

23416

INSIDE

Angus Deaton
on Counting the World's Poor
Comment by Martin Ravallion
Comment by T. N. Srinivasan

•
Yujiro Hayami
on Ecology, History, and
Development

•
Rinku Murgai, Mubarik Ali,
and Derek Byerlee
on Productivity and Sustainability
in the Punjabs

•
Raj M. Desai and
Itzhak Goldberg
on Politics of Russian
Enterprise Reform

THE WORLD BANK

Research Observer

Volume 16 • Number 2 • Fall 2001

OXFORD
UNIVERSITY PRESS

FILE COPY

ISSN 0257-3032

THE WORLD BANK

Research Observer

EDITOR

Shantayanan Devarajan, World Bank

COEDITORS

Gershon Feder, World Bank

Shahid Yusuf, World Bank

EDITORIAL BOARD

Paul Collier, World Bank

Susan Collins, Georgetown University

Angus Deaton, Princeton University

Barry Eichengreen, University of California–Berkeley

Gregory K. Ingram, World Bank

Emmanuel Jimenez, World Bank

Howard Pack, University of Pennsylvania

Michael Walton, World Bank

The World Bank Research Observer is intended for anyone who has a professional interest in development. *Observer* articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the *Observer* considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the Editor. Please direct all editorial correspondence to the Editor, *The World Bank Research Observer*, 1818 H Street, NW, Washington, DC 20433, USA.

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

For more information, please visit the Web sites of the *Research Observer* at www.wbro.oupjournals.org, the World Bank at www.worldbank.org, and Oxford University Press at www.oup-usa.org.

Counting the World's Poor: Problems and Possible Solutions Angus Deaton	125
Comment on "Counting the World's Poor" Martin Ravallion	149
Comment on "Counting the World's Poor" T. N. Srinivasan	157
Ecology, History, and Development: A Perspective from Rural Southeast Asia Yujiro Hayami	169
Productivity Growth and Sustainability in Post-Green Revolution Agriculture: The Case of the Indian and Pakistan Punjabs Rinku Murgai, Mubarik Ali, and Derek Byerlee	199
The Politics of Russian Enterprise Reform: Insiders, Local Governments, and the Obstacles to Restructuring Raj M. Desai and Itzhak Goldberg	219

The World Bank Research Observer (ISSN 0257-3032) is published twice a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009 for The International Bank for Reconstruction and Development / THE WORLD BANK. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Research Observer*, The World Bank, 1818 H Street, NW, Washington, D.C. 20433, USA. E-mail: researchobserver@worldbank.org.

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

Subscriptions: Subscription is on a yearly basis. The annual rates are US\$35 (£27 in UK and Europe) for individuals; US\$80 (£55 in UK and Europe) for academic libraries; US\$100 (£67 in UK and Europe) for corporations. Single issues are available for US\$22 (£17 in UK and Europe) for individuals; US\$50 (£34 in UK and Europe) for academic libraries; US\$63 (£42 in UK and Europe) for corporations. All prices include postage. Individual rates are applicable only when a subscription is for individual use and are not available if delivery is made to a corporate address. *Readers with mailing addresses in non-OECD countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request by writing to the UK address below.* All subscription requests, single issue and back issue orders, changes of address, and claims for missing issues should be sent to:

North America: Oxford University Press, Journals Customer Service, 2001 Evans Road, Cary, NC 27513-2009, USA. Toll-free in the USA and Canada: 800-852-7323, or 919-677-0977. Fax: 919-677-1714. E-mail: jnlorders@oup-usa.org.

Elsewhere: Oxford University Press, Journals Customer Service, Great Clarendon Street, Oxford OX2 6DP, UK. Tel: +44 1865 267907. Fax: +44 1865 267485. E-mail: jnl.orders@oup.co.uk.

Advertising: Helen Pearson, Oxford Journals Advertising, P.O. Box 347, Abingdon SO, OX14 1GJ, UK. Tel/Fax: +44 1235 201904. E-mail: helen@oxfordads.com.

Requests for permissions, reprints, and photocopies: All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without either prior written permission of the publisher (Oxford University Press, Journals Rights and Permissions, Great Clarendon Street, Oxford OX2 6DP, UK; tel: +44 1865 267561; fax: +44 1865 267485) or a license permitting restricted copying issued in the USA by the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 (fax: 978-750-4470), or in the UK by the Copyright Licensing Agency Ltd., 90 Tottenham Court Road, London W1P 9HE, UK. Reprints of individual articles are available only from the authors.

Copyright: Copyright © 2001 The International Bank for Reconstruction and Development / THE WORLD BANK. It is a condition of publication in the journal that authors assign copyright to The International Bank for Reconstruction and Development / THE WORLD BANK. However, requests for permission to reprint material found in the journal should come to Oxford University Press. This ensures that requests from third parties to reproduce articles are handled efficiently and consistently and will also allow the article to be disseminated as widely as possible. Authors may use their own material in other publications provided that the journal is acknowledged as the original place of publication and Oxford University Press is notified in writing and in advance.

Indexing and Abstracting: *The World Bank Research Observer* is indexed and/or abstracted by ABI/INFORM, CAB Abstracts, Current Contents/Social and Behavioral Sciences, Journal of Economic Literature/EconLit, PAIS International, RePEc (Research in Economic Papers), Social Sciences Citation Index, and Wilson Business Abstracts. The microform edition is available through UMI, 300 North Zeeb Road, Ann Arbor, MI 48106, USA.

Paper Used: *The World Bank Research Observer* is printed on acid-free paper that meets the minimum requirements of ANSI Standard Z39.48-1984 (Permanence of Paper).

Postal Information: *The World Bank Research Observer* (ISSN 0257-3032) is published twice a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Send address changes to *The World Bank Research Observer*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009.

Counting the World's Poor: Problems and Possible Solutions

Angus Deaton

As recent discussions have made clear, the apparent lack of poverty reduction in the face of historically high rates of economic growth—both in the world as a whole and in specific countries (most notably India)—provides fuel for the argument that economic growth does little to reduce poverty. How confident can we be that the data actually support these inferences? At the international level, the regular revision of purchasing power parity exchange rates plays havoc with the poverty estimates, changing them in ways that have little or nothing to do with the actual experience of the poor. At the domestic level, the problems in measuring poverty are important not only for the world count but also for tracking income poverty within individual countries. Yet, in many countries, there are large and growing discrepancies between the survey data—the source of poverty counts—and the national accounts—the source of the measure of economic growth. Thus economic growth, as measured, has at best a weak relationship with poverty, as measured.

The World Bank prepares and publishes estimates of the number of poor people in the world. Although these numbers should be taken with a pinch of salt, they are arguably important. In an institution where the reduction of poverty is the paramount objective, some overall yardstick of progress (or the lack of it) is required. The numbers are frequently quoted by politicians and by leaders of international organizations, including the World Bank itself, who believe the numbers are effective for advocacy. Indeed, there is a long history of studies of poverty mobilizing support among the nonpoor for antipoverty policies. So it is important to know whether the world and national poverty counts are sound enough to support these uses.

As recent discussions have made clear, the apparent lack of poverty reduction in the face of historically high rates of economic growth—both in the world as a whole and in specific countries (most notably India)—is providing fuel for the argument that economic growth does little to reduce poverty. How confident can we be that the data actually support these inferences? Are the changes in the poverty counts sufficiently well measured to support conclusions about growth and poverty reduction? Should

the World Bank stand ready to be judged by its success in reducing the current measures of world or even national poverty? If not, can better data collection or better methodologies improve the numbers?

A more fundamental issue is whether the number of people suffering consumption (or income) deprivation is the right yardstick. It is widely agreed that other dimensions of poverty are more important than income poverty, particularly deprivations in health, education, and democratic rights. There are difficult measurement issues associated with all of these other dimensions, especially if we try to combine them into a single measure, such as the Human Development Index. For many purposes, we need to monitor changes over periods of a few years, which is typically quite difficult for nonincome-based measures. For example, the average literacy rate of a population changes too slowly to be much use as a monitoring device in all but the longest of runs. Even so, and independently of measurement issues, the importance of these nonincome dimensions of poverty should be a warning not to overemphasize income poverty at the expense of other measures. It should also be noted that other dimensions of deprivation are just that, "other measures." They are not perfectly correlated with income or consumption poverty so that, just as we cannot subsume health or literacy measures into an income measure, we also should not try to measure income poverty by using health or education as a proxy. Difficulty in measuring income poverty is a poor argument for focusing on other dimensions of welfare and deprivation.

This article discusses a number of problems with the current \$1/day poverty counts, makes some suggestions for improvement, and identifies issues that need further research. World poverty numbers are calculated in two stages. At the first or international stage, a world poverty line is set and used to derive comparable poverty lines for each country. At the second or domestic stage, the poverty lines are used to count the number of poor people in each country, and the totals are added up over countries. It is useful to think about each stage separately, and there is disquieting evidence about both. At the international level, the purchasing power parity (PPP) exchange rates that are used to turn the \$1/day poverty line into national currencies are arguably inappropriate in theory. In practice, their regular revision (to different base years with different relative prices) plays havoc with the poverty estimates, changing them in ways that have little or nothing to do with the actual experience of the poor.

Within each country, at the domestic stage, the World Bank's poverty count comes from household surveys. The incorporation of these unit record data into its statistical base is surely one of the Bank's major achievements over the past 20 years. Yet in many countries—including India, where around a third of the world's poor live—there are large and growing discrepancies between the survey data and the national accounts. Because poverty counts come from the survey data and growth measures come from the national accounts, and because they are evidently measuring differ-

ent things, there is no consistent empirical basis for conclusions about the extent to which growth reduces poverty. That economic growth, *as measured*, has at best a weak relationship with poverty, *as measured*, means little more than would a finding that growth in China had failed to reduce poverty in India; they are simply different things.

International Issues in Counting the Poor

I discuss the international issues first, followed by the domestic ones. For those who are not persuaded that the world poverty counts are useful or believe them to be “essentially propagandist” (Srinivasan 2000), these international issues are irrelevant.

Purchasing Power Parities and the Base Poverty Line

The current concept for world poverty is the number of people who live in households whose daily consumption per head is less than the PPP equivalent of \$1/day in constant 1985 PPP dollars. In the latest revision, this has been updated to \$1.08 in constant 1993 PPP dollars, but is still conveniently referred to as the \$1/day poverty line (Chen and Ravallion 2000). The virtues of this approach are many. It is simple, is easy to remember, and applies equally to all countries. It is denominated in a currency that is familiar to the relatively wealthy people who are the primary users of the measures and who are the primary target for rhetoric based on them. The \$1/day was originally selected as being representative of poverty lines in use in low-income countries; the new \$1.08 is representative of current lines and is thus anchored in actual practice. The \$1.08/day is converted to the 1993 value of local currencies using internal World Bank PPP tables, and then updated to the target year using a local consumption deflator. The resulting poverty line is then used to count the number of people in each poor (technically World Bank part two) country using household survey data. When possible, the raw data are used; when this is not possible, values are interpolations based on published tables from such surveys.

Given world and U.S. inflation between 1985 and 1993, it is somewhat surprising that the international poverty line should have increased by only 8 percent. But the updating was carried out by going back to the country poverty lines and converting back to international dollars. Thus the modest increase comes not from a failure to allow for world inflation, but because the PPP international dollar has strengthened relative to the currencies of the poor countries whose poverty lines are incorporated into the international line.

The PPP numbers come from data collected from 110 countries in 1993 by the International Comparison Project. Earlier PPP numbers came from the Penn World Tables, which covered only 60 countries. The revision from the earlier PPPs to the new PPPs in the most recent revision of the poverty counts, Chen and Ravallion

(2000), resulted in some disconcertingly large changes in poverty counts even for the same country in the same year and using the same survey data. The worst cases are Sub-Saharan Africa, where the poverty rate in 1993 (that is, the same year) increased from 39.1 to 49.7 percent; Latin America, with the poverty rate falling from 23.5 to 15.3 percent; and the Middle East and North Africa, where the poverty rate fell from 4.1 to 1.9 percent. Changes of this size risk swamping real changes, and it seems impossible to make statements about changes in world poverty when the ground underneath one's feet is changing in this way. Note that Chen and Ravallion (2000) recalculate previous counts at the new PPPs, so the issue is *not* that we are comparing current poverty using one PPP conversion with earlier poverty using another conversion. The problem is that, if revisions can have such enormous effects, it is difficult to place much confidence in any set of poverty estimates, either of levels or of changes, even those based on the latest and best set of PPP exchange rates.

For countries whose poverty lines are near the middle of the distribution of consumption or income, poverty rates can be extremely sensitive to changes in the poverty line, including those induced by PPP (or other) errors. World Bank (1997) lists Thailand as having only 0.1 percent of its population living on less than \$1/day at PPP. This virtual elimination of poverty was cited in the *New York Times* by then Chief Economist Joseph Stiglitz as one of the consequences of the Asian economic miracle (Stiglitz 1997), but it is much more likely a tribute to inappropriate PPP conversion. This anecdote suggests that, at a minimum, it would be wise to check the PPP poverty lines for plausibility in the countries to which they apply. Of course, if the lines are wildly implausible, we need some other method.

There are two conceptually different kinds of revisions to PPP exchange rates. First, there are the usual revisions that come with better information, elimination of previous errors, and so on. For example, it is claimed that for China, the new updated PPP rates are much more solidly based than were the old rates. In addition, the PPP exchange rates are now calculated, rather than imputed, for 50 more countries. Second, there are the revisions that come from the change in the base, given that PPP exchange rates depend on the relative prices of commodities in the base year. This second source of revisions ultimately traces back to the conceptual or "index number" problems associated with making international comparisons of well-being. Indeed, one might argue that the \$1/day poverty line, by its very simplicity, is misleading just because it seems to sweep away the problems.

The calculation of the cost of subsistence for a poor Indian versus that for a poor Kenyan or Brazilian is fraught with conceptual difficulties. For example, we could calculate the relative cost of living of an Indian living in Calcutta compared with the (hard-to-imagine) cost of living if that person was moved to São Paulo. We would likely get a very different number from doing the same with a Brazilian transplanted to Calcutta. Suppose that, in country A, the staple food for poor people is beans and in country B it is rice. If we price beans in country B in B's currency and divide by the

bean price in country A in A's currency, we get an implicit "beans" PPP exchange rate. A similar calculation for rice gives a "rice" PPP exchange rate. These will give the same answer (or at least reciprocals of one another) only if both goods are freely traded between the two countries without tariffs, barriers, or transportation costs. These conditions are far from holding true in practice, even approximately.

Many (perhaps most) of the world's poor eat food staples that are grown only a short distance away and whose price is only loosely (if at all) connected to world commodity prices. And the staple food in one country (or even region) is often a relatively rare (and even expensive) luxury somewhere else (for example, sticky rice in Northern Thailand versus the United States). Therefore, the exchange rate obtained from the prices of the same commodity is not very useful for converting poverty lines from one country to another.

In practice, matters are not quite so stark. Actual PPPs price a representative bundle of goods in each country and compare the local cost of the bundle with the U.S. dollar cost of the same bundle. But note that the answer depends on the structure of *relative* prices at the time of the comparison, so that, when PPPs are revised with a new base year, the new exchange rates are not simply a new measure of the *old* concept but a new measure of a *new* concept. World prices of primary commodities are notoriously volatile, and for some countries, primary commodities make up a large share of gross domestic product (GDP). As a result, the PPP exchange rate for Nigeria (for example) relative to the United States will vary with the world price of oil in the base year. This is presumably why the African and Latin American poverty rates are so sensitive to revisions of the PPP base year. Although it might be argued that changes in the world prices of beans and rice ought to change the relative poverty lines of Brazilians and Indians, it is much harder to make the case for changes in the world price of oil. Even in theory, PPP exchange rates as currently defined are not designed to convert poverty bundles (see also Srinivasan 2000).

One way of improving the PPP poverty lines would be as follows. First best would be an international comparison program that focuses not on GDP and its components but on some more appropriate poverty bundle. In this case, the first best is perhaps not very good, if only because even a poverty-bundle PPP would not address the underlying index number issues. A second-best shortcut would be to take the current set of 1993 poverty lines, update them to the present, and check them in each country (or at least those countries that have significant numbers of poor people). In cases where the numbers are implausible (for example, showing no poverty in Thailand), the PPP exchange rate can be calculated using the first-best procedure. The calculation would price out in dollars and in local currency an appropriate poverty bundle, for example the average consumption bundle of the bottom quintile of the population. Although the choice of the bottom quintile is arbitrary (but sensible), the calculations are straightforward given a household expenditure survey, which often collects data on quantities as well as on expenditures. When the survey does not

contain quantity data, they can be calculated from the expenditure data using a set of local consumer price quotes, such as those that are routinely used for the construction of national consumer price indexes.

Armed with a set of PPP poverty lines that have been locally validated and adjusted as necessary, the crucial next step is to hold these fixed (in real terms) and not make further adjustments as PPP exchange rates are revised in response to changes in the base year and its associated relative prices. The benefits of this proposal are that African and Latin American poverty rates are not held hostage to fluctuations in the prices of primary commodities in the base year of PPP calculation, and that world poverty rates are consistently calculated against a fixed target. Of course, it could be argued that the line is no longer the constant PPP \$1/day line, and that the World Bank is failing to update for new information. But this objection is more semantic than real. The lines originally started at \$1/day and can retain the label. They are updated by local prices, just not for changes in world relative prices. The proposal effectively changes the *definition* of PPP exchange rates, whose shifting basis is an embarrassment, not an advantage.

Even if the lines, once set, are to some extent arbitrary, and even if it were the case that a reworking from scratch would lead to new lines, there is much to be said for holding them fixed. Indeed, it is hard not to be impressed by the durability of official poverty lines in countries as diverse as India and the United States, whose poverty lines have survived unchanged (except for inflation correction) for 30 and 40 years, respectively. Once an initial poverty line has been set, and provided it commands public and political support at the time it is set, it seems that it can generate useful estimates of poverty for many years. Moreover, and for both India and the United States, the original nutritional basis for the lines is often cited in their support. The same could be true for the \$1/day lines.

International Comparability without Purchasing Power Parity

There are a number of alternatives to the \$1/day PPP poverty line, although none is without problems. There is a tradition, followed in both the United States and India, of setting poverty lines with reference to what Ravallion (1998) refers to as "the nutritional requirements for good health." This is defined as the level of income (or total consumption) at which, on average, nutritional norms are met. Note that, because the standard is set in terms of income or total consumption, this is *not* the amount necessary to purchase the nutritional norms only. It is the amount spent on the nutritional norms and on other goods by an average household that just meets the norms. An internationally comparable set of poverty lines could be set in this way with reference to a universal calorie norm, say 2,000 calories per person per day. A national household expenditure survey for each country would be the basis of the

calculation of the level of total household expenditure per capita at which this norm is met in expectation.

Again, this method is not quite what it seems, and there are important practical qualifications. First, as with the recommendations for the amended PPP method, the procedure is used *once and once only*. The resulting poverty lines for each country are subsequently held fixed over time. This may seem contradictory; if the calorie method is correct, it should surely be applied consistently over time and space. But it is not correct in the sense of being a “scientifically sound” method of setting the line. Instead, it is a device for generating a socially acceptable poverty line that can then be held fixed over time. The food rhetoric helps build legitimacy for the line, but, judging from India and the United States, maintaining that legitimacy does not require updating, except possibly in the very long term. Successive commissions and expert groups have altered details of the Indian lines, for example, by introducing variation across states (Government of India 1993). In the United States, there have been frequent (sometimes detailed) discussions of reform, particularly National Research Council (1995), but none of the recommendations have been adopted.

Second, it is important *not* to set separate nutritional poverty lines for different regions or sectors within a country. At the same level of per capita total household expenditure, urban people spend less on food, buy more expensive calories, and consume fewer of them. Therefore, applying the method to urban and rural areas using the same calorie target will lead to higher poverty lines in urban than rural areas. Across regions with different income levels, the operation of Engel’s Law has much the same effect, so that poverty lines can move more or less in proportion to average incomes, generating effectively relative, not absolute lines (Ravallion and Bidani 1994; Ravallion and Sen 1996). Although relative lines may make a good deal of sense in other contexts, they are not appropriate for world counts that are based on an explicitly absolute standard.

Domestic Issues in Counting the Poor

Even if there existed perfectly accurate and perfectly relevant international exchange rates, any problems with measuring domestic poverty will carry through to the world totals. There are several such problems.

National Accounts and Household Survey Data

A major source of controversy is the discrepancy between national accounts and survey estimates of consumption, especially when, as in India, the difference is increasing. India is the largest single contributor to the world poverty count, with more than 400 million poor people out of a world total of 1.2 billion. The ratio of National

Sample Survey (NSS) consumption to national accounts statistics (NAS) consumption has fallen from around unity in the 1950s, to 89 percent in 1968–69, to around 75 percent in the late 1970s, to a little more than 50 percent in recent years (see Bardhan 1974; Mukherjee and Chatterjee 1974; Srinivasan and others 1974 for early discussions; and Srinivasan 2000 for a recent update). Although different authors give somewhat different ratios, there is general agreement that the discrepancy is widening and that the difference in per capita growth rates is currently about 2 percent a year. As a result, the rapid postreform growth in GDP and consumption per head measured in the NAS does not show up in the NSS surveys. The latter are now the exclusive basis for the official poverty counts, and they show little or no reduction in official poverty during the reform period.

The Indian example is only the most notable of similar or related phenomena elsewhere, including some industrial countries, such as the United States. Triplett (1997) estimates for the United States that, even after correcting for differences in concept, between 1973 and 1994, per capita consumption in the national accounts has grown 0.4 percent a year faster than per capita consumption in the Consumer Expenditure Survey (CEX). Between 1984 and 1994, the difference in growth rates was a full 1 percent a year. (Note that the CEX data are not used to derive the national poverty counts in the United States.)

Ravallion (2000b) notes a similar drift in other countries and regions, particularly in China (the second-largest single contributor to world poverty) as well as in Latin America, where survey consumption growth is only 70 percent of NAS consumption growth. In Sub-Saharan Africa, consumption growth in the surveys is less well linked to the NAS but is also consistently slower. Although there is no drift in the transition economies of Eastern Europe and Central Asia, it is only because there appears to be no relationship whatever between consumption growth in the NAS and in the surveys.

Ignoring the transition economies, Ravallion cannot reject the hypothesis that the coefficient of the growth rate of NAS consumption on survey consumption is unity in a regression on a cross-section of countries. However, as is appropriate for his purpose, Ravallion's units are countries, not people, and most of the poor people in the world live in countries or regions (East Asia, India, and Sub-Saharan Africa) where survey growth rates of consumption are substantially less than national accounts growth rates. Even when the coefficient is insignificantly different from one, the point estimate is 0.836, indicating substantial slippage, and is lower still in East Asia (0.628), Sub-Saharan Africa (0.645), and South Asia (0.742).

That the drift is so common around the world and that it is generally in the same direction suggest a common cause: clearly, it would be an important research project to try to track it down. For example, does the positive skewness in consumption, together with the plutocratic basis of the NAS versus the democratic basis of the surveys, result in a tendency for average consumption in the surveys to lie beneath the

population mean? Is average consumption from the surveys even lower when inequality is increasing? Some limited calculations show this to be unlikely for reasonable parameters. There are other plausible links between increasing inequality and increasing bias. For example, the rich are more likely to be missed in household surveys—something that is widely suspected in India and has been well documented in the United States (Groves and Couper 1998).

Because growth around the world is almost universally assessed from the NAS, and because the poverty counts come entirely from the survey data, the discrepancies between the two drive a wedge between measured growth and measured poverty reduction. If we were to measure both growth and poverty from the household survey data, the discrepancy would vanish; there has been little or no measured growth, and little or no measured poverty reduction. Similarly, if we were to follow Bhalla (2000a, 2000b) and treat the national accounts estimate of mean consumption as the truth, using it to “correct” the survey data, the discrepancy would also vanish but in the other direction; there would be lots of growth and lots of poverty reduction. Meanwhile, claiming that growth has done little to reduce poverty is comparable to saying that, in spite of rapid growth in China, poverty in India remained the same. The NAS and the surveys evidently measure different things.

Measuring consumption. Which is correct? We don’t know, although it seems safe to say that there are almost certainly errors in both the NAS and the surveys. There is a longstanding prejudice by many economists (although not in India) against surveys and in favor of national accounts that is probably without basis. However, the following points are worth noting.

First, the two concepts are different. In particular, there are items in NAS consumption, perhaps most important consumption of nonprofit organizations and the imputed rental of owner-occupied dwellings, which are not included in the surveys. In India, it is thought that as much as half of the current discrepancy between the two estimates can be accounted for by implicit rents. It is also plausible that these rents are growing more rapidly than other items of consumption, thus contributing to the differential rates of growth.

Second, although expenditure estimates from household surveys are sometimes used in constructing NAS consumption for some commodities, most items are derived as residuals, so that errors and omissions elsewhere in the accounts are automatically absorbed into consumption. A major problem is consumption by businesses, which has to be estimated and subtracted from the production totals to get the consumption of households. Consumption of coarse grains by animals presents a similar problem and is often estimated using a set of multipliers that are difficult to estimate with any precision and that are infrequently updated. It is not clear that the NAS estimates of consumption should be treated as the gold standard to which the survey estimates should ideally correspond. Indeed, Bardhan (1974:117) writes, “it

would be highly improper to judge the goodness of NSS estimates by pointing to their divergence (or lack of it) from such residual estimates.”

Third, there are many well-known problems with data collection from household surveys. Best-practice consumption measures (as in India or Indonesia) use long lists of specific items. The evidence suggests that attempts to shortcut this process lead to underestimation of consumption (Deaton and Grosh 2000). Yet many surveys use such questionnaires, which tend to change slowly over time and may not keep up with the increase in consumption of goods and services that may have been unknown a decade or two ago.

Fourth, consumption surveys miss some households through refusals. Surveys in developing countries typically report low refusal rates (for example, 0 in the Indian NSS) and do not document the extent to which willing respondents have been substituted for unwilling ones. There is some evidence that wealthy households are most likely to refuse to cooperate; for example, it is often impossible for enumerators to gain access to gated communities. If so, it might be conjectured that real income growth, by increasing the fraction of such people, will lead to an increasing underestimation of consumption. Against this, it might be argued that the very poor, especially those without fixed abode, are also missed in household surveys.

Fifth, in most surveys, different kinds of households have different probabilities of being included, either by design or because of problems in the field, as when some types of household are more prone to refuse their cooperation. As a result, the survey results must be weighted to give an accurate representation of the population as a whole, and the calculation of suitable weights depends on the availability of accurate, up-to-date information about the population. Although statistical agencies have a number of tools for ensuring accuracy, it would not be surprising if the weights were to become progressively more inaccurate as the latest census becomes outdated. That said, the divergence between survey and NAS estimates in India shows no evidence of the saw-tooth pattern—growing between censuses and then falling immediately after each census—that would be predicted by such an account.

Sixth, it is tempting to try to use the behavior of other measures of poverty, such as literacy or infant mortality, to judge the relative plausibility of the NAS on one hand and the surveys on the other. There is undoubtedly good research to be done along these lines. Even so, definitive results should not be expected. Health and literacy measures should not be regarded as proxies for income, if only because they frequently change in ways that have little or nothing to do with income. We do not have a reliable link between income and nonincome measures of poverty that would allow us to infer the former from the latter.

Much the same can be said of attempts to look at indicators like the share of the budget devoted to food. Although it is true that the food share drops with increases in income, it is also influenced by other factors, including relative prices, demographic changes, and the distribution of income, to mention only the most obvious. As a re-

sult, factors (such as an increase in inequality) that are possible suspects in driving the discrepancy between the national accounts and the surveys are also likely to affect the relationship between food and total expenditure. Therefore it is difficult to know how much weight to attach to calculations such as that in Bhalla (2000a) that impute the growth in total expenditure from data on food availability. Perhaps more promising in practice, if not in theory, are the measures of living standards derived by Filmer and Pritchett (1998) from principal components analyses of the various asset ownership measures that are included in Demographic and Health Surveys.

What should be done? It is, of course, possible for poverty to remain constant or to increase in the face of positive growth in average consumption if people in the lower tail of the distribution are left behind, and if most or all of the growth accrues to people above the poverty line. That growth should disproportionately favor the better off is entirely plausible in some settings. For example, in India, the sources of growth have been outside of the agricultural sector where most poor people make their living. But differences of definition apart, the NAS and survey estimates of consumption cannot both be right because their means are different and are systematically becoming more different over time. Whatever the role of increasing inequality in moderating the poverty-reducing effects of growth, one of the reasons why measured growth has not reduced measured poverty is the growing statistical discrepancy between the two sets of data on which each is measured.

What to do? There is a school of thought in India among the pro-reformers, as in Bhalla (2000a), which argues that the NSS consumption figures should be scaled up so as to match the NAS, and that the rescaled survey estimates should be used to calculate poverty. (Indeed, this is how the official poverty counts used to be done in India, and no very convincing reason was ever given for the change following the recommendations in Government of India [1993].) Bhalla (2000b) argues that the “means from NAS, distribution from surveys” formula should be adopted for the world poverty accounts and calculates that growth has indeed reduced poverty. But there is a serious lack of evidence for at least some of the assumptions that would validate such a calculation. In particular, the following must be true: (a) the NAS estimates are correct; (b) the survey estimates of the mean are incorrect; and (c), in spite of b, the consumption levels of each household in the surveys are correct up to a multiplicative factor. The last condition is a real stretch.

Most of the plausible accounts of what might have gone wrong with the surveys, such as a progressively larger number of refusals, involve a systematic underestimation of inequality as well as of the mean. It seems particularly implausible that the degree of underestimation in the surveys should be the same for the urban and rural households (see Ravallion 2000a). Indeed, many people in India, particularly those skeptical of the benefits of the reforms, believe that the growth that has taken place in consumption has accrued to people well above the poverty line and to urban more

than rural people. They believe that the people who have most benefited are the most likely to be undersampled by the surveys. Moreover, the consumption of the poor—who do not own valuable housing—is unlikely to be understated by the omission of owner-occupied rents. If these arguments are correct, one might suppose that the NAS modestly overestimates the rate of growth of consumption, that the NSS greatly underestimates it by missing the wealthy, and that the NSS-based poverty figures are essentially correct.

A good statistical principle is that, if two sources of data disagree and we have no reason to favor one over the other, then we should combine them to make a better estimate. This principle would argue for a more modest version of the scaling proposal, in which the survey data are scaled up by some weighted average of the NAS and the survey means, at least after correcting for conceptual differences and coverage. Of course, the data on economic growth would also have to be scaled down accordingly. This proposal would make neither side happy, each of which has its reasons for believing that one set of accounts is biased and the other bias-free, but I can see no very convincing argument against it. Its limits are obvious; it takes no account of reasonable suspicions that the NSS errors are not neutral between rich and poor or between urban and rural. But in the absence of the sort of better information that one might hope would be produced in the longer term, it is surely better than either of the extreme alternatives that currently dominate the debate.

The urgent need is to start a serious program of reconciliation between the NAS and survey data in a few countries, of which India must surely be one. (The Bureau of Labor Statistics and the Bureau of Economic Analysis at Census are talking about setting up a similar program in the United States. Perhaps common cause could be made.) This program must work both with the national accounts and with the survey organizations and must not begin by assuming that one is right and the other wrong. It might be productive to start with a particular commodity or group of commodities, such as cereals, where the production, export, and import data are relatively strong and where there is a great deal of experience with household surveys. It would be possible to mount special supplementary household surveys for the task, which could focus on only a few goods. But it would also be a good idea to audit the performance of the survey agencies, just as the NAS agency is being audited. For example, the Indian NSS used to lead the world in household survey methodology and practice. But that was a long time ago; the organization is no longer as dynamic and open to new practice as it once was. There has been a great deal of progress in survey methodology in the last 20 years, with the result that those statistical offices that used to be the leaders, mostly in South and East Asia, are now the most likely to be behind.

The World Bank is the only international body that has the capacity and organization to run a research program designed to resolve the discrepancies between national accounts and survey data. To its credit, the World Bank has played a major

role, through the Living Standards Measurement Study (LSMS) surveys, in collecting household survey data on a nearly worldwide basis. The next step is to find out why these surveys give such different results from the national accounts. Until it does so, neither the Bank nor anyone else will have a firm platform on which to make statements about the central issue of how growth affects poverty.

Price Indexes

Local consumer price indexes are used to update a real poverty line for comparison with the nominal expenditures or incomes collected in the surveys. Consumer price indexes are of variable quality around the world and poverty counts can be very sensitive to any errors. Consumer price indexes are constructed from two components: a set of prices, collected on a regular schedule from retail shops and markets around the country, and a set of weights, which typically come from a household expenditure survey.

Problems can arise with both components. For example, some countries collect good data in urban centers, but find it much more difficult to collect adequate price data in the countryside, where many poor people live. Urban bias may result in prices that are more relevant for relatively affluent urban workers than for the poor whose poverty line we are trying to calculate. There are also perennial questions about the representativeness of the markets and whether the enumerators observe the prices that people actually pay. In India, the villages where prices are collected are updated infrequently, if at all.

Weights can also be troublesome, especially if not updated sufficiently often. When it was revised in late 1995, the Indian Consumer Price Index for Agricultural Laborers (CPIAL) had weights that were 35 years old. It is usually thought that Laspeyres indexes whose weights are held fixed for too long are increasingly likely to overstate the rate of inflation and, thus, in this context, to cause an underestimation of the rate of poverty reduction. (Note, however, that price indexes are *not* at the root of the discrepancy between NAS and survey data; the divergence is in nominal terms.) When updating its poverty lines, the Indian Planning Commission reweights the components of the CPIAL so as to match more closely the purchases of people near the poverty line, using weights from 1973–74, which are still elderly by international standards. (The World Bank, in its \$1/day poverty calculations, uses a slightly modified version of the CPIAL.) Deaton and Tarozzi (2000) have made independent estimates of price indexes in India, which suggest that even the reweighted CPIAL has been rising too quickly. Their alternative estimates show some decline in poverty in the 1990s, although much less than would be the case if the NSS data were scaled up to agree with the NAS data on consumption. I suspect that India is much better provided with good price indexes than are other countries. Indeed, much of the recent debate on the effects of the East Asian crisis on poverty

in Indonesia has foundered on the unavailability of good price indexes for rural areas (see Thomas and others 1999).

Local price indexes are used not only for updating over time but also for adjusting poverty lines for urban-rural price differences, as well as for regional differences in prices. In the World Bank's calculations of the number of poor in the world, separate urban and rural indexes are used only for India and China. In other countries, a single index does service for everyone, an expedient that must overstate rural relative to urban poverty.

For many years, India had urban and rural poverty lines of Rs 57 and Rs 49, respectively, at 1973–74 prices, a differential of 16 percent. Both lines were updated by the implicit price deflator of consumers' expenditure in the national accounts. Since 1993, the planning commission has adopted different urban and rural poverty lines for each state. The lines are designed to reflect different prices in different parts of the country, and each line is updated by a poverty-reweighted urban or rural price index specific to that state. For a number of reasons, the implicit urban-rural price difference in the current lines, a difference that was never explicitly measured, has risen to around 40 percent on average, with larger or smaller differentials in individual states.

This large urban-rural price difference, for which Deaton and Tarozzi (2000) find no evidence in prices actually paid, yields urban poverty rates that are larger than rural poverty rates in many states. This finding seems implausible to most observers and sits ill with the evidence on other poverty-related indicators. Deaton and Tarozzi recalculate the rates based on their price indexes, with differentials quite close to the once standard 16 percent, and find that urban poverty is reduced from 32.4 to 18.1 percent in 1993–94, which reduces the total number of Indians in poverty by 23 million people. Thus price differences within a country can have major effects on the poverty counts both nationally as well as in their contribution to the world totals.

Problems with Household Surveys

The World Bank is to be commended for building so much primary survey data into its poverty counts. This is a great advance over the situation 20 years ago, and the Bank has been the prime mover in bringing these data to bear on the measurement of poverty. But the surveys have a number of limitations for poverty work that need to be kept in mind, some of which might be ameliorated by more work or different approaches. Here I discuss three of the most important limitations: coverage, reference periods, and income versus consumption.

Coverage. Household income and expenditure surveys are good at collecting data on cash that passes through the household for both incomes and outlays. They are some-

what less good (but probably still good enough) at collecting data on own production and consumption. However, they do not attempt to include in the consumption or income estimates the benefits that people receive from publicly provided goods of one kind or another. Although it would be difficult to collect such information, some relevant data are collected in the surveys (for example, visits to clinics and school attendance) and perhaps more might be.

Problems develop if the poverty counts ignore this information. In particular, consider the debates about structural adjustment and the (widely credited) arguments by nongovernmental organizations and others that the World Bank has impoverished the poor by forcing cuts in public expenditure. It is not very useful, as the Bank sometimes does, to argue that there was no impoverishment because the poverty counts (which take no account of the provision of public goods) are going down. Any program that eliminated government services (closing clinics and schools) and shared the money among the population would reduce the poverty count as the World Bank currently makes it. The Bank has a good deal of experience in benefit incidence analysis, that is, the use of household surveys to allocate the benefits of public expenditures. It would be worth giving serious consideration to adding some of these estimated benefits to the consumption totals before making the poverty counts.

Reference periods. Different household surveys use different recall periods, even for the same types of goods. Even within the LSMS surveys, there is some variation. Given a true flow of expenditures, the rate of expenditure that is reported is sensitive to the reporting period; longer reporting periods lead to more forgetting, at least for common purchases like food. Again, India provides a good example. Based on experiments done long ago by Mahalanobis and Sen (1954), whose results were quite contrary to later findings in the literature, the Indian NSS has always used a 30-day recall period for food (see Deaton and Grosh 2000). Seven or 14 days is a more common standard in modern surveys, and Scott and Amenuvegbe (1990) have argued that estimates are biased downward after even one day. In the last several years, the NSS has experimented by using a 7-day reference period for half the households and a 30-day period for the other half. There were also changes for goods, such as durables and clothing, in which a 365-day reference period was used, again compared with a 30-day period in the standard survey.

Visaria (2000) shows that the shorter recall period causes more (food) expenditures to be reported. Using the official poverty lines and the new data, measured rural poverty in the first six months of 1998 falls from 42.6 to 23.6 percent, and urban poverty from 32.9 to 20 percent, removing about 175 million people from poverty. (The effect would be somewhat larger for the \$1/day international line, which is higher than the Indian official poverty line.) Of course, Visaria's calculations offer no apparent help on the issue of differences in trend between the NAS and the NSS versions of consumption. (Although it is possible that the relatively underreported foods

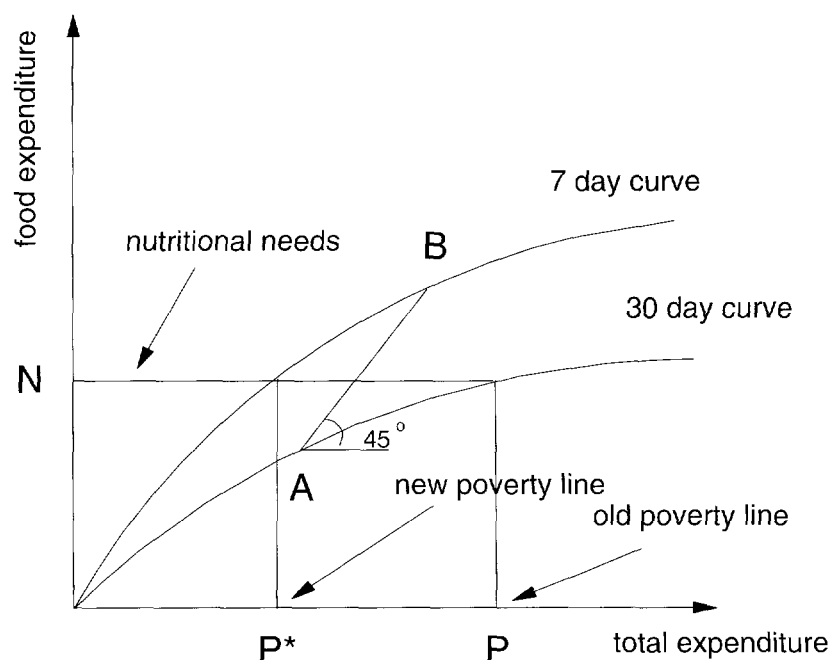
are those that have been growing more rapidly, preliminary analysis of the data does not support the conjecture.) Somewhat surprisingly, the estimates based on the experimental questionnaire (7-day, 30-day, and 365-day recall, as opposed to 30 days for all goods) show *less* variance across households than do the estimates from the traditional questionnaire (Government of India 2000; Deaton 2000). Thus there is no evidence of people being incorrectly characterized as poor simply because they happened to make food purchases outside the shorter recall period.

One question about Visaria's calculations is whether it is legitimate to use the old poverty line with the new reporting periods and the greater food consumption that they reveal. The answer depends on where one thinks the poverty line comes from and the purpose it serves. Because the Indian line was originally set by the nutritional needs method, perhaps the obvious thing to do is to try to recalculate what the poverty line would have been had the original method been followed using data with the shorter reference period. This is not the only possible treatment. Instead, one might argue that the legitimacy and longevity of the line owes more to the general perception that Rs 57 per head in 1973–74 prices is a reasonable poverty line in reference to people's experiences of earning and spending and their notions of what constitutes poverty. If so, there is no reason to revise it in the light of the new survey data, and Visaria's calculations are correct as they stand. Alternatively, one might argue that the line is acceptable only because people accept the headcount ratio that it generates, in which case the survey data are essentially irrelevant.

If we follow the calorie approach, it turns out that the poverty lines should be revised, not upward as intuition might suggest, but *downward*, so that Visaria's calculations, if anything, *overstate* the rate of poverty. The argument is as follows. Pick a level of household per capita total expenditure and consider the average food expenditure over all households at that level, under the new and old survey methodologies. Because the new methodology recovers food expenditures that were previously forgotten, food expenditures rise on average. But total expenditure is both defined and measured as the sum of food and nonfood expenditures, so that total expenditure will rise by the same amount as food expenditure. (For the purpose of the argument, ignore the change in survey procedures for nonfood.)

Figure 1 illustrates. The old (30-day reporting period) Engel curve is the lower curve. For an average household (conditional on expenditure) at point A, food and total expenditure (or per capita) increase by equal amounts along the 45-degree line to B, say. Because the slope of the curve is less than 45 degrees—the marginal propensity to spend on food is less than one—the switch to the seven-day period moves the Engel curve outward to the higher curve. The constant level of nutritional needs, represented by point N, thus requires a shift downward in the poverty line from the old P to the new, lower P*. With the new, better reporting period, more food is reported, so that if we put ourselves back in the shoes of the planning

Figure 1. The Effects of a Shorter Reporting Period on a Nutritional Poverty Line



commission when the poverty line was first calculated, the original poverty line would have generated more calories than the norm, and would therefore have been revised down.

In Deaton (2000), I present evidence on the relevance of figure 1 using data from the 52nd Round of the NSS gathered during 1995–96. The Engel curves relating per capita expenditure on food to total household expenditure per capita are in the order predicted in figure 1, with the seven-day Engel curve lying on or above the 30-day Engel curve. However, the two curves essentially coincide below and up to the poverty line so that, unless the ratio of calories to food expenditure is different in the two questionnaires, there is no reason to revise the original poverty line to meet the new questionnaire design.

In summary, Visaria's reduced poverty counts have as much (or more, judged by international survey practice) legitimacy as the official counts, or indeed the \$1/day calculations. To get some idea of the magnitudes involved, the change in the survey reporting period reduces measured Indian poverty by about as many people as the total number of poor in China. Clearly, there is a large margin of error in the world poverty estimate of 1.2 billion people. In the long run, household surveys should be

brought into reasonable conformity in their reporting periods. In the short run, the counts should not simply add together counts from surveys with different reporting periods without some sort of correction. This will involve sometimes admittedly arbitrary choices, and perhaps the removal altogether of some surveys. But that is better than adding together incommensurate numbers.

Income versus consumption. The surveys used for the world poverty count sometimes collect data on income and sometimes on consumption. The latter is the preferred measure (rightly, in my view), so that the surveys that collect income (38 out of the 91 in Chen and Ravallion [2000], including all the Latin American and Caribbean surveys) have to be “converted” to a consumption basis. This is done by scaling down the income data in the survey by the ratio of national consumption to national income (presumably personal disposable income) in the national accounts. This is plausible, but in my judgment, it is unlikely to lead to good results.

First, there is some inconsistency in using the national accounts data in this situation, but not when there is a full consumption survey, even when the NAS and survey data are mutually inconsistent. The argument is presumably that there is little choice. Even so, it would seem more logical to use some average of NAS and survey data to correct the expenditure survey when available or to make some adjustment to the NAS estimate of the consumption ratio in the case when it is not.

Second, in countries where household surveys collect data on both income and consumption, there is usually little or no evidence of positive saving among poorer households; indeed, the reverse is true. (China seems to be an exception to this generalization.) Even in industrial countries, survey saving rates are frequently negative until well up the income distribution. A frequent argument is that survey respondents understate income relative to consumption. If so, multiplying incomes by a number that is less than one is a “correction” in the wrong direction. More generally, whatever the accuracy of the consumption and income reports themselves, it remains true that, in most countries, multiplying survey income by the national consumption rate would do a very poor job of replicating survey consumption, certainly at the bottom of the income distribution.

Likewise, a proportional correction for saving is not likely to convert income data from a survey into something that looks like consumption data from a survey. In the usual situation, when the poverty line is below the mode of a unimodal distribution, the resulting underestimation of consumption by the poor (relative to a consumption survey) will tend to overstate poverty rates. Countries with income surveys probably tend to be assigned higher poverty rates relative to those with consumption surveys. This contention could be tested for a few countries where surveys collect both kinds of data, or by running a regression across the countries in the database and seeing whether those with income surveys have higher poverty, controlling for GDP and other variables.

Unfortunately, the degree of heterogeneity across countries is unlikely to permit any simple correction based on the results of such a regression.

Third, measured consumption is famously less variable than measured income. This is true in logarithms as well as in levels (the variance of the logarithm of consumption is less than the variance of the logarithm of income). Therefore, even if the proportional saving assumption were right on average, the corrected income data would still be too variable relative to measured consumption. Once again, and provided the poverty line is below the mode, consumption poverty will be overestimated by income poverty.

My guess is that it would not be too difficult to work out some pattern of corrections that, although hard to defend absolutely, would likely still be better than the current procedures. In the long term, the battle for expenditure surveys will be won. Even the holdouts in Latin America are moving in that direction.

Summary, Recommendations, and Directions for Research

The World Bank's worldwide count of the poor starts from a common international poverty line and counts the number of people in each country whose consumption lies below it. The international poverty line, at \$1 or \$2 a day, is converted into domestic currencies using PPP exchange rates. I argue that, although there is much to recommend the \$1/day line, its dependence on PPP exchange rates has a number of unfortunate consequences.

An arguably better procedure for the future would be to hold fixed (in real terms) the current domestic poverty lines, and not to revise them along with changes in PPP exchange rates induced by updating base years. Such a program would work best if the current rates were subjected to detailed, local scrutiny and corrected in a way that would simultaneously give them local credence without major deviation from the \$1/day standard. Such corrections will likely be difficult. They must not be interpreted as a license for individual countries to define their poverty lines with a view to their international political effects. Corrections will be easier if, as I suspect, most anomalies are the result of inappropriate PPP exchange rates, perhaps because interpolation from other countries gives a poor answer or because the bundle of goods being priced is inappropriate for the poor.

Other problems relate less to international comparability than to the calculation of poverty within each country. The proliferation of household surveys, for which the World Bank should take much credit, has greatly enriched the basis for good poverty calculations around the world. But the very availability of expenditure surveys has highlighted an old problem that seems to have recently become much worse, which is a divergence between estimates of average consumption based on surveys

and on national income accounts. The World Bank is probably the only organization that is capable of leading an international research initiative to address these discrepancies. Until it does so, policymakers and commentators will be invited to choose their evidence according to their political predilections. Furthermore, neither the World Bank nor anyone else will be able to make well-supported statements about what is surely a central issue in economic development: whether growth has reduced poverty in the past or is likely to do so in the future.

There is also much scope for the improvement of survey practice and the convergence of household expenditure surveys toward international best-practice standards. The experimentation with recall periods by the Indian National Sample Survey, although it has sown some confusion in the short run, is an excellent example of the sort of methodological work that is needed and which statistical offices are usually reluctant to perform. If nothing else, these experiments have highlighted the extreme sensitivity (175 million people less in poverty, the headcount ratio reduced by half) of poverty measurement to what many would previously have regarded as a technical, if not downright esoteric, issue of survey design.

Here, too, is an important research program. We need to understand much better than we do now how the measurement of consumption is affected by the design of consumer expenditure surveys—not only reference periods but such issues as the detail of the questionnaire, the need for bounding interviews and repeat visits, multiple versus single respondents, whether ownership information on durables is worth collecting, and so forth. Although the World Bank's CIMS has done much to normalize the use of enriched expenditure surveys in many countries, it has done little experimentation and has not greatly contributed to settling design questions (see Deaton and Grosh 2000).

Another important research issue is finding better ways to set the poverty line. Within the standard realm of income (consumption) poverty, there is much to be said for giving greater weight to people's own reports of consumption adequacy. Ravallion has done sterling work in this direction and there is widespread interest around the world in these methodologies (Ravallion 1998; Pradhan and Ravallion 2000). In the United States, for example, the Gallup poll questions people on how much it would take for a family like theirs to "just get by" in their community. The responses yield sensible poverty lines, at least once outliers have been removed. People can also be asked to self-assess their poverty, as in Ravallion's work. If such schemes could be put on a sound footing, they raise the tantalizing prospect of reducing our dependence on extremely expensive, time-consuming, and most likely inaccurate consumption surveys. It is a good deal cheaper to ask one or two questions than 200 or 300.

Recent work in South Africa shows strong correlation between measures of financial, physical, and mental health (Case and Wilson 2000). But there are real concerns about "adaptation." People do not perceive themselves to be better off, even in the face of large increases in real income, essentially because they adapt their expect-

tations to their circumstances. For example, Easterlin's (1995) famous finding shows that subjective levels of happiness did not increase among Japanese consumers between 1958 and 1987 in spite of a fivefold increase in real per capita income. The World Bank might find itself successfully promoting growth with a concomitant reduction in income poverty, but then being negatively assessed because people do not report themselves to be better off. The Washington Consensus derailed by the hedonic treadmill!

Those of us who have been exposed to field experience have been impressed by the prominence of health concerns in what people tell you about their poverty. Income, housing, and jobs tend to predominate when health is normal, but if someone gets sick, is hit by a car, or has a friend or relative who has been raped or murdered, income poverty recedes into the background in people's perceptions. Many millions of people around the world will die from AIDS, with untold misery and deprivation. It would be a terrible thing if the World Bank dealt with this crisis only in terms of its effects on income poverty, which it already shows some signs of doing (PovertyNet 2000).

My view is that the World Bank should back away from its current too-concentrated focus on income headcount numbers. It should emphasize a much wider range of other measures, focusing on deprivations that may be more important than deprivation of income.

Note

Angus Deaton is with the Research Program in Development Studies at Princeton University. This article was written at the suggestion of Nick Stern, to whom the author is grateful for preliminary discussions. Martin Ravallion clarified many issues and commented extensively on an earlier draft. The author is grateful to him, as well as to Richard Adams, Misha Belkindas, Surjit Bhalla, Anne Case, Paul Collier, Jean Drèze, Valerie Kozel, Peter Lanjouw, Karen Mason, Salvatore Schiavo-Campo, Eric Swanson, Pravin Visaria, and Mike Walton for helpful discussions and suggestions. Angus Deaton can be reached via e-mail at deaton@princeton.edu.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Bardhan, Pranab K. 1974. "The Pattern of Income Distribution in India: A Review." In T. N. Srinivasan and P.K. Bardhan, eds., *Poverty and Income Distribution in India*. Calcutta: Statistical Publishing Society, 103–38.

Bhalla, Surjit S. 2000a. "Growth and Poverty in India: Myth and Reality." New Delhi (June). Processed.

———. 2000b. "Trends in World Poverty: Ideology and Research." New Delhi (June). Processed.

Case, Anne, and Francis Wilson. 2000. "Health and Wellbeing in South Africa: Evidence from the Langeberg Survey." Research Program in Development Studies, Princeton, N.J. (December). Processed.

- Chen, Shaohua, and Martin Ravallion. 2000. "How Did the World's Poorest Fare in the 1990s?" World Bank, Washington, D.C. Processed.
- Deaton, Angus. 2000. "Preliminary Notes on Reporting Periods in the Indian NSS, 52nd through 54th Rounds." Research Program in Development Studies, Princeton, N.J.
- Deaton, Angus, and Margaret Grosh. 2000. "Consumption." In Margaret Grosh and Paul Glewwe, eds., *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study*, vol. 1. Washington, D.C.: World Bank.
- Deaton, Angus, and Alessandro Tarozzi. 2000. "Prices and Poverty in India." Princeton University, Program in Development Studies, Princeton, N.J. (July). Available on the World Wide Web at www.wss.princeton.edu:80/~rpds/working.htm.
- Easterlin, Richard A. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* 27:35–47.
- Filmer, Deon, and Iant Pritchett. 1998. "Estimating Wealth Effects without Expenditure Data—Or Tears: With an Application to Educational Enrollments in States of India." World Bank Policy Research Working Paper 1994, Washington, D.C. (October). Processed.
- Government of India. 1993. *Report of the Expert Group on the Estimation of the Proportion and Number of Poor*. Delhi: Planning Commission.
- . 2000. *Choice of Reference Period for Consumption Data*. National Sample Survey Organization, Calcutta, India.
- Groves, Robert M., and Mick P. Couper. 1998. *Nonresponse in Household Surveys*. New York: Wiley.
- Mahalanobis, P. C., and S. B. Sen. 1954. "On Some Aspects of the Indian National Sample Survey." *Bulletin of the International Statistical Institute* 34.
- Mukherjee, M., and G. S. Chatterjee. 1974. "On the Validity of NSS Estimates of Consumption Expenditure." In T. N. Srinivasan and P. K. Bardhan, eds., *Poverty and Income Distribution in India*. Calcutta: Statistical Publishing Society. 139–47.
- National Research Council. 1995. *Measuring Poverty: A New Approach*. Washington, D.C.: National Academies Press.
- PovertyNet. 2000. Newsletter #21 (July). Available on the World Wide Web at www.worldbank.org/poverty/news/jul00.htm#4.
- Pradhan, Menno, and Martin Ravallion. 2000. "Measuring Poverty Using Qualitative Perceptions of Welfare." *Review of Economics and Statistics* 82(3):462–71.
- Ravallion, Martin. 1998. "Poverty Lines in Theory and Practice." ISMS Working Paper No. 133, World Bank, Washington, D.C.
- . 2000a. "Should Poverty Measures Be Anchored to the National Accounts?" World Bank, Washington, D.C. (July). Processed.
- . 2000b. "Do National Accounts Provide Unbiased Estimates of Survey Based Estimates of Living Standards?" World Bank, Washington, D.C. (August). Processed.
- Ravallion, Martin, and Benu Bidani. 1994. "How Robust Is a Poverty Profile?" *World Bank Economic Review* 8(1):75–102.
- Ravallion, Martin, and Binayak Sen. 1996. "Monitoring Poverty in Bangladesh." *Economic Development and Cultural Change* 44(4):761–92.
- Scott, Christopher, and Ben Amenuvegbe. 1990. "Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana." Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper No. 6. World Bank, Washington, D.C.
- Srinivasan, T. N. 2000. "Growth and Poverty Alleviation: Lessons from Development Experience." Department of Economics, Yale University, New Haven, Conn. (November). Processed.

- Srinivasan, T. N., P. N. Radhakrishnan, and A. Vaidyanathan. 1974. "Data on Distribution of Consumption Expenditure in India: An Evaluation." In T. N. Srinivasan and P. K. Bardhan, eds., *Poverty and Income Distribution in India*. Calcutta: Statistical Publishing Society, 148–62.
- Stiglitz, Joseph. 1997. "How to Fix the Asian Economies." *New York Times*, October 31, sec. A, p. 27, col. 2.
- Thomas, D., K. Beegle, E. Frankenberg, and G. Teruel. 1999. "Household Budgets, Household Composition and the Crisis in Indonesia: Evidence from Longitudinal Household Survey Data." RAND Corporation, Santa Monica, Calif. Processed.
- Triplett, Jack E. 1997. "Measuring Consumption: The Post-1973 Slowdown and the Research Issues." *Federal Reserve Bank of St Louis Review* May/June:9–42.
- Visaria, Pravin. 2000. "Poverty in India during 1994–98: Alternative Estimates." Institute for Economic Growth, New Delhi (June). Processed.
- World Bank. 1997. *World Development Report 1997*. Washington, DC: World Bank.

Comment on “Counting the World’s Poor,” by Angus Deaton

Martin Ravallion

There is almost never just one way to measure an economic variable, and poverty is no exception. Judgments are required about the best method—given the data available—at virtually every step, starting with the design and processing of the underlying household survey data. The judgments made should be explicit and subjected to independent review.

In that spirit, Deaton (this issue) provides a usefully critical discussion of the current methods used by the World Bank for measuring poverty. This is welcome and timely. It is welcome in part because it is by Angus Deaton, who for 20 years has been a leading scholar on the theory and methods of welfare measurement using household-level data. And it is timely because of recent concerns about whether our current measurement tools fully reflect how the living conditions of the world’s poor are changing.

I will not address all the points raised by Deaton—avoiding those on which I think there is broad agreement that the World Bank’s current methods could be improved. These include the way that (invariably troublesome) income surveys are handled, the scope for better use of subjective-qualitative welfare data, and the need to better acknowledge the limitations of welfare metrics based solely on consumption of market goods.

This leaves two main issues that are very important for global poverty measurements—namely, how currency conversions are done and whether the poverty measures should be anchored to the national accounts. On both counts, Deaton makes recommendations for changing current methods. Before considering his recommendations, I will briefly describe how the World Bank measures poverty.

How Does the World Bank Measure Poverty?

The World Bank’s poverty measures based on the \$1/day line began with the 1990 *World Development Report* (World Bank 1990). That report provided estimates for developing countries as a whole and by region for 1985, based on 22 household

surveys, one for each of 22 countries, and model-based extrapolations for other countries (Ravallion and others 1991). Since then, the data set has expanded to include nearly 300 surveys spanning about 90 countries, with an attempt to provide consistent comparisons over time. The latest version of the main summary table by region gives estimates for 1987, 1990, 1993, 1996, and 1998 (Chen and Ravallion 2000).

The essential features of the current methodology can be summarized as follows.¹ An international poverty line is chosen as being typical of prevailing poverty lines used in a sample of low-income countries (World Bank 1990; Ravallion and others 1991).² The international poverty line is converted to local currency using the latest available purchasing power parity (PPP) exchange rates for consumption. The best available consumer price indexes are then used to convert the international poverty line in local currency to prices prevailing at the time of the surveys. These poverty lines are applied to distributions of consumption per person (or income if consumption is not available) constructed from nationally representative household surveys. Adjustments to the data are often required for consistency, such as ensuring that population weights are used to obtain an unbiased estimate of the individual distribution of household consumption per person. Calculations are done from the primary data (either micro data or appropriate tabulations).

These aggregate (global or regional) poverty measures based on international poverty lines should not be confused with measures based on national or subnational lines. The vast bulk of the World Bank's poverty monitoring effort does not use PPP exchange rates at all. At the country level, poverty monitoring is based on poverty lines considered appropriate in each country, which are updated over time. The PPP rates are required only for forming regional or global aggregates, for which the judgment is made that any two people with the same command over commodities should be treated the same way no matter where they live. Thus the aggregates ignore individual country poverty lines. The measures based on international poverty lines are presented side by side with national poverty lines in *World Development Indicators* (World Bank 2000) and the Bank's Global Poverty Monitoring Web site, which gives the individual country estimates over time for a range of poverty lines and measures (www.worldbank.org/research/povmonitor/).

Currency Conversions for International Poverty Lines

The practice of using PPP exchange rates in forming global poverty aggregates goes back to the first *World Development Report* on poverty (World Bank 1980). This method is clearly preferable to using official exchange rates, but it has long been recognized that using PPP exchange rates is not ideal. Since the 1990 *World Development Report*, the consumption PPP rates have been used, rather than PPPs for aggregate

output. This avoids some of the problems noted by Deaton. However, PPP exchange rates were not designed for the purpose of making internationally comparable poverty lines, but rather for making comparisons of average national income and consumption. The problem is coming up with something better.

Deaton is concerned about the impact of revisions to the source data on PPP rates for the global poverty measures. The PPP exchange rates have been updated over time, with the latest using the 1993 price data from the UN International Comparisons Project, entailing considerable improvement in the coverage and quality of the underlying data on prices. The underlying distributional data have also been updated regularly. With new data, all preceding calculations are also corrected back in time to ensure internal consistency.

The World Bank always uses the most recent data available for each update (three times a year). That includes all the new survey data (including going back in time) as well as new PPPs. Deaton is right that this produces some sizable revisions. For example, in the latest version of the main regional tabulation (Chen and Ravallion 2000), the share of world poverty attributed to Sub-Saharan Africa has risen relative to previous versions, due partly to revisions in the PPP exchange rates for this region. These revisions have also reduced the global rate of poverty reduction because poverty has increased in Sub-Saharan Africa. However, there has not been a downward revision of the Bank's overall \$1/day head count of poverty for the most recent common year between the latest estimates and the previous estimates, namely, 1993. For that year, the new estimate is 1.31 billion people living below \$1/day (Chen and Ravallion 2000); previously, we estimated 1.30 billion (Ravallion and Chen 1997). However, the regional composition of poverty has changed, with a higher proportion of the poor in Sub-Saharan Africa (reflecting the new data on price levels as well as new survey data for relatively poorer countries).

It is hard to argue that we should ignore new data on either the distributions or the PPP exchange rates. This is one worry about Deaton's proposal of not updating PPPs; it could create persistent drift from accepted international comparisons. I am also worried about the judgments for which Deaton's proposal calls, in making sure that the PPP lines are "plausible." The problem is knowing what is plausible. The researcher might be unduly persuaded by dubious priors—raising concerns about the transparency and objectivity of the measures.

Disparities between Survey Data and the National Accounts

There are two main sources of data on people's average command over commodities. One is the level of private consumption expenditure per capita from the national accounts (NAS). The other is average household expenditure on consumption obtained from a national sample survey of households—typically the same survey used

to measure poverty. Unlike the NAS, surveys also tell us about the distribution of consumption and/or income.

There are differences between the two data sources in terms of coverage, definitions, and methods. For example, as it is measured in practice in most developing countries, "private consumption" in the NAS includes spending by nonprofit enterprises (charities, nongovernmental organizations, and political parties) as well as households. In some countries (such as India), the survey-based consumption measure does not include imputed rents for owner-occupied housing. There are also measurement errors in both data sources. Survey measures entail both sampling and nonsampling errors; underreporting of income and consumption is thought to be a common problem in surveys. There are comparability problems between surveys (between countries and over time), such as due to differences in questionnaire design. NAS consumption is also prone to error, given that it is not directly measured in the NAS but is often inferred residually after accounting for more directly measured channels of output absorption at the commodity level.

It is not surprising that the two data sources do not agree in general. The discrepancy with the NAS is significantly greater for income surveys than consumption surveys (Ravallion 2001). This echoes the long-standing concerns (on both theoretical and practical grounds) about using income surveys for measuring poverty to which Deaton refers. Consumption surveys give numbers that are about as close as one might expect to private consumption in the NAS, on average. But here, too, there are some sizable deviations from the average. For example, for India aggregate consumption in the National Sample Survey (NSS) is about 60–70 percent of consumption in the NAS during the 1990s (depending on the method used to measure consumption in the NAS). Across countries in the 1990s, growth rates in mean consumption from surveys have tended to be lower than those from the NAS. They are only slightly lower as a rule, but more markedly so in some cases, including India and much of Eastern Europe and Central Asia (where there is remarkably little correlation between data from the two sources; see Ravallion 2001).

The discrepancies between consumption estimates from these two sources for India have fuelled a debate about whether the NSS is capturing the full gains to the poor from the country's postreform economic growth in the 1990s. It seems that, in some kind of contagion, these concerns have spilled over to the rest of the developing world. One observer even claimed that the World Bank measured poverty the way it did—based on surveys for measuring the levels of household consumption and income—because it was in the institution's interests to deliberately understate the extent to which economic growth was reducing poverty in the world (Bhalla 2000). As Deaton notes, although it is a current debate in India, this is not a new debate there or elsewhere; for example, the same issue has been raised before in the contexts of concerns about underreporting of income in the (many) income surveys used for poverty measurement in Latin America. The recent perception in some quarters that something

is seriously wrong in the existing poverty monitoring efforts makes Deaton's assessment even more welcome.

In some respects, the fact that this debate has resurfaced in India might be surprising. India had not previously been seen as a country with problematic survey data. Indeed, India's poverty monitoring based on the NSS has been widely seen as a model for other countries (although this perception has faded somewhat over time, as survey methods have advanced elsewhere). We do not have any new information about the survey's design and implementation to make us suspect that something has gone wrong with the NSS in the 1990s, at least prior to the latest (55th) round for 1999–2000. And expectations for rapid poverty reduction, postreform, may well have been optimistic. There are reasons why the type of growth seen in the 1990s, given initial conditions in much of the country, would not yet be evident as greatly reduced poverty (Ravallion 2000b).

It should also be noted that the extent of divergence between the two data sources depends in part on measurement choices. Applying more consistent definitions of "consumption" between the two sources can go some way toward reducing the gap; see, for example, Slesnick (1998) on the United States. Two examples from India illustrate the dependence on methodology.

The first example concerns the methods used by the NAS. The appearance of sizable divergence between NSS and NAS consumption aggregates for India in the 1990s can be deceptive unless one takes account of the changes in methods used by the NAS. When the central statistics office changed its methods to accord more closely with new international standards set in 1993, there was a sizable upward revision in the NAS consumption estimates in the 1990s. Some observers have calculated the ratio of aggregate consumption from India's NSS to that in the NAS ignoring the switch in NAS methods; this gives a deceptively large fall in the ratio in the 1990s. Comparing estimates using the same methods, Sen (2000) finds little sign of a decrease in the ratio of NSS consumption to NAS consumption during the 1990s, although there are signs of a longer-term trend decline in this ratio going back to the 1970s. The bulk of the decline precedes the 1990s and clearly has nothing to do with the reforms.

The second example reinforces Deaton's comments on how much survey design can matter. There were important changes in the consumption questions in the 1999–2000 (55th) round of the NSS, breaking from methods used consistently since the 1950s. The 55th round shows a sizable reduction in poverty in India. However, we do not know how much of this is real and how much is due to the changes in survey design. By one assessment, the conclusion that poverty fell is not robust to attempts to deal with the likely comparability problems raised by the 55th round (Sen 2000). Other assessments suggest that poverty has indeed fallen, although not nearly as much as the 55th round implies. Nor has it been demonstrated that the changes made to the design of the NSS in the 55th round will give more precise consumption

estimates; that is not obvious on a priori grounds. Further experiments on alternative survey methods are needed, as well as a plan for data transition to allow consistent comparisons over time.

Implications for Poverty Measurement

Discrepancies between the surveys and the national accounts have implications for poverty measurement that depend critically on the reasons for the divergence. Deaton is right that further research on this issue is called for. One interpretation assumes that any discrepancy is entirely due to underestimation of mean consumption in the surveys, but that the surveys nonetheless get relative consumption and hence inequality right. If one accepts that interpretation, then the solution is simple: just replace the mean from the surveys with the consumption mean from the NAS and recalculate the poverty measures (as in, for example, Bhalla 2000). A more sophisticated version of the same basic method takes the mean consumption or income components in the surveys and scales them up to accord with the NAS. Although agreeing that the surveys probably miss some of the consumption gains, an alternative interpretation says that the problem has much more to do with consumption by the nonpoor, and that the surveys more likely underestimate the rate of increase in inequality.

It has not been established, and is quite unlikely from what we know, that the discrepancy between the two data sources is entirely due to underestimation of consumption levels in the surveys. But even if that were the sole source of the discrepancy, it is hard to believe that the surveys would be getting the mean wrong but inequality right. Our limited knowledge of the problems of underreporting and noncompliance in consumption and income surveys does not suggest that the problem is likely to be distribution neutral. In consumption surveys, the bulk of the problem is thought to be for the nonpoor. (In income surveys, there is possibly also underestimation of annual income of the rural poor using survey questions based on subannual income recall, given seasonality.) If the problem arises from underestimation of consumption by the nonpoor, then the correct mean for measuring poverty is the mean of the survey distribution, although this underestimates the true mean (Ravallion 2000a). In other words, for measuring poverty, there is an error in the mean that offsets the error in estimating the distribution around the mean. However, it is unlikely that none of the discrepancy arises from underestimation of consumption at the poverty line. At least some of the error is probably passed onto the poverty measure at any given poverty line.

What can be done? If one accepts that some of the discrepancy is due to underestimation of consumption at the poverty line, then in theory there must exist a scaling factor applied to the survey mean that will give the right poverty measure at the given (but also wrong) distribution. Although we can postulate its existence in theory,

we have no basis for setting this scaling factor. Nor is it plausible that the error is a strictly decreasing function of the scaling factor; there will be scaling factors greater than unity that give less accurate measures than using no scaling at all.

Deaton recommends taking a weighted average of the two measures of consumption, one from the survey and one from the NAS. (Then the scaling factor is a linear function of the ratio of NAS consumption to survey mean consumption.) It is not clear why this would be better than no adjustment. Combining the two measures of average consumption may well give a better estimate of the mean. But that is not the statistic in which we are interested. The problem is how to set that weight when the aim is to reduce the error in the poverty measure. As yet, there is no obviously defensible solution in sight, and there is no basis for presuming that anyone's guess will give a better measure than using no adjustment at all.

There is also a question of whether anchoring the poverty measures to the national accounts would foster better poverty monitoring in the future. The World Bank's practices in poverty measurement reflect a judgment that the household surveys should ideally stand as a data source largely independent from the NAS. In this view, the surveys should be the main testing ground for claims made about the impacts of policy or economic changes on the poor.

Yes, there are problems galore in the existing surveys. But, as Deaton notes, the World Bank also has an important role to play in supporting efforts at the country level to improve and standardize the surveys, such as through the ongoing Living Standards Measurement Study and related support in promoting household survey capabilities in developing countries.

Notes

Martin Ravallion is research manager with the Development Research Group at the World Bank.

1. The methods are described in more detail in Ravallion and others (1991); Chen and others (1994); Ravallion and Chen (1997); and Chen and Ravallion (2000).

2. The countries include India, which had been the country on which international poverty comparisons were based for the first *World Development Report* on poverty, namely, World Bank (1980). The \$1/day line is close to the mean for a larger set of low-income countries. The line is updated with revisions to the purchasing power parity exchange rates (as discussed in Chen and Ravallion 2000).

References

- Bhalla, Surjit. 2000. "Trends in World Poverty: Research and Ideology." Available online at www.oxusresearch.com.
- Chen, Shaohua, Guarav Datt, and Martin Ravallion. 1994. "Is Poverty Increasing or Decreasing in the Developing World?" *Review of Income and Wealth* 40:359–76.

- Chen, Shaohua, and Martin Ravallion. 2000. "How Did the World's Poorest Fare in the 1990s?" Policy Research Working Paper 2409, World Bank, Washington, D.C.
- Ravallion, Martin. 2000a. "Should Poverty Measures Be Anchored to the National Accounts?" *Economic and Political Weekly* 34(3):245–52.
- . 2000b. "What Is Needed for a More Pro-poor Growth Process in India?" *Economic and Political Weekly* 35:1089–93.
- . 2001. "Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree?" Policy Research Working Paper 2665, World Bank, Washington, D.C. Available online at <http://econ.worldbank.org>.
- Ravallion, Martin, and Shaohua Chen. 1997. "What Can New Survey Data Tell Us about Recent Changes in Distribution and Poverty?" *World Bank Economic Review* 11(2):357–82.
- Ravallion, Martin, Gaurav Datt, and Dominique van de Walle. 1991. "Quantifying Absolute Poverty in the Developing World." *Review of Income and Wealth* 37:345–61.
- Sen, Abhijit. 2000. "Estimates of Consumer Expenditure and Its Distribution: Statistical Priorities after the NSS 55th Round." *Economic and Political Weekly* 35(51):4499–518.
- Slesnick, Daniel. 1998. "Are Our Data Relevant to the Theory? The Case of Aggregate Consumption." *Journal of Business and Economic Statistics* 16(1):52–61.
- World Bank. 1980. *World Development Report: Poverty and Human Development*. New York: Oxford University Press.
- . 1990. *World Development Report: Poverty*. New York: Oxford University Press.
- . 2000. *World Development Indicators*. Washington, D.C.: World Bank.

Comment on "Counting the World's Poor," by Angus Deaton

T. N. Srinivasan

In his very instructive article, Angus Deaton argues that for international institutions (for example, the World Bank) and the national governments of most poor countries (for example, India), reduction of poverty is or is professed to be the paramount objective of policy. As such, they require some overall yardstick of progress (or lack thereof). According to Deaton, there is a long history of studies of poverty mobilizing support among the nonpoor for antipoverty values. So it is important to know whether the world and national poverty counts are sound enough to support their uses.

Deaton's analysis of the problems with poverty counts and suggestions for improvement, including issues needing further research, are based on two distinct stages in counting the poor (see also Deaton 2000). At the first or international stage, a world poverty line is set and used to derive comparable poverty lines for each country. At the second or domestic stage, the poverty lines are used to count the number of poor people in each country, and the others are added up over countries. He finds disquieting evidence about both stages of counting. The data for poverty counts in the second stage come from household surveys, whereas data on aggregate economic growth are from National Accounts Statistics (NAS). Deaton finds that in many countries there are large and growing disparities between survey data and national accounts so that there is no consistent empirical basis for conclusions about the extent to which growth reduces poverty. It is scandalous that even after nearly half a century of pursuing national and international programs for the eradication of mass poverty, the empirical foundations for assessing the success or failure of the programs and drawing lessons from them are so weak as to be deemed nonexistent. In what follows, I reinforce many of Deaton's arguments.

Deaton's discussion of global counts of the poor presumes that the World Bank and other international development institutions will continue to produce them anyway. He then asks how best to produce them and suggests improvements to current procedures, while concluding that the World Bank and others should in fact back away from such counts. I go further. In my view, global counts have little meaning and even less policy relevance. Abandoning them and focusing on na-

tional and subnational poverty analysis that goes beyond headcounts would be the sensible course to follow. Like Deaton, I draw heavily on the experience of India's statistical system. I focus only on consumption-based poverty lines. The reason is the challenge of defining household income in a theoretically satisfactory manner and collecting data on income based on that definition through household surveys in any country (developed or developing). Deaton (1989) discusses the difficulties in meeting the challenge. Poverty counts based on income-based poverty lines are even more problematic than consumption-based ones.

Unavoidable Arbitrariness in Defining Poverty Lines

The criteria by which the incidence of poverty in any society at any point in time is to be judged are essentially arbitrary. Adam Smith long ago argued that in defining necessities, one must include not just "the commodities which are indispensably necessary for the support of life" but also those which "the custom of the country renders indecent for creditable people, even of the lowest order to be without" (Smith 1937:821–22). To be nonpoor, a person must be able to afford at least the necessities. Following Smith, one should include a "decency" component in them. Obviously, these are subjective and would vary over time and space.

A claim of objectivity and nonarbitrariness is often made for a particular poverty line (that is, a value of consumption expenditure or income per person) by linking it to the required habitual or long-term intake of food (or, more precisely, its energy content) for an individual to be adequately nourished. The use of such a poverty line would identify being poor with not being adequately nourished.¹

Such a claim appears untenable and its use in poverty counts based on household surveys questionable. First, the long-term energy requirement for adequate nutrition of an individual of a given age and gender and performing specified tasks is not a fixed number of kilocalories per day. Second, household surveys collect data on *expenditures* on food and quantities consumed of various items of uncooked and cooked foods by the *household* over a short reference period. Even ignoring measurement errors in translating such data into their energy content, the resulting number at best represents the energy content of food consumed by the household as a whole. It is not possible to deduce *the long-term energy intake of each person in the household from the total energy intake of the household over a short reference period*.

Third, in surveys of the consumption expenditure of households in poor countries, such as India's National Sample Survey (NSS), the estimated energy intakes of very poor households are often found to be too low for survival and those of rich households too high to be consistent with good health. The reason for this in part could be that not all the meals that individuals from poor households consume outside their

homes are included in their food consumption. In addition, some of the food consumed by nonhousehold members is included in the food consumption of rich households. Agricultural workers from poor households often eat meals provided by their employers at the fields where they work. Although the survey enumerators were instructed to ask each household about such meals and include them in (or exclude from) the food consumption of that household, depending on the circumstances, these instructions are unlikely to have been followed fully and meticulously in practice. For these reasons, estimates based on an ostensibly nutrition-based poverty line, and survey-based household consumption are poor measures of the extent of undernutrition and poverty (see Srinivasan 1992).

Deaton rightly recognizes that dimensions of poverty other than income or consumption are as important or more important, particularly deprivations in health, education, and democratic rights, and that there are difficult measurement issues associated with all of these dimensions. The UNDP (1996, 1997) estimates a “capability poverty measure” based on Amartya Sen’s concepts of “capabilities and functionings” and a “human poverty index,” a weighted composite of several nonincome indicators including those incorporated in the capability measure. These measures do not adequately recognize and account for the difficult conceptual and measurement problems that Deaton highlights.

The best and easily understood starting point for deriving a consumption-based poverty line is a “poverty consumption” bundle of goods and services for a representative (in size and age-gender composition) household. It is common to assume that part of the bundle would be provided free or at subsidized prices by the state. Valuing the private component of the bundle at appropriate prices yields the poverty line. This valuation of the private component is in effect what an expert group did for India’s Planning Commission in 1962, in defining poverty lines for rural and urban households in India. There is unavoidable arbitrariness in determining which goods and services (and in what amounts) are to be included in a poverty bundle. Nevertheless, given a poverty bundle for a representative household, appropriate adjustments for differences of any other household in its size and age-gender composition could be made to arrive at a household-specific poverty line. A household would be deemed to be poor if it does not have the resources, measured in terms of either income or total consumption expenditure, to buy the private component of *its poverty bundle at the prices it faces*. In such a definition, a household that can afford to but does not buy its poverty bundle is not deemed poor.

Clearly, if an annual survey collects data on the resources that each household commands and the prices it faces, it is a straightforward matter to estimate the number of the poor. As long as the constituents and the nonprivate component of the poverty bundle remain the same over time and space and surveys in different regions and time periods continue to collect household-specific data on resources and

prices, estimating the poor in each region and time period continues to be relatively uncomplicated.

The price data collected in different household expenditure surveys differ in their coverage, completeness, and representativeness. In some, only expenditure data are collected. In others, total expenditures on and quantities purchased of each commodity are collected so that unit values could be computed. In still others, prices paid are also collected along with quantities purchased. But I know of no survey that collects data on prices actually paid by households in each of their transactions involving purchases included in their consumption. The Indian NSS does collect price data (although it is not transaction specific).

Deaton and Tarozzi (2000) use more than 7 million pieces of price (more precisely, unit value) information from two rounds (1987–88 and 1993–94) of NSS expenditure surveys. In his analysis of the unit value data from the NSS for the state of Maharashtra in 1983, Deaton (1997) found that they matched independently collected market price data.² Survey data confirm that households living in the same region and canvassed at the same time reported paying different prices for the same commodity even after allowance was made for possible differences in quality and other factors. Indeed, such differences raise serious questions about the common assumption in the analysis of household surveys that households purchase homogeneous commodities in competitive markets in which the “law of one price” holds. Apart from such interhousehold variation in prices, there are spatial and intertemporal variations as well. Also, there could be (and often are) interhousehold, interregional, and intertemporal variations in access to the state-provided component of the poverty bundle. I am unaware of any poverty count that allows for these variations.

Clearly, it is impractical to update poverty lines through revaluation of a *given* poverty bundle at prices that are specific to each household, region, and period of time. As such, a common practice is to use some price index to adjust some poverty line (not necessarily one derived from valuing a poverty bundle) at base year (or base region) prices to arrive at a poverty line for a different year (or region). The fact that poor (rural) households face different prices compared with nonpoor (urban) households could be taken into account in such an approach by using poor-specific rural and urban price indexes to update poverty lines (or alternatively to deflate consumption expenditures). For example, in India, because a large proportion of rural poor are believed to be landless agricultural laborers, the consumer price index for agricultural laborers has been used for updating the rural poverty line in official estimates of poverty. A simple average of the consumer price index for industrial workers and that for urban nonmanual workers is used to update the poverty line.³ However, commodity weights used in constructing these indexes are outdated, and the price quotations used are not representative of the relevant transactions. Using commodity weights and unit values based on the household surveys, Deaton and Tarozzi

(2000:34–35) recompute the poverty estimates. They find that, in contrast to the diverging trends in the official estimates, “between 1987–88 and 1993–94, there was no great difference in the rate of decline of urban and rural poverty.” Clearly, the choice of price index matters. But whatever index is used, as Deaton points out, the basic, standard textbook index number problem remains and cannot be wished away.

I have thus far addressed problems encountered in Deaton’s “domestic” stage. Moving to his “international” stage, a poverty bundle common to all regions within a geographically and culturally diverse country such as India, let alone for all countries of the world, cannot be meaningfully defined. (I will return to this issue in the concluding section.) If such a bundle could be defined, then the national poverty line at any point in time would be the value of that bundle at the prices in *local currency* that households face in that nation at that point in time. There is no need for any exchange rate in such a calculation. Deaton is absolutely right in arguing that, because such an internationally accepted bundle does not exist, it does not make sense to simply convert \$1/day to local currency values using purchasing power parity (PPP) exchange rates with commodities weighted by their shares in the consumption of the poor. The reason is that doing so makes poverty lines move around with changes in PPP exchange rates arising from world market price changes that have no relevance to the poor. For example, the poverty line for one country would be shifted by a change in the world price of a commodity that is not consumed by the poor in that country but consumed by the poor in some other country, because such a price change affects the PPP exchange rate. In any case, global poverty counts are based on neither a common global poverty bundle nor conversions to local currency values using PPP exchange rates with commodity weights more relevant to the poor.

As Deaton points out, an international poverty line for base year 1985 was chosen (\$1/day at 1985 PPP dollars) as being representative of poverty lines in use in low-income countries. In making this choice, poverty lines in local currency (consumption expenditures per person per day) in use in 1985 were presumably converted to U.S. dollar terms using the then-available PPP exchange rate for each currency. Because these were apparently found to cluster around one dollar, \$1/day in constant 1985 PPP dollars was seen as representative of the poverty lines then in use.

It should be obvious that even assuming that local-currency poverty lines in 1985 represented the value of a national poverty bundle, it cannot be claimed that the \$1/day at 1985 PPP dollars poverty line is representative of national poverty lines even in that base year. Moreover, as Deaton documents persuasively, revisions of the PPP rate to reflect, on one hand, better and more accurate information, but on the other hand, the change in base year play havoc with the poverty counts. Deaton’s critique is more than enough to persuade any serious analyst that these poverty counts are virtually meaningless.

Poverty Counts and Policymaking

There is no denying that if eradication of mass poverty is the objective of policy, indicators of progress toward achieving the objective are needed. It is well understood that poverty is a multifaceted phenomenon in that being poor does not simply mean not having adequate material resources at one's disposal. It also means not having a voice in making decisions that affect one's life. That is, it means not having as much access as one is entitled to state provision of goods and services, including those of the administrative, educational, healthcare, and legal systems.

Because each of these dimensions captures a different aspect of poverty, a separate indicator is needed. By combining indicators capturing disparate dimensions into a single composite index, such as the UNDP's Human Development Index, one loses some of the information contained in the separate indicators. The extent of the loss depends on the independence of the different dimensions in describing the phenomenon of poverty. If they are not independent at all—in the sense of each indicator being perfectly correlated with every other—then combining them linearly into a composite indicator is not necessary. In fact, any one of the indicators would do just as well in describing poverty. But in this case there is no loss of information either because, given one indicator, the others do not add information. Perfect correlation is very unlikely and information loss in using a composite index is likely to be substantial, so it is most useful to put together several indicators of poverty, one of which would be based on consumption. Having said this, I should hasten to add that conceptual and measurement problems with other indicators are likely to be just as serious, if not more so, as those associated with consumption poverty. In particular, problems in defining and measuring analogues of \$1/day global poverty counts with respect of some consumption dimensions of poverty would be far more daunting.

Poverty indicators, whether multiple or just a single composite, serve three distinct purposes. The first is simply to depict the prevailing situation at one or more points of a time in each country or region and perhaps the world. Such a depiction provides a yardstick or scorecard for the performance of national governments and international agencies, such as the World Bank, in achieving their professed objective of poverty reduction. More important, it is a prelude to a positive analysis of likely determinants of different dimensions of poverty. Clearly, aggregations either of space (for example, subnational regions with distinct characteristics are aggregated into a nation) or of time (for example, aggregation over a long time involving major systematic changes) would most likely mask the influence of different determinants of poverty. The reason is that the relationship between determinants and poverty outcomes almost surely would be very different across the units being aggregated. It is most unlikely that global indicators, such as \$1/day counts of the poor, would be the starting point for any useful analysis of determinants of poverty. However, national

and (even better) subnational poverty counts, even the ones not based on a fixed poverty bundle, are much better starting points for analysis.

The second purpose is the normative use of poverty indicators for policy formulation. For this purpose, certainly global indicators and even national indicators (in large, populous countries, such as India, where subnational entities are diverse in many dimensions) are not very useful. The reason is simple: A policy is likely to be most effective in reaching its targets and achieving its objectives if its locus is where the targets happen to be. Because most policies targeted at the poor are in the domain of subnational (or even lower) units, poverty indicators at higher levels are not helpful in policy formulation. This is not to say that national and international policies are irrelevant, only that their effects on aggregate indicators are best understood through the aggregation of their effects on indicators of poverty at lower levels.

There is a different and difficult problem in the normative use of poverty indicators at any level. Unless one has a framework that describes the mechanisms through which policies affect the determinants of poverty in all its dimensions and their quantitative significance, poverty alleviation policies cannot be well formulated. This is illustrated by the debate on whether more rapid aggregate growth, greater openness to international trade and investment, or more flexible labor markets have in the past or will in the future reduce poverty. Without a framework for analyzing the influence of these policies on the determinants of poverty, one cannot infer anything useful for policy formulation from any observed association (or lack thereof) between specific policies and greater reduction in poverty. This problem would remain even if there were some best way of counting the poor at any level of disaggregation.

The third purpose for a poverty indicator is its value for mobilizing support for aiding the poor and for policies (subnational, national, and international) that presumably could alleviate poverty. Even those who would readily concede that global poverty counts are useless for the first two purposes would think that they might serve the third purpose well. Certainly, saying that in 2000 so many millions of people in the world went to bed hungry or lived on less than \$1/day grabs attention.

Resolving to reduce that number by half by 2020 (or whatever future year) indicates a certain purposefulness for global policymaking. However, it is arguable whether such attention and resolve have concrete value in raising resources for poverty alleviation or changing policies. Long ago, exhortations based on appalling poverty in developing countries led to the target of 0.7 percent of gross domestic product for industrial countries to contribute for development assistance. Some of the richest countries of the world have yet to reach the target. Regardless of whether there would have been greater development if the development assistance target been reached, exhortations based on global poverty counts are unlikely to generate more resources from the rich for poverty alleviation.

Household Expenditure Surveys and National Accounts Statistics

Deaton's discussion of the problems with the data from household expenditure surveys (HES) and NAS is comprehensive, illuminating, and disturbing. My discussion of the sources of differences between the two will be brief and touch on a few important ones. Minhas (1988) provides a more comprehensive and deeper analysis. The HES in some developing countries were initially meant to provide data for commodity weights to be used in consumer price indexes. As such, surveys were canvassed relatively infrequently. In India, the HES component of the NSS from the start was meant to track changes in levels of living. Until 1973–74, the HES component was canvassed every year and again from 1983.

However, survey periods did not always cover a full year; even if they did, the year varied (most often it was July–June, but sometimes it was January–December or October–September). There were also differences in sample design across years, particularly when the main topic of inquiry differed from household consumption expenditure. In such years, the HES component canvassed far fewer households and sometimes with an abbreviated schedule of inquiries. Almost always the ultimate sampling units—namely, households—were drawn from each sample village and urban block after a complete enumeration of households in that village and block. As such, because the sample universe was not the population of households, the fact that population censuses were conducted only every ten years did not affect the sample frame.⁴ However, the number of strata into which the rural and urban areas were first divided and from which villages and blocks were drawn has varied over the years.

Since 1948–49, the NAS have covered the fiscal year April–March. But the basic data from which the NAS were put together have remained a hodgepodge. They range from actual revenues and expenditures based on audited accounts (for example, for governments and organized enterprises) to estimates based on sample surveys (for example, for yields of principal crops or output of some small-scale industries), price imputations, and ratios and norms that were themselves based on often outdated and unrepresentative samples. NAS estimates are known to be subject to error, the magnitude of which has not been assessed since the report of the first national income committee in 1949.

Private consumption expenditure estimates in the NAS are residuals derived from estimates of the domestic availability of each commodity (domestic output plus imports minus exports) left after deducting nonprivate consumption uses (investment including net additions to stocks as intermediate inputs and for consumption by the public sector). Clearly, errors in estimating each of these components and also in their valuation affect the estimate of private consumption expenditure. In addition, the private sector in the NAS includes households as well as nonprofit institutions, but the HES component of the NSS covers households only.

Estimates of consumption expenditure in each annual round of the NSS are based on expenditures during the reference period reported by sample households canvassed at each subround. The expenditures include the value of their purchases from the market and consumption of home-produced goods and services at imputed prices (that is, the relevant retail prices). If the information provided by households were free of errors of any kind, then the total expenditure would be, by definition, the sum over all transactions by the household of the quantities purchased from the market (or consumed out of home production) and the corresponding purchase (or imputed) prices in each transaction. In principle, every transaction of every household should be reflected in the NAS. However, even if the quantities involved in each transaction were covered, it would be likely that the prices used in the NAS to value them would differ from those paid by households.

The NAS, being annual, are deemed free of seasonal effects. In the NSS, an independent sample is canvassed in each subround, and all subrounds together cover the whole year. As such, in principle in the estimate of average annual consumption expenditure from all rounds together, any purely seasonal element in prices and quantities should be absent. But intrayear trend, if any, in prices and quantities will remain. However, in the distribution of monthly consumption expenditure among households in the sample, seasonal as well as trend effects would remain. In other words, the monthly expenditure reported by a household is the sum of the average consumption per month over the relevant year plus the seasonal effect and deviation from trend of that month. Thus the variance of the distribution of reported consumption is higher than that of the monthly average over the year. This means that a poverty count from the reported consumption will be higher than the true count based on average monthly consumption over the year if the poverty line is below the mode of the true distribution.⁵

The length of the reference period could influence the estimates of consumption. Too long a period may create a recall bias in the downward direction. Too short a period may create a bias of telescoping later purchases forward, a bias in the upward direction. Telescoping is more likely with respect to infrequently purchased items, such as appliances and expensive durables, where the act of purchase is remembered but not its precise timing. Because poor households are unlikely to purchase such items, even if nonpoor households telescope some of their purchases, telescoping to any serious extent would not bias poverty counts. In any case, biases of recall and of nonresponse could depend on the level of the aggregate consumption expenditures. In other words, the size of the bias may differ between poor and rich households. However, if the biases applied only to households above the poverty line, poverty counts would be unbiased.

Clearly, biases in the estimate ought to be distinguished from errors that affect the variance of the estimate. Sometimes a sample design is adopted that yields a consistent, although biased estimate of the population mean, but with a lower mean squared

error. The NSS sample design was meant to deliver an unbiased estimate of the aggregate mean consumption expenditure. From this perspective, a sample size that yields an unbiased estimate of mean consumption expenditure at the national level with given precision may be inadequate to yield estimates of adequate precision at subnational levels. The precision of the *sample mean* as an estimate of the *population mean* and the precision of the *sample cumulative distribution* as an estimate of the *population cumulative distribution* have to be distinguished. Most surveys are designed to deliver an estimate of the population mean with a desired precision. The precision of such surveys in estimating the population cumulative distribution, which is the basis for counting the poor in the population, may be very low. In other words, a survey that yields a reasonably precise estimate of the mean consumption expenditure of the population of households may yield a poverty count with a wide margin of error.⁶

Enough has been said to demonstrate that there are prior reasons why estimates based on household surveys would differ from the NAS, even if there were no deterioration in the quality of the system that conducts surveys and puts together the national accounts. The fact that the difference between NSS and NAS estimates in India appear to have increased since the 1960s (for a contrary view, see Sen 2000) is a cause for concern. However, there is no strong evidence in favor of one or the other estimate. As such, there is no convincing rationale for counting the poor through a simplistic procedure using a synthetic distribution with a mean expenditure based on the NAS and a household cumulative distribution derived from the NSS.

Conclusion and Recommendation

I agree with nearly all of Deaton's recommendations. There is an urgent need for a serious research program for reconciliation between NAS and HES data in a few countries, including India. There should be more experimentation in survey practice. In particular, research is needed for a better understanding of how measurement of consumption is affected by the design of surveys, including the length of reference periods, length and detail of questionnaires, length of interviews, repeat visits to the same household, and whether to have more than one respondent from each household.

Most important of all, Deaton's plea for finding better ways to set the poverty line is right on the mark. As should be evident from my comments, I prefer to start from a well-defined poverty bundle. Clearly, if it is to be time and space invariant, it has to be defined in terms of characteristics for healthy life and functioning, depending on an individual's age, gender, work activity, and other relevant attributes. But this is impractical. As an alternative, one may try to define a few poverty bundles in terms of goods and services. The need for more than one arises from the fact of variation in climates and dietary habits if nothing else. Given the poverty bundle appropriate to a subset of the population and well-designed surveys, it would be simple to define

poverty lines that are specific to that subset and time period, based on prices faced by the poor. But this alternative may not be that much more practical than defining a single global poverty bundle. There is no easy way of determining how many bundles would be needed to capture the variations in relevant dimensions. In any case, once there is more than one bundle and an associated poverty line based on it that is appropriate for each region or subset of the population, index number problems reappear if one attempts to construct a global poverty line that is representative of all the regional poverty lines.

Comparability and global representativeness are therefore impossible to achieve. It is clear that the \$1/day at a constant PPP exchange rate poverty line does not satisfy either. I would therefore prefer to abandon the search for the impossible and stick to national poverty lines. But if the politics of resource mobilization for poverty alleviation demand the use of poverty lines that sound comparable, even if in fact they are not, then the best solution is the one suggested by Deaton. It is to keep the current \$1/day poverty lines in local currency terms, adjust them only for local price inflation and abandon the use of PPP exchange rate adjustments in the future. Doing so would at least eliminate the egregious errors arising from the use of periodically revised PPP exchange rates.

Notes

T. N. Srinivasan is Samuel C. Park Jr. Professor of Economics at Yale University and Visiting Fellow, Center for Research on Economic Development and Policy Reform, Stanford University. The author thanks Angus Deaton for his extremely valuable comments on an earlier draft.

1. There are two approaches for deriving such a poverty line. In one, a “poverty bundle” of goods and services is defined to include items of food whose total energy content meets the nutritional requirement. The other is based on the (Engel) curve relating energy content of food consumed by each household to its total expenditures on all items of consumption. This curve, assuming it is upward sloping, shows the level of total expenditure at which the nutritional requirement of the household is met. It constitutes the nutrition-based poverty line. My criticism applies to both. The second approach is to derive *poverty lines that are specific to a region and period* from corresponding nutrition Engel curves. Strictly speaking, if this approach is used once only to derive a base-period poverty line for a region (which is adjusted using price indexes to derive poverty lines for other regions and periods), it ceases to be nutrition-based.

2. I thank Angus Deaton for this information.

3. Deaton pointed out that recent practice might differ from my description.

4. However, if villages and blocks had been chosen with probability proportional to the size of their population in the last decadal census before the date of the survey, then the “multiplier” for each sample village or block in the estimate of the consumption expenditures would not reflect the size of its population at the time of the survey. This could bias the estimate.

5. I thank Angus Deaton for this observation.

6. Deaton suggests the following example to illustrate this. Suppose households make purchases every fifth day for five days’ worth of consumption. *All* reference periods from one day to five days will deliver unbiased estimates of mean consumption. But, except for the reference period of five days, others will yield biased estimates of the headcount ratio.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Deaton, A. 1989. "Saving in Developing Countries: Theory and Review." *World Bank Economic Review*, 3(Supplement):61–96.

———. 1997. *The Analysis of Household Surveys*. Washington, D.C.: World Bank.

———. 2000. "Counting the World's Poor: Problems and Possible Solutions." Program in Development Studies of Economics, Princeton University, Princeton, N.J. Processed.

Deaton, A., and A. Tarozzi. 2000. "Prices and Poverty in India." Program in Development Studies, Princeton University, Princeton, N.J. Processed.

Minhas, B. S. 1988. "Validation of Large Scale Sample Survey Data: Case of NSS Estimates of Household Consumption Expenditure." *Sankhya* Series B 50(3; supplement):1–63.

Sen, A. 2000. "Estimates of Consumer Expenditure and Its Distribution." *Economic and Political Weekly*, December 16, 4499–518.

Smith, Adam. 1937. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Edwin Cannan. New York: Modern Library.

Srinivasan, T. N. 1992. "Undernutrition: Concepts, Measurements and Policy Implications." in S. R. Osmani, ed., *Nutrition and Poverty*. Oxford: Clarendon Press.

UNDP. 1996, 1997. *Human Development Report*. New York: United Nations Development Program.

Ecology, History, and Development: A Perspective from Rural Southeast Asia

Yujiro Hayami

The process by which different ecological conditions and historical trajectories interacted to create different social and cultural systems resulted in major differences in economic development performance within Southeast Asia. In the late 19th century, Indonesia, the Philippines, and Thailand commonly experienced vent-for-surplus development through exploitation of unused lands. Nevertheless, different agrarian structures were created. Indonesia's development was mainly based on the exploitation of tropical rain forest under Dutch colonialism. It resulted in the bifurcation of the rural sector between rice-farming peasant proprietors and large plantations for tropical export crops based on hired labor. In the Philippines, exploitation of the same resource base under Spanish rule resulted in pervasive landlessness among the rural population. Relatively homogeneous landowning peasants continued to dominate in Thailand, where delta plains that were suitable only for rice production formed the resource base for development. These different agrarian structures associated with different social value systems have accounted for differential development performance across the three economies in the recent three decades.

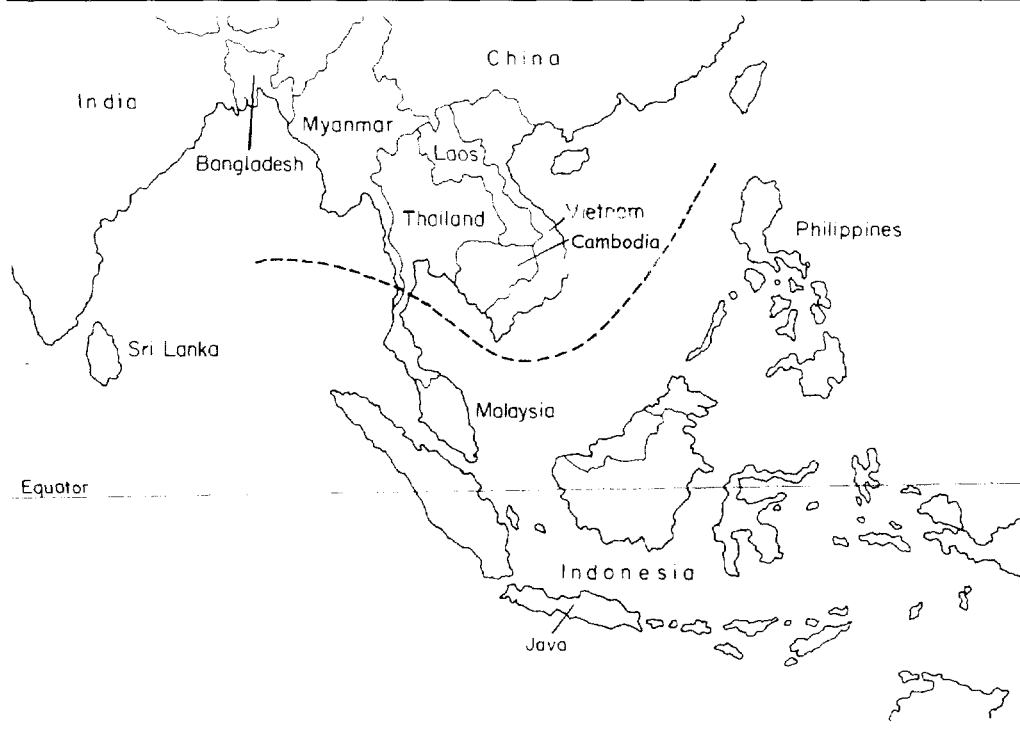
Ecology and history have a fundamental impact on the course of economic development. Nevertheless, few studies have investigated how different ecological conditions and historical trajectories have interacted to forge different social and cultural systems, resulting in major differences in development performance across economies. This article aims to shed light on this process, drawing on agricultural development experiences in Southeast Asia. Relatively high growth performance in agriculture in the past three or four decades has been counted as one of the factors underlying the "economic miracle" of this region (World Bank 1993). Yet economic performance has varied within the region. Variations in the recent agricultural growth performance reflect differences in the agrarian structure, which were created through distinct colonial regimes under diverse ecological environments. As such, this study does not aim to conduct research for predicting the future course of rural development in

Southeast Asia; rather, I intend to provide a broad telescopic guide to such research for this region as well as other regions.

Southeast Asia can be classified into two major ecological zones: the continental zone, including Thailand, Vietnam, and Myanmar, and the insular and peninsular zone (henceforth called the insular zone), including Indonesia, Malaysia, and the Philippines (see figure 1). Major river deltas characterize the continental zone, and tropical rain forests characterize the insular zone. Before the 1860s, when new transportation technology integrated this region with the rapidly industrializing West, people in Southeast Asia lived on wet rice production in small valleys or shifting cultivation in upland forests. Many of the major deltas and thick rain forests were unused for agricultural production. When the region was faced with growing demand from the West for tropical products, this unused land became the basis of vent-for-surplus growth. Deltas were converted into paddy fields for commercial rice production, and rain forests were converted into plantations for export cash crops.

Corresponding to the different crops produced, peasants or small family farms continued to dominate the deltas, whereas the insular areas were bifurcated between peasants cultivating rice in small valleys and coastal plains, and large plantations based on hired labor. The different agrarian organizations were rooted in different

Figure 1. Map of Southeast Asia



ecological conditions and in land policies across different political regimes. For example, the distribution of land ownership became far more skewed in the Philippines under Spanish colonialism than in Indonesia under Dutch colonialism, although both countries were in the insular zone. Differences in agrarian structure formed along different historical trajectories under different ecological conditions and have had far-reaching influences on the performance of agricultural development across Southeast Asia. Several important variables other than the agrarian structure, including government policies, influence rural development. However, the agrarian structure is a major determinant of the political economy of the countries in the region, which exerts critical influences on their policy choices.

This article first outlines the characteristics of resource endowments, agrarian structures, growth in aggregate agricultural output, and changes in the shares of major export commodities in world markets in Indonesia, the Philippines, and Thailand in the past three decades. Second, it reviews the process of vent-for-surplus development in Southeast Asia in the late 19th to the early 20th century, emphasizing the critical roles of major river deltas in the continental zone of Southeast Asia and tropical rain forests in the insular zone. Different agrarian structures evolved in the three economies in the vent-for-surplus development process under different ecological conditions and political regimes. The preemption of uncultivated but cultivable land by the power elite was the major force to have resulted in skewed land distributions. Finally, the article tries to explain differential agricultural growth performance across Indonesia, the Philippines, and Thailand in terms of different agrarian structures. It discusses how governments forged policy choices in terms of the structure of political economy, under unique ecological conditions and unique historical trajectories. Although this study is limited to comparisons within Southeast Asia, its approach may be applicable to comparisons across regions, such as Africa versus Asia, to draw insights on broader development issues.

Recent Developments

Table 1 compares the endowments of land for agricultural production relative to the population and the labor force in Indonesia, the Philippines, and Thailand in 1965 and 1996.¹ Land is measured by area of cropland, which is the sum of areas of arable land (used for annual cropping) and land under permanent cropping, using data from the Food and Agriculture Organization. Arable land is classified into lowland paddy fields and upland annual cropland. Data for these subcategories of arable land have not been enumerated in the Food and Agriculture Organization statistics, which are mainly based on the census of farm households. It was only recently that the data for lowland paddy fields, based mainly on aerial photography, began to be available in the official reports of national statistical agencies for some specific years, which

Table 1. Land Endowments for Agricultural Production in Inodonesia, the Philippines, and Thailand, 1965–96

<i>Indicator</i>	<i>Indonesia</i>			<i>Philippines</i>			<i>Thailand</i>		
	<i>1965</i>	<i>1996</i>	<i>1996/65 ratio</i>	<i>1965</i>	<i>1996</i>	<i>1996/65 ratio</i>	<i>1965</i>	<i>1996</i>	<i>1996/65 ratio</i>
Number of farm workers (1,000) ^a	29,006	47,713	1.64	7,363	12,128	1.65	12,450	20,824	1.67
Cropland ^b									
Total (000ha)	26,000	30,987	1.19	6,660	9,520	1.42	12,600	20,445	1.62
Per capita (ha)	0.24	0.15	0.63	0.21	0.14	0.67	0.41	0.35	0.85
Per farm worker (ha)	0.90	0.65	0.72	0.90	0.78	0.87	1.01	0.98	0.97
Percentage of cropland (%)									
Lowland paddy field ^c	—	27	—	—	32	—	—	53	—
Upland annual crop land ^d	—	31	—	—	22	—	—	31	—
Land under permanent crop	31	42	1.35	38	46	1.21	11	16	1.45

—, Not available.

a. Economically active population.

b. Arable land area plus area under permanent crop.

c. Lowland paddy field areas pertain to 1995 in Indonesia, 1991 in the Philippines and 1993 in Thailand.

d. Arable land area minus lowland paddy field area.

Sources: FAOSTAT database (FAO 1999) except for lowland paddy field areas, which are taken from *1996 Statistical Yearbook of Indonesia* (Indonesia Biro Pusat Statistik 1996) for Indonesia (8,484,000 ha in 1995), *Philippine Statistical Yearbook 1998* (Republic of the Philippines 1991) for the Philippines (3,001,000 ha in 1991), and *Agricultural Statistics 1995/96* (Thailand Ministry of Agriculture 1996) for Thailand (10,934,000 ha in 1993).

are used for calculations in table 1.² The areas of upland annual cropland are measured as differences between arable land and lowland paddy field areas.

The data in table 1 show that in per capita terms, cropland area in Thailand is more than twice as large as in Indonesia and the Philippines, but only marginally larger in per-farmworker terms (in 1996). Between 1965 and 1996, cropland areas increased by about 20 percent in Indonesia, 40 percent in the Philippines, and 60 percent in Thailand. However, in all three economies, the rate of expansion in cropland area was lower than the rate of growth in population and in agricultural labor force. The cropland endowment relative to population decreased by 15 percent in Thailand and by more than 30 percent in Indonesia and the Philippines. Cropland relative to the agricultural labor force remained about the same in Thailand, decreased by 28 percent in Indonesia, and decreased by 13 percent in the Philippines. These data suggest that Thailand has been endowed with relatively favorable conditions for expanding land cultivation, compared with Indonesia and the Philippines.

For the analysis here, the important characteristic that distinguishes Thailand from Indonesia and the Philippines is the high share of paddy field area in total cropland. Relative paddy field area is greater than 50 percent in Thailand, compared with only about 30 percent in Indonesia and the Philippines (in 1996). Indonesia and the Philippines are characterized by high shares of area under permanent crops, amounting to more than 40 percent in contrast to less than 20 percent in Thailand. Permanent crops in these economies consist mainly of tropical trees for cash crop production, such as coffee, coconuts, and rubber. Although data are not available for lowland paddy field area in 1965, the share of tree cropland was as much higher in Indonesia and the Philippines than in Thailand 30 years ago that it is today. The data reflect the ecological difference between the continental zone of Southeast Asia, as represented by Thailand, and the insular zone, as represented by Indonesia and the Philippines. The continental zone has major river deltas almost exclusively used for wet rice production. The insular zone was originally covered mainly by tropical rain forests that could be converted into profitable plantations of tropical cash crops.

The different types of agricultural production corresponding to different environmental conditions gave rise to different agrarian structures in the continental and insular zones. As is common in cereal-producing areas throughout the world, peasants or small family farms make up the organization of production. Thai agriculture, which has been traditionally dependent on rice, has been characterized by the dominance of peasants or small family farms as the organization of production. By contrast, a significant portion of tropical cash crop production has been carried out by plantations or large estate farms dependent on hired labor, although many peasants have also grown cash crops. Table 2 compares the distribution of farm size and the incidence of tenancy across Indonesia, the Philippines, and Thailand for the period before the influence of Philippine land reform became significant.

Table 2. The Distribution of Operational Farm Size and the Incidence of Agricultural Tenancy in Indonesia, the Philippines, and Thailand, 1970s

<i>Indicator</i>	<i>Indonesia 1973</i>	<i>Philippines 1971</i>	<i>Thailand 1978</i>
Average operational farm size (ha)	1.1	3.6	3.7
Percentage of farms and farmland			
Below 5 ha			
Farms	98	85	72
Land area	69	48	39
Above 50 ha			
Farms	0 ^a	0.2	0 ^a
Land area	14	14	0.9
Gini coefficient of land concentration	0.56	0.51	0.45
Percentage of tenanted area in total farmland			
Pure tenancy	2	21	6
Total ^b	24	33	16
Percentage of share tenancy in tenanted land	60	79	29

a. Less than 0.05 percent.

b. Area in pure tenancy farms plus area in owner-cum-tenant farms.

Sources: Hayami and Otsuka (1993) and U.N.-FAO (1971).

In all the three economies in the 1970s, small farms operating below 5 hectares were the majority, comprising 70–100 percent of farms and cultivating 40–70 percent of farmland. Large farms above 50 hectares, which were considered agribusiness plantations, were negligible in number; those estate farms operated on 14 percent of the land area in Indonesia and the Philippines and less than 1 percent in Thailand. Thus the agricultural sector in the insular zone was bifurcated between peasants subsisting on small parcels of land and large plantations with hired labor under the hierarchy of management. The continental zone was characterized by the unimodal distribution of self-employed family farms. These plantations were privately owned and managed in the case of the Philippines; those in Indonesia were mostly state enterprises expropriated from Dutch planters after independence. The incidence of tenancy also varied widely. Tenancy was distinctively higher in the Philippines than in the other two countries, especially in terms of percentage of area under pure tenancy. In fact, the central focus of this article is on the ecological factors and historical processes that resulted in such differences in agrarian structures in Southeast Asia.

Table 3 compares agricultural growth performances in Indonesia, the Philippines, and Thailand between 1965 and 1995. In terms of total agricultural output, the growth rate was about the same in Indonesia and Thailand, but in per capita and per-farmworker terms, Indonesia's growth rates were somewhat higher. The Philippines had the lowest growth rates for those three measures. In terms of output per hectare of cropland, growth in the Philippines and Thailand was comparable and much slower than in Indonesia. The slow growth of land productivity in Thailand

Table 3. Growth of Agricultural Production in Indonesia, the Philippines, and Thailand, 1961–95

Country, indicator	Index			Growth rate (annual percent)		
	1961–65	1976–80	1991–95	1961–65 to 1976–80	1976–80 to 1991–95	1961–65 to 1991–95
Indonesia						
Total	100	157	309	3.0	4.5	3.8
Per capita	100	111	165	0.7	2.6	1.7
Per farm worker	100	138	232	2.1	3.5	2.8
Per ha ^a	100	157	263	3.0	3.4	3.2
Philippines						
Total	100	177	239	3.8	2.0	2.9
Per capita	100	116	111	1.0	–0.3	0.3
Per farm worker	100	134	156	2.0	1.0	1.5
Per ha ^a	100	137	165	2.1	1.2	1.7
Thailand						
Total	100	190	277	4.3	2.5	3.4
Per capita	100	123	140	1.4	0.9	1.1
Per farm worker	100	145	199	2.5	2.1	2.3
Per ha ^a	100	129	163	1.7	1.6	1.6

a. Per hectare of cropland (arable land plus land under permanent crop).

Source: FAO (1999).

resulted partly because of a major expansion of the cultivation frontier in the North-east, which was characterized by poor soil and unstable rainfall, and partly because of the relatively low rate in the diffusion of modern, high-yielding rice varieties. It was difficult to grow short-stature modern varieties in the flood-prone areas of the Chao Phraya Delta as well as in the drought-prone areas of the Northeast. In addition, farmers were slow to adopt modern varieties because of their low value in Thailand's export markets.

In terms of both environmental conditions and relative resource endowments, traditional comparative advantage in agricultural production in Thailand lay in rice and that of Indonesia and the Philippines lay in tropical cash crops. Table 4 shows that Thailand is a major rice exporter (the world's largest), and its world market share continued to rise between 1961–65 and 1991–95. Indonesia and the Philippines remained net importers of rice, although their import margins were significantly reduced owing to the success of the Green Revolution. This success was especially great in Indonesia, accounting for the high rate of growth in aggregate agricultural output despite the relatively slow growth in cropland area (tables 1 and 3).

Surprisingly, Thailand has become an exporter of several tropical cash crops that have declined in importance for Indonesia and, more conspicuously, for the Philippines. Sugar represents a typical example. Thailand was a net importer of sugar before

Table 4. Shares of Net Exports in World Total Export Value of Selected Agricultural Commodities in Indonesia, the Philippines, and Thailand, 1961–95

Commodity, country	Share in world market (percent)		
	1961–65	1976–80	1991–95
Rice			
Indonesia	–11.0	–17.6	–3.5
Philippines	–3.1	0.5	–0.3
Thailand	19.4	18.7	26.1
Maize			
Indonesia	–0.1	–0.1	–0.8
Philippines	0	–0.2	–0.1
Thailand	3.6	2.9	0.3
Sugar ^a			
Indonesia	0.3	–1.4	–1.1
Philippines	7.3	4.0	0.6
Thailand	0.2	2.7	6.9
Coffee ^b			
Indonesia	1.0	4.5	4.9
Philippines	0	0.3	0
Thailand	–0.1	0	0.8
Coconut oil			
Indonesia	0	–1.0	15.8
Philippines	39.8	69.4	60.4
Thailand	–0.1	–0.2	0
Palm oil			
Indonesia	17.8	14.5	13.9
Philippines	–0.9	–0.1	–0.1
Thailand	0	–0.6	0
Rubber			
Indonesia	23.3	23.7	27.9
Philippines	–0.3	0.1	0.3
Thailand	8.8	12.9	32.6
Pineapple ^c			
Indonesia	0	0	7.8
Philippines	12.9	20.7	15.0
Thailand	0	16.9	45.9
Banana			
Indonesia	0	0	0.1
Philippines	0	8.0	5.4
Thailand	0.1	0.1	0

a. Sugar, raw equivalent.

b. Coffee, green and roast.

c. Canned pineapple.

Source: FAO (1991) database.

World War II and was barely self-sufficient in the early 1960s. Nevertheless, Thailand rose to the third-largest exporter in the world (next to Brazil and Australia) in the 1990s. By contrast, Indonesia and the Philippines, two traditional exporters of sugar in Asia, have nearly completely lost significance in the international market. By the 1990s, Thailand exceeded Indonesia in the export of rubber and exceeded the Philippines in the export of pineapple products. Indonesia was able to achieve a major increase in its world market share of coffee and also to maintain high shares of palm oil and rubber. After the 1970s, the Philippines lost out in world competition in most of the tropical cash crops in which it had a traditional comparative advantage. The market share data reveal the strengthened competitive position of Thai agriculture and the dwindled position of Philippine agriculture, although comparative cost data are not available.

Indonesia and Thailand belonged to high-performing economies in the East Asian economic miracle throughout the three decades ending in the outbreak of the financial crisis in 1997 (World Bank 1993). The Philippine economy staggered, especially during the so-called lost decade of the 1980s. Gross national product (GNP) per capita in Thailand, which was about the same as that of the Philippines in the 1970s, became twice as large by the early 1990s. During the same period, Indonesia's GNP per capita increased from only about one-half that of the Philippines to about the same level. It should be reasonable to expect that the different performance of agriculture in the three countries was a significant factor underlying the different GNP growth rates.

The Basis of Vent-for-Surplus Development

A basic framework of my perspective is the classification of Southeast Asia into continental and insular zones, characterizing the former by major river deltas and the latter by tropical rain forests. Of course, such a characterization is subject to the hazard of oversimplification, disregarding wide ecological variations within each region.³ In Thailand, for example, the major delta of Chao Phraya River encompasses only a part of the Central Plain, one of four regions in Thailand. The North is characterized by small river valleys amid hills and mountains rising toward the China–Laos–Myanmar border, where irrigated rice farming can be practiced by tapping the small streams on which early Thai dynasties were built. The Northeast bordering on the Lao People's Democratic Republic is characterized by an undulated plateau with sporadic rainfall and poor soil, which had remained as the last frontier until Lao migrants settled recently by means of risky rainfed farming. The South toward the border of Malaysia has an environment similar to the insular zone, which was originally covered by rain forest.

Ecological variations within the insular zone are equally large. Especially pronounced is the difference between Java (and Bali) and the Outer Islands, such as

Kalimantan and Sumatra, in Indonesia. Though much of the latter area is typically covered by tropical rain forest, the environment of Java is categorically different, as it is characterized by volcanic slopes with fertile soil and steady water supply, which make much of this area uniquely suited for irrigated rice farming. The environment of the Philippines is largely similar to that of the Outer Islands in Indonesia but mixed with volcanic terrain similar to that of Java.

Despite the large variations within each region, in general major river deltas characterize the environment of the continental zone, and tropical rain forests characterize the insular zone. These two land types formed the basis of economic development in Southeast Asia from the late 19th century to the early 20th century. Myint's (1965, 1971) so-called vent-for-surplus theory focuses on the process of development of "empty land" with low population density, large tracts of unused land, and abundant natural resources, typically found in Southeast Asia and East Africa at the outset of Western colonization. When these economies were integrated into international trade, unused natural resources (hitherto having had no value to indigenous people) began to command market value. Those resources could produce primary commodities to meet the high export demand of Western economies. In this way, previously unused resources became a source of economic development. The deltas of major rivers, such as Chao Phraya in Thailand, Irrawaddy in Myanmar, and Mekong in Vietnam, became the basis of vent-for-surplus development in the continental zone of Southeast Asia. The rain forest provided the basis for the development of the insular zone.

Major river deltas in the continental zone are flat and low relative to sea level, so that the surface is almost completely submerged by floods in the rainy season, although it dries up in the dry season with no reservoir to store water. As such, floodplains in the major deltas had defied human settlement until the mid-19th century, literally remaining empty land. Major civil engineering work made it possible to control flooding so that the deltas were transformed into habitable and agriculturally productive land. In Thailand the water control work took the form of developing the network of canals connected with the Chao Phraya River. The canals guide floodwater more evenly over wider areas for rice production. The canal banks provide flood-proof spaces on which farmers can settle.

The government of enlightened King Mongkut (Rama IV of the Chakri dynasty) initiated canal construction in the Chao Phraya delta. Construction began shortly after the government and Great Britain signed the so-called Bowring Treaty in 1855, which opened the kingdom to trade with the West. Soon rapid increases in foreign demand for Thai rice, which significantly raised both the price of rice and the value of rice land, induced the mobilization of private investment. A group of influential courtiers and wealthy Chinese traders established the Siam Canals Land and Irrigation Company, a major builder of private canals. The company secured concession in 1889 to dig canals in a vast tract of swampy land northeast of Bangkok, under

the clause that the company was allowed to hold ownership over reclaimed land along the canals. Its operation was managed by the Chinese business elite and construction work was heavily based on Chinese migrant laborers hired for wages, unlike the *corvée* labor that was used in the king's prior projects.⁴ However, the farmers who settled in reclaimed land as tenants were Thai who had migrated from other regions.

There is little doubt that the opening of the Chao Phraya Delta for rice production was the basis of vent-for-surplus growth, which moved Thailand's economy toward specialization in rice production in the late 19th century. Although comparable data are not available for earlier years, the area planted in rice in the Central Plain was as large as 6.8 million rai (1 rai = 0.16 hectare) or 85 percent of total rice area in the kingdom in the 1905–09 period. That was larger than the national total of 5.8 million rai in 1850 (Ingram 1971). The opening for cultivation of the Irrawaddy Delta in Myanmar and the Mekong Delta in Vietnam were similarly important for those economies during the comparable period.

The tropical rain forests in Southeast Asia's insular zone played a comparable role in vent-for-surplus development. Since long before the mid-19th century, rain forests had been the source of supply of valuable products for trade, such as cinnamon, cloves, bird nests, deer horns, and hides. However, the high incidence of malaria and other tropical diseases defied human settlement inside thick tropical forests at low elevation. Typically, native people lived on sea coasts and occasionally entered the forests for collection and extraction of natural products for sale to foreign traders or their agents who sailed to their coasts. In the late 19th century, Western capital and entrepreneurship began to convert the forests into plantations of tropical export crops, heavily relying on migrant labor from China. Before this period, Western colonial powers had attempted to collect tropical products from the natives by taxation and other coercive means.

The concurrent exploitation of continental deltas and insular forests occurred with the greater integration of Southeast Asia into the world economy in the late nineteenth century. Much greater integration than before resulted from the establishment of the free trade regime under the hegemony of Britain and the revolution in ocean transportation. The free trade system was imposed on native economies by force, either directly by Britain in its colonies or indirectly by forcing liberalization on local sovereign and other colonial powers. In the Bowring Treaty, Thailand not only conceded extraterritoriality to Britain but also lost financial autonomy. Export and import duties were fixed at the flat rate of 3 percent *ad valorem*, and internal taxes, such as exercise taxes, transportation tolls, and even land taxes, were not allowed to change by the will of the kingdom alone. Advisers from Britain carefully monitored public finance (Ingram 1971). As such, the kingdom of Siam for several decades after signing the Bowring Treaty was almost a protectorate of Great Britain. Great Britain imposed the free trade system on Thailand, both internationally and internally, in a way similar to Great Britain's other colonies, such as Burma and Malay.

Furthermore, Great Britain pressed the other Western colonies to adopt the free trade system. For example, the British occupation of Manila in 1762–64 during the Seven Years' War broke Spain's monopoly of the reexport trade in Manila of Chinese goods to Mexico by galleon ships, opening up the port of Manila to shipments and commerce by traders in other nations. Great Britain's continued pressure underlay the successive opening of other ports in the Philippines until the mid-19th century (Larkin 1972).

The reason behind Great Britain's strong drive for free trade was its high manufacturing production capacity, which established the country as the "Workshop of the World" after the Industrial Revolution. British industries sought markets for their products and sources of raw material supplies. Having established the modern factory system that could produce industrial products at lower cost than local cottage industries in the tropics, Great Britain found it advantageous to trade its manufactured commodities for tropical agricultural products and minerals. Thus trade replaced the forced collection of tropical commodities through tax and other means, which were commonly practiced by earlier colonialists, such as the Spanish conquistadors in the Philippines and the Dutch East India Company in Indonesia. Soon other Western nations followed Great Britain in industrialization and in trade.

Corresponding to its expanded industrial production capacity, the West's demand for raw materials for processing—such as cotton, rubber, and tin—became very large. Moreover, as the level of income and wages rose, tropical delicacies—such as pepper, coffee, and tea, hitherto limited to the high-income elite—became common items on the tables of ordinary working people.

This tendency was further strengthened by major innovations in ocean transportation, consisting of the introduction of the steamship and the opening of the Suez Canal in 1869. These two innovations combined reduced the transportation cost of commodities from Bangkok to major ports in Europe, such as London, below that from Bangkok to Thailand's old capital, Chiang Mai (Ingram 1971). Without such innovations, it would have been impossible for bulky commodities like rice produced in Southeast Asia to find markets as far as Europe (Furnivall 1948).

Although the innovations in ocean transportation reduced the prices of commodities from Southeast Asia in the West, they also reduced the prices of Western commodities in Southeast Asia to a large extent. Thus under the liberal trade regime in the late 19th century, industrial commodities flowed into Southeast Asia, out-competing local handicraft industries. Deindustrialization became a common feature in Southeast Asia (Resnick 1970). Thailand, which had been an exporter of cotton products before the 1850s, quickly became a major importer (Ingram 1971). Indigenous labor shifted from manufacturing to primary production for export. This shift, together with the migration of labor from China and India, provided the basis for exploiting unused natural resources, such as major river deltas and tropical rain forests, for vent-for-surplus development.

An example that clearly illustrates the impact of opening to international trade and specialization in primary production can be seen in the development of sugar production in Negros, Philippines. Prior to the opening of nearby Iloilo City as an international port in 1855, Negros Island was sparsely populated, and much of its area was uncultivated. When the port opened, the island was rapidly transformed into sugar plantations. Concurrently, local weaving industries surrounding Iloilo, which had hitherto made textiles a major export item from this region, were brought into havoc by the inflow of cheap British cloth (McCoy 1982).

In the global trade system created in the late 19th to the early 20th centuries, the exchange was *not simply* between industrial commodities in the West and primary commodities in Southeast Asia. Rice produced in the continental zone was originally brought to Europe as cheap food for industrial laborers (some was reexported to Latin America). Later, as plantations developed in the insular zone, demand for rice as the basic subsistence need for plantation laborers expanded at a speed that could not be met by local supply. Correspondingly, the share of rice exported from the continental zone to the insular zone within Southeast Asia increased. Thus triangular trade flows emerged in this period—rice produced from the continental zone was brought to the insular zone, and tropical cash crops produced in the insular zone by laborers fed on the imported rice were exported to Europe in exchange for industrial products. In this triangular trade flow, comparative advantage dictated regional specialization. For example, the sugar industry, which appeared to be a promising industry for export in Thailand at the onset of trade opening, was soon destroyed by imports from Indonesia and the Philippines (Ingram 1971).

In this way, vent-for-surplus development in Southeast Asia based on the exploitation of hitherto unused land resources was reinforced by comparative advantage within the region, which was largely determined by ecological conditions. Comparative advantage based on natural resource endowments was also reinforced by colonial policies on farmland as well as public investments in physical and institutional infrastructure. For example, the strong sugarcane research program organized by the Dutch colonial government significantly contributed to strengthening the international competitiveness of the sugar industry in Indonesia (Evenson 1976).

Evolution of Agrarian Systems

This process of vent-for-surplus development would have influenced the formation of agrarian structures in Indonesia, the Philippines, and Thailand (table 2). Thailand is characterized by the unimodal distribution of peasants or family farms, with an insignificant number of large estate farms or plantations and a relatively low incidence of tenancy. Indonesia and the Philippines are characterized by bifurcation between the peasant sector, which grows mainly subsistence crops, and the planta-

tion sector, which grows tropical cash crops. Relative to the other two countries, the incidence of tenancy is high in the Philippines and, combined with the bifurcation of agricultural production, implies that the share of landless population in the rural sector is highest in the Philippines.

It is common to explain the persistence of the peasant mode in contrast to the emergence of the plantation system in terms of different technological requirements for production between subsistence food crops and export cash crops. However, in my perspective, the bifurcated farm-size distribution and the problem of landlessness in Southeast Asia (as well as in other parts in the world) stemmed essentially from "preemption of land" by colonial and domestic elite rather than technological factors for agricultural production.

Conditions of the Plantation System⁵

A conventional explanation for the establishment of the plantation system is based on the scale economies inherent in the production of tropical export crops (Baldwin 1956). However, few crops are subject to sufficiently strong scale economies at the farm level to make the use of plantation organization necessary (Pim 1946; Wickizer 1951, 1960; Lim 1968; Hayami and others 1990).⁶ In fact, there are examples of every so-called plantation crop being grown successfully by peasants somewhere in the world.

Significant increasing returns emerge only at the levels of processing and marketing activities. The vertical integration of a large estate farm with a large-scale central processing and/or marketing system is called for because of the need to supply farm-produced raw materials on a timely schedule. A typical example is fermented black tea. The manufacturing of black tea at a standardized quality for export requires a modern machine plant into which fresh leaves must be fed within a few hours after plucking (Wickizer 1951, 1960). The need for close coordination between farm production and processing underlies the traditional use of the plantation system for black tea manufacture. By contrast mainly peasants in China and Japan produce unfermented green tea. Even for the manufacture of black tea, it is not imperative to use the plantation system. This is evident in the case of Taiwan, where smallholders produce both black and green tea with small-scale equipment. The large fermentation plant has been used in plantations as a device enforcing the work schedule and standardizing product quality for the export market. In fact, farm production by smallholders based on the system of contract farming has developed relatively recently in Kenya (Lamb and Muller 1982).

In the case of bananas for export, harvested fruit must be packed, sent to the wharf, and loaded on a refrigerated boat within a day. A boatful of bananas that can meet the quality standards of foreign buyers must be collected within a few days. Therefore, the whole production process from planting to harvesting must be precisely

controlled so as to meet the shipment schedule. Although the plantation system has a decisive advantage for this export product, bananas for domestic consumption are usually produced by peasants.

Plantations have no significant advantage over peasants for the crops for which centralized processing and marketing are not necessary. Typical examples are cocoa and coconuts. The fermentation of cocoa and the drying and smoking of coconuts to make copra can be handled in small lots with no large capital requirement beyond small indigenous tools and facilities. These crops are grown predominantly by peasants.

Sugar is frequently cited as a classic case of scale economies stemming from the need for coordination between farm production and large-scale central processing (Binswanger and Rosenzweig 1986). Efficient operation of a centrifugal sugar mill requires the steady supply of a large amount of cane over time. It requires coordination of production from planting to harvesting and processing. However, this coordination need not be as stringent as it is for tea and bananas. The rate of sugar extraction decreases as the processing of cane is delayed, but this loss is in no way comparable to the devastating damage to the quality of tea and bananas for export that may result from delayed processing. Sugarcane can be hauled from relatively long distances and stored for several days. Therefore, the need for vertical integration is not as large, and the sugar mill can achieve the necessary coordination through contracts with cane growers on the time and the quota of cane delivery. In fact, Australia, Taiwan, and more recently Thailand have developed an efficient sugar industry with smallholders.

Another explanation for the use of the plantation system is the advantage of large estate farms in accessing capital. Binswanger and Rosenzweig (1986) argue that this gives plantations an advantage with regard to tree crops characterized by long gestation periods from planting to maturity. However, the opportunity costs of labor and capital applied to formation of the tree capital are not necessarily high for peasants. Typically, they plant the trees in previously unused land. If such land is located near their residence, they open new land for planting by means of family labor at low opportunity cost during the idle season for the production of food crops on farm land already in use. Often, when peasants migrate to frontier areas, they slash and burn jungles and plant subsistence crops such as maize, potatoes, and upland rice, together with tree seedlings. Such complex intercropping is difficult to manage with hired labor in the plantation system because of the inherent difficulty in monitoring the work of hired wage laborers over spatially dispersed and ecologically variable farm operations (Brewster 1950; Binswanger and Rosenzweig 1986; Hayami and Otsuka 1993).

Therefore, even in the export boom of tropical cash crops under colonialism from the nineteenth century to the early 20th century, the plantation system failed to make inroads in regions where the indigenous population had established family farms (Lewis 1970). Western traders found it more profitable to purchase tropical agricul-

tural commodities from peasant producers in exchange for imported manufactured commodities, than to produce the tropical crops themselves by means of the plantation system.

The establishment of plantations in less developed economies increased as the demand for tropical products by the industrialized nations continued to rise and the regions physically suited for the production of these products had no significant peasant population that could produce and trade their commodities. Opening frontier land for the production of new crops entailed high capital outlays. Virgin land had to be cleared and developed and physical infrastructure, such as roads, irrigation systems, bridges, and docking facilities, had to be constructed. Capital, in the form of machinery and equipment, had to be imported and redesigned to adapt to local situations. Laborers were imported from the more populous regions and trained in the production of these crops.

The establishment of plantations thus requires huge initial capital investment. For the investors to internalize gains from investment in infrastructure, farm size must be large. Viewed from this perspective, it follows that the plantation system evolved not because it was generally a more efficient mode of productive organization than the peasant mode. Instead, the system was adopted because it was the most effective type of agricultural organization for extracting the economic benefit accruing from the exploitation of sparsely populated virgin areas, typically in the process of vent-for-surplus development.

From this perspective, it is easy to understand why the same crop is grown mainly by peasants in one place and mainly by plantations in another. For example, for sugarcane production, the peasant mode is more common in the old settled areas of Luzon, and the plantation system predominates in the newly opened Negros, both in the Philippines (Hayami and others 1990). Usually the share of peasants in the production of export cash crops rises when the initial land-opening stage is over and infrastructure is decently established with increased population density (Booth 1988).

Although the plantation system had an economic advantage in the vent-for-surplus stage, plantations could not have been established if governments had not granted concessions to hold large tracts of virgin land for the exclusive use of plantations. Typically colonial governments granted concessions to Western planters. For example, the Dutch colonial government had traditionally tried to prevent alienation of farmland from indigenous peasants by regulating against land purchase by foreigners, including the ethnic Chinese. However, in the late 19th century, demand for tropical cash crops rose sharply. The colonial government passed the Agricultural Land Law of 1870, which granted Dutch planters long-term contracts to lease wild land. The land was *de jure* owned by the government (although it was *de facto* used by native tribes). Although this new institutional arrangement should have accelerated the development of "empty land" for cash crop production, it served as an instrument to preempt land for the elite, closing smallholders' access to land. Similar

public land-leasing arrangements were also practiced under the American colonial administration with frontier land in the Philippines, especially in Mindanao, which became the basis of large plantations under the management of multinational corporations (Hayami and others 1990).

Land Preemption and Tenancy

The incidence of land tenancy is closely related to the preemption of land. Of course, a land tenancy relationship can emerge as a practice among peasants in the absence of preemption. If external forces did not disturb a rural community, land tenure institutions would evolve gradually from communal to private ownership. Corresponding to the growing relative scarcity of land under mounting population pressure, it becomes necessary to intensify the utilization of land, typically from shifting cultivation with long furrow to that with short furrow, to annual cropping, and further to multiple cropping per year involving irrigation (Boserup 1965).

The process of agricultural intensification requires major investment for improving land infrastructure, from removing stones and roots from newly opened land, to land leveling and terracing, and further to irrigation and drainage. To secure incentives for such investment, it becomes necessary to give land users the right to use their land exclusively. Thus land tenure institutions normally evolve from communal ownership to private ownership, involving various steps from periodical re-allotment of communal land among community members, to life-long usufruct rights, to usufruct rights inheritable by heirs, and further to private property rights amenable for market transactions.

Land tenancy arrangements gradually develop as an institution to increase production efficiency by improving combinations of land and labor (including entrepreneurship) as individual land tenure becomes longer and more exclusive. When a farmer finds his family labor short for cultivation of a land parcel on which a long-term usufruct is established (because of sickness or some other reason), he may rent out a part of it to someone whose land endowment is short relative to their labor endowment. It is a Pareto improvement if the latter pays to the former a rent equivalent to the marginal productivity of the land. At the same time, land tenancy associated with private property rights on land can work as an institution to increase inequality in income distribution and social hierarchy within a community. A farmer endowed with superior muscular power or entrepreneurship may rent more land and increase income and may eventually buy the land. As he eventually accumulates more land than his family labor can efficiently cultivate, he may rent out a part of his land to someone who has become landless for whatever reason. Increased income from rent revenue added to farm income may motivate him to purchase more land for renting. This process should progress faster as the relative scarcity of land rises under increased population pressure.

Such autonomous evolution of land property rights and tenancy relationships does not usually result in the large-scale absentee landlordism observed in several developing economies. Rather it tends to create stratification of peasantry along a continuous spectrum between landlord-cum-owner and owner-cum-tenant farmers. Although land tenancy is commonly practiced, a majority of farmland continues to be under owner cultivation, and both noncultivating landlords and pure tenants are the minority.

Such an agrarian structure is typically found in the peasant sector in Indonesia. Unlike other colonial powers, the Dutch did not try to impose Western institutions, such as private property rights in land. Rather they preserved or even strengthened traditional community institutions and organizations. The Agrarian Law of 1870 granted long-term lease of wild public land to foreign planters, but did not allow them to purchase or rent cultivated land from native peasants individually. Instead, sugar planters were allowed to lease rice land through contracts with the heads of villages. The contracts normally extended for less than 20 years. The lessee was allowed to occupy only one-third of the village land, which had to be rotated over three crop seasons. This rotation was designed to prevent planters from gaining a permanent hold on village land. Periodic reallocation of village land under the direction of village headmen strengthened traditional tendencies toward communal landholding (Pelzer 1945).

The situation in the Philippines provides a sharp contrast. At the time of conquest, the Spaniards introduced the notion of legal title to land (McLennan 1969). They applied to the Philippines the same principle they applied to other new territories—that all the lands except those officially proved to be private or communal possessions belonged to the Spanish crown. The Crown's property rights were established over vast areas of uncultivated land, including areas used as commons by native people. Much of the royal domain was granted to conquistadors and monastic orders, such as Augustinian and Franciscan friars. This institutional development in the early Spanish era represented a wholesale preemption of usable land, closing access by native people. Later, the population increased, and foreign demand for Philippine products increased through trade liberalization. Large landholdings created from earlier royal grants became plantations in the upland areas and rice haciendas with tenant labor in the lowland areas.

Native peasants were even deprived of the opportunity to establish ownership by opening new lands for cultivation. For example, the inner part of Central Luzon had been covered by jungle and used only for cattle ranching. When it was finally converted into large rice haciendas in the late 19th century, many peasants migrated from the north, believing that they had settled in no-man's-land. After the peasants opened the jungle, landowners' agents visited and notified the peasants that they had to pay rent as tenants on haciendas (Hesters and Mabun 1924).

Pervasive landlordism in the Philippines was also rooted in the relatively free land transactions under the Spanish regime. Chinese and Chinese mestizos, who engaged

in internal trade along littorals, where native peasants held traditional land rights, acquired land through moneylending using land as collateral. A common arrangement was that the borrower continued to cultivate his land as a sharecropper of his creditor during the loan period. If the borrower became unable to repay the loan at the end of the period, the land title shifted to the creditor and the borrower usually continued sharecropping (McLennan 1969). The scale of landholding accumulated in this commercial process in the coastal area was typically much smaller than that of haciendas in the inner part of Central Luzon (Hayami and Kikuchi 1981). Thus, before the Marcos land reform in the 1970s, rice area in the Philippines was predominantly cultivated by share tenants, typically owning no land of their own. The pervasive landlordism in the rice sector and plantations in the cash crop sector that characterized the traditional agrarian structure in the Philippines were both rooted in the preemption of land in the Spanish period.

In Thailand, preemption occurred in the vent-for-surplus stage through the granting of land concessions to private canal builders in the Chao Phraya Delta. As a result, the incidence of tenancy is significant in the Central Plain, especially in the Rangsit area northeast of Bangkok, where the private company intensively dug canals. Yet, taking Thailand as a whole, tenancy is of minor importance compared with Indonesia and the Philippines, partly because of relatively abundant land endowment and, more important, because of government policy.

The country's ancient custom was to give every man the right to take as much land from the state as he and his family could cultivate, which was normally considered to be 25 rai (equivalent to 4 hectares). This institution was maintained even after opening trade with the West. The Consolidated Land Act of 1908 did not specify an exact area of land, but gave people the right to take as much land as they could profitably cultivate. In practice these areas ranged between 20 and 50 rai. The Land Act of 1936 specified 50 rai as the maximum that one could take. These laws kept access to land wide open for ordinary Thai, making the situation diametrically different from that of the Philippines.⁷ Both of these Thai laws incorporated the old custom that the cultivator could receive title to the land only after he cultivated it for three years. This clause together with land taxation, which applied to not only cultivated but also uncultivated holdings, discouraged holding land idle for speculation (Ingram 1971).

The basic factor underlying the major difference in land policy between the Philippines and Thailand was the difference in the culture or the value system between the Spanish colonial rulers and the rulers of the independent kingdom. The Dutch colonial rulers tried to preserve traditional village institutions, thereby avoiding alienation of land from peasants in Indonesia. Their motivation might have been to maintain social stability for the sake of extracting tropical agricultural products from this colony at minimum administrative cost, as argued by Furnivall (1944, 1948).

It is also important to observe that the preemption of rice land through canal construction in Thailand resulted in the emergence of large-scale landlordism, but not

in the formation of plantations. This was similar to the case of the Philippines. In Thailand, large holdings of landlords were usually subdivided into small parcels for rice cultivation by the family labor of landless peasants under tenancy contracts. The owners of large tracts of rice land established titles through land preemption, such as obtaining concessions for canal digging in the Chao Phraya Delta. They preferred tenancy to plantation operations, perhaps at least in part because of the difficulty of standardizing tasks of rice production and, hence, of monitoring the efforts of workers.

Another reason the owners preferred tenancy may have been because paddy is storable. Unlike black tea and bananas, paddy does not require close coordination between farm production and processing/marketing. Although rice milling and marketing for export involved significant scale economies, the operators could secure adequate supply of paddy through ordinary market transactions. As the result, they could dispense with efforts to vertically integrate farm production with processing and marketing by means of the plantation system or the contract farming system. Therefore it may not be unreasonable to postulate the counterfactual hypothesis that, if the nature of rice-milling technology were such as to require close coordination with paddy production, large rice plantations would have been established in the Rangsit area where territorial concessions were granted to private canal builders.

Outside the newly opened delta area, the practice of tenancy is fairly common in the old settled North region. The agrarian structure in the north of Thailand, which did not experience preemption, is similar to that of the peasant sector in Indonesia. It is characterized by a continuous spectrum from landlord-cum-owner to owner-cum-tenant farmers.

Agrarian Structure and Agricultural Growth Performance

This section discusses whether the different agrarian structures that emerged along different historical paths under different ecological conditions explain, at least in part, the different agricultural growth performance across Indonesia, the Philippines, and Thailand. The analysis focuses on two questions. First, why did Indonesia and the Philippines, which had strong comparative advantage in tropical cash crops such as sugar before World War II, lose ground to Thailand in world market competition in recent years (table 4)? Second, why was the agricultural output growth of the Philippines so slow relative to growth in Indonesia and Thailand (table 3)?

Decline in the Advantage of Plantations

The previous section argued that the efficiency of plantations relative to the peasant system is high in the initial opening-up process of land-abundant and labor-scarce economies. However, several negative aspects of plantations grow large as tropical

economies shift from the land-abundant to the land-scarce stage after the completion of the opening-up process. Correspondingly, the relative advantage of the peasant system increased.

Negative aspects of the plantation system. First, the plantation system tends to substitute capital for labor because of the inherent difficulty in supervising wage laborers in spatially dispersed and ecologically diverse farm operations. In addition, plantations have relatively easy access to both the private credit market and concessional loans from the government. The substitution of capital for labor is socially inefficient in many developing economies, which are characterized by the abundant endowment of labor relative to capital.

Second, agricultural land tends to be cultivated less intensively in the plantation system, which employs mainly wage labor and usually practices monoculture. Complicated intercropping and crop-livestock combinations are more difficult to manage in the command system, implying that both the labor input and income per hectare are lower on the plantations.⁸ This is a source of inefficiency in the plantation system where land becomes scarce relative to labor under the pressure of population growth. By contrast, small-sized family farms tend to cultivate land more intensively.

Third, plantations usually specialize in a single crop. This monocrop bias reduces the flexibility of productive organizations to respond to changing demand by shifting to the production of other crops. Moreover, continual cropping of a single crop tends to result in soil degradation and an increase in the incidence of pests. Counter-application of fertilizer and chemicals causes serious stress on the environment and human health and incurs high costs.

Fourth, the specialization of plantation workers in specific tasks inhibits the development of their managerial and entrepreneurial capacity (Baldwin 1956; Myint 1965; Beckford 1972).

Fifth, the plantation system is a source of class conflict between laborers and managers/capitalists. The presence of a plantation enclave in rural economies where the peasant mode of production predominates has often strained relationships in rural communities. Therefore, in terms of the criterion of social stability, the plantation system is no match for the system of relatively homogeneous small producers that own small assets.

Increased advantage of the peasant system. Although Southeast Asia had traditionally been endowed with relatively abundant land resources ready for exploitation, frontiers for new land opening were progressively closed under the explosive population growth that characterized developing economies after World War II. It seems reasonable that the advantage of the plantation system declined and that of the peasant system increased correspondingly. Therefore it is not surprising that Thai agriculture, which predominantly consisted of smallholders, began to perform better than

that of Indonesia and the Philippines, which were characterized by large plantation sectors.⁹ Major increases in the export of nonrice agricultural commodities, such as rubber, kenaf, and cassava tips from Thailand, were totally based on the production of smallholders. To a significant extent, the expansion in the production of nonrice export crops in Thailand was supported by the existence of open land frontiers, which enabled relatively fast increases in the area under cultivation (table 1). The important point in the present context, however, is that the exploitation of cultivation frontiers was carried out by smallholders and not by plantations.

Relative increases in the efficiency of the peasant system were not limited to Thailand. The dramatic rise in Indonesia's share in world coffee and cocoa markets was entirely based on smallholders (Akiyama and Nishio 1996). The production of coconut oil in the Philippines, for which the country was able to maintain its high world market share, was extracted predominantly from copra made by smallholders, although some coconut plantations continued to operate, especially in Mindanao.

The advantage of the plantation system is better coordination between large-scale marketing/processing and farm-level production. Yet the disadvantage of the peasant system in this aspect could be overcome by organizing contract farming. In contract farming, an agribusiness firm manages processing and marketing, but contracts for the supply of farm products with peasant farmers. The firm provides technical guidance, credit, and other services to peasants in return for their pledged production to the firm. In this way, the system can take advantage of peasants in farm production without sacrificing scale economies in processing and marketing. An advantage of this system is that it taps not only the manual labor but also the management ability of rural people in developing economies. Thailand used this system when it began production of canned pineapple relatively recently; Thailand has surpassed the Philippines, formerly the world's leading exporter, whose production is based on large plantations in Mindanao.¹⁰

The Dilemma of Land Reform

The Philippine government has attempted to mitigate the social unrest rooted in pervasive landlessness and landlordism. Redistributive land reform extends back to the American colonial regime. However, the framework of the reform applied in the past four decades was established by the Agrarian Land Reform Code of 1963, which was enacted under President Macapagal (Hayami and others 1990).

The major thrust of the code was the creation of owner-cultivatorship on rice and corn land. This involved two steps. First, "Operation Leasehold" converted share tenancy to leasehold tenancy with rent fixed at the rate of 25 percent of the average harvest for three normal years preceding the operation. Second, "Operation Land Transfer" transferred land ownership to tenants. In the latter operation, the government expropriated land in excess of landlords' retention limit (75 hectares). It com-

pensated the landlords with 10 percent of the land value in cash and the rest in interest-free redeemable Land Bank bonds. The land was resold to the tenants for annual amortization payments within 25 years.

The code was amended in 1971 under President Marcos to extend land reform to the whole nation, with automatic conversion of all share tenants to leaseholders. The 1971 Code was enforced by Presidential Decrees No. 2 and No. 27 under the martial law proclaimed in 1972. The landlord's retention limit was reduced successively from 75 to 7 hectares. The period of amortization payments was shortened to 15 years. It is easy to enumerate the shortcomings of the land reform programs in the Philippines, yet there is no denying that large haciendas in Central Luzon were broken down. Most tenants established their status as leaseholders or amortizing owners, although sizable areas remain under the direct administration of landlords.

The beneficiaries of land reform have captured a large economic surplus. Rice yields have increased significantly due to the development of irrigation and the application of new varieties and fertilizers, while rent and amortization payments have been fixed. Thus, land reform has been successful in transferring much of the economic return to land from absentee landlords to former sharecroppers. However, the reform has created serious income inequality within village communities. The income of landless laborers has not risen (or may have declined) because the strong population pressure on land has prevented their wages from rising despite agricultural productivity increases.

Major distortions in resource allocations occurred because reforms were applied in a discriminatory manner to a certain sector of agriculture. By limiting program application mainly to tenanted land, the reforms created a strong incentive for landlords to evict their tenants and cultivate their land directly. However, labor inputs and, hence, agricultural output and labor income per hectare are usually higher in small family farms than in large farms. This is because of the inherent difficulty large farms have in supervising wage laborers. Therefore, the exemption of land under the direct administration of landlords had the effect of reducing labor input per hectare below an optimum level, thereby reducing the income of the laborers.

The impact of the regulations on tenancy contracts (especially the prohibition of share tenancy and the control of land rent) were equally serious. They reduced the incentive of large landholders to rent out their land in small parcels, resulting in a reduction in social product and labor income. This behavior applied not only to landlords but also to land reform beneficiaries. As the income of former sharecroppers, who were converted into leaseholders or amortizing owners, rose significantly, many of them retreated from arduous farmwork, leaving it to landless laborers. Yet they hesitated to subrent their holdings to landless laborers because their formal titles might be transferred. The land reform laws said that formal titles would be transferred to sublessees if they could prove to the agrarian reform office that they were the actual tillers of the land. Thus land reform beneficiaries have to continue to cultivate

their land based on hired labor, even if they are not able to work because they are sick, old, or engaged in nonfarm activities. The reforms inevitably resulted in inefficient combinations of land and labor.

Significant negative effects of land reform on agricultural production efficiency also occurred outside the rice and corn sector. The cash crop sector has not been covered by reform programs. The Comprehensive Agrarian Reform Law of 1988 intended to cover the cash crop sector, but it has not been significantly implemented. However, plantation owners fear that their land will eventually be expropriated. It is only natural that they have stopped investing in improvements in their land infrastructure, including planting and replanting trees. Some landowners even preferred to keep their land idle rather than use it for agricultural production. This was often the case in frontier regions like Mindanao.

Such fears and lack of investment might underlie, to a significant extent, the low rate of expansion in cropland area in the Philippines as compared with Thailand (table 1). The poor performance of the Philippines in competition for world export market shares is at least partly rooted in the uncertainty of the planters of tropical cash crops concerning the future course of land reform (table 4).

Toward Political Economy

Different agrarian structures developed in Indonesia, the Philippines, and Thailand along different historical paths and under different ecological conditions. From the late nineteenth to the early twentieth century, development of the three economies followed a typical pattern along the vent-for-surplus theory. Vent-for-surplus development is based on the exploitation of unused natural resources corresponding to their integration into the world market. Thailand represents the continental zone of Southeast Asia, whose resource base for vent-for-surplus development was the major delta of the Chao Phraya River. Indonesia and the Philippines represent the insular zone, whose resource base was the tropical rain forest. This difference in the resource base underlay the major difference in farm-size distribution—the unimodal distribution of peasants or family farms in Thailand as compared with the coexistence of peasants and large estate farms or plantations specializing in tropical export crops in Indonesia and the Philippines.

Different land policies, especially with respect to preemption of unused land by the elite, under different political regimes resulted in major differences in the pattern of land ownership. The preemption was wholesale in the Philippines under Spanish colonialism. It provided the base for the highly skewed land distribution, which was characterized by the bifurcation between noncultivating landlords and sharecroppers in lowland rice areas and between plantation owners and wage laborers in upland areas. In Indonesia, the preemption took place as the Dutch colonial govern-

ment granted long-term lease of uncultivated public land to foreign planters. However, the government tried to prevent the alienation of cultivated land from native peasants to avoid social instability. As a result, the peasant sector continued to consist mainly of landlord-cum-owner and owner-cum-tenant cultivators, and both noncultivating landlords and the pure landless remained a minority. In Thailand the preemption occurred through the grant of concessions for private canal building. However, the incidence of tenancy did not become serious because the government of the independent kingdom preserved the traditional institution of giving land to anyone who could open and cultivate it. Relatively homogeneous landowning peasants continued to dominate the rural sector of Thailand.

It appears that major differences in the agrarian structure have been significant factors in the differences in agricultural growth performance across the three economies in recent years. As frontiers for opening new land for cultivation were progressively closed, the initial advantage of the plantation system in large-scale land development began to be outweighed by its disadvantage in monitoring hired labor. At the same time, the advantage of the peasant system increased with respect to the use of family labor needing no supervision. This tendency seems to be manifest in Thailand's growing share in world exports of tropical cash crops in recent years. Furthermore, the land reform programs in the Philippines that called for reducing inequality in the distribution of land ownership have made land markets inactive, resulting in major distortions in resource allocations and serious underinvestment in agriculture.

Many factors in addition to the agrarian structure have contributed to the differential performances in agriculture. For example, one factor that is commonly cited in explaining the poor growth performance of Philippine agriculture is the prolonged continuation of the country's industrial protection policy geared for import substitution. Under this policy regime, the agricultural sector was penalized by high tariffs on manufactured commodities and overvalued exchange rates (Ariff and Hill 1985; Bautista 1987; Intal and Power 1989). Another factor was the state's trade monopoly on sugar and coconut products, which was heavily tinted with cronyism in the late stage of the Marcos regime (Hayami and others 1990). The political instability in the 1980s from the downfall of the Marcos administration throughout the succeeding Aquino regime discouraged both domestic and foreign investments in agriculture as well as in other sectors.

By contrast, Thailand began to shift from the import-substitution to the export-oriented strategy in the 1970s, almost two decades ahead of the Philippines (Warr 1993; Warr and Nidhiprabha 1995). In addition, the government of Thailand intervened little in the activities of private traders in agricultural marketing. The government's control of trade was largely limited to indirect measures, such as the imposition of an export tax on rice (the rice premium). A large number of traders of various sizes, ranging from small collectors of farm produce at the village level to large ex-

porters to the foreign market, were well coordinated through free competition. They made up a highly efficient channel to deliver overseas demands to farmers. Their activities were facilitated by major public investment in infrastructure, especially highways. This free trade system, supported by the government's provision of public goods, created remarkable diversification of agricultural resources to new export crops. The agricultural diversification was achieved while Thailand continued to strengthen its competitive position in rice production (Siamwalla and others 1990; World Bank 1987).

It is unlikely that such differences in government policy are independent of differences in the agrarian structure and value system in society that are deeply rooted in different ecological conditions and historical paths. For example, Hara (1994) advances a hypothesis on the reason why import-substitution industrialization was pursued more strongly for a longer period in the Philippines than in Thailand and other Association for Southeast Asian Nations economies. He argues that in the Philippines the business elite who benefited from industrial protection originated from the landed oligarchy. Therefore little countervailing power was mobilized against the industrial protection policy. By contrast, the rural countervailing power against industrial protection was comparatively high in Indonesia, Malaysia, and Thailand because the urban business elite were predominantly ethnic Chinese. Also, the rather harmonious division of labor that developed between Thai farmers and Chinese traders in the Siam Kingdom may have prevented the modern Thai government from adopting antimarket and antitrader interventions.

Another example is the remarkable success of the Green Revolution in Indonesia. The country's high growth of land productivity was, to a large extent, based on the strong support of the rice sector during the three decades of the Soeharto administration. It invested in irrigation, agricultural research and extension, plus subsidies for inputs and credits. This support was effective in overcoming the "Dutch disease effects" that seriously damaged agriculture in some oil-producing countries, such as Nigeria in the 1970s and early 1980s (Hayami 1997). It seems that Soeharto's policy choice was not independent of the tradition in Indonesia of protecting peasants as the stabilizing block of society.

In the Philippines in the 1970s, the Masagana-99 Program promoted the Green Revolution by means of distributing to farmers packages of new seeds, subsidized fertilizers, and other modern inputs, in a manner similar to Indonesia's Bimas Program. However, in the absence of "peasant fundamentalism" in the Philippines comparable to that of Indonesia, the Masagana Program lasted only about a decade (Hayami and Kikuchi 2000). Thus rice self-sufficiency in the Philippines, which had been achieved during the 1970s, could not be sustained in the 1980s. Indonesia rose from the world's largest importer of rice in the 1970s to achieve self-sufficiency in the 1980s. However, the country again became a major importer with the recent economic crisis associated with the downfall of the Soeharto regime. Unlike Indonesia and the

Philippines, in Thailand's liberal trade regime, the role of government policy in promoting fertilizer application was not very significant and effective (World Bank 1987).

For now, such political-economy theorizing is largely conjectural. Yet the agrarian structure of a nation and the value system in its society have been created along a unique historical trajectory, under unique ecological conditions. These should have a far-reaching influence on the organization of political economy and, hence, on policy choices. The positive analysis of this relationship presents a major challenge for future research. The analysis might be extended beyond the comparison within a region, as attempted here, to comparisons across regions. It might shed light on major questions in world development. For example, it might help in analyzing why Africa lags behind in achieving innovations in agricultural technology comparable to the Green Revolution in Asia.

Notes

Yujiro Hayami is the director of the Graduate Program at the Foundation of Advanced Studies on International Development and an adjunct professor at the National Graduate Institute for Policy Studies, Tokyo. He wishes to gratefully acknowledge helpful comments from Takamasa Akiyama, Robert Allen, Gershon Feder, and Yair Mundlak, as well as technical assistance from Kei Kajisa and Yue Yaguchi.

1. For a more comprehensive assessment of agricultural growth performance in Asia including Southeast Asia, see the five-volume report of the Asian Development Bank's project titled "Study of Rural Asia." Especially relevant to the context of this section are its overview (Asian Development Bank 2000), volume one by Rosegrant and Hazell (2000) and volume two by Kosa-ard and Rekasem (2000). Another major study specifically addressed to Southeast Asian agriculture is in progress by the Development Economics Research Group at the World Bank, titled "Dynamism of Rural Sector Growth: Policy Lessons from East Asian Countries."

2. Huke and Huke (1997) estimate paddy field areas in Indonesia, the Philippines, and Thailand in the mid-1990s as being 9,441,000, 3,456,000, and 9,806,000 hectares, respectively. However, they do not specify the years for these data. The substitution of Huke and Huke's estimates for the data used in table 1 does not change the conclusion of this article.

3. Descriptions of ecological and environmental conditions in Southeast Asia in this article are mainly based on Takaya (1985).

4. Though the corvée obligation was replaced by tax in kind or money, slavery was also phased out gradually over the reigns of Kings Mongkut and Chulalongkorn, ending in its abolishment in 1905 (Ingram 1971; Feeny 1982). The elimination of slavery and the corvée should have been an important factor for allocating a greater share of Thai labor to rice cultivation.

5. This section draws heavily on Hayami (1994, 1996).

6. Absence of scale economies in agriculture is also attested by the estimation of aggregate production functions based on inter-country cross-section data (Hayami and Ruttan 1985).

7. All forestlands were *de jure* state-owned, but were *de facto* open-access, except valuable teak forests that were an important source of the kingdom's revenue (Feeny 1999).

8. Official statistics often record that yields per hectare of cash crops such as coffee and rubber are higher in plantations than in small holders. However, these statistics do not take into account various products intercropped with principal cash crops by small holders, whereas monoculture is the common practice of plantations.

9. In addition to this disadvantage, the plantation sector in postindependence Indonesia that expropriated the estates of Dutch planters seems to have suffered from inefficiency common to state enterprises. Several attempts to cure this problem include the "nuclear estate" scheme by which a state plantation acts as a marketing/processing center with a demonstration farm for technical extension, along which smallholders are organized in a manner similar to contract farming. These attempts have often been marred by the direct application of the technology and practice of plantations without due understanding of the conditions of smallholders (Barlow and Tomich 1991). The case of Indonesia represents a contrast to the relatively high efficiency of plantations in Malaysia under private entrepreneurship. Private plantations in Malaysia are also well supported by the cooperative research and extension system that has been organized since the colonial period.

10. However, it needs a high degree of entrepreneurship and managerial skill to organize and operate the efficient contract farming system. It is not easy to enforce contracts with a large number of smallholders concerning the quantity, quality, and time of their product delivery to processing plants and/or marketing centers. Insufficient ability and effort of agribusiness firms in this regard have often resulted in failure in the operation of contract farming. Thus the performance of contract farming has so far been mixed even in Thailand (Siamwalla 1992). The same applies to other areas, including Africa where it is reported that contract farming organized by government agencies is usually inefficient (Jaffee and Morton 1995).

References

- Akiyama, Takamasa, and Akihiko Nishio. 1996. *Indonesia's Cocoa Boom: Hands-Off Policy Encourages Smallholder Dynamism*. Policy Research Working Paper 1580, World Bank, Washington, D.C.
- Ariff, Mohamed, and Hal Hill. 1985. *Export-Oriented Industrialization: The ASEAN Experience*. Sydney: Allen and Unwin.
- Asian Development Bank. 2000. *Rural Asia beyond the Green Revolution*. Manila: Asian Development Bank.
- Baldwin, Robert E. 1956. "Patterns of Development in Newly Settled Regions." *Manchester School of Economics and Social Studies* 24(May):161–79.
- Barlow, Colin, and Thomas P. Tomich. 1991. "Indonesian Agricultural Development: The Awkward Case of Smallholder Tree Crops." *Bulletin of Indonesian Economic Studies* 27(December):29–53.
- Bautista, Romeo M. 1987. *Production Incentive in Philippine Agriculture: Effects of Trade and Exchange Rate Policies*. Washington, D.C.: International Food Policy Research Institute.
- Beckford, George L. 1972. *Persistent Poverty: Underdevelopment in Plantation Economies in the Third World*. New York: Oxford University Press.
- Binswanger, Hans P., and Mark R. Rosenzweig. 1986. "Behavioral and Material Determinants of Production Relations in Agriculture." *Journal of Development Studies* 22–23(April):503–39.
- Booth, Anne. 1988. *Agricultural Development in Indonesia*. Sydney: Allen and Unwin.
- Boserup, Ester. 1965. *The Conditions of Agricultural Growth*. London: Allen and Unwin.
- Brewster, John M. 1950. "The Machine Process in Agriculture and Industry." *Journal of Farm Economics* 32(February):69–81.
- Evenson, Robert E. 1976. "International Transmission of Technology in the Production of Sugarcane." *Journal of Development Studies* 12(January):208–31.
- FAO. 1999. *FAOSTAT 98 CD-ROM*. Rome: FAO. Available online at <http://apps.fao.org/>.
- Feeny, David. 1982. *The Political Economy of Agricultural Productivity: Thai Agricultural Development, 1880–1975*. Vancouver: University of British Columbia Press.

- . 1999. "Thailand and Japan: Why Was Japan First?" In Yujiro Hayami and Masao Aoki, eds., *The Institutional Foundations of East Asian Economic Development*. London: Macmillan, pp. 413–42.
- Furnivall, John S. 1944. *Netherlands India: A Study of Plural Economy*. Cambridge: Cambridge University Press.
- . 1948. *Colonial Policy and Practice: A Comparative Study of Burma and Netherlands India*. Cambridge: Cambridge University Press.
- Hara, Yonosuke. 1994. *Tonan Ajia Shokoku no Keizai Hatten* (Economic Development of Southeast Asian Countries). Tokyo: Institute of Oriental Culture, University of Tokyo.
- Hayami, Yujiro. 1994. "Peasant and Plantation in Asia." In Gerald M. Meier, ed., *From Classical Economics to Development Economics*. New York: St Martin's Press.
- . 1996. "The Peasant in Economic Modernization." *American Journal of Agricultural Economics* 78(December):1157–67.
- . 1997. *Development Economics: From the Poverty to the Wealth of Nations*. Oxford: Oxford University Press.
- Hayami, Yujiro, and Masao Kikuchi. 1981. *Asian Village Economy at the Crossroads*. Tokyo: University of Tokyo Press; Baltimore, Md.: Johns Hopkins University Press.
- . 2000. *A Rice Village Saga: Three Decades of Green Revolution in Rice the Philippines*. London: Macmillan; New York: Barnes and Noble; Los Baños, Philippines: International Rice Research Institute.
- Hayami, Yujiro, and Keijiro Otsuka. 1993. *The Economics of Contract Choice*. New York: St. Martin's Press.
- Hayami, Yujiro, and Vernon W. Ruttan. 1985. *Agricultural Development: an International Perspective*, rev. ed. Baltimore, Md.: Johns Hopkins University Press.
- Hayami, Yujiro, M. Agnes Quisumbing, and Lourdes S. Adriano. 1990. *Toward an Alternative Land Reform Paradigm: A Philippine Perspective*. Quezon City, Philippines: Ateneo de Manila University Press.
- Hesters, Evette D., and Pablo Mabun. 1924. "Some Economic and Social Aspects of Philippine Tenancies." *Philippine Agriculturist* 12(February):367–444.
- Huke, Robert E., and Eleanor H. Huke. 1997. *Rice Area by Type of Culture: South, Southeast, and East Asia, a revised and updated database*. Manila: International Rice Research Institute.
- Indonesia Biro Pusat Statistik. 1996. *Statistical Yearbook of Indonesia 1996*. Jakarta: Badan Pusat Statistik.
- Ingram, John C. 1971. *Economic Change in Thailand, 1850–1970*. Stanford, Calif.: Stanford University Press.
- Intal, Ponciano S. Jr., and John H. Power. 1989. *Trade, Exchange Rate, and Agricultural Pricing Policies in the Philippines*. Washington, D.C.: World Bank.
- Jaffee, Steven, and John Morton. 1995. *Marketing Africa's High-Value Foods*. Dubuque, Iowa: Kendall/Hunt Publishing Co.
- Kosa-ard, Mingsarn Santikarn, and Benjavan Rekasem. 2000. *The Growth and Substitutability of Agriculture in Asia*. Oxford: Oxford University Press.
- Lamb, Geoffrey, and Linda Muller. 1982. *Control Accountability, and Incentives in a Successful Development Institution: The Kenya Tea Development Authority*. Washington, D.C.: World Bank.
- Larkin, John. S. 1972. *The Pampangans: Colonial Society in a Philippine Province*. Berkeley: University of California Press.
- Lewis, W. Arthur., ed. 1970. *Tropical Development, 1880–1913: Studies in Economic Progress*. London: Allen and Unwin.

- Lim, Yongil. 1968. "Input of Tea Industry on Growth of the Ceylonese Economy." *Social and Economic Studies* 17(December):453–67.
- McCoy, Alfred W. 1982. "A Queen Dies Slowly: The Rise and Decline of Iloilo City." In Alfred D. McCoy and E. D. C. de Jesus, eds., *Philippine Social History: Global Trade and Local Transformations*. Quezon City: Ateneo de Manila University Press. pp. 297–358.
- McLennan, Marshall S. 1969. "Land and Tenancy in the Central Luzon Plain." *Philippine Studies* 17(October):651–82.
- Myint, Hla. 1965. *The Economics of Developing Countries*. New York: Praeger.
- . 1971. *Economic Theory and Underdeveloped Countries*. New York: Oxford University Press.
- Pelzer, Karl J. 1945. *Pioneer Settlement in the Asiatic Tropics*. New York: American Geographical Society.
- Pim, Alan. 1946. *Colonial Agricultural Production*. London: Royal Institute of International Affairs; New York: Oxford University Press.
- Republic of the Philippines. 1991. *Philippine Statistical Yearbook 1991*. Manila: National Economic and Development Authority.
- Resnick, Steven A. 1970. "The Decline in Rural Industry under Export Expansion: A Comparison among Burma, Philippines and Thailand, 1870–1938." *Journal of Economic History* 30(March):51–73.
- Rosegrant, Mark, and Peter B. R. Hazell. 2000. *Transforming the Rural Asian Economy: The Unfinished Revolution*. Oxford: Oxford University Press.
- Siamwalla, Ammar. 1992. "Myths, Demons and the Future of Thai Agriculture." Paper presented at *Year-end Conference on Thailand's Economic Structure: Toward Balanced Development*. Bangkok: Thailand Development Research Institute.
- Siamwalla, Ammar, Suthad Sethaboonsarng, and Direk Patanasirawat. 1990. *Thai Agriculture: Resources, Institutions and Policies*. Bangkok: Thailand Development Research Institute.
- Takaya, Koichi. 1985. *Tonan Ajia no Shizen to Tochiriyō* (Nature and Land Use in Southeast Asia). Tokyo: Keiso Shobo.
- Thailand Ministry of Agriculture. 1996. *Agricultural Statistics of Thailand 1995/96*. Bangkok: Division of Agricultural Economics, Office of the Under Secretary of State, Ministry of Agriculture.
- U.N.-FAO. 1971. *1970 World Census of Agriculture: Analysis and Comparison of the Results*. Rome: U.N.-FAO.
- Warr, Peter G., ed., 1993. *Thai Economy in Transition*. Cambridge: Cambridge University Press.
- Warr, Peter G., and Bhanupong Nidhiprabha. 1995. *Thailand's Macroeconomic Miracle: Stable Adjustment and Sustained Growth*. Washington, D.C.: World Bank.
- Wickizer, Vernon D. 1951. *Coffee, Tea, and Cocoa: An Economic and Political Analysis*. Stanford, Calif.: Stanford University Press.
- Wickizer, Vernon D. 1960. "The Smallholder in Tropical Export Crop Production." *Food Research Institute Studies* 1(February):49–99.
- World Bank. 1987. *Thailand: Agro-Industrial Diversification: Issues and Prospects*. Washington, D.C.: World Bank.
- . 1993. *The East Asian Miracle*. Oxford: Oxford University Press.

Productivity Growth and Sustainability in Post–Green Revolution Agriculture: The Case of the Indian and Pakistan Punjab

Rinku Murgai, Mubarik Ali, and Derek Byerlee

This article attempts to determine the long-term productivity and sustainability of irrigated agriculture in the Indian and Pakistan Punjab by measuring trends in total factor productivity for production systems in both states since the advent of the Green Revolution. These measurements over time and across systems have resulted in three major findings. First, there were wide spatial and temporal variations between the two Punjab. Although output growth and crop yields were much higher in the Indian Punjab, productivity growth was higher by only a small margin. Moreover, the lowest growth in productivity took place during the initial Green Revolution period (as opposed to the later intensification and post–Green Revolution periods) and in the wheat–rice system in both states. The time lag between adoption of Green Revolution technologies and realization of productivity gains is related to learning-induced efficiency gains, better utilization of capital investments over time, and problems with the standard methods of productivity measurement that downwardly bias estimates, particularly during the Green Revolution period. Second, input growth accounted for most of the output growth in both Punjab during the period under study. Third, intensification, especially in the wheat–rice system, resulted in resource degradation in both Punjab. Data from Pakistan show that resource degradation reduced overall productivity growth from technical change and from education and infrastructure investment by one-third. These findings imply the need for policies that promote agricultural productivity and sustainability through public investments in education, roads, and research and extension; and that reduce resource degradation by decreasing or eliminating subsidies that encourage intensification of inputs.

The Indo-Gangetic Plain of northern India and Pakistan has one of the largest concentrations of poor people in the world. The agricultural sector, which employs more than half the area's 500 million inhabitants, has long been considered key to food

security and poverty alleviation for this population. Beginning in the mid-1960s, Green Revolution technologies were introduced in the area, including high-yielding modern varieties of rice and wheat, the area's two major crops. This change was supported by investment in irrigation and market infrastructure. As a result, the area experienced a dramatic increase in agricultural production, especially in India's Punjab State and Pakistan's Punjab Province.

But despite this promising beginning, the further intensification of input use since the adoption of Green Revolution technologies has provided lower marginal returns (Byerlee 1992); and the continued intensification of cropping has sometimes caused degradation of the resource base in the form of salinization, overexploitation of groundwater, physical and chemical deterioration of the soil, and pest and disease problems (Fujisaka and others 1994; Siddiq 1994). Consequently, there is now great concern about the potential for productivity growth in irrigated Green Revolution systems and their sustainability over the longer term.

The Debate about Agricultural Performance

Despite the evidence of sharply lower growth rates for food grain yields, there is considerable controversy about aggregate performance of the agricultural sector. In particular, good performance in nonfood crops—such as cotton in Pakistan; oilseeds, fruits, and vegetables in northwest India; and livestock in both countries—may have offset the slowdown in food grains.

Moreover, crop yields are only a measure of partial factor productivity, whereas overall agriculture sector performance is generally measured by total factor productivity (TFP). The TFP approach compares an index of output changes with an index of input changes, thus making it possible to attribute residual output growth to technical progress, changes in input quality, and changes in the physical and economic environment. Experience from industrialized countries suggests that, over the longer run, TFP in the agriculture sector should grow at 1.5 to 2 percent a year, and that one- to two-thirds of that growth will be due to investment in research and extension.

Recent estimates of TFP for agriculture in Pakistan and northwestern India provide conflicting conclusions. For Pakistan, two studies indicate negative TFP growth in the post-Green Revolution period, especially in the Punjab (Azam and others 1991; Ali and Velasco 1994). By contrast, another study (Khan 1994) concludes that TFP in Pakistan grew sharply in the period 1980–92, at an annual rate of 2.1 percent, suggesting that the agricultural sector in that country performed well in recent years. For northwest India, there is little recent evidence on aggregate TFP growth, although the few available studies generally indicate that it was positive (Kaur 1991; Sidhu and Byerlee 1991; Kumar and Rosegrant 1994; Evenson and others 1999). These conflicting results are due in part to the studies' widely varying coverage of inputs

and outputs, methods of valuing inputs, index procedures used to estimate TFP, and levels of disaggregation.

In addition, there is little quantitative evidence of the impact of resource degradation on productivity growth. Thus, with the body of information that now exists, it is difficult to accurately assess productivity growth and sustainability under intensification, and to reach definite policy conclusions about how best to ensure food security and alleviate poverty in the Indo-Gangetic Plain.

Objectives of the Study

The main objectives of this study are, first, to provide comparable estimates of TFP growth in the Indian and Pakistan Punjab since the advent of the Green Revolution and, second, to relate productivity trends to changes in resource quality. The article is organized as follows. The next section outlines the methodology. We then describe major trends in the agriculture sector, especially those related to production performance, input use, and resource degradation. We then present estimates of TFP at the state level and by cropping system. In each case, we decompose output growth into the contributions from growth in input use and growth in TFP. In the subsequent section, we use detailed data on resource degradation in Pakistan to further decompose productivity trends into the effects of technology, resource degradation, human resources, and infrastructure. The final two sections discuss policy implications and summarize our main findings.

Methodology

Our approach is to estimate growth in TFP for three periods corresponding to different phases of technical change: (1) The Green Revolution period itself (1966–74), when input-responsive modern varieties of wheat and rice were widely adopted, leading to an immediate and dramatic increase in production; (2) the input intensification period (1975–85 in India, 1975–84 in Pakistan), when the use of fertilizers and capital inputs increased rapidly; and (3) the post-Green Revolution period (1986–94 in India, 1985–94 in Pakistan), when input use leveled off (Byerlee 1992).

We base our calculations (box 1) on district-level data on all inputs, outputs, and prices, collected from statistical agencies and secondary sources in both states. The data cover the period 1961–94 in India and 1966–94 in Pakistan. Input categories include land, labor, water, machinery, draught animals, fertilizer, and pesticide costs. To minimize aggregation bias in TFP, inputs of different qualities are valued by the price of each quality type. Land is divided into irrigated and unirrigated, labor into skilled and unskilled (based on the rural literacy rate in each district), water into canal

Box 1. Calculating Change in Total Factor Productivity as a Result of the Green Revolution

Of the several ways to measure TFP using different rules for aggregating outputs and inputs (Alston and others 1995), we use the chain-linked Tornqvist-Theil index, because it provides an exact measure of technical change for the linear homogenous translog production function with Hicks-neutral technical change (Diewert 1976). TFP is obtained by taking the difference between the growth rates of the aggregate output and input indices:

$$TFP \approx \ln(TFP_t / TFP_{t-1}) = [\ln(Q_t / Q_{t-1}) - \ln(X_t / X_{t-1})] = \sum_i 1/2 (R_{it} + R_{it-1}) \ln(Q_{it} / Q_{it-1}) - \sum_j 1/2 (S_{jt} + S_{jt-1}) \ln(X_{jt} / X_{jt-1})$$

where Q_t is the aggregate output index, X_t is the aggregate input index, and R_{it} and S_{jt} are the revenue share of output i and cost share of input j at time t , respectively.

and tubewell, and fertilizer into individual nutrient sources (nitrogen, phosphorous, and potassium).

Outputs are aggregated into an output index using district-specific farm harvest prices for crops and market center-specific prices for livestock products.

These data are used to estimate TFP separately for different agro-ecological zones, defined in terms of cropping systems, to avoid the problem of aggregation across heterogeneous regions. This approach enables direct comparison of productivity trends and helps determine whether productivity slowdown and environmental degradation are associated with particular cropping systems and ecologies. In India, the districts are divided into three cropping systems: wheat-rice, wheat-cotton, and wheat-maize. In Pakistan, they are divided into wheat-rice, wheat-cotton, wheat-mixed summer crops (often maize or sugarcane), and wheat-mungbean (or wheat-fallow). The district-level data are then aggregated to quantify TFP growth in terms of the dominant cropping pattern. For Pakistan, where we collected test results from 1971–94 on the quality of groundwater and soil (organic matter, phosphorous content, and soluble salts), productivity growth is also econometrically decomposed into the effects of technology, resource degradation, human resources, and infrastructure. Disaggregated data on resource quality for the Indian Punjab and for the pre-1971 years in Pakistan were not available.

Trends in Production, Input Use, and Resource Degradation

Table 1 describes the production record in the two Punjabs for the three major crops—wheat, rice, and cotton—during the Green Revolution, intensification, and post-Green Revolution periods. During the first period, modern wheat and rice varieties were widely and rapidly adopted in both states. Wheat production increased by more than 7 percent annually, with yield increases accounting for slightly more than half

Table 1. Yield Performance of Major Crops, Indian and Pakistan Punjabs

	<i>Indian Punjab</i>	<i>Pakistan Punjab</i>
Growth rate in yields (%)		
Wheat	[3.6]	[2.2]
Green Revolution	4.7	5.1
Intensification	2.6	1.1
Post-Green Revolution	2.5	2.1
Rice	[4.1]	[-0.3 ^{ns}]
Green Revolution	9.4	4.2
Intensification	2.3	-1.6
Post-Green Revolution	0.7 ^{ns}	-1.4
Cotton	[1.6]	[3.6]
Green Revolution	0.4 ^{ns}	-0.6 ^{ns}
Intensification	0.1 ^{ns}	2.8 ^{ns}
Post-Green Revolution	7.3	8.0
Average yields (kg/ha)		
Wheat		
Green Revolution	2,004	1,246
Intensification	2,750	1,605
Post-Green Revolution	3,643	1,902
Rice		
Green Revolution	1,609	1,320
Intensification	2,777	1,366
Post-Green Revolution	3,246	1,215
Cotton		
Green Revolution	347	288
Intensification	316	267
Post-Green Revolution	504	601

Note: The figures in brackets [] indicate growth rate in the parameter value during overall study period.

ns = not significantly different from zero at 10 percent.

Sources: Ali and Byerlee (forthcoming) and Murgai (forthcoming).

that growth; rice production also grew rapidly, especially in the Indian Punjab. In the post-Green Revolution period, however, yield growth rates in both states decreased to an average of 2 percent a year for wheat and became stagnant or negative for rice, creating concerns that the Green Revolution may not be sustainable.

Adding to these concerns, the gap between yields in the two states widened over time, even though the two states had similar cropping patterns and ecologies. Wheat yields in India during the post-Green Revolution period were nearly double those in Pakistan, although they share similar agro-climatic conditions. Rice yields in India were also much higher during the post-Green Revolution period, although this was known to be partly due to Pakistan's specialization in low-yielding, highly valued Basmati rice. However, cotton yields in Pakistan were higher in both level and growth rate.

The data indicate that the differences between the states in production performance are associated with differences in input use and cropping intensity (table 2). In the first period, the adoption of modern varieties stimulated rapid input intensification in both Punjabs. In India, fertilizer use jumped from 33 to 156 kg of nutrients per hectare of cropped area between the first and third periods; labor use gradually declined; and the use of mechanical power (tractors, harvesters, and threshers) increased from 4.3 to 41 hours per hectare. Pakistan followed the same patterns, although cropping intensities and the use of fertilizer and machinery were considerably lower in all periods.

The data also show considerable degradation of the water and soil resource base in both states. There are indications that the wheat-rice cropping system in the Indian

Table 2. Changes in Cropping Intensity, Irrigated Area, and Input Use, Indian and Pakistan Punjabs

	<i>Indian Punjab</i>	<i>Pakistan Punjab</i>
Cropping intensity (%)	[1.0]	[0.8]
Green Revolution	138.0	117.0
Intensification	158.0	126.0
Post-Green Revolution	174.0	136.0
Irrigated area (%)	[1.7]	[0.3]
Green Revolution	68.6	81.9
Intensification	81.9	85.1
Post-Green Revolution	90.8	86.3
Fertilizer (kg/ha)	[13.1]	[12.6]
Green Revolution	33.0	14.1
Intensification	99.2	48.3
Post-Green Revolution	155.9	86.1
Machines (hrs/ha)	[12.8]	[12.0]
Green Revolution	4.3	1.5
Intensification	15.3	5.7
Post-Green Revolution	41.0	14.8
Tubewells (##/1,000ha) ^a	[11.5]	[3.9]
Green Revolution	27.3	8.2
Intensification	80.6	16.2
Post-Green Revolution	104.4	26.0
Labor (days/ha)	[-1.4]	[-0.9]
Green Revolution	84.2	85.0
Intensification	75.9	98.7
Post-Green Revolution	64.6	71.1

Note: The figures in brackets [] indicate growth rate in the parameter value during overall study period.

ns = not significantly different from zero at 10 percent.

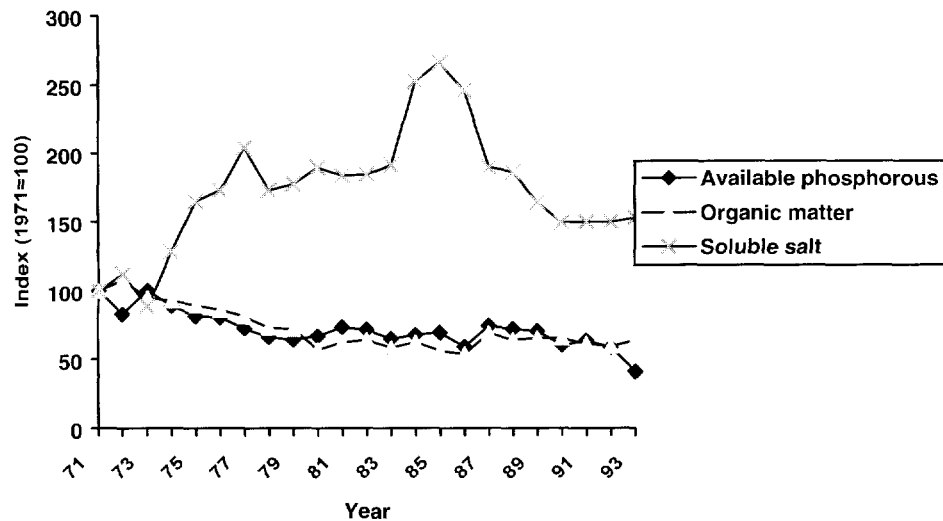
a. Tubewell numbers are not directly comparable. Tubewells in Pakistan are much larger.

Sources: Ali and Byerlee (forthcoming) and Murgai (forthcoming).

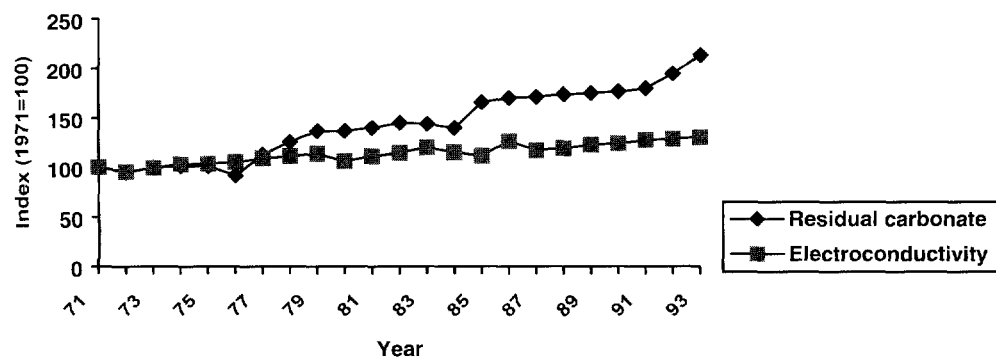
Punjab was hurt by a steep decline in the water table, while rising water levels in the wheat-cotton zone led to severe waterlogging in the wheat-cotton zone. Data from the Pakistan Punjab also confirm a serious problem of waterlogging and salinity, due in part to deterioration in the quality of tubewell water (reflected in a significant increase in residual carbonate and electroconductivity of groundwater). Soil quality in Pakistan (in terms of available soil organic matter and phosphorus) also deteriorated, particularly in the wheat-rice zone (figures 1a and 1b).

Figure 1. a. Indices of Trends in Soil Quality; b. Indices of Trends in Tubewell Water Quality

a.



b.



Source: Ali and Byerlee (forthcoming).

Trend in TFP

Our analysis shows that the Indian Punjab experienced a better overall growth rate than Pakistan, with the states having, respectively, 5 and 3.2 percent annual growth rates in output and 1.9 and 1.5 percent annual growth rates in TFP. But despite these different growth rates, both states experienced similar lags between the adoption of modern varieties and the realization of TFP gains. Contrary to expectations, our evidence shows that TFP in both states barely increased during the Green Revolution period itself, when output growth was most rapid. Instead, TFP gains accelerated during the input intensification period, after the adoption of modern varieties had essentially been completed. Table 3 summarizes state-level trends in aggregated output and input use, and TFP growth in both Punjabs.

We propose three possible reasons for this lag.

- First, empirical evidence from areas of Asia that experienced rapid Green Revolution-induced change suggests that when new technologies were first adopted, inefficiency was fairly high (about 30 percent). In general, high levels of technical inefficiency are due mainly to deficiencies in information and technical skills (Ali and Byerlee 1991), and these were probably serious factors in both Punjabs, where poorly educated farmers switched, in a single generation, from traditional agriculture to complex multiple cropping systems dependent on significant levels of modern inputs.

The increase in technical efficiency a few years later, during the second period, can be attributed to learning by doing, as farmers gained experience using the

Table 3. Growth Rate of Partial Factor Productivity and TFP Indices, Indian and Pakistan Punjabs

	Partial Factor Productivity			Total Factor Productivity		
	Land (%)	Labor (%)	Water (%)	Output (%)	Input (%)	TFP (%)
Indian Punjab	[4.4]	[4.9]	[-8.1]	[5.0]	[3.0]	[1.9]
Green Revolution	4.9	5.4	-22.1	5.9	4.6	1.3
Intensification	4.5	3.8	0.6 ^{ns}	4.3	2.4	1.8
Post-Green Revolution	2.9	5.6	0.6 ^{ns}	4.5	3.0	1.5
Pakistan Punjab	[2.4]	[2.5]	[-1.4]	[3.2]	[1.9]	[1.5]
Green Revolution	2.0	-3.9	-5.9	3.2	4.6	-1.4 ^{ns}
Intensification	2.3	3.0	-1.3	2.8	1.4	1.4
Post-Green Revolution	2.9	6.2	1.5	3.9	1.0	2.9

Note: The figures in brackets [] indicate values during overall study period.

ns = not significantly different from zero at 10 percent.

Sources: Ali and Byerlee (forthcoming) and Murgai (forthcoming).

new technologies; and also to an increase in human capital as education levels rose in both states. Indeed, evidence from India suggests that Green Revolution technological change directly increased the returns to education by spurring greater private investment in schooling, particularly in states such as the Punjab (Foster and Rosenzweig 1996).

- A second reason for the lag in realizing TFP gains is that the introduction of modern varieties was accompanied by significant capital investment by small farmers, particularly in tubewells, during the Green Revolution period. The tubewells are likely to have been underutilized in the short run, until the newly created capacity could be exploited. Such underutilization of quasi-fixed inputs generally leads to an underestimation of TFP gains during the period when investment costs are incurred, but TFP gains become observable as excess capacity is absorbed (Berndt and Fuss 1986).
- Third, low TFP growth during the Green Revolution relates in part to limitations of the conventional method of productivity measurement when technical change is biased toward saving one or more factors (Murgai forthcoming). When technical change is biased in this sense, it is impossible to separate the contribution of technical change from that of factor accumulation, because part of the contribution of technical change is captured in changes in the factor shares used to aggregate inputs. In the case of land- and labor-saving technologies, conventional TFP calculations underestimate the contribution of technical change to growth, particularly during the Green Revolution period.

In contrast to the first two periods, TFP growth slowed slightly in India—from 1.9 to 1.5 percent—during the third period, while in Pakistan it rose sharply, to 2.9 percent, due partly to Pakistan's gains in the livestock sector and partly to its large increase in cotton yields (from 267 kg/ha during intensification to 601 kg/ha during the post-Green Revolution period) and mungbean production. The strong performance of cotton and mungbean followed the introduction of modern varieties of both crops and a sharp increase in pesticide use, although the latter has proven technically and environmentally unsustainable in recent years.

In India, by comparison, irrigated cotton was a much less important crop (Channugam 1994; Kurosaki 1999). Until the reforms of 1994, quantity restrictions on cotton kept India's domestic price below world prices, and nontariff barriers restricted cotton imports (World Bank 1996). Varietal research for irrigated cotton in India consequently lagged, although rain-fed cotton has been very successful in the central part of the country. In addition, because of price policies favoring production for the domestic market, the Indian Punjab produced only short-staple cotton for local consumption, and the long-staple varieties produced by Pakistan for export were never adapted to Indian conditions.

Growth performance varied not only between the states but within each state as well, with the performance of different cropping systems varying widely (table 4). The wheat-rice systems, which benefited from the development of modern semi-dwarf crop varieties, had the lowest rates of TFP growth. In the Indian and Pakistan wheat-rice zones, TFP grew at 1.4 and 0.1 percent a year, respectively. These results confirm widespread concerns that continuous double cropping of cereals, especially wheat and rice, which require very different soil and water management practices, is an unsustainable cropping pattern (Pingali and Rosegrant 1994; Byerlee and Siddiq 1994; Cassman and Pingali 1995; Ali 1996). Deterioration in soil and water quality seemed especially serious in the wheat-rice system, as evidenced by indicators of soil and water quality disaggregated by system in Pakistan (Ali and Byerlee forthcoming). The better performance of this system in the Indian Punjab reflects, in part, the focus on early maturing coarse rice for local consumption, as well as concerted efforts to arrest resource degradation (e.g., through widespread use of gypsum to combat secondary salinity from tubewells).

The wheat-cotton systems in both states, by contrast, had much higher rates of TFP growth. However, the high pesticide use in this system has led to environmental and health damage, which is not taken into account in the growth-accounting method using for estimating TFP.

TFP growth was also relatively high in Pakistan's wheat-mungbean system, where production growth rates were maintained in the post-Green Revolution period due to the release of early maturing mungbean varieties, which allowed a more sustainable cereal-legume rotation. In the wheat-mungbean zone, yield growth rates for wheat were among the highest in the state.

Table 4. Growth Rate of Output, Input, and TFP indices, by Cropping Systems

	Output (%)	Input (%)	TFP (%)
Indian Punjab			
Wheat-rice	5.1	3.7	1.4
Wheat-cotton	5.0	2.5	2.5
Wheat-maize	3.6	1.2	2.4
Overall	5.5	3.0	1.9
Pakistan Punjab			
Wheat-rice	2.4	2.3	0.1
Wheat-cotton	4.1	2.1	1.9
Wheat-mungbean	4.3	2.3	2.0
Wheat-mix	2.6	1.6	1.0
Overall	3.2	1.9	1.5

Sources: Ali and Byerlee (forthcoming) and Murgai (forthcoming).

Decomposition of Productivity Growth

Temporal and spatial differences in productivity growth, taken together, highlight the potential roles of technological change, infrastructure and human capital, and resource degradation in determining TFP growth. Indeed, previous studies have related TFP growth in the Indian and Pakistan agricultural sectors to technological change (tied explicitly or implicitly to research investments), extension systems, infrastructure investments, human capital endowments, and policy reform (Rosegrant and Evenson 1992; Kumar and Mruthyunjaya 1992; Fan and others 2000; Pingali and Heisey 2001). However, there has, until now, been little effort to determine the quantitative impact of resource degradation on productivity growth.

Part of the problem has been the difficulty of agreeing on how to measure the impact of resource degradation on TFP. Some have argued that TFP measurement should incorporate changes in resource quality and externalities, such as water pollution (Herdt and Lynam 1992), while others measure TFP only with conventional inputs and consider resource stocks as a technical constraint that influences trends in TFP (Squires 1992). We prefer the second approach, for three reasons. First, attempts to account for market failure and social costs, such as resource degradation, in estimates of TFP violate the theoretical basis of those estimates (Byerlee and Murgai forthcoming). Second, it is difficult, in practice, to value changes in resource quality, even where these changes can be physically quantified. Finally, in the medium term covered by this article, farmers may not be able to observe resource degradation, and therefore it is exogenous rather than endogenous to decisionmaking (see Policy Discussion). Keeping these considerations in mind, we assess resource degradation concerns in detail by estimating a cost function for Pakistan that relates costs of production to human and physical infrastructure development, technological change, and resource quality.¹ (As noted above, disaggregated data on resource quality for the Indian Punjab were not available.) We use the adult literacy rate to capture the effect of changes in labor quality, and the inverse of the distance of a village from the nearest metal road to quantify the effect of improvement in physical infrastructure. The effect of technological change is proxied by two variables: (1) the proportion of area sown to modern wheat varieties, and (2) cropping intensity, which is a proxy for adoption of modern summer crop varieties that shorten the growing period and thus allow for early planting of winter crops. The effect of resource degradation is estimated through measures of soil and water quality (phosphorus, organic matter, and soluble salts for soil or electroconductivity for water). Region-specific time-trend dummies are included to capture the remaining unspecified effects of technological change, resource degradation, and change in resource productivity not included in the function.

The coefficients estimated in the cost function are used to decompose productivity trends into the effects of technological change, improvements in human resources

and infrastructure, and natural resource degradation. The decomposition was performed by multiplying the negative of the coefficient in the cost function by the system-level rate of change (in percent) per year for each variable included in the cost function.

Based on this analysis, we found that, for the Pakistan Punjab, technological change and improvements in human and physical infrastructure together produced an average growth of 0.94 percent per year, with each accounting for about half the total—and that resource degradation, in aggregate, lowered growth by 0.53 percent per year (table 5).² The combined effect of technological change, improvements in human and physical infrastructure, and resource degradation was negative in the wheat-rice system (i.e., there was an overall increase in unit cost), and the contribution of technological change was highest in the wheat-cotton and wheat-mungbean systems. These results confirm the pattern found in the TFP estimates reported in the previous sections.

Soil and water degradation reduced productivity in all regions, highlighting the effect of natural resource variables on productivity. In the wheat-rice system, resource degradation more than canceled the productivity-enhancing contributions of technological change, education, and infrastructure. The unspecified “other factors” captured by coefficients on the regional time-trend variable also reduced productivity quite strongly in all but the wheat-cotton system. These omitted variables include resource degradation factors that were not measured, such as the development of pest complexes due to inappropriate use of pesticides and to monocropping of cereals. More research is needed to identify the management practices causing such a decline. Moreover, as massive public investment to control waterlogging and salinity is not included in the cost function, which relates only to private costs and returns, the effect of these factors is probably underestimated.

Policy Discussion

Concerns about food security following the food crisis of the mid-1960s led the governments of India and Pakistan to concentrate resources in irrigated areas, such as the Punjab, which held promise for the greatest crop yield increases (Sims 1993). The subsequent gap in agricultural performance and in sources of growth between the two states, despite quite similar agro-ecological potential, seems to relate to differences in nonprice policies toward agriculture, since price incentives, a major policy instrument in both countries, were fairly similar.

At the core of the price policies was a strategy of massive subsidization of fertilizer, credit, power, and irrigation inputs. In India, subsidies on the four major inputs grew at 9 to 12 percent a year in real terms between 1981 and 1993, accounting for between 2.2 and 2.7 percent of gross domestic product by the end of that period (Gulati

Table 5. Decomposition of Total Change in Crop Sector Productivity in Pakistan's Punjab, 1971–94
(percent per year)

Region	Soil and water quality deterioration					Total	Technological change (CI and MV ^b)	Public investment (roads and literacy)	Net effects
	Water electro- conductivity	Soil phosphorous	Soil organic matter	Total soil-soluble salts	Other factors ^a				
Overall Punjab	–0.007	–0.049	–0.137	–0.027	–0.314	–0.534	0.497	0.443	0.406
Wheat-rice	–0.011	–0.087	–0.212	–0.058	–0.925	–1.293	0.339	0.371	–0.582
Wheat-mung	0.001	–0.004	–0.066	0.000	–0.557	–0.626	0.527	0.464	0.366
Wheat-cotton	0.003	–0.064	–0.109	0.001	0.467	0.298	0.727	0.498	1.523
Whcat-mixed	–0.024	–0.027	–0.155	–0.057	–0.306	–0.568	0.350	0.418	0.199

Note: The rate of productivity growth in each production system was evaluated by multiplying the negative of the coefficient in the cost function with the system level rate of change per annum in each factor.

a. Based on the system-specific trend coefficients (converted to percentage terms) in the cost function.

b. CI = cropping intensity, MV = percent area in modern varieties of wheat

Source: Ali and Byerlee (forthcoming).

and Sharma 1995). Subsidies were also an important element of public spending in Pakistan, but with a stronger bias toward large farmers (Sims 1986, 1993). Inputs subsidies in both countries were maintained well beyond an initial period when they might have been economically justified to overcome farmers' risk aversion and to support learning by doing. This was because once established, they were politically difficult to remove. In the Indian Punjab, where small and medium farmers who dominate the electorate were major beneficiaries, the subsidies became even more entrenched than in Pakistan.

At the same time, however, output prices on basic food grains were taxed to maintain low food prices. These controls on food marketing kept grain prices below world prices—a problem compounded by overvalued exchange rates and tariff protection of the nonagricultural sector in both states. Thus, despite the high subsidies, there was a large net transfer of resources out of agriculture. Effective protection rates for food grains over much of this period averaged between –33 and –50 percent (Faruquee 1995; Gulati and Kelley 1999).

In any case, similar pricing strategies in the two states, along with empirical evidence on agricultural supply response and input elasticities from the Asian context, suggest that the differential performance between the states is more likely the result of nonprice factors. This view is supported by other studies showing that investments in rural infrastructure, human capital, and research and extension play a dominant role in influencing supply and productivity growth (Binswanger and others 1993; Fan and others 1999; Rao 1989).

Based on comparisons of public expenditures in agriculture in the two states and on the importance to growth of education and infrastructure, as evident from the cost function analysis, India would be expected to have more rapid input use and productivity growth than Pakistan, where the share of public resources allocated to agriculture was lower (Choudhry and Faruquee 1995). Indeed, India did have more rapid productivity growth, due to massive public investment in rural infrastructure, human capital, and research and extension, and to the subsequent better quality of those factors. India's greater investment in rural infrastructure meant, in particular, that by the mid-1980s, all rural villages in the Indian Punjab were electrified, the density of the road network was well above that of the Pakistan Punjab, and more than 90 percent of the cropped area was irrigated (Fan and others 1999). In Pakistan, by comparison, investment in education and rural infrastructure was much lower (Mujahid-Mukhtar 1991; Rosegrant and Evenson 1992; Faruquee 1995). India also had a relatively better developed network of agricultural research centers and universities, especially in the northwest (Mruthyunjaya and Ranjitha 1998).

Given the relative importance, however, of nonprice incentives for growth, the trend of expanding subsidies at the expense of productivity-enhancing investments raises concerns about the sustainability of growth over the long term if current price-based policies continue (Fan and others 1999). In Pakistan, the total share of public

resources allocated to agriculture has been declining, and irrigation-related expenditures have been particularly impaired, with a 4 percent annual rate of decrease (Ahmad and Kutcher 1992). In particular, continuous underinvestment in operational and management costs has seriously reduced the efficiency of the irrigation system (Chaudhry and Ali 1989), resulting in such problems as regular breaches, excessive seepage, and limited water supplies for the tail reaches of distributaries. Research spending also fell in real terms in the 1990s and accounts for a falling share of the agriculture budget (Choudhry and Faruquee 1995). Across the border in India, even though the share of public resources devoted to agriculture has risen steadily since the early 1980s, subsidies to agriculture have increased three times faster than other expenditures (World Bank 1996).

Apart from crowding out productivity-enhancing expenditures, input subsidies have also been a major cause of overcapitalization, inefficient use of inputs, and a shift in cropping patterns toward water- and fertilizer-intensive crops, thus contributing, in India, to soil degradation, salinity problems, and overexploitation of groundwater (Joshi and Tyagi 1994; Vaidyanathan 2000). In Pakistan, the subsidy on canal water prices has led to inefficient use of water and has contributed to the waterlogging and salinity problem (Ahmad and Kutcher 1992). In addition, the flat rate structure of water and electricity prices, together with a subsidy on tubewell drilling without regulation of the number of tubewells, has distorted the efficient use of water (Johnson 1989). Resource degradation is not, in itself, a reason for policy intervention if it is internalized in producer decisionmaking. In this case, however, distorted policies have led private and social costs to diverge.

Removal of price distortions in the form of input subsidies would be a major step toward arresting resource degradation and encouraging more sustainable systems. It would also encourage higher productivity growth by freeing resources for high-priority public investments in rural infrastructure, education, and research and extension. Since the early to mid-1990s, both Pakistan and India have initiated steps to reduce price distortions in the agricultural sector, but these efforts have not yet been accompanied by an increase in public investments in agriculture.³

Arresting resource degradation will also require a concerted effort on several other fronts. First, the shortage of public investment funds has led to considerable underinvestment in establishing new drainage systems, which are central to resolving the waterlogging and salinity problems. It may be argued that irrigation is a private good and should gradually be handed over to private markets, but drainage is a public good and will remain so, at least beyond the farmfield—and it is beyond the farmfield that most drainage investments are needed (World Bank 1994). New institutional structures for irrigation are now being piloted in Pakistan; these approaches include the devolution of water management to farmers' organizations, and the establishment of public utilities to operate and price water further up the system (Bandaragoda and Firdousi 1992). These measures, too, may help resolve resource degradation problems.

Second, halting resource degradation will require that research systems, which have been oriented toward developing technologies based on packages of modern inputs, place more emphasis on input-efficient and environmentally friendly practices. This will require considerable location-specific research on such themes as integrated pest and nutrient management and cropping systems. It will also require diversifying rotations to include legumes and the use of conservation tillage. Many such practices are information intensive and will require much greater information dissemination and extension efforts. Research systems in both states have shifted direction toward these new priorities in the 1990s.

Finally, a large number of institutions in the two Punjabs have overlapping mandates to address soil and water management problems in irrigated agriculture, and their efforts are poorly coordinated. In the Pakistan Punjab alone, for example, nearly a dozen institutions are working on salinity problems. Information about land and water problems is also institutionally dispersed, as is policymaking (John Mellor Associates and Asianics Agro-Development International 1994). It is important, therefore, to establish a central agency in each Punjab to regularly provide farmers and policymakers with current information on the status of land and water resources in irrigated areas.

Conclusion

In examining the critical issue of long-term productivity and sustainability of irrigated agriculture in the Indian and Pakistan Punjabs, this study confirms previous findings that India experienced much higher and more rapid growth of yields for food crops. However, the results suggest that most of India's higher growth was due to the more rapid growth of inputs. Though overall productivity growth in the Indian Punjab was higher, it was not by a large margin.

The gap in input use and agricultural performance between the two states, despite quite similar agro-ecological potential, seems to relate to differences in nonprice policies that encouraged much faster growth of input use in India than in Pakistan. But while investment (both public and private) plays a central role in productivity growth, there was, in both Punjabs, a considerable lag between investment in infrastructure and Green Revolution inputs and the realization of productivity growth. This seems to relate in part to learning by doing and investment in human capital, which take time to produce improvements in technical efficiency; and in part to the better utilization of lumpy capital investments over time, especially tubewell capacity. For policymakers, this lag suggests that a long-term commitment is needed to realize complementarities between investment in technologies and supporting infrastructure.

The results of this study also raise serious concerns about the long-term sustainability of intensive irrigated Green Revolution systems due to resource degradation. For Pakistan, this study provides the first quantitative evidence of the impact of resource deg-

radation, which is estimated to reduce productivity growth by one-third overall, and in the case of wheat-rice, to practically cancel the effect of technological change. These results, combined with the stagnation of cereal output in recent years, highlight the urgent need for measures that will arrest the problem of resource degradation and maintain the Punjabs' most valuable assets—their irrigated land base.

Notes

Rinku Murgai is an economist in the Development Economics Research Group of the World Bank; Mubarik Ali is an agricultural economist at the Asian Vegetable Research and Development Center; and Derek Byerlee is a lead economist in the Rural Development Department of the World Bank. Rinku Murgai can be reached via e-mail at rmurgai@worldbank.org, Derek Byerlee at dbyerlee@worldbank.org, and Mubarik Ali at mubarik@netra.avrdc.org.tw.

This study summarizes the results of a study sponsored by the World Bank, "Total Factor Productivity Growth in Post-Green Revolution Agriculture of Pakistan and Northwest India." More detailed discussions of the India and Pakistan results can be found in Murgai (1997) and Ali and Byerlee (*forthcoming*), respectively. The authors thank the Research Support Board for funding and two anonymous reviewers for valuable comments.

1. We selected the dual (cost function) approach over the primal (production function) approach for econometrically decomposing productivity growth, because the former has a number of advantages (Alston and others 1995). The use of factor prices, rather than their quantities, as explanatory variables avoids problems of simultaneity that arise when input choices are endogenous with output. Factor prices are more likely to be behaviorally exogenous to a producer. In addition, the dual approach allows estimation of a system of equations comprising the cost function and the system of factor share equations, which results in greater efficiency.

2. Total annual productivity growth estimated through the econometric analysis is 0.41 percent for 1971–94, lower than the 1.30 percent estimated through the index number approach for the corresponding period, but with the same ranking by production system. There are several possible reasons for the difference in productivity growth obtained using the index number and econometric approaches: (a) the TFP growth rate (primal) is computed with input levels held constant, whereas the cost function rate (dual) is computed with input level adjusted optimally to technological change (Antle and McGuckin 1993); (b) the productivity measure obtained from the cost function is net of factor substitution, whereas the index number estimate includes the substitution effect (Ray 1982); and (c) not all the variables related to technological change could be included in the cost function, which therefore might have underestimated technological progress.

3. Pakistan has eliminated its fertilizer subsidy, and India has removed credit subsidies but has yet to eliminate fertilizer subsidies. In India, as part of a broader liberalization program that started in 1991, the irrigation subsidy has also been marginally reduced, but at the expense of the quality and reliability of water delivery. Efforts to reduce the subsidy on rural power have been unsuccessful; in fact, the subsidy has increased by 14 percent a year in real terms between 1991 and 1995 (World Bank 1996). Output markets for the main crops have also been liberalized in both countries, leading to an increase in the prices of rice and cotton in India, and of food grains in Pakistan.

References

The world "processed" describes informally reproduced works that may not be commonly available through library systems.

- Ahmad, M., and G. P. Kutcher. 1992. "Irrigation Planning with Environmental Considerations: A Case Study of Pakistan's Indus Basin." Technical Paper 166. World Bank, Washington, D.C. Processed.
- Ali, M. 1996. "Quantifying the Socioeconomic Determinants of Sustainable Crop Production: An Application to Wheat Cultivation in the Tarai of Nepal." *Agricultural Economics* 14:45–60.
- Ali, M., and D. Byerlee. 1991. "Economic Efficiency of Small Farmers in a Changing World: A Survey of Recent Evidence." *Journal of International Development* 3(1):1–27.
- . Forthcoming. "Productivity Growth and Resource Degradation in Pakistan's Punjab: A Decomposition Analysis." *Economics Development and Cultural Change*
- Ali, M., and L. E. Velasco. 1994. "Intensification-Induced Resource Degradation: The Crop Production Sector in Pakistan." Working Paper 94-14. International Rice Research Institute, Manila. Processed.
- Alston, J. M., G. W. Norton, and P.G. Pardey. 1995. *Science under Scarcity: Principles and practice for Agricultural Research Evaluation and Priority Setting*. Ithaca, N.Y. and London: Cornell University Press.
- Antle, J. M., and T. McGuckin. 1993. "Technological Innovation, Agricultural Productivity, and Environmental Quality." In G. A. Carlson, D. Zilberman, and J. A. Miranowski, eds., *Agricultural and Environmental Resource Economics*. New York: Oxford University Press.
- Azam, Q. T., E. A. Bloom, and R. E. Evenson. 1991. "Agricultural Research Productivity in Pakistan." Islamabad: Pakistan Agricultural Research Council and New Haven, Conn.: Yale University Press.
- Bandaragoda, D. J., and G. R. Firdousi. 1992. "Institutional Factors Affecting Irrigation Performance in Pakistan: Research and Policy Priorities." International Irrigation Management Institute, Colombo, Sri Lanka.
- Berndt, E. R., and M. A. Fuss. 1986. "Productivity Measurement with Adjustments for Variation in Capacity Utilization and Other Forms of Temporary Equilibrium." *Journal of Econometrics* 33:7–29.
- Binswanger, H. P., S. R. Khandker, and M. R. Rosenzweig. 1993. "How Infrastructure and Financial Institutions Affect Agricultural Output and Investment in India." *Journal of Development Economics* 31:337–66.
- Byerlee, D. 1992. "Technical Change, Productivity, and Sustainability in Irrigated Cropping Systems of South Asia: Emerging Issues in the Post-Green Revolution Era." *Journal of International Development* 4:477–96.
- Byerlee, D., and A. Siddiq. 1994. "Has the Green Revolution Been Sustained? The Quantitative Impacts of the Seed-Fertilizer Technology in Pakistan Revisited." *World Development* 22(9):1345–61.
- Byerlee, D., and R. Murgai. Forthcoming. "Sense and Sustainability Revisited: The Limits of Total Factor Productivity Measures of System Sustainability." *Agricultural Economics*.
- Cassman, K. G., and P. L. Pingali. 1995. "Extrapolating Trends from Long-Term Experiments to Farmers Fields: The Case of Irrigated Rice Systems in Asia." In V. Barnett, R. Payne, and R. Steiner, eds., *Agricultural Sustainability: Economic, Environmental and Statistical Considerations*. New York: John Wiley and Sons.
- Chanmugam, N. 1994. "Factors Influencing the Gap in Crop Productivity between the Indian and Pakistani Punjab." M.S. thesis, Cornell University, Ithaca, N.Y.
- Chaudhry, A. M., and M. Ali. 1989. Measuring Benefits to Operation and Maintenance Expenditure in the Canal Irrigation System of Pakistan: A Simulation Analysis." *Agricultural Economics* 3:199–212.
- Choudhry, Y., and R. Faruquee. 1995. "Pakistan: Public Expenditure in Agriculture." Report IDP-153. World Bank, Washington, D.C.
- Diewert, W. E. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4:115–46.
- Evenson, R. E., C. E. Pray, and M. W. Rosegrant. 1999. *Agricultural Research and Productivity Growth in India*. Research Report 109. International Food Policy Research Institute, Washington, D.C.

- Fan, S., P. Hazell, and S. Thorat. 1999. *Government Spending, Growth and Poverty: An Analysis of Interlinkages in Rural India*. Research Report 110. International Food Policy Research Institute, Washington, D.C.
- . 2000. "Government Spending, Growth and Poverty in Rural India." *American Journal of Agricultural Economics* 82(4): 1038–1051.
- Faruquee, R. 1995. "Structural and Policy Reforms for Agricultural Growth: The Case of Pakistan." Agricultural and Natural Resources Division, South Asia Department World Bank, Washington, D.C.
- Foster, A. D., and M. R. Rosenzweig. 1996. "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution." *American Economic Review* 86(4):931–53.
- Fujisaka, S., L. Harrington, and P. Hobbs. 1994. "Rice-Wheat in South Asia: Systems and Long-Term Priorities Established through Diagnostic Research." *Agricultural Systems* 46:169–87.
- Gulati, A., and A. Sharma. 1995. "Subsidy Syndrome in Indian Agriculture." *Economic and Political Weekly* 30(39):A93–A102.
- Gulati, A., and T. Kelley. 1999. *Trade Liberalization and Indian Agriculture*. New Delhi: Oxford University Press.
- Herd, R. W., and J. K. Lynam. 1992. "Sustainable Development and the Changing Needs of International Agricultural Research." In D. R. Lee, S. Kearl, and N. Uphoff, eds., *Assessing the Importance of International Agricultural Research for Sustainable Development*. Ithaca, N.Y.: Cornell University Press.
- John Mellor Associates and Asianics Agro-Development International. 1994. *Institutional Reforms to Accelerate Irrigated Agriculture*. Government of Pakistan and World Bank, Washington, D.C.
- Johnson, R. 1989. "Private Tubewell Development in Pakistan's Punjab: Review of Past Public Programs: Policies and Relevant Research." International Irrigation Management Institute, Colombo, Sri Lanka.
- Joshi, P. K., and N. K. Tyagi. 1994. "Salt-Affected and Waterlogged Soils in India: A Review." Joint Study of the Indian Council of Agricultural Research. New Delhi and the International Food Policy Research Institute, Washington, D.C.
- Kaur, P. 1991. *Returns to Investment on Agricultural Research and Extension*. Ph.D. thesis. Punjab Agricultural University, Ludhiana.
- Khan, M. H. 1994. "The Structural Adjustment Process and Agricultural Change in Pakistan in the 1980s and 1990s." Simon Fraser University, Vancouver. Processed.
- Kumar, P., and M. Rosegrant. 1994. "Productivity and Sources of Growth for Rice in India." *Economic and Political Weekly* 29(53):183–88.
- Kumar, P., and Mruthyunjaya. 1992. "Measurement Analysis of Total Factor Productivity Growth in Wheat." *Indian Journal of Agricultural Economics* 47(3): 451–58.
- Kurosaki, T. 1999. "Agriculture in India and Pakistan, 1900–95: Productivity and Crop Mix." *Economic and Political Weekly* (December):A160–A168.
- Mruthyunjaya, and P. Ranjitha. 1998. "The Indian Agricultural Research System: Structure, Current Policy Issues and Future Orientation." *World Development* 26:1089–102.
- Mujahid-Mukhtar, E. 1991. "The History of Agricultural Development and the Role of Policy: A Comparison of the Two Punjab (1849–1981)." *Indian Economic Journal* 38:39–61.
- Murgai, R. 1997. "Diversity in Economic Growth and Technical Change: A Districtwise Disaggregation of the Punjab and Haryana Growth Experience." Ph.D. diss., University of California, Berkeley.
- . Forthcoming. "The Green Revolution and the Productivity Paradox: Evidence from the Indian Punjab." *Agricultural Economics*.
- Pingali, P. L., and M. W. Rosegrant. 1994. "Confronting the Environmental Consequences of the Green Revolution in Asia." Discussion Paper 2, Environment and Production Technology Division, International Food Policy Research Institute, Washington, D.C.

- Pingali, P.L. and P. W. Heisey. 2001. "Cereal Crop Productivity in Developing Countries: Past Trends and Future Prospects." In J. M. Alston, P. G. Pardey, and M. J. Taylor, eds., *Agricultural Science Policy: Changing Global Agendas*. Baltimore: Johns Hopkins University Press, 56–84.
- Rao, M. R. 1989. "Agricultural Supply Response: A Survey." *Agricultural Economics* 3:1–22.
- Ray, S. C. 1982. "A Translog Cost Function Analysis of U.S. Agriculture." *American Journal of Agricultural Economics* 64:490–98.
- Rosegrant, M. W., and R. E. Evenson. 1992. "Agricultural Productivity and Sources of Growth in South Asia." *American Journal of Agricultural Economics* 74(3):757–61.
- Siddiq, A. 1994. "Sustainability of the Indus Basin: Impact of Tertiary Salinity/Sodicity on Wheat Productivity, Damage Assessment and Future Public Policy." Ph.D. diss., Department of Agricultural Economics, University of Illinois, Urbana-Champaign.
- Sidhu, D. S. and D. Byerlee. 1991. "Technical Change and Wheat Productivity in Post-Green Revolution Punjab." *Economic and Political Weekly* 26(52):A159–A166.
- Sims, H. 1986. "The State and Agricultural Productivity: Continuity versus Change in the Indian and Pakistan Punjab." *Asian Survey* 26(4):483–500.
- . 1993. "Economic Imperatives, Political Risks and Modes of Action: Agricultural Policy Implementation in India and Pakistan." *Journal of Developing Areas* 27:145–62.
- Squires, D. E. 1992. "Productivity Measurement in Common Property Resource Industries: An Application to the Specific Coast Trawl Fisheries." *RAND Journal of Economics* 23(2):221–36.
- Vaidyanathan, A. 2000. "India's Agricultural Development Policy." *Economic and Political Weekly* 35(20):1735–41.
- World Bank. 1994. *Pakistan Irrigation and Drainage: Issues and Options*. Report 11884-PAK. World Bank, Washington, D.C.
- . 1996. *India Country Economic Memorandum—Five Years of Stabilization and Reform: The Challenges Ahead*. Report 15882-IN. World Bank, Washington, D.C.

The Politics of Russian Enterprise Reform: Insiders, Local Governments, and the Obstacles to Restructuring

Raj M. Desai and Itzhak Goldberg

Russia and other countries in the Commonwealth of Independent States that have implemented voucher privatization programs have to account for the puzzling behavior of insiders—manager-owners—who, in stripping assets from the firms they own, appear to be stealing from one pocket to fill the other. This article suggests that asset stripping and the absence of restructuring result from interactions between insiders and subnational governments in a particular property rights regime, in which the ability to realize value is limited by uncertainty and illiquidity. As the central institutions that govern the Russian economy have ceded their powers to the provinces, regional and local governments have imposed a variety of distortions on enterprises to protect local employment.

To disentangle these vicious circles of control, this article considers three sets of institutional changes: adjustments to the system of fiscal federalism by which subnational governments would be allowed to retain tax revenues generated locally; legal improvements in the protection of property rights; and the provision of mechanisms for restructuring and ownership transformation in insider-dominated firms. The aim of these reforms would be to change the incentives that local governments, owners, and investors face; to convince subnational governments that a more sustainable way of protecting employment lies in protecting local investment; to raise the cost of theft and corruption by insiders and local officials; and to allow investors to acquire controlling stakes in viable firms.

Russia implemented real sector reforms—including the privatization of 15,000 enterprises through vouchers—with the goal of making the transition to the market credible, irreversible, and rapid. The objective of the mass privatization program was to establish a critical mass of profit-seeking corporations that were no longer dependent on state support for their survival, and a class of owners willing to invest in their enterprises and manage their restructuring. However, the mass privatization program has brought few tangible benefits to enterprises privatized in voucher auctions. In-

effective corporate governance, little new investment, and the distortions of continuing government intervention have limited the amount of restructuring these firms could undertake in practice. A diverse constellation of indicators suggests that, among firms in transition countries of comparable wealth, Russian firms are ranked at the bottom in terms of output, factor productivity, profitability, and cost-efficiency. Moreover, recent surveys of Russian industrial sectors conclude that the policies of subnational governments tend to favor low-productivity incumbents, protecting them from takeovers and more productive new entrants (McKinsey & Co. 1999; EBRD 1998; Commander and Tolstopiatenko 1996; Earle and Estrin 1997).¹

In this article, we suggest that the absence of restructuring and investment are the consequences of interactions between insiders (manager-owners) and subnational governments within a particular property rights system. Two interrelated features of this system are consequential to problems in the real sector in Russia. First, subnational governments still use enterprises to protect local employment. In the Russian regions, governments face a tradeoff between enforcing cash-based tax collections and maximizing tax revenues, on the one hand, and extracting social benefits in the form of excess employment on the other. If the manager-owners were to maximize the value of the firm, it would be left vulnerable to expropriatory policies that alter the net return on investment—for example, costly regulations, tariffs, or restrictions on layoffs. Anticipating that, managers choose instead to maximize the private benefits of control, and no restructuring is accomplished. The subnational governments, knowing that the taxable revenue of the firm will have been reduced as a result of cash-flow diversion, respond by collecting revenues in kind and enacting policies that force firms to maintain employment levels. These subnational governments, then, are as vested in maintaining the status quo as are enterprise insiders. The governments shield insiders from takeover attempts, obstruct the enforcement of outsiders' property rights, and perpetuate the enterprises as a source of private benefits for the managers, and as a source of social and political benefits for the region.²

Second, control rights are not automatically granted along with ownership rights, due primarily to the unenforceability of investment contracts in the Russian economy. In economies with fully enforceable property rights, investment decisions would be contractually accompanied by the transfer of control rights. Where control rights and cash-flow rights are separable, the transfer of control becomes, to a large extent, a matter subject to considerable uncertainty and managerial discretion. Outsider investors therefore have little confidence that they will achieve effective control even if they buy (nominal) majority stakes in a company, or that subnational governments will resist imposing distortions once the investment takes place.

The following section details the system of enterprise insider ownership that evolved as a result of the mass privatization program. The next section examines the role of subnational governments in using enterprises to protect employment and derive other benefits. It describes recent attempts by subnational governments to reassert control

over local enterprises. The section on disentangling vicious circles outlines three sets of institutional changes needed to transform the incentives that subnational governments, managers, and investors face: (1) the elimination of tax sharing between local, regional, and federal levels; (2) the establishment of a credible system of protection of property rights; and (3) the provision of mechanisms for restructuring and ownership transformation in insider-dominated firms.

Insider Control

Property rights over enterprises were allocated *de facto* during the Soviet era. Well before the collapse of the Soviet Union, state-enterprise directors benefited—even profited—from many of the rights associated with “ownership.” The outcome of the privatization program adopted in 1992 reflected these preexisting institutional constraints (McFaul 1995). Following the 1989 Law on State Enterprises, entrepreneurial state-enterprise directors set up the numerous Gorbachev-era entities—cooperatives, collectives, and joint ventures—that fed off large state enterprises. Article 7 of the law gave directors (and employees) in an enterprise the right to lease its assets; ultimately, these leased assets could be purchased at once or in installments through the cooperative or collective (Grigoriev 1992). Subsequent decrees in 1989 and 1990 on lease and leasing relations expanded the ability of employees to propose lease arrangements. Although leases were to be awarded on a competitive basis, no competitions were held in practice, and nearly all leases were awarded by enterprise insiders to themselves (Frydman and others 1993). In 1991–92, the number of lease enterprises increased from 2,400 to almost 9,500.

The intent of the law, of course, was to decentralize the elaboration of annual plans. Indeed, at the time, some viewed lease buyouts as an embryonic form of privatization, which, if regulated, might engender a smooth (stable, controlled) transition (Filatochev and others 1992). However, the buyouts were little more than profit-sharing arrangements for the management and employees. Buyouts were often set up for arbitrage purposes to secure inputs at subsidized prices, sell goods at uncontrolled prices, and pocket the difference. Directors thus reaped the benefits of their control rights while the costs and liabilities associated with ownership remained socialized.

When the Russian privatization law was passed in 1992, all lease enterprises whose agreements were entered into prior to the passage of the law were treated in a separate category of enterprise. In these circumstances, redemption was carried out as specified in the lease. But in an untold number of cases, the redemption terms were unclear or unspecified.³ In these situations, lease enterprises were privatized with other enterprises, but leaseholders were given priority in purchasing shares. Thus it is difficult to estimate the overlap between enterprises privatized by voucher and enterprises that were part of the leasing program.

Unlike enterprise directors in such countries as Czechoslovakia and Poland, where state-owned enterprise reform laws enabled the Communist Party to remove underperforming managers as late as 1989, the Soviet Union had no such opportunity. Here, enterprise insiders—managers and employees—had both the incentive and the power to defend themselves against President Yeltsin's attempts to reallocate property rights through rapid privatization. In the Soviet Union, workers did not constitute a distinctive group with interests opposed to management. This was because of the paternalistic relationship between managers and workers, the absence of labor markets, the lack of sector-based trade unions, and the direct provision of social services by enterprises. Workers and managers were traditionally more or less united in their common struggle to bargain for lower production quotas (Willerton 1992; McFaul 1993). Consequently, workers made no claims against property rights distinct from managers. Together, workers and managers constituted a pressure group that dedicated itself toward ensuring that enterprise insiders received the numerous privileges that ultimately made their way into the Russian mass privatization program. As former Prime Minister Gaidar himself later conceded, "Beginning in May and June [1992] it was impossible to stand up to the pressure of the industrial lobby."⁴ Thus, the designers of the Russian mass privatization program confronted the following political reality: Preexisting institutional arrangements that distorted free markets could not be wiped clean by governmental fiat.

A central difference between the mass privatization program in Russia and similar programs in Czechoslovakia, the Slovak Federal Republic, and Poland was the tradeability of Russian vouchers. Under the three variants of the Russian program, insiders could purchase up to 51 percent of enterprise shares. However, the tradeability of vouchers allowed insiders to purchase additional shares up to 65 percent on average (Blasi and Shleifer 1996). Tradeable vouchers, then, may have turned out to be an instrument of insider control, rather than the intended means of consolidating outsider ownership.

Despite conflicting evidence on managerial turnover in the past year, there is general agreement that managers and salaried employees continue to hold approximately 50–60 percent of shares in privatized enterprises (Radygin 1999a). Estimates of ownership structure have varied across different enterprise surveys, which are difficult to compare over time. Table 1 presents estimates of stock ownership in privatized Russian enterprises. The greatest variation tends to be for employees and managers as owners, perhaps reflecting the empirical (as well as theoretical) difficulties in separating the two.

Nevertheless, one of the critical shortcomings of the Russian mass privatization program was that it did not enable an institutional concentration of ownership through financial intermediaries, such as investment funds, as did the Czech and Polish programs (Goldberg and others 1997). The Russian designers believed that consolidation would occur spontaneously as the investment funds, acquiring vouch-

Table 1. Estimates of Ownership in Russian Enterprises, 1994–99

(percentage of share capital)

<i>Owner</i>	<i>1994</i>	<i>1995</i>	<i>1996</i>	<i>1997</i>	<i>1999^a</i>
Employees	44–56	39–49	26–43	23–40	36
Managers	9–17	10–17	12–18	12–36	15
Domestic legal entities ^b	7–11	20–23	23–25	22–24	23
Domestic individuals	3–6	9–11	8–12	11–13	16
Foreign owners	1–2	1–2	1–2	4–5	8
State agencies	12–20	11–13	9–11	7–14	3

Note: Estimates vary considerably depending on when, during the course of the year, the survey was taken, and on the sample size.

a. No survey was undertaken in 1998.

b. Includes investment funds, all other firms (including holding companies and financial-industrial groups), and banks.

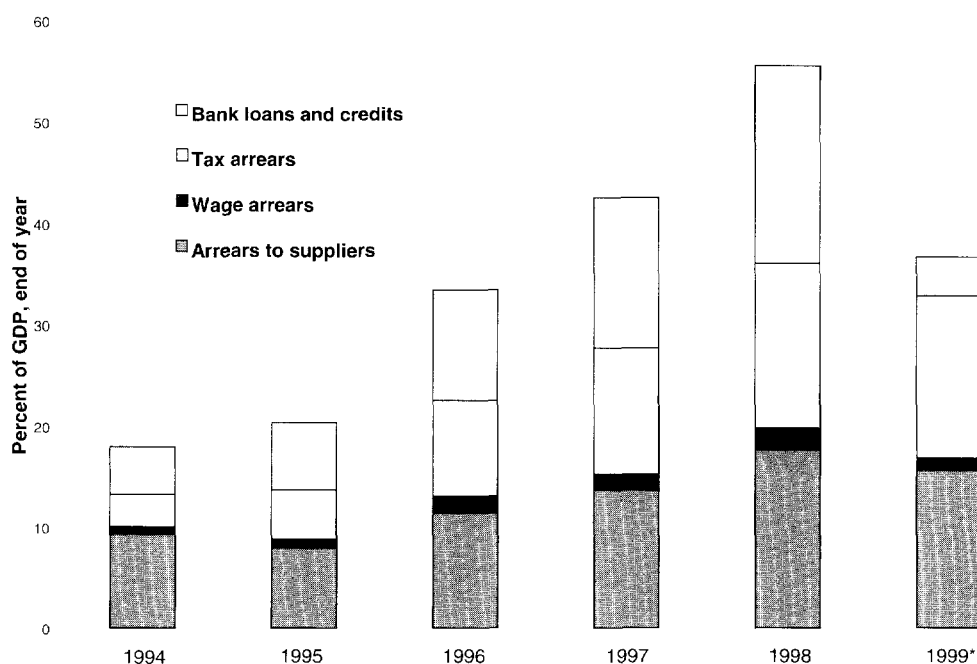
Sources: Figure ranges are taken from the following: for 1994, Blasi and Shleifer (1996) and Radygin (1995); for 1995, Blasi (1997) and Radygin (1996); for 1996, Blasi and others (1997), Radygin and others (1998), and Institute for Economics in Transition (1996); for 1997, Djankov (1998) and data from Institute for Economics in Transition; and for 1999, Radygin (1999b).

ers from employees, would use them to gain sufficient blocs in privatized enterprises. More important, table 1 shows that domestic corporations—including Russian holding companies and industrial groups—and other foreign and domestic outsiders increased their shares in enterprises, on average, from around 20 percent to more than 40 percent between 1994 and 1999.

Even in those enterprises in which management-held shares were nominally in the minority, managers had at their disposal a variety of techniques by which their de facto control exceeded their nominal share ownership. Where outsider shareholders were fragmented, control was effectively ceded to managers. Employees were the largest category of shareholder in most enterprises put through the mass privatization program, yet many of their shares are nonvoting. In addition, managers have often taken to imposing (illegal) bans on selling shares to outsiders, limits on share ownership, and implicit threats (dismissal, wage cuts, and limitations on access to social assets) on workers who violate these rules. Faced with the prospect of takeovers, managers typically issued new shares (also often illegally) to dilute the power of outsiders.

Meanwhile, although the fiscal situation of the Russian government has been improving since early 1998, the enterprise sector has remained in a low-investment, low-productivity trap. Large enterprises have been able to build up significant wage, supplier, and tax arrears, while bartering whatever goods they did produce in their place. According to *Goskomstat's* (the state statistical agency's) own figures, by mid-1999, overdue accounts receivables on enterprise balance sheets amounted to 40 percent of gross domestic product (GDP), while barter reached approximately 50 percent of all industrial sales (RECEP 1999). Figure 1 shows the evolution of arrears and

Figure 1. The Structure of Russian Enterprise Debts, 1994–99



*As of March 1999.

Source: RECEP (various issues)

the composition of the main enterprise liabilities in industry, agriculture, construction, and transportation. Total enterprise arrears have more than doubled as a share of GDP between 1995 and 1998. Most firms have resorted to arrears as a source of trade credit at a time when the volume of bank credit has been shrinking.

Why Insiders Do Not Maximize Value

Every system of corporate governance is a structure of control rights such that suppliers of finance assure themselves that their investment will not be squandered by those to whom it is entrusted. In the classic agency perspective, agency costs arise from (1) monitoring expenditures by the principals; (2) resources spent by insiders to guarantee to outside shareholders that shirking shall be limited, or bonding expenditures by the agents; and (3) the residual loss, or the reduction in the value of the firm that obtains when the owner dilutes his ownership stake (Jensen and Meckling 1976; Williamson 1988). Firms that are owned by managers and salaried employees, from this perspective, should face smaller agency costs of corporate control than

those held by outsiders. There should be no need for a governance system in enterprises wholly owned by those who initiate and implement decisions, provide the financing, and bear the residual risks (and returns) associated with those decisions. Thus we are faced with having to account for the puzzling behavior of insider-managers in Russia—as well as other postsocialist environments in which privatized firms are dominated by insiders—who, in stripping assets from the firms they own, appear to be stealing from one pocket to fill the other.

The central problem of corporate governance in Russia, however, does not involve the protection of minority shareholders or other financiers. Here, manager-owners do not face sufficient incentives to restructure the firm and maximize its value over the long run for three possible reasons. First, due to the problems of delineating property rights, manager-owners perceive titles as uncertain, temporary, and potentially subject to expropriation. With short time horizons, their expected gain from increasing value and share appreciation or dividends is typically less than the value of stripped assets. Second, maximizing value is a reasonable long-term objective only if value can be realized. Manager-owners have limited confidence that if they increase value it can be realized by selling their shares. Insider control is so entrenched that outside investors have little confidence in taking effective control of the firm even if they acquire a significant share. Given this illiquidity of secondary markets, managers have little incentive to increase value and might instead sell the *actual assets themselves* as a way of alienating their control rights. Third, because in many cases the managers do not have formal claim to shares of *employees* (although managers control employees' voting rights), distribution of dividends and realization of capital gains creates the potential for severe conflicts between employees and managers. In addition, barter allows opaque accounting to hide income from employees (as well as from the authorities).

Asset Stripping and Diversion of Cash Flows

Although the original rationale for quick privatization was to prevent asset stripping by managers in state-owned enterprises, manager-owners have significantly degraded enterprise assets. Instead of increasing the value of firms through reinvestment, enterprise manager-owners have typically extracted income streams from firms at the expense of minority shareholders. The managers have diverted cash flows to offshore accounts and shell corporations, concentrated losses among subsidiaries held by outsiders (rather than evenly distributing them between an insider-owned holding company and subsidiaries), and delayed the payment of dividends. Because dividends are taxable and have to be shared with other stockholders, manager-owners are more inclined to withdraw cash flows from their enterprises through fictitious expenses or theft (Goldberg and Desai 1998).

In recent years, the Russian government has made efforts to halt or limit this kind of asset degradation through the enhanced supervisory powers of the Federal Com-

mission on the Securities Market (FCSM) or through the empowerment of the anti-monopoly office. These efforts have been successfully blocked by a coalition of enterprise insiders and subnational governments, together with their allies in the Duma (the Russian parliament). These beneficiaries of partial reform have effectively frozen in place the rents (and the opportunities for theft) generated by distortions in the Russian economy, namely, the absence of further reforms of the property rights and corporate governance regime (Hellman 1998). It has been suggested that enterprises in the interim period between corporatization and privatization are the most prone to asset stripping. However, in the Russian economy, one of the notable pathologies of the control of insiders and subnational governments over enterprises is that asset stripping and cash flow diversions occur under private ownership.

Insider Control and Barter

Tax arrears constitute subsidies by subnational governments to firms; however, there is reason to believe that insider control drives noncash settlements. The Russian system of taxation and payments (for example, the freezing of bank accounts) has been frequently cited as a key rationale for barter (Hendley and others 1999b; Commander and Mumssen 1998; Gaddy and Ickes 1998). If this explanation were true, Russia, in this sense, would have been something of an anomaly. Although it is likely to promote financial disintermediation (cash being less traceable than bank transfers and checks), excessive and arbitrary taxation in middle-income countries does not necessarily prompt barter. In Russia, the State Tax Service has taken to deducting tax payments directly from the ruble bank accounts of debtor companies, and these same firms have used offshore banking whenever cash transactions have been necessary.

Our analysis leads to an additional explanation for the prominence of barter in the Russian economy, namely, that barter is a consequence of the failure of corporate governance. In this sense, barter is both a means of avoiding the payment of private or public debts in cash *and* a way of concealing the real state of affairs not only from tax authorities but from minority shareholders (indeed, even passive shareholders, such as employees). Noncash settlements also enable manager-owners to degrade assets in a less transparent environment.⁵

Therefore, noncash transactions are an instrument by which manager-owners hide transactions from employees and from outsiders—both minority shareholders and tax authorities. Those who argue that changes in enterprise ownership will have no effect until hard budget constraints and payments discipline are ensured have the sequence reversed. In the Russian economy, both nonpayment and barter—in addition to allowing nonviable enterprises to remain in operation—are used by enterprise insiders to preserve their power through opaqueness and fraud. Payments discipline cannot be improved unless there are, in parallel, significant changes in the incentives of manager-owners and subnational governments.

Subnational Government and Enterprise Control

Although the Russian federal government has perhaps given up attempts to intervene in company management, local and regional governments have continued to exert a strong influence over the actions of key enterprises, regardless of whether they have been formally privatized. Indeed, subnational governments show an unwillingness to draw sharp distinctions between public and private property, or to impose hard budget constraints on large enterprises. The governments want to avoid unemployment and to preserve rent-seeking opportunities for vested interests. Thus there is a continuing source of ambiguity in (and unenforceability of) property rights and ownership.

The central organs of the Russian state have lost their powers of regulation and control to the provinces; therefore, municipal, *oblast* and *okrug* (subnational government units), and autonomous-republic administrations have seized these powers. This represents a bid to maintain the economic operations of major local enterprises as well as to protect the local workforce from economic change. One recent study concludes that the devolution of economic control to the regions has preserved the subnational administrative and hierarchical structures of the Soviet system, as well as the power and influence of those who manage them (Ericson 1999). Attempts by subnational governments to protect employment by preventing firms from changing production lines or employment levels constitute a significant distortion of enterprise operations. Anticipating such distortions, managers have often engaged in a “preemptive” diversion of cash flows.

Moreover, in certain regions (especially in one-company towns), the nexus of interdependence between government officials and enterprise management typically grants management quasi-governmental powers. These powers include influence over the apparatus that executes and enforces the law in a given locality, making the enforcement of shareholder rights difficult despite a panoply of recent legislation intended to protect shareholders (Law on Joint-Stock Companies 1996, Company Law 1996, Privatization Law 1997, Investor Protection Act 1999). In those cases where there is significant residual shareholding by the state, there is often tacit cooperation between government representatives on supervisory boards and management opposed to takeovers, consolidation of outsider ownership, or enforcement of the ownership prerogatives of outsiders. These, of course, are long-standing allegations for which systematic evidence is difficult to marshal. Studies have detailed several incidents of managerial obstruction, foot-dragging, and noncompliance (Johnson 1997; Black and others 2000).

The Effects of Tax Bargaining and Tax Sharing

For the largest Russian enterprises, tax payments have been bargained ex post. Tax payments with goods and through the use of offsets have allowed firms to economize

on cash payments to tax authorities. Domestically, these firms have been driven further into the barter economy to avoid leaving a cash trail for the state to pick up. Because subnational governments have been unwilling to enforce the liquidation of insolvent firms, profitable companies as well are encouraged to masquerade as bankrupt ones. Thus both profitable and insolvent Russian firms have incentives to avoid taxes.

Subnational governments in the Russian Federation share all revenues collected, thus drawing taxes from a common pool of taxable resources. Regional and local governments have little incentive to preserve the tax-paying ability of local firms, but rather are more likely to preempt federal tax collectors by accepting services in lieu of tax payments—something federal authorities cannot do easily.

The fact that goods or services used to pay taxes may not have an immediate use constitutes an implicit subsidy granted by the tax collector to the firm and may explain the increasing use of this practice among regional and local governments interested in maintaining employment levels and the social benefits that firms grant employees. Moreover, bartered goods, in-kind payments, and promissory notes are usually greater in price than their cash equivalents (Commander and Mumssen 1998).⁶ Pricing noncash instruments at a premium can serve as an implicit subsidy. Tax authorities issuing tax offset notes are often complicit with inflated pricing because it enables the state to increase its revenues. The managers are interested in inflating revenues because it magnifies the perceived size of the enterprise, which may prove useful as the firm bargains with subnational governments.

Subnational Governments and Enterprise Restructuring

Incumbent managers of loss-making enterprises may have given up on the long-term route of deriving wealth through increasing the value of the enterprise. Selling their shares is a limited option because of the costliness of exchanging ownership rights in the Russian economy. Although the rational short-term strategy of the manager-owners is asset stripping, the longer-run strategy is underutilization of the remaining assets, barter, and accumulation of arrears. Subnational governments tolerate this to maintain employment levels in their regions. The employment incentive also explains the liberal forbearance granted by tax authorities toward enforcing tax debt collection from these firms. Eventually, the enterprise may be viewed as an instrument of social policy in a particular region, as well as a real estate bequest to the heirs of the manager-owners.⁷

Subnational governments actively obstruct enterprise restructuring by perpetuating soft budget constraints and in some cases by imposing formal limitations on enterprise activities. Subsidies can take the form of preferential tax treatment, implicit regionally channeled subsidies through discounts on utility bills, and favored status in public procurement, all of which are intended to prevent companies from shut-

ting down and laying off employees. This puts potentially productive companies at a cost disadvantage, blocking investments and growth on their part.⁸

More directly, regional and municipal governments may effectively ban companies from laying off excess workers. Local authorities have the means to discipline disobedient managers, for example, by subjecting them to troublesome fire, safety, health, and other inspections.⁹ Finally, local officials may require bribes or otherwise extract rents that can remain hidden from other levels of government and, unlike taxes, are not shared.

Renationalization

Regional and municipal governments have reasserted their claims to property rights in the wake of the August crisis. Since mid-1998, de facto renationalizations of previously privatized property have taken place among several well-known corporations. The Belgorod iron ore combine, Alkar Aluminum in Sverdlovsk, the Krasnoyarsk Metallurgical Works, the Mikhailovsk Iron Works, Tatneft, Kamaz, Avtovaz, Zil, and Moskvitch all underwent partial renationalization by the end of 1998. In 1999, further takeovers occurred in the regions of Sverdlovsk, Ulyanovsk, Krasnoyarsk, Voronezh, Primorye, Chelyabinsk, and Moscow (Ericson 1999).

In oil-producing regions, shares of several oil companies found their way into regional governments or regional government-owned companies—including Komineft (Yamalo-Nenetsky) and ANKH (Irkutsk). First, several regional investment vehicles under the protection of local governments were set up to consolidate local or regional government holdings in important local industries. In some cases, these were simply regional government-owned holding companies, which may have attempted to increase their shares through a capital increase. In other cases, the subnational governments restructured debts owned by corporations by converting those debts into equity. Second, companies invoked a 1996 presidential decree on wage and tax arrears, repaying their tax debts to regional and federal budgets by issuing new stocks. In one case—YUKOS—newly issued shares wound up in the hands of a private owner (in this case, the Menatep Group), and the revenues from those shares were transferred to the budget. In all the other cases, government shares have increased in companies under this arrangement (Radygin 1999a).

Disentangling the Vicious Circles

Attempts to reform the property rights system in the Russian Federation have often involved the following dilemma: The enforcement of the rules needed to improve corporate governance is typically in the hands of those who derive significant rents from ensuring that these very rules are *not* enforced. Therefore, the success of any

reforms will hinge on the extent to which the incentives of the major actors are transformed. In this section, we analyze some prospective reform agendas intended to improve the institutional underpinnings of the relationship between federal and subnational authorities in the corporate-legal environment and of the market for corporate control.

Fiscal Federalism and Subnational Incentives

Analyses of the role of federalism in the development of markets confirm the major role that regional and local governments can play in the growth of the private sector or in its obstruction. Russian subnational governments have impeded the growth of the private sector through a combination of taxation, regulation, and corruption (Shleifer and Treisman 2000).¹⁰ “Spontaneous decentralization” in Russia left its fiscal system without stable or transparent rules for tax sharing and expenditure assignments and failed to provide subnational governments with incentives to support their local tax base or to limit rent extraction from local firms (Lavrov 1996).

In general, administrative and fiscal decentralization in Russia has produced several adverse effects. First, both regional and local governments—municipalities and *rayons* (districts)—face few incentives to limit their own extraction of rents from local firms (Blanchard and Shleifer 2000; Frye and Shleifer 1997). Second, capture by large enterprises, whose continued operations were vital to subnational economies, led subnational governments to protect them from competition by new entrants (McKinsey & Co. 1999). Third, without clearly specified rules, revenue sharing remains highly discretionary, and revenues retained at subnational levels are subject to yearly local-regional-federal bargains (Freinkman and others 1999). Local governments, in particular, benefit the least from these ad hoc arrangements: Most of their tax revenue is shared with regional governments, and any increase in local revenues will be almost entirely offset by changes in shared revenues. Fourth, low tax retention rates among local governments have created few incentives to support enterprise restructuring, to limit the extraction of rents from local firms, or to prevent capture by local firms (Zhuravskaya 2000; Shleifer 1997).¹¹

Comparison with China suggests striking differences between Chinese and Russian fiscal federalism in terms of the powers of the central government and the mechanisms of intergovernmental finance. In China during the reform period, a central government under the tight control of the Communist Party was in a stronger position to reward and punish local officials than was the case in the politically fragmented Russian Federation. The Chinese Communist Party retains the power to appoint and fire governors through the Party-State’s system of promotions and punishments. Consequently, top provincial officials are bureaucratically integrated with the central government (Huang 1996). Moreover, unlike the Russian state, China’s central government did not see the need to renegotiate revenue sharing agreements with its

regions. In addition, since 1979, the Chinese provinces have assumed primary responsibility for governing the provincial economies. Local expenditures are closely linked to revenue generated, and thus the provinces have been limited in their ability to borrow. Consequently, revenue-sharing arrangements in China are characterized by far less center-province bargaining and by greater local retention of tax revenues (Jin and others 1999).

Russian federal authorities have recently embarked on a campaign to bring wayward regions back in line and to recentralize several powers that had devolved to regional governors. In the short run, a more direct approach may be to eliminate tax sharing between the three levels of government. The current system of revenue sharing has fostered a race between federal, regional, and local authorities to collect revenues from the same taxable base—through multiple overlapping taxes, special levies, or extortion. This is a “tragedy of the commons” of sorts in that none of the three levels of government has an incentive to halt its own extraction of rents from tax-paying firms or to protect the corporate sector from these distortions. Eliminating tax sharing and allowing local budgets to be the sole beneficiaries of taxes paid by local enterprises would end the overlapping claims of different layers of government to the common tax pool. It would also provide the subnational authorities with stronger incentives for maintaining a favorable investment climate in their localities. In particular, local budgets should be the single beneficiary of taxes paid by local enterprises.¹²

Property Rights Enforcement and the Legal Environment

In recent years, the legal basis of the Russian property rights regime has been substantially improved by two laws designed to prevent the abuse of power in corporations. For example, the Company Law of 1996 attempts to break the power of the general directors by prohibiting them from serving simultaneously as the chairman of the board. To prevent insider trading, the Company Law restricts the general director from conducting important transactions without the consent of the board. The Company Law also offers a number of legal recourses for minority shareholders. These include the right of appeal against any resolution adopted by the annual general meeting, the ability to seek the annulment of major transactions, and the right to sue for invalidation of a transaction involving conflicts of interest.

The Investor Protection Act of 1999 promotes the equitable treatment of shareholders by strengthening disclosure requirements for joint stock companies. It also grants the securities commission FCSM greater freedom to initiate court proceedings on behalf of individual shareholders. As a result of the Investor Protection Act, the other Russian self-regulated organizations—the Professional Association of Registrars, Transfer Agents, and Depositories (PARTAD) and the National Association of Securities Market Participants (NAUFOR)—can investigate the violation of shareholder rights by registrars, depositories, custodians, or brokers. Although Russian legisla-

tion does not specifically provide for class action suits, a joint complaint by shareholders is possible during appeals procedures and by filing a collective claim through the FCSM. Although the FCSM has successfully brought a number of cases to court and has won more than 80 percent of them, it still has relatively few weapons for forcing compliance with existing laws compared with its counterparts in Organisation for Economic Co-operation and Development (OECD) countries.

Nevertheless, governmental foot-dragging (particularly at the subnational level) and the absence of impartial, independent adjudication have repeatedly obstructed the enforcement of the growing body of sophisticated Russian corporate laws (Carothers 1998).¹³ The adoption of these laws and expanded protection for investors have not prevented asset stripping or the defrauding of minority shareholders and creditors. Thus the credibility of legislation can be severely undermined when weak states introduce well-intentioned but unenforceable laws; indeed, such laws may be counterproductive (EBRD 1999). For example, both the Company Law and the Investor Protection Act have been thwarted by the widespread use of offshore corporate vehicles, which have diverted assets or cash flows from companies to affiliated parties (that is, management or controlling shareholders) hidden from the principal regulators (Jesover and Nestor 2000).

Efforts to strengthen the rule of law as it relates to enterprise restructuring should focus not only on the form and content of the laws themselves but on increasing government compliance, judicial independence, and safeguards against corruption. In part, the solution may simply require increased funding. Currently, the shortfalls in federal funding for the judiciary are supplemented (inadequately) by funds from *oblast* governments. This encourages courts that often favor local interests, which are frequently unable to enforce judicial decisions and are left vulnerable to corruption. But ultimately, long-term improvement in formal enforcement mechanisms will come to rely on institutional building efforts within the Russian systems of courts—especially the *arbitrazh* courts, which adjudicate cases of a commercial nature—and bailiffs.

The establishment of a genuine system of administrative review in the Russian Federation would serve as a check on executive authority and would be a step toward ensuring predictability and transparency in governmental decisionmaking. Well-designed, legally based administrative reviews require the content and time frame of decisions to be made public and clearly specified in advance. They require government bodies to supply explanations when public inquiries are made and to provide mechanisms for recourse. Such mechanisms can include rights of appeal, judicial review through administrative courts or tribunals, or external complaint and mediation through the use of ombudsmen. Such reforms are expensive and painstaking but integral to curbing abuses of public office.

The bailiff corps, which is charged with enforcement of the decisions of the courts, is typically underfunded in transition countries. Some countries, such as Slovakia

and Latvia, have semi-privatized or are in the process of privatizing enforcement based on French-style civil codes, according to which bailiffs behave much like public notaries. The bailiffs are certified by the state and charge government-regulated fees, but they are permitted to advertise and find their own customers.

Enterprise Restructuring and Ownership Transformation

Loss-making private or privatized firms should be properly rationalized by market mechanisms that either promote restructuring from within or allow coalitions of outsiders to come together and dislodge incumbent owners. We examine two such mechanisms in turn.

Creditor-led restructuring. In Russia, improved bankruptcy procedures that have been in place since 1998 should have facilitated the reorganization of insolvent companies. Two fundamental changes to Russian bankruptcy procedures were put in place following the passage of the 1998 Bankruptcy Act. First, the law removes the requirement of the previous law that courts undertake independent valuations of debtors' assets and liabilities to determine whether a debtor was insolvent. Regional courts often refused to base insolvency on cash flows (basing them on net worth instead), thus preventing creditors from being paid. Second, creditors faced no incentives to cooperate in their petitions against a debtor, resulting in coordination problems. The new law bases insolvency on cash flows, that is, an inability to service debts over a prescribed period. In addition, the new law requires a creditors' committee to implement interim management of an insolvent firm and to share in court costs.

Any creditor, including a government authority that owns tax debts, can initiate a bankruptcy petition against a delinquent firm; however, in practice there are four critical problems in relying on bankruptcy procedures to initiate ownership transformation. First, traditional creditors may be simply unwilling or unable to initiate bankruptcy. Due to the tight industrial and financial links between suppliers and customers in monopolistic settings, it is unlikely that suppliers would bankrupt their main customers. Moreover, in Russia as in most countries in the Commonwealth of Independent States, total liquid liabilities in the financial sector comprise less than 15 percent of GDP (compared with 50–70 percent in Central Europe and the Baltic states). Thus banks are in a fragile position for initiating debt conciliation. Most outstanding liabilities are not to banks, but rather to governments, workers, suppliers, and nonbudgetary funds.

Second, given the requirement of creditor coordination, it is impossible to ensure that an investor can take ownership of a bankrupt firm. The creditors are given equity shares pro rata, and they must agree to a sale to a third party. Bankruptcy is collective in nature; therefore, all creditors must share the costs of the bankruptcy proceedings and all must agree on external management and other matters.

Third, ownership transformation based on bankruptcies will ultimately rely on court procedures and on the capacity of the judicial system to ensure success. These are significant burdens to entrust to an already overloaded system.

Fourth, bankruptcy is a disruptive and often lengthy process. If an amicable agreement is not forthcoming, the proceedings might encourage the antipathy of regional authorities opposed to bankrupting a company to change ownership.

In view of these problems, using arrears as leverage to induce restructuring or the sale of assets may be a more appropriate solution than bankruptcy. The major creditors in Russia are the state-owned utilities (UES and Gazprom), local gas and power distributors, and tax authorities. However, neither voluntary restructuring nor sale is feasible as long as subnational governments continue to subsidize loss-making enterprises by tolerating tax arrears, energy companies continue supplying enterprises that do not pay, and insiders can costlessly divert the income streams of their enterprises.

Investor-led restructuring and markets for control. Alongside bankruptcy rules and other mechanisms that permit holders of debts to acquire control of loss-making enterprises, market economies usually possess a variety of other means by which control rights in enterprises can be transferred to new owners. The development of these markets for corporate control and takeovers in the Russian Federation would constitute major progress in releasing loss-making enterprises from their role as a source of private rents for owners and political benefits for public officials.

Experience suggests that investors are more willing to contribute capital when the web of firm obligations is untangled and settled, and when the acquisition of controlling stakes in enterprises is ensured. Along these lines, government-supported mechanisms can be established to resolve certain outstanding liabilities, while at the same time offering residual stakes and converted tax debt to equity to new investors. The objective behind such a move would be to create new majority owners by converting tax debt into equity that would be sold immediately to external investors in competitive auctions. These converted shares, along with residual shares owned by government authorities, would constitute significant percentages of share capital in several cases, and would effectively dilute the controlling shares of insiders. Box 1 proposes an example of such a mechanism through investment-based ownership transformation.

There are two major pitfalls to building takeover markets by government fiat: investor capture and government capture. The risk of investor capture is simply that takeover markets will, once again, open the door to current enterprise insiders eager to expand their control over productive assets in the Russian economy. As with the loans-for-shares program, there is a risk that firms acquired in takeovers will go to investors with hidden connections to the enterprise who have submitted inflated bids, restricting competition for corporate control. Government capture will occur if takeovers result in the renationalization of private enterprises. Piecemeal takeovers—by

Box 1. Investment-Based Ownership Transformation: A Pilot for the Russian Regions

1. A regional government receives a technical assistance/institutional building loan from a multilateral donor to hire an investment brokerage. The investment brokerage selects a pool of companies based on the following criteria: (i) basic viability; (ii) a minimum stock of tax arrears; and (iii) an agreement between the regional tax authorities and the management-owners to dilute the latter's share in ownership in exchange for part of the proceeds from the transaction. The regional government, interested in collecting revenue from overdue and uncollectible arrears, convinces management, *under threat of bankruptcy*, to accept the dilution.
2. The regional government, according to the procurement rules of the multilateral lender, conducts an open tender to select an investment brokerage, in effect, to reprivatize the selected pool of companies. The formula for compensation of the investment brokerage is based on a fixed fee that covers the expenses of valuing the whole pool, in addition to a success fee for each company successfully resold.
3. The selected investment brokerage conducts a tender for the sale of selected companies. The value of these companies will provide a benchmark tax debt-to-equity conversion rate, and the bids will, in effect, represent the market valuation of this rate.
4. The investment brokerage, as part of the sale of the enterprise shares, will assist in the conversion process of the outstanding tax debts to equity. The conversion would take place only as part of the transaction.
5. Given the layoffs likely following the investment-based ownership transformation, such a scheme should be supported by a compensatory program to assist the unemployed and to reduce their dependence on enterprise-based social resources (for example, a severance pay program financed by an international financial institution).

In this example, all debts are held by tax authorities; the example leaves out liabilities to suppliers, lenders, and labor. Most important are the arrears of the companies to the local gas and power distribution companies, which are often controlled by regional governments. In principle, all of these liabilities could be converted to equity and auctioned in a similar fashion, with the proceeds from the sale being divided among all the claimants pro rata. The control of local governments over suppliers—mainly the distribution companies—could help in coping with the creditors' coordination problem (referred to in the section on enterprise restructuring and ownership transformation). It is advisable to convert all claims and sell them simultaneously to avoid the unintentional dilution of a buyer's expected ownership.

We have avoided distinguishing between levels of government. Although a large proportion of total tax arrears is currently held with the federal government, the probability of collection of these is lower than that of the debts to the regional governments. The investment-based ownership transformation will have to include agreements between the federal and regional creditors about offsetting these debts as part of the fiscal relationship between the two.

Source: Desai and Goldberg (2000).

which a governmental body converts tax debts into shares, then awaits an acceptable investor—can also encourage government administrators to restrict bidding to favored parties or to otherwise bias the auction in the interim.

Investor-led restructuring and takeover markets require that the property rights of investors be protected from expropriation or other arbitrary distortions. An important component of this protection is that owners of private property and buildings also hold the right to the land on which these assets are located. Land privatization

will raise the investment potential and collateral value of privatized enterprises and consequently will prevent governments from exercising control through their ownership of the land. In the same vein, rules for expropriation, forfeiture, and nationalization should be critically examined to limit state intervention in private ownership.

Conclusion

In the absence of property rights enforcement, there are stronger incentives for theft than for restructuring. In manager-owned firms in the Russian Federation, managers can benefit far less from the appreciation of their stock than from siphoning off profits. Moreover, the complex of mutual dependence between subnational authorities and enterprises encourages local and regional governments to rely on enterprises for social benefits rather than as sources of local development. In the Russian Federation, these two groups face few incentives to promote new investment or to respect the rights of shareholders or minorities. The Russian enterprise dilemma can be partially resolved through a combination of institutional reforms aimed at changing the incentives of subnational governments, improving enforcement, and providing mechanisms for enterprise restructuring and for outsiders to acquire controlling interests in firms.

The proposition that conventional liberal reform will provide universal benefits in the long run has not persuaded vested interests that have acquired control over the political process to change the status quo. Ultimately, the sustainability of the reforms suggested here will depend on the presence of constituencies that can lobby, persuade, or otherwise circumvent the coalitions that support existing economic distortions. Whether such constituencies capable of demanding reform are emerging in the Russian Federation will become clearer as the aftermath of the financial crisis and the policies of President Putin's government unfold. Here we mention some conditions under which two constituencies—investors and subnational governments—could be effective.

Investors are either reformed enterprise insiders or outsiders. As we have suggested, once insiders recognize that the subsidization of loss-making enterprises is not sustainable, they may consider selling assets or restructuring voluntarily to raise finance. At the same time, outside investors must be confident that they can acquire sufficient control in enterprises. Several types of outside shareholders—investment companies and nonbanking financial institutions, domestic enterprises, and foreign investors—are likely to respond favorably to the strengthening of enforcement regimes. Several investment companies, for example, have become active players in the Russian takeover market. The better-known brokerages facilitate mergers and acquisitions on behalf of foreign and local investors. There is some indication that shareholders in Russian firms have begun to increase their activism in Russian boardrooms. In this regard, it is encouraging that empirical evidence suggests that outsiders' shares in

Russian enterprises are steadily rising (see table 1), and that companies increasingly rely on the resources of the legal system.¹⁴

Subnational governments comprise another constituency. Changing the fiscal incentives for regional and local governments, in particular, allowing greater control over locally generated tax revenues, will raise the costs of rent extraction from firms in two ways. First, it will encourage governments to lift formal and informal restraints on restructuring. Second, these governments are likely to take steps to protect investors in local firms, to comply with and enforce laws, and, consequently, to limit theft by management. In this way, subnational authorities will recognize that a more sustainable way of protecting employment lies in providing managers with incentives to increase the value of their enterprises and attract outside investment.

The critical role that regional governments play in limiting enterprise restructuring in Russia should no longer be ignored. However, it is important to keep in mind the potentially positive role the regions could play in reviving growth in the real sector. The proposals in this article are intended to be a step in this direction.

Notes

Raj M. Desai is with the Edmund A. Walsh School of Foreign Service at Georgetown University. Itzhak Goldberg is with the Private and Financial Sector Development Department, Europe and Central Asia Region, at the World Bank. Raj M. Desai can be reached via e-mail at desair@georgetown.edu and Itzhak Goldberg at igoldberg@worldbank.org.

1. Russia's labor productivity has fallen from 30 percent of the level in the United States in 1991 to 19 percent in 1999.

2. For a game-theoretic elaboration of this argument, see the annex in Desai and Goldberg (1999).

3. Leases would run between 5 and 15 years, and a redemption clause simply allowed the lessor to purchase the property at the end of the term. But in many cases, the property sharing arrangements on the conclusion of the lease were extremely vague—typically based on principles of self-government and worker participation, but leaving unspecified the governance structure of the enterprise.

4. The "industrial lobby," in Gaidar's retelling, consisted of the Russian Union of Industrialists and Entrepreneurs; its political affiliate, the Civic Union; the Russian Assembly of Social Partnership (an umbrella trade union); and, most important, the legislators in the Congress of People's Deputies. See *New Times* (1993).

5. In a typical scheme, for example, managers could transfer assets from company A, in which minority shareholders have a stake, to company B (for example, a management-controlled pocket company), in which they do not in the following way. Managers of A require A to sell some output to B at a controlled price. B pays A in goods rather than in cash, then sells the output on domestic or international markets at normal prices. Under these conditions, managers are able to turn A into a cost center and B into a profit center as in a normal case of transfer pricing. But the addition of interfirm barter without the premium allows managers to hide A's true cost of production and to exaggerate the amount transferred to B.

6. According to Commander and Mumssen (1998:18), in addition, the fact that goods used to pay taxes may not have an immediate use constitutes an implicit subsidy granted by the tax collector to the firm.

7. Without a property tax on unutilized property offset against a positive profits tax for profitable enterprises, there are few disincentives to hold enterprise real estate for this purpose.

8. Obsolete (subscale and/or inefficient in their use of energy) steel and cement plants avoid shut-downs by paying for only a fraction of their energy bills—their largest cost component. Because these

companies are often the major employers in a town, municipal and regional officials go to great lengths to keep them operating. Regional governments channel implicit federal energy subsidies to these companies by letting arrears to federal suppliers (Gazprom and UES) accumulate at the local gas and electricity distribution companies. These energy distribution companies are often under effective control of the regional governments, making their bankruptcy practically impossible. Subsidies slow down recovery in many manufacturing sectors by preventing upgrading investments and industry consolidation in and around viable industrial assets (McKinsey & Co. 1999).

9. The Novgorod region of Russia is a rare positive example of what can be done by regional governments in today's Russia. It managed to attract more foreign direct investment than almost any other Russian region, including nearby St. Petersburg, by removing red tape, facilitating access to land, and offering tax holidays to investors. As a result, the region has enjoyed economic growth since 1995, and over half of industrial output comes from productive foreign companies.

10. Russia has failed "to construct a democratic, fiscally stable, federal order on the ruins of a Communist state," and the significant power accorded regional potentates has provided much space for regional and local patron-client relationships (Shleifer and Treisman 2000).

11. Shleifer (1997) finds that Russian *oblasts*, compared with Polish regions, received much higher portions of their total revenue from the central government, prompting Russian regions to devote greater efforts to bargaining for federal transfers rather than to fostering growth in the local private sector and the local tax base.

12. At the time of this writing, there have been several indications that the Russian government is inclined to follow these reforms. See Economist Intelligence Unit (2000).

13. It has been argued, in the context of the recent wave of efforts to promote "rule of law" in developing and transition countries, that there are three basic categories of rule of law reform (Carothers 1998). The first type concentrates on efforts to reform or rewrite the laws themselves. The second type emphasizes reform in the basic legal institutions, especially the judiciary and the bar. The third type aims at increasing government compliance with the law and true judicial independence. Most international rule-of-law aid to date has focused on the first two types of reform; the third—the most difficult and important—has received much less attention. Against this background, the rule of law in the context of enterprise restructuring in Russia would focus on laws, institutions, and government compliance (including implementation of laws and prevention of corruption) as they relate to corporate governance and enterprise restructuring.

14. A recent study finds that improvements in contracting procedures, monitoring, and enforcement provide benefits for both transacting partners, coordinating expectations during implementation and making the outcome of disputes more predictable (Hendley and others 1999b).

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Black, Bernard, Reiner Kraakman, and Anna Tarasova. 2000. "Russian Privatization and Corporate Governance: What Went Wrong?" *Stanford Law Review* 52:1731–1808.

Blanchard, Olivier J., and Andrei Shleifer. 2000. "Federalism with and without political centralization: China versus Russia." Working Paper No. 7616. National Bureau of Economic Research (NBER).

Blasi, Joseph. 1997. "Corporate Ownership and Corporate Governance in the Russian Federation." In Ira Lieberman, Stipon Nestor, and Raj M. Desai, eds., *Between State and Market: Mass Privatization in Transition Economies*. Washington, D.C.: World Bank in association with the OECD.

Blasi, Joseph, and Andrei Shleifer. 1996. "Corporate Governance in Russia: An Initial Look." In Roman Frydman, Cheryl W. Gray, and Andrzej Rapaczinski, eds., *Corporate Governance in Central Europe and Russia*. Budapest: Central European University Press.

- Blasi, Joseph, Maya Kroumova, and Douglas Kruse. 1997. *Kremlin Capitalism: The Privatization of the Russian Economy*. Ithaca, N.Y.: ILR Press.
- Carothers, Thomas. 1998. "The Rule of Law Revival." *Foreign Affairs* 77(2):95–106.
- Commander, Simon, and Christian Mumssen. 1998. "Understanding Barter in Russia." Working Paper 37. EBRD, London.
- Commander, Simon, and Andrei Tolstopiatenko. 1996. "Enterprise Restructuring and the Structure of Compensation: Why Is Unemployment Low in the FSU?" Policy Research Discussion Paper, Policy Research Department, World Bank, Washington, D.C.
- Desai, Raj M., and Itzhak Goldberg. 1999. "The Vicious Circles of Control: Regional Governments and Insiders in Privatized Russian Enterprises." Policy Research Working Paper 2287. World Bank, Washington, D.C.
- . 2000. "Stakeholders, Governance, and the Russian Enterprise Dilemma." Finance and Development. IMF, Washington, D.C.
- Djankov, Simion. 1998. "Enterprise Restructuring in Russia." In Harry Broadman, ed., *Russian Enterprise Reform: Policies to Further the Transition*. World Bank Discussion Paper No. 400. World Bank, Washington, D.C.
- EBRD. 1998. *Transition Report 1998*. London: EBRD, p. 130.
- Earle, John, and Saul Estrin. 1997. "After Voucher Privatization: The Structure of Corporate Ownership in Russian Manufacturing Industry." Discussion Paper No. 1736. Center for Economic Policy Research (U.K.).
- Economist Intelligence Unit. 2000. *Russia Country Briefing*. London: Economist Intelligence Unit.
- Ericson, Richard. 1999. "The Post-Soviet Russian Economic System: An Industrial Feudalism?" Working Paper 140. SITE, Stockholm.
- Filatovchev, Igor, Trevor Buck, and Mike Wright. 1992. "Privatization and Buy-Outs in the U.S.S.R." *Soviet Studies* 44(2):265–82.
- Freinkman, Lev, Daniel Treisman, and Stepan Titov. 1999. *Subnational Budgeting in Russia: Preempting a Potential Debt Crisis*. Washington, D.C.: World Bank.
- Frydman, Roman, Andrzej Rapaczinski, John S. Earle, and others. 1993. *The Privatization Process in Russia, Ukraine, and the Baltic States*. CEU Privatization Reports 0968–5278, v. 2. New York: Central European University Press, pp. 20–22.
- Frye, Timothy, and Andrei Shleifer. 1997. "The Invisible Hand and the Grabbing Hand." *American Economic Review* 87(May):354–58.
- Gaddy, Clifford G., and Barry W. Ickes. 1998. "Russia's Virtual Economy." *Foreign Affairs* 77(September–October):53–67.
- Goldberg, Itzhak, and Raj Desai. 1998. "Corporate Governance in Russia." In Harry Broadman, ed., *Russian Enterprise Reform: Policies to Further the Transition*. World Bank Discussion Paper 400. World Bank, Washington, D.C.
- Goldberg, Itzhak, Gregory Jedrzejczak, and Michael Fuchs. 1997. "IPO-PLUS: A New Approach to Privatization." Policy Research Working Paper 1821. World Bank, Washington, D.C.
- Grigoriev, Leonid. 1992. "Ultior Property Rights and Privatization: Even God Cannot Change the Past," in Anders Åslund, ed., *The Post-Soviet Economy: Soviet and Western Perspectives*. New York: St. Martin's Press.
- Hellman, Joel. 1998. "Winners Take All: The Politics of Partial Reform in Post Communist Transition." *World Politics* 50(2):203–34.
- Hendley, Kathryn, Peter Murrel, and Randi Ryterman. 2001. "Law Works in Russia: The Role of

- Legal Institutions in the Transactions of Russian Enterprises." In Peter Murrell, ed., *Assessing the Value of Law in Transition Economies*. Ann Arbor: University of Michigan Press. Forthcoming.
- Huang, Yasheng. 1996. *Inflation and Investment Controls in China: The Political Economy of Central-Local Relations during the Reform Era*. New York: Cambridge University Press.
- Institute for Economics in Transition. 1996. *Annual Report 1996*. Moscow: IET Press.
- Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3(4): 305–60.
- Jesover, Fianna, and Stilpon Nestor. 2000. "Synthesis Report of the 2nd OECD Meeting of the Corporate Governance Roundtable for Russia on Shareholder Rights and Equitable Treatment." Moscow, 24–25 February.
- Jin, Huihui, Yingyi Qian, and Barry Weingast. 1999. "Regional Decentralization and Fiscal Incentives: Federalism, Chinese Style." Stanford University, Stanford, Calif. Processed.
- Johnson, Juliet. 1997. "Russia's Emerging Financial-Industrial Groups." *Post-Soviet Affairs* 13: 333–65.
- Lavrov, Alexei. 1996. "Fiscal Federalism and Financial Stabilization." *Problems of Economic Transition* 39(1): 83–94.
- McFaul, Michael. 1993. "Russian Centrism and Revolutionary Transformation." *Post-Soviet Affairs* 9: 196–222.
- . 1995. "State Power, Institutional Change, and the Politics of Privatization in Russia." *World Politics* 472: 210–43.
- McKinsey & Co. 1999. *Unlocking Economic Growth in Russia*. Washington, D.C: McKinsey Global Institute.
- New Times. 1993. "Gaidar v. Gaidar." *New Times* 17, p. 18.
- Radygin, Alexander. 1995. "The Russian Model of Mass Privatization: Governmental Policy and First Results." In *Privatization, Conversion, and Enterprise Reform in Russia*. Boulder, CO: Westview.
- . 1996. "Securities Market Development and its Relationship to Corporate Governance." Paper presented at OECD conference on capital markets.
- . 1998. "Dominant Shareholders and Performance of Industrial Enterprises." *Russian Economic Barometer* 1: 8–41.
- . 1999a. "Ownership and Control of Russian Industry." Paper presented at the OECD-USAID Conference on Corporate Governance in Russia, Moscow, 31 May–2 June 1999.
- . 1999b. "Post-Privatization Corporate Governance in Russia: Singular Path or Typical Transition Trajectory." Paper presented to the conference on Eastern Transition Trajectories, Grenoble, December 1999.
- Russian-European Center for Economic Policy (RECEP). 1999. *Russian Economic Trends*. Moscow: RECEP.
- Russian-European Center for Economic Policy (RECEP). Various issues. *Russian Economic Trends*. Moscow: RECEP.
- Shleifer, Andrei. 1997. "Government in Transition." Working Paper. Harvard Institute for Economic Research. Cambridge, Mass..
- Shleifer, Andrei, and Daniel Treisman. 2000. *Without a Map: Political Tactics and Economic Reform in Russia*. Cambridge, Mass.: MIT Press.
- Willerton, John. 1992. *Patronage and Politics in the U.S.S.R.* Cambridge: Cambridge University Press.
- Williamson, Oliver. 1988. "Corporate Finance and Corporate Governance." *Journal of Finance* 43: 567–91.
- Zhuravskaya, Ekaterina. 2000. "Incentives to Provide Local Public Goods: Fiscal Federalism, Russian Style." Working Paper. Institute for the Study of Economies in Transition, Stockholm.

Introduce a Friend or Colleague to

OXFORD
UNIVERSITY PRESS

THE WORLD BANK RESEARCH OBSERVER

Simply photocopy this page, fill in the name and address, and Oxford University Press will send a **FREE** sample copy of *The World Bank Research Observer* without obligation!

Name: _____

Address: _____

City: _____

State: _____

Zip: _____

Country: _____

0104infrfa/wbrofr



Contact Oxford University Press

In the Americas:

Oxford University Press
2001 Evans Road, Cary NC 27513
Tel: (919) 677- 0977 Fax: (919) 677-1714
E-mail: jnlorders@oup-usa.org

Elsewhere:

Oxford University Press
Great Clarendon Street, Oxford OX2 6DP, UK
Tel: +44 (0) 1865 267907 Fax: +44 (0) 1865 267485
E-mail: jnl.orders@oup.co.uk

If you're not already subscribing to *Oxford Review of Economic Policy*, here's what you're missing...

OXFORD
REVIEW OF
ECONOMIC
POLICY

Editor
J. A. K. J. J. J.



Oxford Review of Economic Policy is a refereed journal with an international authorship. Each issue of *OXREP* concentrates on a current theme in economic policy, and comprises an assessment and a number of articles.

Articles are presented in a non-technical manner, and are aimed at a wide audience including government, business, and policy-makers as well as academics and students.

OXREP is now FREE ONLINE!
access full-text at no charge:
www.oxrep.oupjournals.org

Does your library subscribe to *Oxford Review of Economic Policy*?

Why not use this form to recommend a subscription?

To recommend a subscription to *OXREP*, please photocopy this page, complete the information below, and send it along to the humanities collection development librarian.

I have reviewed this journal and would like to recommend a library subscription to *OXREP*, published by Oxford University Press. Thank you.

Name _____

Title/Department _____

Date _____

Signature _____

Oxford Review of Economic Policy

ISSN 0266-903X 2001

Volume 17 (4 issues)

Institutions: \$270/£160

I wish to subscribe to *Oxford Review of Economic Policy*

Name (please print) _____

Address _____

Zip/Post code _____ Country _____

☐ I enclose the correct amount
(check payable to Oxford University Press)

☐ Please bill me

Please charge my MCard / Visa / AmEx / Diners (circle one)

Card number _____

Start date _____ Exp. date _____

Signature _____

0106ecopol/form01



2001 Subscription Rates Volume 17, 2001 (4 issues)

- ☐ **Special Offer** for First Time Individual Subscribers: \$55/£30
- ☐ Individuals: \$79/£45
- ☐ Institutions: \$270/£160
- ☐ Students (Please provide copy of current ID): \$44/£24

Single issues are available upon request. Air-expedited delivery fee is included for subscriptions outside North America. Individual subscriptions must be paid for by personal check or credit card and sent to private addresses only. Canadian residents add 7% GST.

To begin your subscription, call 1-800-852-7323 (USA & Canada). Or call 919-677-0977. Fax: 919-677-1714.

Outside the Americas, call +44 (0)1865 267907 or fax +44 (0)1865 267485.

In the Americas, please return to:

Oxford University Press, Journals Marketing (fma),
2001 Evans Road, Cary, NC 27513, USA.

Elsewhere, please return to:

Oxford University Press, Journals Marketing (fma),
Great Clarendon St., Oxford, OX2 6DP, UK.

Journals in Economics from OXFORD UNIVERSITY PRESS

Journal of International Economic Law

Journal of International Economic Law

Dedicated to encouraging thoughtful and scholarly attention to the relation of law to international economic activity.

www.jiel.oupjournals.org

Contributions to Political Economy

Contributions to Political Economy

A forum for academic discussion of original ideas and arguments drawn from lines of thought associated with the works of classical political economists Marx, Keynes, and Sraffa.

www.cpe.oupjournals.org

Contemporary Economic Policy

A journal of Western Economic Association International, *CEP* publishes quality research and analysis on policy issues of widespread concern.

www.cep.oupjournals.org

Oxford Economics ONLINE!

Subscribers can BROWSE TABLE OF CONTENTS and VIEW SAMPLE ISSUES. Register free of charge for the Table of Contents Alerting Service and an e-mail will be sent every time a new issue appears!

For further details go to: www.economics.oupjournals.org

Oxford Economic Papers

A general economics journal, *OEP* publishes refereed papers in economic theory, development and history, as well as applied economics and econometrics.

www.oep.oupjournals.org

Economic Inquiry

Economic Inquiry

A journal of Western Economic Association International, this peer-reviewed, general economics journal publishes original research from top scholars in the field.

www.ei.oupjournals.org

The Review of Financial Studies

The Review of Financial Studies

Provides significant new research in financial economics, striving to establish a balance between theoretical and empirical studies.

www.rfs.oupjournals.org

Contact Oxford University Press

0104econfa/wbrofa

In the Americas:

Oxford University Press

2001 Evans Road, Cary, NC 27513

Tel: (919) 677-0977 Fax: (919) 677-1714

E-mail: jnlorders@oup-usa.org

Elsewhere:

Oxford University Press

Great Clarendon Street, Oxford OX2 6DP, UK

Tel: +44 (0) 1865 267907 Fax: +44 (0) 1865 267485

E-mail: jnl.orders@oup.co.uk

THE WORLD BANK

1818 H Street NW
Washington, DC 20433, USA
World Wide Web: <http://www.worldbank.org/>
E-mail: researchobserver@worldbank.org



ISBN 0-19-850910-3