

# Effects of Colombia's Social Protection System on Workers' Choice between Formal and Informal Employment

*Adriana Camacho*

*Emily Conover*

*Alejandro Hoyos*

The World Bank  
Development Economics Vice Presidency  
Partnerships, Capacity Building Unit  
August 2013



## Abstract

This paper examines whether the Colombian government's expansion of social programs in the early 1990s, particularly the publicly provided health insurance, discouraged formal employment. Using household survey data and variation across municipalities in the onset of interviews for the SISBEN, the instrument used to identify beneficiaries for public health insurance,

it shows robust and consistent estimates of an increase in informal employment of approximately 4 percentage points. Similar results are obtained using an alternative dataset, consisting of a panel of individuals interviewed for the first and second SISBEN. The findings suggest that marginal individuals optimized when deciding whether to participate in the formal sector.

---

This paper is a product of the Partnerships, Capacity Building Unit, Development Economics Vice Presidency. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at [econover@hamilton.edu](mailto:econover@hamilton.edu).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# Effects of Colombia's Social Protection System on Workers' Choice between Formal and Informal Employment\*

Adriana Camacho

Emily Conover

Alejandro Hoyos

JEL classification codes: I11, I18, O17.

Keywords: Social Protection, Incentives, Informal Sector, Health Insurance, Social Protection, Colombia.

---

\*Adriana Camacho is an Associated Professor in the Economics Department and CEDE at Universidad de Los Andes, Bogotá, Colombia; her email address is [adcamach@uniandes.edu.co](mailto:adcamach@uniandes.edu.co). Emily Conover (corresponding author) is an Assistant Professor in the Economics Department at Hamilton College; her email address is [econover@hamilton.edu](mailto:econover@hamilton.edu). Alejandro Hoyos is a Ph.D. Student in Economics at the University of Chicago; his email address is [ahoyos@uchicago.edu](mailto:ahoyos@uchicago.edu). The authors are very grateful to Anna Aizer, Kathryn Anderson, Donald Conover, Alain de Janvry, Amy Damon, Chang-Tai Hsieh, David Levine, Edward Miguel, Enrico Moretti, Suresh Naidu, Elisabeth Sadoulet, Dean Scrimgeour and two anonymous referees. Thanks to participants at the NEUDC conference, seminario CEDE, the Hamilton-Colgate economic seminar series, ICESI seminar participants, LACEA Conference, AEA meetings, IZA-World Bank Conference, and IEN-LACEA.

Expansions of government-funded social programs generate controversies regarding the perverse incentives that these programs may create. Opponents claim that despite their intended benefits, people become dependent on such programs and under-report or hide their income to ensure their eligibility. Informal employment is another possible unintended consequence of social programs. In this paper, we examine the impact of Colombia's public health insurance, the Subsidized Regime (SR), on workers' choice of formal or informal employment.

Informal employment, which is common in many developing countries, is particularly high in Latin America. Galiani and Weinschelbaum (2012) report an average formality rate of 50% between 2002-2004 across a group of 12 Latin American countries. In addition, a World Bank report indicates that during the 1990s, informality rates increased in several Latin American countries (Perry et al. 2007). In Colombia, the average informality rate from 1990–2005 was 51%.<sup>1</sup> From a social welfare perspective, informality is often seen as undesirable since the informal sector encompasses constrained and low productivity firms, where workers have unstable jobs, are unprotected by labor laws, and are unlikely to receive benefits such as health insurance or pensions.

For some workers, informal employment is not always undesirable. Maloney (2004) indicates that informality status may be preferred by some workers because it grants them greater independence and work schedule freedom. Informality may also be favored if the cost of taxes or social security contributions exceeds workers' valuation of the services they provide. This situation is exacerbated when the government provides informal workers with comparable free or discounted services, which is the case for the SR in Colombia (Levy 2008; Pagés and Madrigal 2008)<sup>2</sup>.

The upward trend in informality rates in some Latin American countries coincides with the

---

<sup>1</sup>This rate is calculated using a benefit-based definition that refers to an employee who does not receive social security benefits through employment. In 2003, the International Conference of Labor Statisticians (ICLS) established the benefit-based definition of informality as the official one.

<sup>2</sup>See Loayza (1996) and De Soto (1989) for characterizations of the informal labor market and its determinants.

expansion of social program provisions.<sup>3</sup> The cost of health insurance through employment in Colombia is often cited as anecdotal evidence of a powerful disincentive for formal employment. However, prior to this paper, no systematic empirical evidence has been presented to confirm or reject this hypothesis.<sup>4</sup>

To determine whether the SR expansion generated incentives for informal employment, we use two datasets and different identification strategies. The first comes from repeated cross-sections of the Colombian Household Survey. Using individual-level data and controlling for municipality and year effects, we estimate whether there was an increase in informal employment after the municipality started conducting the interviews needed to determine eligibility for the SR (known as SISBEN interviews). These interviews were implemented at different points in time across municipalities, which allows us to disentangle the effects of the expansion of the health insurance provision from other labor market reforms and macroeconomic shocks occurring at the national level during the same period of time. In section III, we show that pre-existing informality rates are uncorrelated with the timing of the SISBEN, and we find robust and consistent effects of an increase in informal employment of approximately 4 percentage points. This finding corresponds to approximately 8% of the labor force not entering the formal sector and an increase in the government's health expenditures of approximately 11%.

Our second dataset is a panel of people interviewed in the first and second SISBEN. Using information about an individual's eligibility status (as determined by her Poverty Index Score from the first SISBEN)<sup>5</sup> and individual fixed effects, we corroborate the increase in informal employment found with the household survey dataset. Additional estimations that account

---

<sup>3</sup>See Grosh (1994) for a review of 30 social programs in 11 Latin American countries.

<sup>4</sup>For example, in a news article reported by *Presidencia de la República* (Feb. 2006), the Minister of Social Protection indicated that people's valuation of the public health insurance program (the Subsidized Regime (SR)) was high enough to create incentives to discourage formal employment. A study by Santamaria et al. (2008) reports that almost half of SR recipients (49%) indicated that they were not willing to switch to formal employment because it would mean losing their (SR) benefits.

<sup>5</sup>We use data from 1994–1997 for the first SISBEN and from 2003–2007 for the second SISBEN. The algorithm used to determine eligibility changed from the first to the second SISBEN.

for SR eligibility and heterogeneous effects by household characteristics also indicate consistent results of an increase in informal employment.

The rest of the paper is organized as follows: in the next section, we provide background information on the health sector reform. In section II, we review articles that examined the effects of other reforms that occurred in Colombia during the period of our study as well as studies of informal employment in other Latin American countries. In the data section, we describe the datasets used in the analysis. In Section IV, we present the empirical results for informal employment using the two datasets and different identification strategies. We conclude in Section V.

## **I Background**

In 1990, Colombia adopted a new constitution that emphasized social program spending for the poor. Law 100 of 1993 reformed the social security system, including pensions, workplace accidents, complementary social services, and health insurance. The health sector reform aimed to achieve universal health insurance coverage through the creation of two regimes: a Contributive Regime (CR) and a Subsidized Regime (SR).

The main difference between the SR and all other changes to the social security system is that all but the SR were implemented nationally in January 1994. Unlike the SR, the other changes did not require the SISBEN. Using the variation in the implementation dates of the SISBEN across municipalities and year fixed effects as well as a discontinuous eligibility threshold for the SR, we are able to identify the effects that are due to the SR and not due to the other national changes that occurred when Law 100 came into effect.

The CR, which was implemented in the whole country after the passing of the law, made health insurance through employment mandatory regardless of occupation. The SR was de-

signed for the unemployed poor and was implemented at different points in time in different municipalities. Eligibility for the SR was determined using a Poverty Index Score calculated from answers to the SISBEN, which is a questionnaire collecting individual- and household-level demographic and dwelling information (Castañeda et al. 2005).<sup>6</sup>

The SISBEN was conducted at the municipal level in a similar way to the population census, with door-to-door interviews in lower-income neighborhoods. This was an unprecedented system, and nothing of this magnitude had ever been implemented in Colombia. As explained by officials, not all municipalities had the resources or infrastructure in place to conduct a census of the poor (SISBEN), so implementation dates varied across municipalities. Importantly, these dates are not correlated with the pre-reform informality rates in the municipalities. To identify the results, we use this variation in implementation dates and the eligibility indicator from when the system was introduced.

The government agency that designed the SISBEN indicates that although the instrument was designed to select beneficiaries for different social programs, during the period of our study, it was almost exclusively used for the SR, the largest program in terms of beneficiaries and costs (DNP, 2003).<sup>7</sup> This exclusive use was so extensive that initially, many people thought that being interviewed for the SISBEN was equivalent to becoming eligible for the SR. In reality, households were eligible for the SR only if they had a Poverty Index Score below a certain threshold, as determined by the answers to the SISBEN.

If a worker is formally employed, the worker and the employer are required to make contributions to the CR regardless of the worker's occupation or the SISBEN score of his household. Independent workers can also enroll in the CR by paying 5% of their wages.<sup>8</sup> Employers who do

---

<sup>6</sup>In the first SISBEN, the Poverty Index Score was generated using a principal components methodology that took into account demographic and dwelling characteristics at the individual and household levels. It was designed to capture long-term living conditions and was assigned at the household level.

<sup>7</sup>For details on how the SISBEN has been used, see Camacho and Conover (2011) and Camacho and Conover (2013).

<sup>8</sup>From 1995–2007, the cost to non-independent workers was 12% of a person's wages, 25% of which were paid

not comply with the legislation can be sued by the workers. In exchange for these contributions, employees and their direct dependents (spouse and children, or one parent) have access to a range of health services and medications, known as the *Plan Obligatorio de Salud* (POS).<sup>9</sup>

The SR is financed with a 1% transfer from the CR and with local and central government funds. All members in a household eligible for the SR, regardless of their relationship to the household head, have free access to a package of services and medications known as the *Plan Obligatorio de Salud Subsidiado* (POS-S).<sup>10</sup> This package is less comprehensive than the CR's POS.

Giedion and Uribe (2009) explain that beneficiaries of both the contributive and subsidized regimes have pre-established packages of services and benefits. The CR includes all levels of care, whereas the SR package “covers most low-complexity care and catastrophic illnesses but provides only limited coverage for most hospital care and provides no short-term disability coverage” (Giedion and Uribe 2009). In particular, the SR and CR offer similar coverage for public health education and outreach, preventive care services, basic dental care, and catastrophic care. The CR provides more generous coverage for outpatient services, inpatient services and maternity care and provides sick leave (not covered by the SR).

To prevent dual enrollment, a worker cannot be covered by both the CR and SR. Participation in the CR disqualifies a worker from the SR even if he has a score below the eligibility threshold. The structure of the system raised questions about whether it generated incentives for informality because for those eligible for the SR who were willing and able to engage in informal employment, it reduced health insurance costs for both the employer and employee. In this paper, we explore whether some workers who were eligible for the SR (because of a Poverty

---

by the employee and 75% by the employer. Health and pension contributions were not bundled for the period of our study.

<sup>9</sup>In English, POS stands for “Obligatory Health Plan.” For more information on the details of the POS and POS-S, see Tafur (1998).

<sup>10</sup>In English, POS-S stands for “Subsidized Obligatory Health Plan.”

Index Score below the threshold) and were formally employed (receiving the CR) opted out of making mandatory contributions to the CR and became informal, either by remaining employed and stopping their contributions to the CR or by switching to a non-contributory job.

## II Related Literature

From 1990–2005, the Colombian labor market underwent several legislative changes and economic reforms. Several studies analyze the impact of a change in labor taxes (Kugler 1999; Kugler and Kugler 2009), but these studies do not specifically address the effect of the health insurance provision on informal employment. Mondragón et al. (2010) find that an increase in non-wage costs and in the minimum wage has a significant and positive effect on informal employment and a negative effect on informal wages. Gaviria and Henao (2001) report that a strategy used by households faced with a loss of employment is to participate in informal sector activities. Using data from the National Statistical Agency (DANE), these authors report an upward trend in informal employment<sup>11</sup> between 1996 and 2000 for both men and women (p.28). Unlike the studies above, we do not look at changes in labor taxes or the minimum wage, but we analyze the effects of the SR on informal employment. This analysis is important given that the cost of the SR is estimated to be 1.2% of GDP and because new reforms to the system are currently being considered in Colombia.<sup>12</sup>

Other researchers have recently examined similar questions in other Latin American countries. For instance, in Argentina, Gasparini et al. (2009) find evidence of increased informality generated by a large conditional cash transfer program.

In Mexico, using the roll-out of Seguro Popular in different cities over time and controlling

---

<sup>11</sup>The definition of informality they use is as follows: workers in firms with fewer than 10 employees, non-professional independent workers, domestic workers, and unpaid family workers who work at least 15 hours

<sup>12</sup>One of the reforms that was implemented recently makes the POS and POS-S packages equal.

for macroeconomic variables and (sometimes) state-specific trends or year dummies, Campos-Vazquez and Knox (2008), Bosch and Campos-Vazquez (2010), Azuara and Marinescu (2011), and Aterido et al. (2011) find conflicting results (no overall effect, a negative effect in the creation of formal jobs, a decrease in formality among small sub-groups, or a non-trivial reduction in the inflow of workers into formality) of subsidized health insurance to the poor. In Uruguay, using an extension in health benefits to dependents of private sector workers, Bérigolo and Cruces (2010) find that the health reform had a sizable effect on the type of employment.

Most studies use an identification strategy similar to ours in measuring the effects on informality by exploiting the variation in the roll-out or implementation of a particular large-scale social program, including municipality/state and time effects to capture other fluctuations in the labor market. Unlike our paper, however, previous studies do not identify individual-level eligibility, which we use in addition to the variation across municipalities.

In a more recent paper on Colombia, Calderón-Mejía and Marinescu (2011) use a similar identification strategy to examine the effect of two labor reforms implemented in 2003 and 2006–2007. The reform of 2003 imposed a requirement to use the same base income to contribute to both health insurance and pensions for self-employed workers. In contrast, the reform of 2006–2007 unified the payments to health insurance and pensions; the roll-out of this reform was performed by firm size. They also use the same identification strategy, taking advantage of the roll-out by firm size and including time, municipality, and firm size fixed effects to capture other changes in the labor market.

We consider the informality model developed by Galiani and Weischelbaum (2012) to be an adequate representation of the relation between informal-formal labor markets and the effects of having a health insurance package. Our empirical results complement their findings and are consistent with the predictions of their model.

Taken together, these studies indicate that given the high rates of informality in Latin

America and the desire of many governments to expand social program provisions, a better understanding of the distortions generated by them can help to inform policy decisions. Our contributions to this literature are threefold. First, to our knowledge, no other studies have considered the effect of the SR expansion on informality in Colombia using individual-level data. The health reform was a major reform; health insurance coverage in the population increased from levels below 30% prior to the reform to more than 90% today. Second, our findings are derived from two different methodologies and datasets, and the results are consistent for different robustness checks. Third, the reform took place almost two decades ago, and one of our datasets covers years before and after the reform, allowing the labor market to adjust to the reform and allowing us to test for effects in the longer term.

### III Data

#### Household Survey Data

We use repeated cross-sections of the Colombian Household Survey for periods before and after the implementation of the health care reform. Specifically, the informality modules included the second trimester every two years from 1990–1996 and once every year between 1997 and 2005.<sup>13</sup> These cross-sections are representative samples of the population living in Colombia’s ten largest cities.<sup>14</sup> Informality is defined as employees between 12 and 65 years who do not contribute to health insurance through employment. This definition of informality is most appropriate for our study, which evaluates the effect of a health insurance expansion.<sup>15</sup>

Table 1 shows the descriptive statistics for total, informal, and formal workers in columns 1, 2, and 3, respectively. There is a higher proportion of men in the labor force (56%). Men

---

<sup>13</sup>In Spanish: *Encuesta Nacional de Hogares* (ENH) 1990–2000 and *Encuesta Continua de Hogares* (ECH) 2001–2005.

<sup>14</sup>Bogotá, Cali, Medellín, Barranquilla, Bucaramanga, Pereira, Manizales, Villavicencio, Pasto, and Cúcuta.

<sup>15</sup>Alternative definitions of informality are based on the firm size, occupation, or education level of workers.

are also more likely than women to have informal employment. The average age of workers in Colombia is approximately 36 years. The highest proportion of workers is in the 25 to 34 years of age category. Informal workers are less educated than formal workers. Furthermore, 37% of formal workers have achieved some university education, whereas the corresponding number for informal workers is only 10%. We find that 39% of formal workers are married, in contrast to only 26% of informal workers. Cohabiting is more common among informal workers. The average household size and the proportion of other relatives and non-relatives living in the household are higher for informal workers.

We create a variable for the proportion of potential beneficiaries, which we use to test for heterogeneous effects in section IV. We define potential beneficiaries as household members who are not in the nuclear family, including married children and their spouses, grandchildren, grandparents and their spouses, siblings of the household head, and other relatives. These household members would be eligible for the SR but not the CR. We also define a group of vulnerable people within the household as the proportion of children under 1 year of age and the proportion of elderly members (65 years and older). An informal worker has, on average, a higher proportion of potential beneficiaries in the household and the same proportion of vulnerable members.

## **The Onset of SISBEN Interviews**

To determine whether the SR increased informal employment, we use variation across municipalities and compare the levels of informal employment before and after the onset of SISBEN interviews. We use the SISBEN onset dates as a proxy for the introduction of the SR because the law required that to become enrolled in the SR, people needed to have been interviewed for the SISBEN. Specifically, we use the SISBEN onset date as recorded in the SISBEN database

when more than 1% of the total interviews were conducted in that year,<sup>16</sup> as confirmed by the Secretary of Health of each municipality, or with alternative data sources, such as newspaper articles, as specified in Appendix Table S1, which reflects the onset dates for the SISBEN by municipality.

In all of the specifications, we include year and municipality fixed effects to ensure that economy-wide changes, such as those in the minimum wage, taxes, or job security legislation, and time-invariant municipal characteristics do not influence our results. We are able to do this because there is variation in the implementation dates across municipalities. We also have specifications with controls at the individual level (including age group, schooling, marital status, sex, relation to the household head, and public sector employment); the household level (household composition characteristics specified in each table, proportion employed, average education, and average age); and labor market controls (municipal-level information for population, working-age population, economically active population, and employment; and state-level GDP). The labor market controls, which vary by state and year (and are thus unique to each municipality in our sample), should capture any differential effect of macroeconomic shocks across regions. Identification comes from assuming that the variable of interest is capturing variation due solely to the onset of the SISBEN interviews and that the characteristics likely to influence the informality rates of those municipalities that implemented the SISBEN interviews later do not differ from municipalities that implemented earlier.

We group the municipalities according to the onset in SISBEN interviews in early (1994) and late (1995) adopters. We explore the relationship between earlier and later adopters in Figure 1

---

<sup>16</sup>We do this to avoid coding an onset date that could be due to a data-entry error. For instance, in two municipalities, the onset date appears in 1994, but less than 0.01% of the total interviews for those municipalities were recorded in that year. For one of the municipalities, we confirmed the onset date with a newspaper article. For the other, we re-ran the results, changing the onset date to 1994. The results remained consistent with the results presented in Table 3. In a few cases, these dates differ from those recorded in the SISBEN database because municipalities initially conducted some interviews that were later overwritten with updated interviews. Because we are interested in the onset of interviews, we use the earliest date given by the officials.

and Appendix Table S2. The figure shows an overall decline in informality rates over the period that we study. This decline, however, confounds general trends in health insurance enrollment, macroeconomic conditions, and the effects of the SR. Our empirical strategy will disentangle these effects to identify the effects of the SR.

Figure 1 also shows that in the years prior to the onset of the SISBEN interviews (shaded area), the informality rates of early and late adopters were almost identical. The informality rates diverge after the onset of the interviews in 1994 for the early adopters. Notably, the decline in the informality rate is lower than that observed among the late adopter group of municipalities. In 1996, after the late adopters start, their decline in informality rates also slows. Appendix Table S2 shows the summary statistics using data prior to the onset of the SISBEN interviews. We find that although the largest cities started conducting SISBEN interviews earlier, importantly, the proportion of informal workers on the eve of the first SISBEN interviews is almost the same across the groups.<sup>17</sup> To address the possibility of any labor market or demographic municipal-specific changes, in our analysis, we include labor market, demographic, and education controls.

## Panel Using Individual-Level SISBEN Data

As an alternative data source and to corroborate our results, we use a panel dataset at the individual level that we constructed from the first and second SISBEN. The first SISBEN was conducted between 1994 and 2003, and the second was conducted between 2003 and 2007.

Because of manipulation concerns after 1998 (Camacho and Conover 2011), we use data for

---

<sup>17</sup>We tested whether the early and late adopters had different pre-adoption informality rates by running a municipal-level regression of the informality rate in the pre-period on year and interactions between year and adoption date. An F-test of the interaction terms jointly equal to zero indicates that we failed to reject the null hypothesis of a common trend among the two groups of municipalities. This test, however, uses only 20 observations.

the first SISBEN from 1994–1997. Eligibility for the SR was based on the Poverty Index Score obtained from the first SISBEN, and we use individual fixed effects to determine whether the informality rates observed in the second SISBEN are higher for eligible individuals.

To construct the panel, we match individual-level information from the first and second SISBEN. We restrict our sample to working-age individuals between 12 and 65 years of age. We match the datasets in two ways. (1) We use the Colombian national identity card number. This is a unique ID number initially assigned to all Colombian citizens at age 7 that changes when they turn 18. This is the more conservative match, which we call the “ID match”. (2) We expand the match using names and dates of birth to include people who may have changed ID numbers when they turned 18. In Colombia, people use two last names; the first name is generally from the father, and the second is from the mother. When we match names, we use the first and second given names and the first and second last names, allowing for spelling mistakes and for phonetic distances no greater than 3 of 10 characters using the Leveinshtein criteria.<sup>18</sup> When we use dates of birth, we allow for the possibility of recording error by using both the exact date (day, month, year), either the same date and month of birth and within four years of birth, or the same year of birth and within two months and two days of birth. We call this the “Expanded match”. We report the results using both matches.

Summary statistics for the panel dataset that we construct using both the ID and expanded matches are reported in Table 2. To find a match, the individual must be in the labor force and have been interviewed for the SISBEN in both periods. Thus, we expect some workers from the first SISBEN (such as older people) to exit the labor force, and we do not see young workers who entered the labor force in the second SISBEN. Because we are observing the same individuals, some of the demographic characteristics, such as age and education, will change (increase) over

---

<sup>18</sup>Leveinshtein distances measure the similarity between two strings or two vectors of strings. The number indicates the number of characters that need to be changed in one string to equal the other string, normalized by the length of the string.

time. In particular, there is a slight increase in the levels of education achieved and a reduction in the proportion of single individuals. There is a 9%-10% reduction in informality rates between the first and second SISBEN. In our empirical strategy we look at the evolution of informality rates among those eligible for the SR relative to people who are not eligible.

## IV Effect of the Subsidized Regime on Informal Employment

### Survey Data Results

Assuming that the reform was effectively in place when the municipality began conducting SISBEN interviews, we estimate the impact of the reform on informal employment. We construct an indicator variable, *post*, that aligns the starting date of SISBEN interviews for each city and is equal to one on and after the year when the municipality began conducting interviews. In this way, we are able to disentangle the effects of national-level reforms from health reforms. Specifically, we estimate probit models with the following specification:

$$E[inf_{ihjt}] = \Phi[\alpha + \theta post_{jt} + \gamma_j + \sigma_t + X'_{jt}\beta_1 + Z'_{ihjt}\beta_2 + T'_{hjt}\beta_3 + S'_{ihj}\beta_4], \quad (1)$$

where the sub-indices correspond to *i* for an individual, *h* for a household, *j* for a municipality, and *t* for a year. Here, *inf* corresponds to an indicator variable for being informal, defined as a worker who does not contribute to health insurance through employment;  $\gamma$  represents municipality effects; and  $\sigma$  represents year effects, which capture any reform and shock at the national level. *X* is a vector of controls that vary across municipalities and over time, such as population, working-age population, economically active population, employment, and state-level GDP. These controls serve to capture variation due to business cycles and labor market fluctuations. *Z* corresponds to a vector of individual controls, such as age group, schooling,

marital status, relation to the household head, and whether the person works in the public sector.  $T$  corresponds to a vector of household controls, such as proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the SR, proportion employed, average education, average age, and socio-economic strata level.  $S$  captures controls for the employment sector, as described in Appendix Table S3.

We are interested in  $\theta$ , the parameter for the indicator variable  $post$ . A positive  $\theta$  indicates an increase in informal employment after the onset of the SISBEN interviews. We report the estimates of equation 1 in panel A of Table 3, where we show specifications including different sets of controls in each column listed in the note of the table. The estimates show a positive and significant effect of  $post$ , with an increase in informality between 4.1 and 5 percentage points.<sup>19</sup> All specifications are clustered at the municipality-year level.<sup>20</sup> Next, we construct a proxy for the Poverty Index Score to assess whether the positive increase in informal employment is concentrated among the people who are eligible for the SR.

The variables needed to construct the Poverty Index Score calculated in the SISBEN are not available in the household survey, so we use a geographical socio-economic strata (SES) indicator as a proxy. When using the SISBEN dataset, we find that this indicator is a strong predictor of eligibility (significant at the 1% level). We identify as eligible people living in SES levels 1 and 2, and, for consistency with the SISBEN implementation process, we limit the sample to people living in strata levels targeted by the SISBEN (1-3). We estimate the following regression:

$$\begin{aligned}
 E[inf_{ihjt}] &= \Phi[\alpha + \theta post_{jt} + \eta elig_{ihjt} + \delta post_{jt} * elig_{ihjt} \\
 &+ \gamma_j + \sigma_t + X'_{jt}\beta_1 + Z'_{ihjt}\beta_2 + T'_{hjt}\beta_3 + S'_{ihj}\beta_4],
 \end{aligned} \tag{2}$$

---

<sup>19</sup>To explore whether the effects are driven by self-employed workers or employees, we split the sample by each of these two groups, and we find consistent results across the two groups.

<sup>20</sup>We thank Josh Angrist for providing direction on the appropriate clustering level. The results clustered at the municipal level are consistent with those presented here but are less significant for panel C.

where we follow the same definitions used in equation 1 and *elig* is an indicator variable that takes a value of one for individuals who are eligible for the SR according to the proxy for the Poverty Index Score. The variable of interest is  $\delta$ . A positive  $\delta$  reports the relative increase in informality after the onset of the SISBEN interviews among the people who are eligible for the SR relative to those who are ineligible. The results are reported in panel B of Table 3. The table shows an additional increase in informality between 3.0 and 4.2 percentage points. These findings are consistent with the increase in informality reported in the previous results.

### **Heterogeneous Effects by Household Composition**

Features of the system, such as restricting benefits to certain family members, can generate incentives for some workers to opt out of the CR. We now explore whether these features, together with different household characteristics, affect the decision to become informal.

Family composition may affect the incentives for workers to become or remain informal because the CR benefits a more restricted group of family members: (1) if married/cohabiting for more than two years, children and a spouse/partner who is not directly enrolled in the CR; and (2) if single or without children, a parent who is a dependent. Coverage for any additional family member has an additional cost. This per capita cost is set by the government and is known as *Unidad de Pago por Capitación* (UPC). The SR, in contrast, allows the enrollment of any member of the household, regardless of family links. The less costly and less restrictive enrollment rules for family members or dependents in the SR might encourage some people to seek this type of health care coverage. Specifically, the SR is likely to be attractive to households with large extended families rather than smaller nuclear families.

If this is the case, we expect attenuated effects on households where the proportion of potential beneficiaries for the SR (as defined in section III) is lowest and stronger effects where the

proportion of potential beneficiaries is relatively high. To capture this effect, we estimate the following equation:

$$E[inf_{iht}] = \Phi[\alpha + \theta_1 post_{jt} + \theta_2 prop.ben_{jt} + \theta_3 post * prop.ben_{jt} + \gamma_j + \sigma_t + X'_{jt}\beta_1 + Z'_{iht}\beta_2 + T'_{hjt}\beta_3 + S'_{ihj}\beta_4]. \quad (3)$$

The coefficient of interest is  $\theta_3$ , which captures the change in informal employment after the implementation of the SISBEN in households with a higher proportion of potential beneficiaries. Panel C of Table 3 shows that the effects are large, robust to different controls, and significant, showing that as the proportion of potential beneficiaries increases in the household after the implementation of the SISBEN, the probability of being informal increases by an additional 3 to 4.4 percentage points.

Another feature of the system is the differences in packages provided to CR and SR enrollees. As mentioned in the introduction, the CR provides a more complete package of services and medicines than the SR. There may be heterogeneity in how the different packages of services offered by the CR and SR affect workers. It is possible that people who are eligible for both regimes seek different systems depending on their anticipated health needs. Households with vulnerable members, such as newborns, children, and the elderly, might prefer the more comprehensive CR. In contrast, healthy young people who do not foresee a high need for health services would prefer the SR. For them, the differential cost of the CR would be higher than the loss in terms of the quality and comprehensiveness of its health package.

We use household characteristics as a proxy for health insurance valuation. In particular, we want to see whether there are attenuated effects for households with a higher proportion of vulnerable members because, anticipating higher health care usage, these households are likely to value the CR more relative to households without vulnerable members. This is relevant for

policy because there was a recent discussion about whether the two regimes should provide the same package of services and medicines.

To test this hypothesis, we define a variable that calculates the proportion of children under one year of age and the proportion of elderly members. We specify a model similar to equation 3, but we substitute vulnerable members for the proportion of potential beneficiaries. We find (in results that are not shown) that as the proportion of vulnerable members increases, the probability of becoming informal decreases after the implementation of the SISBEN.

### **Household-Level Regressions**

Finally, to account for the joint decision of members within the household, we estimate whether the proportion of informal employees in the household (of the number of economically active household members) increases after the onset of SISBEN interviews among the eligible households. We find a robust and consistent effect of an increase in the proportion of informal employees in the household of 3.3 to 4.8 percentage points and of 2.2 to 2.5 percentage points among those eligible for the SR (Appendix Table S4, panels A and B, respectively).<sup>21</sup>

### **Panel Data Results**

Using a panel dataset that we construct by matching individuals observed in the first and second SISBEN, as described in section III, we estimate an individual fixed effect regression that captures the change in informality due to the SR.

We observe each individual twice. Because people would only find out if they were eligible for the SR after the implementation of the first SISBEN, we estimate for the eligible population (as determined by the first SISBEN), relative to the non-eligible, the probability of being informally

---

<sup>21</sup>Consistent with our findings, a modified version of Galiani and Weinschelbaum's (2012) informality model suggests that the introduction of the SR could be heterogeneous depending on the household composition.

employed in the second SISBEN, using the following specification:

$$E[inf_{it}] = \Phi[\alpha + \pi elig_{i,t-1} + S'_t \beta_2 + \gamma_i] \quad (4)$$

where  $\pi$  is the coefficient of interest, which, if positive, implies that the probability of being informal increased among the eligible people relative to the non-eligible in different sub-samples close to the threshold.  $S$  is the control for the year in which the first SISBEN was conducted, and  $\gamma_i$  is the individual fixed effect.

The corresponding regression results are reported in Table 4. Using the ID matching, the table shows that informality among the eligible population increased by 4.0 to 5.5 percentage points (panel A). The results are similar but slightly lower in magnitude (2.6 to 4.0 percentage points) when using the Expanded match, as reported in panel B. The corresponding Figure 2 shows, on the left graph, the informality rates relative to the eligibility threshold as determined in the first SISBEN using data from the first (solid line) and second (dash line) SISBEN. The graphs on the right are derived from data in the left graphs. These graphs indicate the difference in informality rates for the post minus the pre period relative to the eligibility threshold. These graphs show that the decline in informality rates is lower for people to the left of the threshold (those eligible for the SR), consistent with the direction and magnitude of the results estimated when using the variation in implementation dates across municipalities.

## V Conclusion

In this paper, we explore whether an expansion of non-contributory health insurance for the poor discouraged workers from entering the formal sector. Although this study was conducted using data from Colombia, it explores the broader question of how linking social benefits to

employment could generate strategic behavior by some individuals and create distortions in the labor market. As is common in empirical work, especially work on developing countries, the ideal dataset needed to conduct this analysis does not exist. Thus, to strengthen our findings, we use two different datasets that, though not perfect, provide consistent results for the question we address.

Using the onset of the SISBEN interviews and household survey data, we estimate that informal employment was approximately 4 percentage points higher than it would have been in the absence of the reform, or roughly 8% of the labor force not entering the formal sector.

Using an alternative dataset that we constructed using matched individuals in the first and second SISBEN and individual fixed effect regressions, we estimate the effect of the SR on informality to be approximately 3 to 4 percentage points when considering people 10 points around the eligibility threshold. These results are similar in magnitude to those obtained with the survey data. The sign of the coefficient suggests that some marginal individuals were optimizing when deciding whether to participate in the formal sector. Our empirical results are consistent with an informality model developed by Galiani and Weinschelbaum (2012).

It is important to consider that our estimates could combine both the effect of the implementation of the SR and those of a change in the enforcement of laws that prohibit informal employment (see Almeida and Carneiro 2012 and Ronconi 2010 for papers that discuss enforcement for labor and social security regulations in Latin America).<sup>22</sup> If enforcement changed with the passing of the law and not with the implementation of the SISBEN, then the variation across municipalities and the year fixed effects, as well as the discontinuous change in the eligibility threshold, should allow us to disentangle the effects of the SR from those of enforcement. If enforcement changed with the implementation of the SISBEN and at the eligibility threshold, then our estimates combine both effects. Looking at newspaper article counts of labor inspections

---

<sup>22</sup>We thank a referee for bringing the issue of enforcement to our attention.

by adoption date, we do not find any evidence that enforcement changed with the adoption of the SISBEN, but we do see an overall increase across all municipalities. This increase should be captured by the year fixed effects, and any increase in enforcement not captured by the year fixed effects would downward bias our estimates.

There are documented health benefits from the creation of the SR (Camacho and Conover 2013; Miller et al. 2009). A complete welfare analysis would also account for the benefits to workers (in terms of flexibility and independence) of becoming informal. In this study, we do not attempt to measure these benefits, but we find that the implementation of the SR resulted in a non-trivial increase in informal employment. To cover the new informal workers and their families, the government's health budget increased by approximately 11%. Overall losses in productivity that result from differences in the informal and formal sectors are quantified in studies by Cárdenas and Rozo (2009) and Hamman and Mejía (2011), which indicate that the output per worker in the formal sector relative to the informal sector is approximately 1.85 to 1.95. In a back-of-the-envelope calculation, using their ratios and our results, the increase in informality corresponds to a reduction of 3.8% of GDP. Note that this is likely an upper bound because it does not account for workers who were previously formal but remain in the same job with an agreement with their employer to opt out of making health insurance contributions, resulting in no loss of productivity.

We emphasize, however, that expanding health insurance coverage need not translate into higher informality rates. Rather, attention should focus on incorporating changes in the system to discourage people from entering the informal sector. One such change was passed by the government in 2005<sup>23</sup> when, in recognition of people not wanting to enter the formal sector for fear of permanently losing their SR benefits, the government allowed people to save their SR

---

<sup>23</sup>See ACUERDO 304 de 2005, available at <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=18274>

slot and reactivate it if they lost their formal employment within a year of enrolling in the CR. Further changes of this type would mitigate the increase in informal employment documented here.

## References

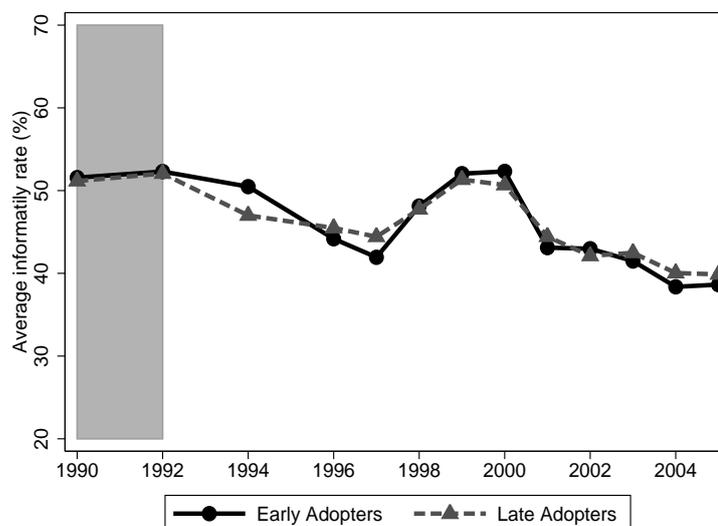
- Almeida, Rita and Pedro Carneiro. 2012. "Enforcement of Labor Regulation and Informality." *American Economic Journal: Applied Economics* 4 (3):64–89.
- Aterido, Reyes, Mary Hallward-Driemeier, and Carmen Pagés. 2011. "Does Expanding Health Insurance Beyond Formal-Sector Workers Encourage Informality? Measuring the Impact of Mexico's Seguro Popular." Policy Research Working Paper Series 5785, The World Bank. URL <http://ideas.repec.org/p/wbk/wbrwps/5785.html>.
- Azuara, Oliver and Ioana Marinescu. 2011. "Informality and the Expansion of Social Protection Programs." MPRA Paper 35073, University Library of Munich, Germany. URL <http://ideas.repec.org/p/pramprapa/35073.html>.
- Bérgolo, Marcelo and Guillermo Cruces. 2011. "Labor Informality and the Incentive Effects of Social Security: Evidence from a Health Reform in Uruguay." IDB Publications 62318, Inter-American Development Bank. URL <http://ideas.repec.org/p/idb/brikps/62318.html>.
- Bosch, Mariano and Raymundo M. Campos-Vázquez. 2010. "The Trade-offs of Social Assistance Programs in the Labor Market: The Case of the "Seguro Popular" program in México." Serie documentos de trabajo del Centro de Estudios Económicos 2010-12, El Colegio de México, Centro de Estudios Económicos. URL <http://ideas.repec.org/p/emx/ceedoc/2010-12.html>.
- Calderón-Mejía, Valentina and Ioana Marinescu. 2011. "The Impact of Colombia's Pension and Health Insurance Systems on Informality." Tech. Rep. 62338, IDB working paper.
- Camacho, Adriana and Emily Conover. 2011. "Manipulation of Social Program Eligibility." *American Economic Journal: Economic Policy* 3 (2):41–65.

- . 2013. “Evaluating the Effects of Public Health Insurance on Child Health in Colombia.” *Economic Development and Cultural Change* 61 (3):633–58.
- Campos-Vázquez, Raymundo M. and Melissa A. Knox. 2010. “Social Protection Programs and Employment: The Case of México’s “Seguro Popular” Program.” Serie documentos de trabajo del Centro de Estudios Económicos 2010-14, El Colegio de México, Centro de Estudios Económicos. URL <http://ideas.repec.org/p/emx/ceedoc/2010-14.html>.
- Cárdenas, Mauricio and Sandra Rozo. 2009. “Firm Informality in Colombia: Problems and Solutions.” *Desarrollo y Sociedad* (63):211–43.
- Castañeda, Tarsicio, Kathy Lindert, Bènedicte de la Brière, Luisa Fernandez, Celia Hubert, Osvaldo Larrañaga, Mónica Orozco, and Rosana Viquez. 2005. “Designing and Implementing Household Targeting Systems: Lessons from Latin America and the United States.” The World Bank, Social Protection Unit Discussion Paper 0529.
- De Soto, Hernando. 1989. “The Other Path: The Invisible Revolution in the Third World.” *New York: Harper and Row* .
- DNP, Misión Social and UNDP. 2003. *Quién se Beneficia del Sisben? Evaluación Integral*. Departamento Nacional de Planeación.
- Galiani, Sebastian and Federico Weinschelbaum. 2012. “Modeling Informality Formally: Households and Firms.” *Economic Inquiry* 50 (3):821–38.
- Gasparini, Leonardo, Francisco Haimovich, and Sergio Olivieri. 2009. “Labor Informality Bias of a Poverty-Alleviation Program in Argentina.” *Journal of Applied Economics* 12 (2).
- Gaviria, Alejandro and Martha Luz Henao. 2001. “Comportamiento del Desempleo en los

- Últimos Años y Estrategias de los Hogares para Enfrentarlo.” *Coyuntura Social de Fedesarrollo* 24:23–38.
- Giedion, Ursula and Manuela V. Uribe. 2009. “Colombia’s Universal Health Insurance System.” *Health Affairs* 28 (3):853.
- Grosh, Margaret E. 1994. *Administering Targeted Social Programs in Latin America: From Platitudes to Practice*. World Bank Publications.
- Hamman, Franz and Luis-Fernando Mejía. 2011. “Formalizando la Informalidad Empresarial en Colombia.” Working Paper 676, Banco de la República Colombia.
- Kugler, Adriana. 1999. “The Impact of Firing Costs on Turnover and Unemployment: Evidence from the Colombian Labour Market Reform.” *International Tax and Public Finance* 6 (3):389–410.
- Kugler, Adriana and Maurice Kugler. 2009. “Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia.” *Economic Development and Cultural Change* 57 (2):335–58.
- Levy, Santiago. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Brookings Institution Press.
- Loayza, Norman V. 1996. “The Economics of the Informal Sector: a Simple Model and some Empirical Evidence from Latin America.” *Carnegie-Rochester Conference Series on Public Policy* 45:129–62.
- Maloney, William F. 2004. “Informality Revisited.” *World Development* 32 (7):1159–78.
- Miller, Grant, Diana Pinto, and Marcos Vera Hernández. 2009. “Risk Protection, Service Use,

- and Health Outcomes under Colombia's Health Insurance Program for the Poor." Open access publications from university college london, University College London.
- Mondragón, Camilo, Ximena Peña, and Daniel Wills. 2010. "Labor Market Rigidities and Informality in Colombia." *Economía* 11 (1):65–101.
- Pagés, Carmen and Lucía Madrigal. 2008. "Is Informality a Good Measure of Job Quality? Evidence from Job Satisfaction Data." Research Department Publications 4603, Inter-American Development Bank, Research Department. URL <http://ideas.repec.org/p/idb/wpaper/4603.html>.
- Perry, Guillermo, William F. Maloney, Omar S. Arias, Pablo Fajnzylber, and Andrew D. Mason. 2007. *Informality: Exit and Exclusion*. World Bank Publications.
- Ronconi, Lucas. 2010. "Enforcement and Compliance with Labor Regulations in Argentina." *Industrial and Labor Relations Review* :719–36.
- Santamaría, Mauricio, Fabián García, and Ana Virginia Mujica. 2008. "Los Costos No Laborales y el Mercado Laboral: Impacto de la Reforma de Salud en Colombia." Working Paper 43, Fedesarrollo.
- Tafur, Luis A. 1998. *La Salud en el Sistema General de Seguridad Social*. Centro Editorial Catorse.

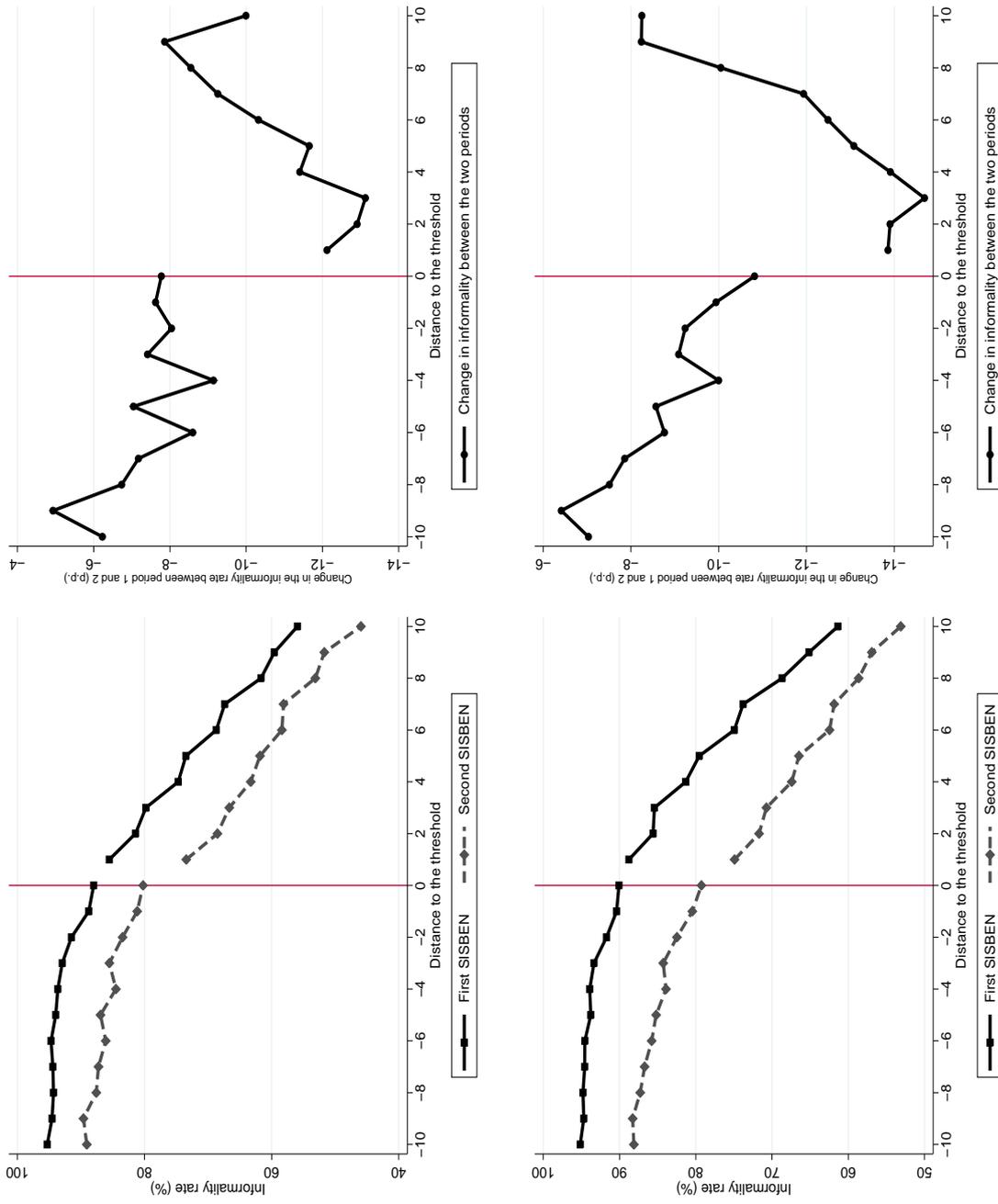
Figure 1: Informality Rates by Onset of SISBEN interviews



*Source:* ENH-ECH weighted household survey data from 1990–2005.

*Note:* Solid lines indicate informality rates for the “early adopters” (i.e., the group of municipalities with the onset of the SISBEN in 1994). Dashed lines indicate informality rates for the “late adopters” (the group of municipalities with the onset of the SISBEN in 1995). The gray shaded region corresponds to the period prior to the adoption of the SISBEN in any municipality, showing a common trend between the two groups.

Figure 2: Difference in Informality Rates Using SISBEN Panel Data (Top panel: ID match; Bottom panel: Expanded match).



Source: Panel dataset created by authors using the first and second SISBEN.

Note: Left graphs indicate informality rates relative to the eligibility threshold as determined in the first SISBEN using data from the first (solid line) and second (dashed line) SISBEN. Right graphs are derived from data from the left graphs and indicate informality rates for the second SISBEN minus informality rates for the first SISBEN relative to the eligibility threshold.

Table 1: Summary Statistics for Demographic Variables, % unless noted

	Total (1)	Informal (2)	Formal (3)
Male	56.36	57.76	55.18
Age			
12 to 17	2.82	5.53	0.54
18 to 24	16.68	19.15	14.59
25 to 34	31.12	27.67	34.02
35 to 44	26.00	24.30	27.43
45 plus	23.39	23.35	23.42
Years (mean)	35.69	34.99	36.28
Education			
None	1.86	3.36	0.60
Some primary	9.41	15.30	4.45
Completed primary	15.67	21.93	10.40
Some secondary	23.16	29.66	17.68
Completed secondary	25.15	20.06	29.44
Some university	9.13	4.80	12.78
Completed university	15.61	4.89	24.65
Marital status			
Cohabitation	21.29	26.00	17.33
Married	32.79	25.74	38.73
Divorced/Widow(er)	12.61	14.07	11.38
Single	33.31	34.18	32.57
Household size			
Mean (people)	4.63	4.88	4.41
Less than or equal to 3	29.95	27.01	32.43
4 to 5	43.53	41.66	45.11
6 to 10	24.79	28.83	21.39
More than 10	1.72	2.50	1.06
Relation			
Household head	45.62	43.50	47.41
Spouse	16.46	16.28	16.62
Child	24.88	25.17	24.63
Other relative	8.67	9.89	7.64
Non-relative	4.36	5.15	3.70
Household composition			
Prop. potential beneficiaries to SR	14.98	15.98	14.14
Prop. vulnerable: kids < 1 yr and elderly	7.22	7.22	7.22
Observations (weighted)	66,951,730	30,613,842	36,337,888

*Source:* Pooled ENH-ECH weighted household survey data from 1990–2005. Authors' calculations.

*Note:* See Section III for details.

Table 2: Characteristics on the Eve of SISBEN Interviews–Panel Dataset

	ID Match		Expanded Match	
	1st SISBEN	2nd SISBEN	1st SISBEN	2nd SISBEN
Informality	0.84	0.75	0.83	0.73
Male	0.80	0.80	0.78	0.78
Household size	4.20	3.92	4.21	3.86
Age				
Mean (years)	35.37	43.95	34.01	42.16
12-17	0.00	0.00	0.02	0.00
18-24	0.12	0.00	0.16	0.02
25-34	0.38	0.16	0.37	0.22
35-44	0.33	0.39	0.30	0.38
45+	0.17	0.45	0.15	0.39
Education				
Some primary	0.28	0.25	0.23	0.22
Completed primary	0.28	0.29	0.29	0.29
Some secondary	0.28	0.22	0.33	0.27
Completed secondary	0.11	0.16	0.11	0.16
Some College	0.01	0.02	0.01	0.02
Completed College	0.00	0.01	0.00	0.01
Marital status				
Cohabiting	0.36	0.35	0.38	0.39
Married	0.37	0.39	0.31	0.33
Divorced or widow(er)	0.10	0.12	0.10	0.12
Single	0.17	0.14	0.21	0.15
Observations	58,418	58,418	169,337	169,337

*Source:* Panel dataset created by authors using the first and second SISBEN.

*Note:* Information for people in the labor force who appear in the first and second SISBEN. See section III for details.

Table 3: Changes in Informality after the Onset of SISBEN Interviews

Dependent variable:	Informal employment				
	(1)	(2)	(3)	(4)	(5)
Panel A: Informality after the Onset of SISBEN Interviews					
Post	0.041** (0.019)	0.050*** (0.019)	0.047** (0.021)	0.045** (0.020)	0.046** (0.021)
Pseudo R-squared	0.025	0.025	0.191	0.197	0.244
Observations	343,688	343,688	342,435	342,435	342,435
Panel B: Informality after the Onset of SISBEN Interviews Among the Eligible to the SR					
Post*Eligible	0.031** (0.014)	0.030** (0.014)	0.039*** (0.011)	0.041*** (0.011)	0.047*** (0.011)
Eligible	0.131*** (0.011)	0.132*** (0.011)	0.024*** (0.008)	0.008 (0.008)	0.001 (0.008)
Post	0.024 (0.022)	0.040** (0.020)	0.029 (0.021)	0.025 (0.020)	0.020 (0.020)
Pseudo R-squared	0.042	0.042	0.169	0.174	0.228
Observations	259,941	259,941	258,873	258,873	258,873
Panel C: Informality After the Onset of Interviews by Proportion of Potential Beneficiaries in the Household					
Post*Prop.	0.044** (0.020)	0.039* (0.021)	0.039** (0.018)	0.035* (0.018)	0.034 (0.021)
Potential Beneficiaries	0.039** (0.016)	0.043*** (0.016)	-0.012 (0.015)	-0.020 (0.016)	-0.011 (0.019)
Prop. Potential Beneficiaries	0.034* (0.019)	0.044** (0.019)	0.040* (0.021)	0.039* (0.020)	0.040* (0.021)
Post	0.026 (0.019)	0.026 (0.019)	0.191 (0.021)	0.197 (0.020)	0.244 (0.021)
Pseudo R-squared	0.026	0.026	0.191	0.197	0.244
Observations	343,688	343,688	342,435	342,435	342,435
Municipality effects	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes
Labor controls		Yes	Yes	Yes	Yes
Individual controls			Yes	Yes	Yes
Household controls				Yes	Yes
Sector/SES controls					Yes

Source: ENH-ECH weighted household survey data from 1990-2005.

Note: Robust standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. Marginal effects reported. The results are clustered at the municipal-year level. Labor controls include population, working age population, economically active population, employment, and state-level GDP. Individual controls include age group, schooling, marital status, sex, relation to the household head, and whether working in a government job. Household controls include the proportion of children under five years in the household (panels A and B), proportion of elderly (panels A and B), proportion of potential beneficiaries to the SR (panels A and B), proportion employed, average education, and average age. Employment sector controls correspond to the economic sectors as defined in Appendix Table S3. Potential beneficiaries is the number of household members other than a spouse or unmarried children. Column (5) of Panels A and C includes both sector- and SES-level controls. Panel B regressions use only information for strata levels less than 4 to make them comparable to the interviews targeted by the SISBEN and do not include SES controls in column (5).

Table 4: Informality Using a Panel Dataset of Individuals

Dependent variable:	Informal Employment		
Panel A: ID Match, using national ID card number			
Eligible	0.040*** (0.004)	0.055*** (0.005)	0.052*** (0.011)
R-squared	0.060	0.069	0.071
Observations	116,836	68,838	20,542
Panel B: Expanded Match, using IDs, names and last names, and date of birth			
Eligible	0.026*** (0.002)	0.040*** (0.003)	0.033*** (0.006)
R-squared	0.056	0.071	0.072
Observations	338,666	194,844	56,780
Distance to the eligibility threshold	+/- 10 pts	+/- 5 pts	+/- 1 pt

*Source:* Panel dataset created by authors using the first and second SISBEN.

*Note:* Data for the first SISBEN are from 1994–1998. Data for the second SISBEN are from 2003–2007. Robust standard errors in parentheses. Year controls and individual fixed effects included. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

## S1 Appendix

Table S1: SR Onset Dates and Household Survey Dates by Municipality

City	SR Onset	ECH survey onset	Sources
Medellín	Jan-1995	1996	Secretary of Health and newspaper articles
Barranquilla	Jan-1994	1994	SISBEN dataset
Bogotá, D.C.	Jan-1994	1994	SISBEN dataset
Manizales	Mar-1994	1994	SISBEN dataset
Villavicencio <sup>+</sup>	Jan-1995	1996	SISBEN dataset
Pasto	1995	1996	Secretary of Health
Cúcuta	Feb-1994	1994	SISBEN dataset
Pereira	Feb-1995	1996	SISBEN dataset
Bucaramanga	May-1995	1996	Secretary of Health and newspaper articles
Cali	Jan-1995	1996	SISBEN dataset and newspaper articles

Source: The ECH household survey module used is conducted in April-June of the specified year. Other sources as described in Table.

Note: <sup>+</sup> The onset dataset in the SISBEN database for this municipality is April 1994, but less than 1% of the interviews were conducted in 1994, thus making it likely that the few interviews recorded in 1994 could be due to a coding error.

Table S2: Characteristics on the Eve of SISBEN Interviews, pre-1994

Adopters:	Earlier (1994)	Later (1996)
Labor variables (%)		
Informal	51.96	51.64
Working age population	70.62	70.93
Active population	63.05	61.49
Employed	90.02	87.51
Demographic variables		
Population	4,077,867	1,366,368
Age	33.77	34
Proportion male	59.08	59.46
Education variables (%)		
No education	2.08	2.26
Some primary	10.85	14.75
Completed primary	17.94	19.11
Some secondary	28.00	28.39
Completed secondary	20.73	18.95
Some college	7.95	7.06
Completed College	12.44	9.46

Source: ENH-ECH weighted household survey data from 1990 and 1992, prior to the onset of the SISBEN interviews.

Table S3: Appendix –Average Informality Rate by Economic Sector pre 1992

Economic Sector	Percent Informal
Personal and Household Services	79.32
Other Mining	73.23
Construction	71.34
Manuf. of Wood and Wood Products, Including Furniture	69.08
Restaurants and Hotels	66.77
Metal Ore Mining	62.76
Wholesale & Retail Trade	60.73
Transport and Storage	57.96
Recreational and Cultural Services	52.27
Agriculture and Hunting	52.18
Textile, Wearing Apparel and Leather Industries	51.42
Manuf. of Food, Beverages and Tobacco	41.66
Manuf. of Non-Metallic Mineral Products	40.79
Manuf. of Metal Products, Machinery and Equipment	40.40
Fishing	38.08
Activities not adequately defined	34.75
Real estate and Business Services	34.49
Basic Metal Industries	33.63
Manuf. of Paper Products, Printing and Publishing	32.60
International and Other Extra-Territorial Bodies	24.83
Coal Mining	22.89
Manuf. of Chemicals, Petroleum, Coal and Plastic Products	21.80
Social and Related Community Services	21.01
Forestry and logging	19.02
Sanitary and Similar Services	17.03
Insurance	12.66
Communication	12.18
Water Works and Supply	11.70
Crude Petroleum and Natural Gas Production	11.62
Electricity, Gas and Steam	11.09
Financial Institutions	8.33
Public Administration and Defense	5.46

Source: ENH Household Survey, 1990 and 1992. Authors' calculations.

Table S4: Household Level Regressions

Dependent variable:	Prop. Informal in HH		
	(1)	(2)	(3)
Panel A: After the Onset of SISBEN Interviews			
Post	0.037** (0.015)	0.048*** (0.018)	0.033** (0.015)
R-squared	0.046	0.046	0.183
Observations	201,347	201,347	201,321
Panel B: After the Onset of SISBEN Interviews Among the Eligible to the SR			
Post*Eligible	0.025** (0.012)	0.025** (0.012)	0.022** (0.009)
Eligible	0.107*** (0.010)	0.108*** (0.010)	0.033*** (0.008)
Post	0.026 (0.016)	0.046** (0.018)	0.031** (0.015)
R-squared	0.064	0.064	0.166
Observations	153,624	153,624	153,600
Municipality effects	Yes	Yes	Yes
Year effects	Yes	Yes	Yes
Labor controls		Yes	Yes
Household controls			Yes

Source: ENH-ECH weighted household survey data from 1990-2005. Author's calculations.

Note: Robust standard errors in brackets. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. Marginal effects reported. Results are clustered at the municipal-year level. In panel B results only use information for strata levels less than 4, to make it comparable to the interviews targeted by the SISBEN. Labor controls include: population, working age population, economically active population, employment and state level GDP. Household controls include: proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the SR, proportion employed, average education, and average age.