and neither labor reform nor public pension sector reform has been possible politically. Still, the country is looking to improve the quality of its public performance, and tax reform is again on the agenda.

Conclusions

We have discussed here that to understand informality it is critical to understand several interaction mechanisms between the state and the citizens. Individuals explicitly or implicitly decide whether and how to engage with the mandates and institutions of the state, weighing costs and benefits and state enforcement capabilities. These decisions are conditioned by social norms shaped by how agents individually and collectively perceive and define a relationship with the state. As documented here, cross-country evidence based on opinion surveys and investment climate surveys, as well as on a few comparative studies, suggests that perceptions of government effectiveness and of the performance of services like the judiciary system in Latin America are below those observed in other regions. And, as summarized here and discussed extensively in the literature, informality is negatively related to institutional quality indicators. This environment is consistent with a social norm that is not conducive to complying with regulations.

Taxation, which lies at the heart of a social contract, is one of the areas that has been most studied regarding the role of social norms. As discussed here, tax morale—a social norm about a citizen's willingness to pay taxes—seems to be correlated with several measures of state performance and with informality, both in Latin America and globally. Low willingness to pay taxes and incentives to operate informally are related to perceptions—supported by reality—of incompetence of the state and lack of fairness in the use of public resources. Moreover, the high and persistent levels of inequality, and the prevalence of a structure of taxes and transfers that is not efficient in leveling the playing field and improving the equality of opportunities, fuel a perception of state ineffectiveness. Many among both the rich and the poor, through different mechanisms, may find it convenient to opt out, which leads to higher informality.

Tax collection is low, given the region's level of development, which reveals that higher levels of taxation might be needed to move toward the path of development. But this begs the question, How and why should taxes be raised when much of the region perceives that the performance of the state is bleak? On one hand, even if the objective of raising taxes is taken at face value, the main challenges are to

expand the tax base, incorporate more citizens in the formal economy, and increase tax compliance. The rich in Latin America contribute a much larger share of the tax collection than what is observed in richer regions, and further increases in tax rates might not be socially tolerable or economically efficient. Furthermore, in attempting to increase tax collection, the region faces the challenge of reducing exemptions, which often is a reflection of state capture. On the other hand, Latin American governments would have to improve their performance—both the quality of their expenditures and the mechanism for citizens to monitor them. Recent mainstream literature is less pessimistic about the effect of taxes (and therefore the size of government) on growth, suggesting that Latin America may be on the side of the curve where taxes, as a whole, may be growth enhancing through the public-goods channel. This implies that increases in both taxes and expenditures in Latin America might be growth enhancing, but only if government effectiveness increases dramatically. This will facilitate establishing the conditions for a gradual change in individual and collective beliefs and attitudes regarding the real and perceived relationship between citizens and the state, which is essential to start reducing the high levels of informality in the region. The strong interest in initiatives to improve government effectiveness suggests that many countries are steering in the right direction.

Reducing informality is a daunting task and a critical development challenge for the region. It requires not only increasing overall productivity and growth in the economy and improving regulations in labor and product markets, but also pursuing a long-term agenda that could move countries faster to a better equilibrium. In other words, it is an agenda that includes building a better social contract from which fewer people are excluded and in which there are fewer incentives to opt out from it. As has been described throughout the chapters of this report and in the ever-expanding body of literature on informality, the policy agenda in the areas of labor, credit, business services, cost of registration, taxes, business regulation, property rights, and access to judiciary service, among others, is critical to making progress in fostering access to the formal economy. Policies in these areas, if designed in a consistent and integrated fashion, can help reduce informality as part of the concomitant process of increasing productivity and incomes. But those policies should be part of a road map so that partial steps might be taken over time, using political windows of opportunity for reform.

Notes

1. The following two sections draw heavily on Saavedra and Tommasi (2007).

2. This discussion also relates to the fact that the willingness to contribute to finance welfare support is heavily influenced by perceptions about how deserving the recipients are, by how “close” the person feels to the recipients (ethnically, culturally, and so forth), and by perceptions about the adequacy of state services and the like (Fong, Bowles, and Gintis 2005; Lindert 2004).

3. This is an example of opting out, which is discussed in more detail later in the chapter.

4. The investment climate surveys used for these calculations contain 11 countries from Latin America (Bolivia, Brazil, Chile, Costa Rica, Ecuador, El Salvador, Guatemala, Guyana, Honduras, Nicaragua, and Peru). The sample contains 42 percent of small firms (20 employees or fewer), 38 percent of medium-size firms (21–100 employees), and 20 percent of large firms (more than 100 employees).

5. Data from the World Bank’s Worldwide Governance Indicators (WGI) database are used extensively in this chapter. See Kaufmann, Kraay, and Mastruzzi (2006). The Government Effectiveness Index is a measure of the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government’s commitment to policies.

6. The Control of Corruption Index from WGI is a measure of perceptions, where corruption is defined as the exercise of public power for private gain, with higher values corresponding to less corruption.

7. They argue that the liberal reforms implemented in the mid-1980s were an attempt to reduce the regulatory intent and, at the same time, increase regulatory capacity.

8. It should be noted that this last variable can be interpreted plainly as the size of the state, which may or may not be correlated with enforcement capacity.

9. Share of informal economy (informality definition 1 throughout the text) is the Schneider measure of the average size of the informal economy. The share of self-employed (informality definition 2) is the ILO measure of self-employed workers in the workforce. The share of person without access to pension is computed based on WDI and Centro de Estudios Distributivos Laborales y Sociales for 14 Latin American countries (informality definition 3). These definitions are similar to those discussed in chapter 1.

10. This is probably capturing Latin American higher levels of inequality because the coefficient is not robust to the inclusion of that variable.

11. This increase is after several years of declining tax revenues caused by low growth, high inflation, and ineffective collection efforts.

12. Tax morale information is collected in the World Values Survey and the *Latinobarometro*, which are representative surveys collected in 80 (worldwide) and 17 (Latin American) countries. At the national level, these surveys have a sample size of at least 1,000 observations per country. Both the values survey and *Latinobarometro* ask the question, “Could you tell me in a scale from 1 to 10 (where 1 is never justified and 10 is totally justified) whether you think it is justified to

cheat on taxes if you have the chance?” Tax morale variables used are expressed as the percent responding that it is never justified.

13. Methodological difficulties in the ability to identify and calculate precisely social interaction parameters, in general, have been discussed in detail by Manski (2000). Glaeser, Sacerdote, and Scheinkman (2003) review the scarce recent empirical work that tries to identify social multiplier effects. Their results suggest that social interactions may be large. They posit that, if one person’s proclivity to certain behavior influences his or her neighbor’s behavior, policy changes to address that behavior will have a direct effect and an indirect one through social influence. “The presence of positive spillovers or strategic complementarities creates a social multiplier where aggregate coefficients will be greater than individual coefficients” (p. 2). Empirically, the estimated ratio of aggregate coefficients (estimated, say, at the city level) to individual coefficients is the social multiplier. The writers apply this approach to three settings. They find that, among Dartmouth University students, one predetermined variable had a bigger impact on joining a fraternity at higher levels of aggregation. Using crime data, they find a very large social multiplier; and, using data on wages and human capital, they find a social multiplier at higher levels of aggregation. Another good overview is provided in Gintis et al. (2005).

14. In game theory parlance, the strategies of players A and B are said to be *strategic complements* if when player A increases a component of his or her strategy, player B will want to do so also. In other words, the cross derivative of the payoff function of B with respect to the action of A and his or her own action is positive. Strategic complementarities, if strong enough, lead to multiplicity of equilibria.

15. The prospect of shame or potential stigma and guilt has a similar effect. The more likely an individual believes it is that he or she will be condemned by others if caught, the more likely he or she is to refrain from evading (Grasmick and Scott 1982; Kahan 2005).

16. The authors make these calculations with and without pensions. These figures exclude pensions because of the problems that exist in calculating how much people who receive the transfer had paid during their working lives. In any case, the inclusion of pensions—even if precisely calculated—will imply that social spending would be even more biased toward the rich.

17. For further discussion, see Goñi, López, and Servén (2006).

18. High quality of institutions may explain the fact that, in Chile, high income inequality is not translated into lower trust in the state, low tax morale, and, consequently, high informality.

19. This section draws heavily on Saavedra and Tommasi (2007).

20. In the specific case of pensions, privatization from pay-as-you-go systems has not always implied an increase in trust in the system and uncertainty about future benefits, and political—and, in many countries, economic—risk has not diminished (Kay 2003; Spiller and Tommasi forthcoming); so “private options” may imply alternatives, such as old-age savings accounts in foreign financial institutions, other assets, or reliance on family networks for the provision of protection in old age.

21. In many cases, like Argentina—and unlike Chile—(Bergman 2002, 2003), there is a negative peer-pressure effect in which tax evasion has wide social acceptance, particularly among the better-off populations.

References

- Ahmed, E., J. B. Rosser, and M. Rosser. 2004. "Income Inequality, Corruption, and the Non-Observed Economy: A Global Perspective." Photocopy. Available at <http://cob.jmu.edu/rosserjb/NOE.Inequality.Global.doc>.
- Alesina, A., and E. Glaeser. 2004. *Fighting Poverty in the US and Europe: A World of Difference*. Oxford, UK: Oxford University Press.
- Alm, J., and J. Martinez-Vazquez. 2007. "Tax Morale and Tax Evasion in Latin America." Photocopy. Georgia State University, Atlanta.
- Alm, J., I. Sánchez, and A. De Juan. 1995. "Economic and Noneconomic Factors in Tax Compliance." *Kyklos* 48 (3): 3–18.
- Alm, J., and B. Torgler. 2006. "Culture Difference and Tax Morale in Europe and the United States." *Journal of Economic Psychology* 27 (2): 224–46.
- Andreoni, J., B. Erard, and J. Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature* 36: 818–41.
- Beramendi, P., and A. Díaz-Cayeros. 2006. "Distributive Tensions in Developing Federations." Paper presented at the 102nd Meeting of the American Political Science Association, Philadelphia, August 31–September 3.
- Bergman, M. 2002. "Who Pays for Social Policy? A Study on Taxes and Trust." *Journal of Social Policy* 31 (2): 289–305.
- . 2003. "Tax Reform and Tax Compliance: The Divergent Paths of Chile and Argentina." *Journal of Latin American Studies* 35: 593–624.
- Biebesheimer, C., and M. Payne. 2001. "IDB Experience in Justice Reform." SDD Technical Paper 101, Inter-American Development Bank, Washington, DC.
- Bird, R., J. Martinez-Vazquez, and B. Torgler. 2006. "Societal Institutions and Tax Effort in Developing Countries." In *The Challenges of Tax Reform in the Global Economy*, ed. J. Alm, J. Martinez-Vazquez, and M. Rider, 283–338. New York: Springer.
- Bird, R., and E. Zolt. 2005. "Redistribution via Taxation: The Limited Role of the Personal Income Tax in Developing Countries." Working Paper 0508, International Tax Program, Institute for International Business, Joseph L. Rotman School of Management, University of Toronto.
- Birdsall, N., and R. Menezes. 2005. "Beyond the Washington Consensus: A New Job-based Social Contract in Latin America." *Foreign Affairs en Español* 5 (3).
- Boylan, D. 1996. "Taxation and Transition: The Politics of the 1990 Chilean Tax Reform." *Latin American Research Review* 31 (1): 7–31.
- Breceda, K., J. Rigolini, and J. Saavedra. 2007. "Taxes and Transfer in Latin America." Photocopy. World Bank, Washington, DC.
- Brewer, M., A. Duncan, A. Shepard, and M. Suárez. 2005. "Did Working Families' Tax Credit Work? The Final Evaluation of the Impact of In-work Support on Parents' Labor Supply and Take-up Behavior in the UK." Photocopy. Institute for Fiscal Studies, London.
- Centeno, M. A., and A. Portes. 2006. "The Informal Economy in the Shadow of the State." In *Out of the Shadows: Political Action and the Informal Economy in Latin America*, ed. P. Fernández-Kelly and J. Shefner. State College, PA: Pennsylvania State University Press.
- Chong, A., and M. Gradstein. 2007. "Inequality and Informality." *Journal of Public Economics* 9: 159–79.
- Chu, K., H. Davoodi, and S. Gupta. 2000. "Income Distribution and Tax, and Government Social Spending Policies in Developing Countries." Working Paper 214, World Institute for Development Economics Research, United Nations University, Helsinki.
- Cowell, F. 1990. *Cheating the Government: The Economics of Evasion*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Davis, Lewis S. 2007. "Explaining the Evidence on Inequality and Growth: Informality and Redistribution." *The B. E. Journal of Macroeconomics* 7 (1) article 7. Available at: <http://www.bepress.com/bejm/vol7/iss1/art7>.
- de Ferranti, D., G. E. Perry, F. H. G. Ferreira, and M. Walton. 2004. *Inequality in Latin America and the Caribbean: Breaking with History?* Washington, DC: World Bank.
- Detragiache, E. 2001. "The Labor Income Tax Credit in an International Perspective." Kingdom of the Netherlands Country Report 01/96, International Monetary Fund, Washington, DC.
- Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1): 1–37.
- Duryea, S., A. Olgiati, and L. Stone. 2006. "The Under-Registration of Births in Latin America." Working Paper 551, Inter-American Development Bank, Washington, DC.
- ECLAC (Economic Commission for Latin America and the Caribbean). 2006. *Shaping the Future of Social Protection: Access, Financing, and Solidarity*. New York: United Nations.
- Engel, E., A. Galetovic, and C. Raddatz. 1999. "Taxes and Income Distribution in Chile: Some Unpleasant Redistributive Arithmetic." *Journal of Development Economics* 59: 155–92.
- Enste, D. H., and F. Schneider. 2000. "Shadow Economics: Size, Causes, and Consequences." *Journal of Economic Literature* 38: 77–114.
- Eyzaguirre, H. 1996. "Institutions and Economic Development: Judicial Reform in Latin America." DPP Strategy and Policy Working Paper 103, Inter-American Development Bank, Washington, DC.
- Falk, A., and U. Fischbacher. 2005. "Modelling Strong Reciprocity." In *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, ed. H. Gintis, S. Bowles, R. T. Boyd, and E. Feher, ch. 6. Cambridge, MA: Massachusetts Institute of Technology Press.
- Farrell, D. 2006. "Tracking the Informal Economy." *Business Week*, May 8; available at <http://www.mckinsey.com/aboutus/mckinseynews/tacklingeconomy.asp>.
- Feld, L., and B. Frey. 2007. "Tax Compliance as the Result of a Psychological Tax Contract: The Role of Incentives and Responsive Regulation." *Law and Policy* 29 (1): 102–20.
- Filgueira, F. 2005. "Welfare and Democracy in Latin America: The Development, Crises and Aftermath of Universal, Dual, and Exclusionary Social States." Prepared for the United Nations Research Institute for Social Development Project on Social Policy and Democratization, New York.
- Fong, C., S. Bowles, and H. Gintis. 2005. "Reciprocity and the Welfare State." In *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, ed. H. Gintis, S. Bowles, R. T. Boyd, and E. Feher, ch. 10. Cambridge, MA: Massachusetts Institute of Technology Press.
- Friedman, E., S. Johnson, D. Kaufmann, and P. Zoido-Lobaton. 2000. "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics* 76: 459–93.

- Gintis, H., S. Bowles, R. T. Boyd, and E. Feher, eds. 2005. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman. 2003. "The Social Multiplier." *Journal of the European Economic Association* 1 (2–3): 345–53.
- Gómez, J. L. R. 2005. "The Political Economy of Nicaragua's Institutional and Organizational Framework for Dealing with Violence." Photocopy. Crisis States Research Center, Universidad Centroamericana, Managua, Nicaragua.
- Gomez-Sabaini, J. 2005. "Evolucion y situacion Tributaria Actual en America Latina: Una Serie de Temas para la Discusion." Presentation to CEPAL, Santiago de Chile, Chile, September.
- Goñi, E., J. H. López, and L. Servén. 2006. "Fiscal Reform for Social Equity in Latin America." Photocopy. World Bank, Washington, DC.
- Grasmick, H., and W. Scott. 1982. "Tax Evasion and Mechanisms of Social Control: A Comparison with Grand and Petty Theft." *Journal of Economic Psychology* 2 (3): 213–30.
- Greenstein, R. 2005. "The Earned Income Tax Credit: Boosting Employment, Aiding the Working Poor." Photocopy. Center on Budget and Policy Priorities, Washington, DC. Available at <http://www.cbpp.org/7-19-05eic.pdf>.
- Guerrero, I., L. F. López-Calva, and M. Walton. 2006. "The Inequality Trap and Its Links to Low Growth in Mexico." Working Paper 298, Stanford Center for International Development, Stanford, CA.
- Hirschman, A. O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations and States*. Cambridge, MA: Harvard University Press.
- Hoffman, S. D., and L. S. Seidman. 2003. *Helping Working Families: The Earned Income Tax Credit*. Kalamazoo, MI: Upjohn Institute.
- Holt, S. 2006. *The Earned Income Tax Credit at Age 30: What We Know*. Brookings Institution Research Brief. Washington: Brookings Institution Press.
- Holtz, J., and J. K. Scholz. 2003. "The Earned Income Tax Credit." In *Means-Tested Transfer Programs in the United States*, ed. R. A. Moffitt. Chicago: University of Chicago Press.
- ICEF (Instituto Centroamericano de Estudios Fiscales). 2005. "Diez Claves para un Pacto Fiscal." *Boletín de Estudios Fiscales* 1.
- IDB (Inter-American Development Bank). 2005. *The Politics of Policies: Economic and Social Progress in Latin America*. Washington, DC.
- . 2006. "Hidden in Plain Sight: Undocumented Citizens in Latin America." *Ideas for Development in the Americas* 9; available at <http://www.iadb.org/news/articledetail.cfm?Language=EN&artid=2697&artType=WS>.
- Johnson, S., D. Kaufmann, and A. Shleifer. 1997. "The Unofficial Economy in Transition." *Brookings Papers on Economic Activity* 2: 159–239.
- Johnson, S., D. Kaufmann, and P. Zoido-Lobaton. 1998. "Regulatory Discretion and the Informal Economy." *American Economic Review* 88 (2): 387–92.
- Kahan, D. 2005. "The Logic of Reciprocity: Trust, Collective Action, and Law." In *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, ed. H. Gintis, S. Bowles, R. T. Boyd, and E. Feher, ch. 12. Cambridge, MA: Massachusetts Institute of Technology Press.
- Kaufmann, D., A. Kraay, and M. Mastruzzi. 2006. "Governance Matters V: Governance Indicators for 1996–2005." Available at <http://ssrn.com/abstract=929549>.
- Kay, S. J. 2003. "Pension Reform and Political Risk." Prepared for the LASA XXIV International Congress, Dallas, March 27–29.
- Kuznets, S. 1955. "Toward a Theory of Economic Growth." In *National Policy for Economic Welfare at Home and Abroad*, ed. R. Lekachman, Garden City, NY: Doubleday.
- Levi, M. 1998. "A State of Trust." In *Trust and Governance*, ed. V. Braithwaite and M. Levi, 77–101. New York: Russell Sage Foundation.
- Lindert, K., E. Skoufias, and J. Shapiro. 2006. "Redistributing Income to the Poor and the Rich: Public Transfers in Latin America and the Caribbean." Social Protection Discussion Paper 0605, World Bank, Washington, DC.
- Lindert, P. 2004. *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century*. Cambridge, UK: Cambridge University Press.
- Lledo, V., A. Schneider, and M. Moore. 2004. "Governance, Taxes and Tax Reform in Latin America." IDS Working Paper 221, Institute of Development Studies, Brighton, UK.
- Loayza, N. 1996. "The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129–62.
- Loayza, N., A. M. Oviedo, and L. Servén. 2005. "The Impact of Regulation on Growth and Informality—Cross-country Evidence." Policy Research Working Paper 3263, World Bank, Washington, DC.
- Loayza, N., and J. Rigolini. 2006. "Informality Trends and Cycles." Policy Research Working Paper 4078, World Bank, Washington, DC.
- Lora, E. 2006a. "El Futuro de los Pactos Fiscales en América Latina." Paper prepared for the forum Políticas Económicas para un Nuevo Pacto Social en América Latina, Fundación CIDOB, Barcelona, October.
- . 2006b. "Trends and Outcomes of Tax Reform." In *The State of State Reforms in Latin America*, ed. E. Lora. Washington, DC: World Bank.
- Lora, E., M. Cardenas, and V. Mercer-Blackman. 2005. "The Policy Making Process of Tax Reform in Latin America." Prepared for the Inter-American Development Bank Workshop on State Reform, Public Policies, and Policymaking Processes, Washington, DC.
- Manski, C. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14 (3): 115–36.
- Martinez-Vazquez, J., and B. Torgler. 2005. "The Evolution of Tax Morale in Modern Spain." Working Paper 2005–33, Center for Research in Economic Management and the Arts, Basel, Switzerland.
- Mendez, J. E., G. O'Donnell, and P. S. Pinheiro, eds. 1999. *The (Un) Rule of Law and the Underprivileged in Latin America*. North Bend, IN: University of Notre Dame Press.
- Meyer, B. 2001. "Taxes, Welfare, and Work by Single Mothers." *NBER Reporter Online*, available at www.nber.org/reporter/fall01/meyer.html (accessed February 2007).

- . 2002. "Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare, and Hours Worked." *American Economic Review* 92 (2): 373–79.
- Perry, G., O. Arias, H. López, W. F. Maloney, and L. Servén. 2006. "Poverty Reduction and Growth: Virtuous and Vicious Circles." World Bank, Washington, DC.
- Rodgers, D. 2004. "Disembedding the City: Crime, Insecurity, and Spatial Organisation in Managua, Nicaragua." *Environment and Urbanization* 16 (2): 113–24.
- Ronconi, L. 2006. "Notes of the Judiciary and Informality." Photocopy.
- Saavedra, J., and M. Tommasi. 2007. "Informality, the State and the Social Contract." Photocopy. Universidad de San Andrés, Buenos Aires, Argentina.
- Schelling, R. 1978. *Micromotives and Macrobehavior*. New York: Norton.
- Schmitt, J. 2005. "Is It Time to Export the U.S. Tax Model to Latin America?" *Challenge* 48 (3): 84–108.
- Scholz, J. T., and M. Lubell. 1998. "Trust and Taxpayers: Testing the Heuristic Approach to Collective Action." *American Journal of Political Science* 42: 398–417.
- Sheffrin, S., and R. Triest. 1992. "Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance." In *Why People Pay Taxes: Tax Compliance and Enforcement*, ed. J. Slemrod. Ann Arbor, MI: University of Michigan Press.
- Sokoloff, K., and E. Zolt. 2005. "Inequality and the Evolution of Institutions of Taxation: Evidence from the Economic History of the Americas." Paper presented at NBER (National Bureau of Economic Research) Political Economy workshop. Available at http://www.yale.net/polisci/info/conferences/Political_Economy_Post-Industrial_Societies/Sokoloff.pdf.
- Spicer, M. W., and L. A. Becker. 1980. "Fiscal Inequity and Tax Evasion: An Experimental Approach." *National Tax Journal* 33: 171–75.
- Spiller, P., and M. Tommasi. Forthcoming. "Public Policy in Argentina." In *The Institutional Foundations of Public Policy in Argentina*, ed. P. Spiller and M. Tommasi. Cambridge, UK: Cambridge University Press.
- Tanzi, V. 2000. "Taxation in Latin America in the Last Decade." Working Paper 76, Center for Research on Economic Development and Policy Reform, Stanford University, Stanford, CA.
- Tanzi V., and H. Zee. 2000. "Tax Policy for Emerging Markets: Developing Countries." Working Paper WF/00/35, International Monetary Fund, Washington, DC.
- Torgler, B. 2005. "Tax Morale in Latin America." *Public Choice* 122: 133–57.
- World Bank. 2005. World Development Indicators. Available at: www.worldbank.org.
- . 2006. *World Development Report 2006: Equity and Development*. Washington, DC.
- . 2007. "Job Creation in Latin America and the Caribbean: Recent Trends and the Policy Challenges." Photocopy. World Bank, Washington, DC.
- Young, H. P. 2006. "Social Norms." Paper prepared for the *New Palgrave Dictionary of Economics*. Basingstroke, UK: Palgrave Macmillan.

ECO-AUDIT

Environmental Benefits Statement

The World Bank is committed to preserving endangered forests and natural resources. The Office of the Publisher has chosen to print *Informality: Exit and Exclusion* on recycled paper with 30 percent post-consumer waste, in accordance with the recommended standards for paper usage set by the Green Press Initiative, a nonprofit program supporting publishers in using fiber that is not sourced from endangered forests. For more information, visit www.greenpressinitiative.org.

Saved:

- 42 trees
- 3,508 lbs. of solid waste
- 29,117 gallons of wastewater
- 8,493 lbs. of net greenhouse gases
- 58 million BTUs of energy



Informality is a manifestation of a flawed relationship between citizens and their state. While the term means different things to different people, it almost always has negative connotations—including inferior working conditions, low productivity businesses, unfair competition, and disrespect for the rule of law. It is not surprising then that the rise observed in informality in Latin American and Caribbean countries over recent decades has garnered attention as a potential brake on improvements in growth and social welfare, and as a force corrosive to the integrity of society.

Informality: Exit and Exclusion analyzes informality in Latin America, exploring root causes and reasons for and implications of its growth. The authors use two distinct but complementary lenses: informality driven by “exclusion” from state benefits or the circuits of the modern economy, and driven by voluntary “exit” decisions resulting from private cost-benefit calculations that lead workers and firms to opt out of formal institutions. They find both lenses have considerable explanatory power to understand the causes and consequences of informality in the region. *Informality: Exit and Exclusion* concludes that reducing informality levels and overcoming the “culture of informality” will require actions to increase aggregate productivity in the economy, reform poorly designed regulations and social policies, and increase the legitimacy of the state by improving the quality and fairness of state institutions and policies. Although the study focuses on Latin America, its analysis, approach, and conclusions are relevant for all developing countries.

Informality: Exit and Exclusion will be of value to professionals and academics studying labor market, social protection, tax, microenterprise development, and urban public policies, and to those working in government, international organizations, research institutions, and universities.



THE WORLD BANK



ISBN 0-8213-7092-8