

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

Adriana Camacho

Emily Conover

Alejandro Hoyos

This version: June 2010

First version: May 2008

Abstract

We examine whether the Colombian government's expansion of publicly provided health insurance in the early nineties created incentives for people to become informal. Using variation in the onset of the Census of the Poor across municipalities and over time, we find robust and consistent estimates of an increase in informal employment between 2 and 5 percentage points after the expansion. As robustness, we also construct a panel of individuals interviewed for the first and second Census of the Poor and find similar results. These results may be of interest to other countries considering similar social program expansions.

JEL No. I11, I18, O17.

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

*Authors are listed in alphabetical order. Contact information: Adriana Camacho, Economics Department, Universidad de Los Andes, email: adcamach@uniandes.edu.co. Emily Conover, Economics Department, Hamilton College, email: econover@hamilton.edu. Alejandro Hoyos, Research Assistant, Inter-American Development Bank, email: ale-hoyo@uniandes.edu.co. We are grateful to Anna Aizer, Kathryn Anderson, Amy Damon, Chang-Tai Hsieh, David Levine, Edward Miguel, Enrico Moretti, Suresh Naidu, Dean Scrimgeour and Lisa Verdon. Thanks to participants at the NEUDC 2009 conference, seminario CEDE, and the Hamilton-Colgate economic seminar series. Comments welcome. All errors are ours.

1 Introduction

Expansion of government funded social programs often generates controversies regarding the perverse incentives that these programs could create. Opponents to these programs often claim that people become dependent on them and under report or hide their income to ensure their eligibility. Informal employment is another possible unintended consequence of social programs.

Informal employment, common in many developing countries, is particularly high in Latin America. Galiani and Weinschelbaum (2007) 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 was 47% from 1985-2005 using a benefits based definition.¹ From a social welfare perspective, informality is undesirable since the informal sector often encompasses constrained and low productivity firms, where workers are unprotected by labor laws, and are unlikely to receive benefits such as health insurance or pensions.

However, it is unclear from the worker's perspective whether informal employment is 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 the workers' valuation of the services they provide. This situation is exacerbated when the government provides informal workers with comparable free services (Levy, 2008; Pagés and

¹A benefit based definition of informality refers to an employee who does not receive health and pension benefits through employment. In 2003, the International conference of Labor Statisticians (ICLS), established the benefit based definition of informality as their official one.

Madrigal, 2008) ².

The upward trend in informality rates in some Latin American countries coincides with the expansion of social program provisions.³ We examine the impact of Colombia's health insurance reform of 1993 on workers' choice of employment in the formal or informal sector. While there is plenty of anecdotal evidence regarding potential disincentives for formal employment generated by the social protection system and the public provision of health insurance in Colombia, no systematic empirical evidence has been presented to confirm or reject this hypothesis.⁴

The health sector reform of December 1993 aimed to achieve universal health insurance coverage by the creation of two regimes: a Contributive (CR) and a Subsidized Regime (SR). The CR made health insurance through employment mandatory, regardless of occupation. Eligibility for the SR was determined with a poverty index score, calculated from the answers to the Census of the Poor. In other words, people did not know whether or not they were eligible for the SR, until they had been interviewed by the Census of the Poor.⁵

To determine whether the provision of health insurance for the poor through the Subsidized Regime was an incentive for leaving formal employment and taking up informal employment we use the variation in the onset of the Census of the Poor across municipalities and over time. We show that at the onset of the Census of the Poor, informality rates are uncorrelated with the timing of the Census of the Poor. Implementation of the Census of

²See Loayza (1996) and de Soto's (1989) for characterizations of the informal labor market and its determinants.

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

⁴For 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 so high as to be creating incentives to discourage formal employment. A study by Santamaria et al. (2008) reports that almost half of SR recipients, 49%, indicated that they are not willing to lose their (SR) benefits by switching to a formal job.

⁵The Census of the Poor is known as the SISBEN in Colombia. SR eligibility was determined using a complex algorithm that took into account demographic and dwelling characteristics at the individual and household level. For more details on the calculation of the SISBEN score see Castañeda (2005).

the Poor by groups of municipalities at different points in time allow us to disentangle the effects of the expansion of health insurance provision from other labor market reforms and macroeconomic shocks occurring during the same period of time for the whole country, such as an increase in the minimum wage.

Our main data set comes from repeated cross-sections of the Colombian Household Survey for periods before and after implementation of Census of the Poor interviews. Using individual level data, and controlling for municipality and year effects, we estimate whether or not there was an increase in informal employment after the municipality started conducting interviews. Given that we are evaluating the effects of government provision of health insurance, we use the definition of informality that accounts for employees between 12 and 65 years old who do not contribute to health insurance through employment.⁶

We find robust and consistent effects of an increase in informal employment between 2 and 5 percentage points. For robustness, we estimate alternative specifications using the variation in fraction of interviews conducted within a municipality and over time, and specifications that account for eligibility for the SR. These results also indicate an increase in informal employment. Using an alternative data set that we constructed with a panel of individuals interviewed in the first and second Census of the Poor, we corroborate the increase in informal employment. Using information about an individual's eligibility status, as determined by the poverty index score from the first Census of the Poor, we find consistent results of an increase in informal employment of approximately 1.5 percentage points. These results indicate that some marginal workers are optimizing when deciding not to participate in the formal sector. The Colombian government is currently considering expanding the

⁶This is a *benefit based* definition of informality, where formal workers are defined as all individuals in the labor force that receive health insurance through their employer and informal workers are those who do not.

package of benefits provided by the SR to make it equivalent to that of the CR. The results of this study indicate that the distortionary effects of this policy change would increase informality even more. Further consideration of these policies and additional changes to the system could help mitigate these distortions.

The paper is organized as follows: in section 2 we provide information on the reforms and review articles that looked at the effects of other reforms that took place in Colombia during the period we study. We also discuss specific reasons why Colombia's social protection system may be generating differential incentives for workers to become or remain informal. In Section 3 we describe the data sets used in the analysis. In Section 4 we present the empirical results for informal employment. We conclude in Section 5.

2 Literature and Background

Between 1990 and 2005 the Colombian labor market underwent several legislative changes and economic reforms. There are several studies that analyze the impact of a change in labor taxes, but they do not specifically address the effect of public health insurance provision on informal employment. For instance, using a panel of manufacturing plants Kugler and Kugler (2005) estimate that a 10% increase in payroll taxes reduces formal employment between 4-5%. They also estimate a reduction in formal wages between 1.4-2.3%.

Kugler (1999) has also studied the incidence of job security regulations implemented in the early nineties on exit rates into and out of unemployment. She finds that these regulations increased labor market flexibility and pro-cyclical relationship between employment and business cycles. Mondragon et al. (2009) 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.

A study by Cárdenas and Bernal (2003) documents a sharp increase in unemployment rates since 1995. They highlight the increase in labor costs due to the changes in legislation as a main factor affecting labor demand. Gaviria and Henao (2001) report that a strategy used by households faced with loss of employment is to participate in informal sector activities. Using data from the National Statistical Agency (DANE), they report an upward trend in informal employment 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 Subsidized Regime (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 there are reforms to the system currently being considered in Colombia. Other countries are also thinking of implementing publicly provided health insurance, and understanding the magnitude of the distortions generated by the system can help inform policy decisions.

The structure of the health sector has raised questions of whether the creation of the Subsidized Regime unintentionally generated incentives for informality, since for those willing and able to engage in informal employment it reduced health insurance costs to both the employer and employee.

If a worker is employed, then the worker and the employer are required to make contributions to the CR regardless of occupation. Employers who do not comply with the legislation could be fined. Employees pay 1/4 of the cost, while employers are responsible for 3/4. Independent workers can also enroll in the Contributive Regime by paying the full cost.⁷ 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

⁷From 1995 to 2007 the cost was 12% of a person's income.

Plan Obligatorio de Salud (POS).⁸

Eligibility for the SR is not determined by employment status, but rather by a poverty index score designed to capture long term living conditions. It is constructed using information from the Census of the Poor. Thus, a worker can be eligible to the SR and be employed. An eligible worker for the SR could potentially opt-out of making mandatory contributions to the CR through employment. The Subsidized Regime was designed to reach people who did not have access to the Contributive Regime. It is financed with a 1% transfer from the Contributive Regime, and with local and central government funds. Households eligible for the Subsidized Regime have free access to a package of services and medications known as the *Plan Obligatorio de Salud Subsidiado* (POSS).⁹ This package is less comprehensive than the one offered by the Contributive Regime, but unlike the Contributive Regime, it covers every member in an eligible household regardless of their relationship to the household head.

A definition of informal employment that is widely used in the literature is the *benefit based* definition. It encompasses workers who do not get mandated labor protection law coverage, or who more broadly lack social security benefits, such as health insurance or pensions. There is no information about whether the individual makes pension contributions through employment for every year that we study (1984-2005). Thus we use information on whether a worker makes contributions to health insurance through employment, before and after the reform.¹⁰ This definition is particularly relevant for the question that we address, since we test whether the expansion of public health insurance provision created incentives for people to opt-out of the mandatory health insurance contribution for employees.

⁸In English POS stands for “Obligatory Health Plan.” For more information on the detail of the POS and POSS see Tafur (1998).

⁹In English POSS stands for “Subsidized Obligatory Health Plan.”

¹⁰We checked for years in which both questions on whether the individual makes contributions to health insurance and pensions are available, and found correlations between 73-77%.

There are some features of the system that can generate incentives for some workers to opt-out of the CR. Family composition may affect the incentives for workers to become or remain informal. The Contributive Regime benefits only a restricted group of family members: (1) if married/cohabiting for more than 2 years, children and a spouse/partner who is not directly enrolled in the Contributive Regime; (2) if single or without children, a parent who is a dependent. Coverage for any additional family member, in the Contributive Regime, has an additional cost. This per capita cost is set by the government and is known as *Unidad de Pago por Capitación* (UPC). The Subsidized Regime, on the other hand, allows 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 Subsidized Regime might encourage some people to seek this type of health care coverage. Specifically, the Subsidized Regime is likely to be attractive to households with large extended families rather than smaller nuclear families. In section 4.2 we explore whether there are stronger effects in households with large extended families.

There could be heterogeneity in how the different packages of services offered by the Contributive Regime and Subsidized Regime affected 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 like newborns, children and elderly might prefer the more comprehensive Contributive Regime. On the other hand, healthy young people who do not foresee a high need for health services would prefer the Subsidized Regime. For them, the differential cost of the Contributive Regime would be higher than the loss in terms of the quality and comprehensiveness of its health package. In section 4.2 we look at whether the effects are attenuated in households with a higher proportion of vulnerable members such as

young children and elderly.

3 Data

3.1 Household Survey Data

We use repeated cross-sections of the Colombian Household Survey, for periods before and after implementation of the health care reform.¹¹ These cross-sections are representative samples of the population living in the following ten cities: Bogotá, Cali, Medellín, Barranquilla, Bucaramanga, Pereira, Manizales, Villavicencio, Pasto and Cúcuta. For the analysis we use the informality modules included in the second trimester every two years from 1986 to 1996, and once every year between 1997 and 2005. Informality is defined as employees between 12 and 65 years who do not contribute to health insurance through employment. This is the definition of informality more appropriate for our study which evaluates the effect of a health insurance expansion.¹²

An advantage of using the Colombian Household Survey is that these data are unlikely to be manipulated since they are not used to determine eligibility to social programs. In fact, Camacho and Conover (2009) show that the Census of the Poor, which is used to determine eligibility to different social programs, exhibits evidence of manipulation after 1998, while data from a household survey do not show evidence of manipulation.

Table 1 shows descriptive statistics for total, formal and informal workers in columns 1, 2 and 3 respectively. There is a higher proportion of men in the labor force (58%). Men are also more likely than women to have an informal employment. The average age

¹¹In Spanish: *Encuesta Nacional de Hogares* (ENH) 1984-2000, and *Encuesta Continua de Hogares* (ECH) 2001-2005.

¹²Alternative definitions of informality are based on firm size, occupation or education level of workers. Using pension rather than health coverage contribution to define informality is likely to yield very similar results given that as reported by Bernal (2009) 94% of workers who contribute to health also contribute to pensions.

of workers in Colombia is approximately 35 years. The highest proportion of workers is in the 25 to 34 age category. Informal workers are less educated than formal workers. 32% of formal workers have achieved some university education, while the corresponding number for informal workers is only 8%. 42% of formal workers are married, while 28% of informal workers are married. Cohabiting is more common among informal workers. Average household size is higher for informal workers. Other relatives and non-relatives within a household account for 14% of the household members.

We defined five groups of vulnerable people within the household: proportion of children below 1, 2 and 5 years; proportion of elderly in the household and proportion of potential beneficiaries. An informal worker has on average a higher proportion of kids in the household but a lower proportion of elderly.

Table 2 describes the sample corresponding to each municipality in column 1, the average informality rate, and shows how informality and formality rates are distributed across these municipalities in columns 2 and 3 respectively. We also report information for the labor controls that are included in the regressions.

Table A1 reports informality rates by economic sector, organized by informality level on the eve of the onset of the Census of the Poor, 1992. Personal and household services is the economic activity with the highest rate of informality (79%), while Public administration and defense has the lowest rate of informality (1%).

3.2 The Onset of the Census of the Poor Interviews

To determine whether the Subsidized Regime increased informal employment, we use variation across municipalities and compare the levels of informal employment before and after the onset of Census of the Poor interviews. In all of the specifications we include year and

municipality controls to ensure that economy wide, such as changes in the minimum wage, taxes or job security legislation, and time invariant municipal characteristics are not influencing our results. We are able to do this because there is variation in implementation dates across municipalities. We also have specifications with individual and household level controls. Identification comes from assuming that the variable of interest is capturing variation due solely to the onset of the Census of the Poor interviews, and that the characteristics of the municipalities which implemented the Census of the Poor interviews later do not differ from municipalities that implemented earlier.

We explore the relationship between earlier (1994) and later (1996) adopters of the Census of the Poor interviews, in Table 3. The Table shows summary statistics using 1992 data, prior to the onset of the Census of the Poor interviews, and grouping the municipalities in earlier and later adopters. We find that although the largest cities started conducting Census of the Poor Interviews early, the proportion of informal workers on the eve of the first Census of the Poor interviews is almost the same across the groups. The order of the distribution across sectors is the same across the groups. And although there are differences in the education distribution across the groups, we report results controlling for these demographic characteristics.

3.3 Panel using Census of the Poor Data

As a robustness check for our results we use a panel dataset at the individual level that we constructed from the first and second Census of the Poor. These Census, administered by the local government, collect household and individual level data used to calculate a poverty index score. The poverty index score is used to select beneficiaries for government social programs including health subsidies. The first Census of the Poor was conducted between

1994 and 2003, and the second between 2003 and 2005. The algorithm used to construct the poverty index score changed between the first and second Census. We used information on eligibility based on the poverty index score obtained from the first Census, and determine whether informality rates observed in the second Census of the Poor are higher for individuals who are barely eligible for the Subsidized Regime, relative to individuals who are barely not eligible.

Summary statistics for the panel dataset that we construct relative to the labor force participants in the first and second Census of the Poor data sets are reported in Table 4. To find a match, the individual must be in the labor force and interviewed for the Census of the Poor in both periods. Thus, we expect some workers from the first Census of the Poor (such as older people) to exit the labor force. Since we are observing the same individuals, some of the demographic characteristics, such as age and education will change (increase) over time. In particular, there is a slight increase in the levels of education achieved, a reduction in the proportion of single individuals, and an increase in the average proportion of children and elderly in the household. The proportion of males in the matched sample (73%) is considerably higher than in the complete sample (67%). A possible explanation is that women's participation in the labor force is lower and less stable (e.g. maternity leave). In the matched sample there is a 7% reduction in informality rates between the first and second Census of the Poor. In our empirical strategy we look at the evolution of informality rates among the eligible to the SR relative to people who are not eligible.

4 Empirical Analysis

4.1 Effect of the Subsidized Regime on Informal Employment

Assuming that the reform was effectively in place when the municipality started doing Census of the Poor interviews, we estimate the effect of the reform on informal employment. We capture this by constructing an indicator variable *post* which aligns the starting date of Census of the Poor interviews for each city, and is equal to 1 on and after the year when the municipality started conducting interviews. This way we are able to disentangle the effects of national level reforms from the health reforms. Specifically, we estimate probit models with the following specification:

$$\text{probit}(\text{inf}_{iht}) = \alpha + \theta \text{post}_{jt} + \gamma_j + \sigma_t + X'_{jt}\beta_1 + Z'_{iht}\beta_2 + T'_{hjt}\beta_3 + S'_{ihj}\beta_4 + \epsilon_{iht} \quad (1)$$

Where the sub-indices correspond to: *i* for individual, *h* for household, *j* for municipality, *t* for year. *inf* corresponds to an indicator variable for being informal, defined as a worker who does not contribute to health insurance through employment. γ represents municipality effects, while σ represents year effects. *X* is a vector of controls that vary across municipality 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. *Z* corresponds to a vector of individual controls, such as age group, schooling, marital status, relation to the household head and whether or not working in activities which are likely to be seasonal.¹³ *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 Subsidized Regime, proportion employed, average education, average age, and socio economic strata level. *S* captures controls for employment sector as described in Table A1.

¹³Specifically for this variable we generated an indicator which took a value of 1 for: independent, self-employed and domestic workers; and a value of 0 for: employers, government workers, or family worker without remuneration.

ϵ correspond to the idiosyncratic error term, assumed to be mean zero and with a constant variance.

We are interested in the value of θ , the parameter for the indicator variable *post*. If $\theta > 0$ then there is an increase in informal employment after the onset of the Census of the Poor interviews. We report estimates in Table 5, where we show specifications including different sets of controls in each column. Estimates show a positive and significant effect of *post*, with an average increase in informality of 4.2 percentage points. All specifications are clustered at the municipality level.

Our identification comes from assuming that the variable *post* is capturing only the variation due to the onset of the Census of the Poor interviews. Year and municipality controls ensure that economy wide changes and time invariant municipal characteristics are not influencing our results. As a robustness check for the effect of the Subsidized Regime on informality we use the variation in the proportion of interviews conducted within a municipality and over time, as recorded in the first Census of the Poor database. With the idea that as more people had access to the Subsidized Regime, they could more easily choose to opt-out of the Contributory Regime which is mandatory through formal employment. To do this, we specify a regression where the variable of interest is *prop_inter*. This variable captures the cumulative proportion of Census of the Poor interviews that have been conducted, allowing us to use the variation within a municipality and over time. It takes a 0 value in periods prior to the onset of Census of the Poor interviews, values between 0 and 1 in years when interviews were conducted but not completed, and a value of 1 in years when all the Census of the Poor interviews were completed. We estimate probit models with the

following specification:

$$\begin{aligned} \text{probit}(inf_{iht}) &= \alpha + \kappa \text{prop_inter}_{jt} + \gamma_j + \sigma_t + X'_{jt}\beta_1 \\ &+ Z'_{iht}\beta_2 + T'_{hjt}\beta_3 + S'_{ihj}\beta_4 + \epsilon_{iht} \end{aligned} \quad (2)$$

Where all other variables are defined as in equation 1. The coefficient of interest is κ , the parameter for the proportion of Census of the Poor interviews conducted. If $\kappa > 0$ then a higher proportion of Census of the Poor interviews conducted in a municipality is associated with an increase in informal employment. We report results from this regression in Table 6. The table shows an average increase in informality of 2 percentage points. All specifications are clustered at the municipality level.

4.2 Informal Employment by Household Composition

In this section we explore the relative effects of offering the Subsidized Regime by household composition groups. We use household characteristics as a proxy for health insurance valuation. In particular we want to see whether we find attenuated effects for households with vulnerable members since, as described in Section 2, the Contributive Regime provides a more complete package of services and medicines than the Subsidized Regime. We expect that households with a higher proportion of vulnerable members are likely to value the Contributive Regime more relative to households without vulnerable members. To test this hypothesis, we define a variable *vulnerable* which calculates the proportion of children under 1 year of age and the proportion of elderly members.

We estimate equations 1 after dividing the data in three groups: the first group does not have vulnerable members in the household; the second group has vulnerable members and the proportion of vulnerable members is less than or equal to the median observed in the data; the third group has vulnerable members and the proportion of vulnerable members is

higher than the median. We report the coefficient estimates for the *post* variable. Results are reported in Table 7. Results show that the effects of the Subsidized Regime are strongest in households with no vulnerable members, with an average effect of 4.3 percentage points.¹⁴

During the period we study, the Subsidize Regime, but not the Contributive Regime, covers every member in an eligible household regardless of their relationship to the household head. Thus we expect attenuated effects on households where the proportion of potential beneficiaries for the Subsidized Regime is lowest, and stronger effects where the proportion of potential beneficiaries is relatively high. We define potential beneficiaries as household members 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 Subsidized Regime but not the Contributive Regime. Again, we divide the sample into three groups: first a group without potential beneficiaries; second a group with potential beneficiaries but with a proportion less than or equal to the median; and finally a group with potential beneficiaries but with a proportion higher than the median. Results are reported in Table 8. The Table shows that the effects are attenuated in the sample without potential beneficiaries, and disappear once we include controls for sectors. The increase in informality is largest for the groups above the median number of potential beneficiaries, with an average effect of 5 percentage points.

4.3 Robustness

4.3.1 Estimation accounting for Eligibility

Using the Household Survey data set, in this section we construct a proxy for the Poverty Index Score to assess whether the positive increase in informal employment is concentrated

¹⁴These results are consistent with those we get when we divide the sample by the proportion of children less than one year of age, and the proportion of elderly independently.

among the people who are eligible for the Subsidized Regime.

The poverty index score calculated from information in the Census of the Poor is made of four components: demographic, education, utilities and housing. The poverty index score is not available in the household survey data, but we use the geographical socio economic strata indicator as a proxy.¹⁵ We identify as eligible people living in SES level less than 3. Using data for people living in strata levels 1-3 (those targeted with household interviews in the Census of the Poor), we estimate the following regression:

$$\begin{aligned} \text{probit}(inf_{iht}) &= \alpha + \theta post_{jt} + \eta elig_{iht} + \delta post_{jt} * elig_{iht} & (3) \\ &+ \gamma_j + \sigma_t + X'_{jt}\beta_1 + Z'_{iht}\beta_2 + T'_{hjt}\beta_3 + S'_{ihj}\beta_4 + \epsilon_{iht} \end{aligned}$$

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 according to the proxy for the poverty index score are eligible for the Subsidized Regime. The variable of interest is δ . A positive δ will report the relative increase in informality, after the onset of the Census of the Poor interviews, among the eligible people for the Subsidized Regime relative to the ineligible. Results, using the indicator for *post* and *prop_inter* for the proportion of interviews, are reported in Tables 9 and 10. The tables show an average increase in informality of 8 percentage points. These findings are consistent with the increase in informality reported in the previous section.

4.3.2 Alternative Estimation Using a Panel Dataset of Individuals

Using a panel dataset that we construct by matching individuals observed in the first and second Census of the Poor, we estimate the Subsidized Regime's effect on the probability of being informally employed while accounting for eligibility. For consistency and comparability

¹⁵Using the Census of the Poor data set we found that socio economic strata is a strong predictor of eligibility at the 1% significance level.

with the analysis conducted using ECH data, we keep all the observations regardless of when the people were interviewed for the first Census of the Poor, even though Camacho and Conover (2009) show that there is evidence of manipulation in the allocation of the poverty index score for people interviewed after 1998. In regressions using only people who were interviewed after 1998 (not shown) we get similar results to the ones reported here with the full sample. We observe each individual twice. We assume that it takes people time to learn about the program and whether or not they are eligible for the Subsidized Regime. For the eligible population relative to the non-eligible, we estimate the probability of being informally employed in the second Census of the Poor, controlling for the informality rates observed in the first Census of the Poor. Specifically we estimate:

$$\begin{aligned}
inf_{iht} &= \alpha + \pi elig_{ihj} + f_1(score_i | score_i \leq \overline{score}) \\
&+ f_2(score_i | score_i > \overline{score}) + X'_j \beta_1 + S'_t \beta_2 + T'_{hjt} \beta_3 + Z'_{iht} \beta_4 + \epsilon_{iht}
\end{aligned} \tag{4}$$

Where π is the coefficient of interest that captures the individual effects including the eligibility indicator equal to 1 for people with Poverty Index Scores below the threshold, determined from information collected in the first Census of the Poor. If π is positive this implies that the probability of being informal increased for the eligible people. \overline{score} is the threshold score used to determine eligibility for the social programs including the health insurance. We control for a quintic function of the score before and after the threshold and we also report local linear regression results using an optimal bandwidth selection algorithm suggested by Fan and Gijbels (1996) and reported by McCrary (2008).

X are state (*departamento*) effects; S are controls for the year in which the first Census of the Poor was conducted; 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

Subsidized Regime, proportion employed, average education and average age; Z is a vector of individual controls, such as age group, schooling, marital status, sex and whether or not working in activities which are likely to be seasonal.

Informality rates exhibit a small discontinuity only when using information from the second Census of the Poor, as seen in Figures 1 and 2. This indicates that at the onset of the interviews there was no discontinuity around the eligibility threshold, but by the time the second Census of the Poor was implemented, some people around the threshold opted for informal employment.

The corresponding regression results to equation 4 are reported in Table 11. The table shows that informality among the eligible population increased by 1.5 percentage points. The lower magnitude of these results relative to those previously reported can be explained by the fact that the data sets represent different populations. The ECH is representative of the largest cities in the country, where average informality rates are around 50%; the panel from the Census of the Poor is taken from the population in the left tail of the Colombian income distribution where the average informality rates are around 85%. The direction and percentage increase in informality is consistent with those estimated using the variation in implementation dates across municipalities.

5 Conclusion

In this paper we explore whether an expansion of non-contributory health insurance for the poor had the unintended consequences of increasing informal employment. Although this study was conducted using data from Colombia, it explores a broader question of how linking social benefits to employment could generate strategic behavior from some individuals and create distortions in the labor market. Using the onset of the Census of the Poor inter-

views in each municipality, we estimate that informal employment increased approximately 4 percentage points.

In an alternative model we estimated the effect of the Subsidized Regime on informality by using the variation in the proportion of interviews conducted within a municipality and over time. We use this variation because as more people had access to the Subsidized Regime, the relative value of the Contributory Regime declined, and thus we see more people opting for informal employment. We find that with an increase in the proportion of interviews conducted, informal employment increased approximately 2 percentage points. These results are of similar magnitude to those obtained when using an alternative panel data set. The sign of the coefficient suggests that some marginal individuals seem to have been optimizing when deciding whether or not to participate in the formal sector.

The results presented here imply that the broad expansion of public health insurance in Colombia contributed to increasing informal employment. In a back-of-the envelope calculation, we estimate that the direct costs of the distortion to the government were between 0.04% and 0.07% of GDP (or approximately 0.5% of government social expenditures and 10% of the SR budget) in SR expenditures. This amount does not seem particularly large, but it is important to keep in mind that does not include indirect costs such as possible reductions in productivity or loss of other tax revenues.

References

- Bernal, Raquel S.** 2009. “The Informal Labor Market in Colombia: Identification and Characterization.” *Desarrollo y Sociedad* (63): 145–208.
- Camacho, Adriana and Emily Conover.** 2009. “Political Manipulation of Social Program Eligibility: Implications for Empirical Research.” Working paper, documento cede no. 2009-19, Universidad de Los Andes.
- Cárdenas, Mauricio and Raquel Bernal.** 2003. “Determinants of Labor Demand in Colombia: 1976-1996.” NBER Working Papers 10077, National Bureau of Economic Research.
- Castañeda, Tarsicio.** 2005. “Targeting Social Spending to the Poor with Proxy Means Testing: Colombia’s SISBEN System.” The World Bank, Social Protection Unit Discussion Paper 0526.
- DeSoto, Hernando.** 1989. “The Other Path: The Informal Revolution.” *New York: Harper and Row* .
- Fan, Jiangqing and Irene Gijbels.** 1996. *Local Polynomial Modelling and Its Applications*. New York: Chapman and Hall.
- Galiani, Sebastian and Federico Weinschelbaum.** 2007. “Modeling Informality Formally: Households and Firms.” de Estudios, Centro and Distributivos, Laborales y Sociales 47, Universidad de La Plata.
- 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.
- Grosh, Margaret E.** 1994. *Administering Targeted Social Programs in Latin America: From Platitudes to Practice*. World Bank Publications.

- 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.** 2005. “Labor Market Effects of Payroll Taxes in a Middle Income Country: Evidence from Colombia.” Nber working papers, National Bureau of Economic Research.
- 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: 129–162.
- Maloney, William F.** 2004. “Informality Revisited.” *World Development* 32(7): 1159–1178.
- McCrary, Justin.** 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics* 127(2): 698–714.
- Mondragón, Camilo, Ximena Peña, and Daniel Wills.** 2009. “Labor Market Rigidities and Informality in Colombia.” Mimeo.
- Pagés, Carmen and Lucía Madrigal.** “Is Informality a Good Measure of Job Quality? Evidence from Job Satisfaction Data.” *Inter-American Development Bank Working Paper* 654.
- Perry, Guillermo, William F. Maloney, Omar S. Arias, Pablo Fajnzylber, and Andrew D. Mason.** 2007. *Informality: Exit and Exclusion.* World Bank Publications.
- Santa María, Mauricio, Fabián García, and Ana Virginia Mujica.** 2008. “El Mercado Laboral y la Reforma a la Salud en Colombia: Incentivos, Preferencias y Algunas Paradojas.” Working paper, Fedesarrollo.

Tafur, Luis A. 1998. *La Salud en el Sistema General de Seguridad Social*. Centro Editorial Catorse.

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

	Total (1)	Informal (2)	Formal (3)
Gender			
Male	58.39	59.12	57.60
Age			
12 to 17	3.71	6.54	0.63
18 to 24	17.97	20.24	15.50
25 to 34	30.71	27.26	34.47
35 to 44	24.95	22.97	27.10
45 plus	22.66	22.99	22.30
Mean (years)	35.15	34.50	35.85
Education			
None	2.65	4.34	0.81
Some primary	12.86	19.26	5.89
Completed primary	17.72	23.23	11.72
Some secondary	24.78	28.97	20.22
Completed secondary	22.81	16.63	29.53
Some university	7.68	3.91	11.79
Completed university	11.51	3.67	20.04
Marital status			
Cohabitation	19.03	22.56	15.18
Married	34.74	28.32	41.73
Divorced / widow(er)	12.20	13.33	10.96
Single	34.04	35.79	32.13
Household size			
Mean (number of people)	4.97	5.20	4.73
Less or equal to 3	25.90	23.73	28.27
4 to 5	40.90	38.72	43.28
6 to 10	30.57	34.12	26.69
More than 10	2.63	3.44	1.75
Relation			
Household head	45.18	42.68	47.91
Spouse	15.06	14.93	15.20
Child	26.15	26.59	25.68
Other relative	8.96	9.77	8.08
Non-relative	4.64	6.03	3.13
Household composition			
Prop. kids < 5 in HH	8.89	9.45	8.27
Prop. kids < 2 in HH	4.23	4.50	3.94
Prop. kids < 1 in HH	2.74	2.89	2.58
Prop. elderly in HH	4.14	4.04	4.25
Prop. potential beneficiaries to SR	15.83	16.52	15.07
Observations (unweighted)	439,653	229,247	210,406
Observations (weighted)	77,511,936	36,115,122	41,396,814

Note: Pooled ENH-ECH Household Survey data from 1984-2005. Authors' calculations. See Section 3 for details.

Table 2: Summary Statistics for Labor Variables and Municipalities, %

	Total (1)	Informal (2)	Formal (3)
<hr/>			
Municipality			
Medellín	12.16	9.52	15.04
Barranquilla	11.42	13.24	9.43
Bogotá	19.22	16.79	21.87
Manizales	7.11	6.16	8.15
Villavicencio	7.37	7.70	7.01
Pasto	8.56	10.53	6.41
Cúcuta	8.52	11.51	5.26
Pereira	6.59	5.99	7.26
Bucaramanga	6.46	6.61	6.29
Cali	12.59	11.94	13.29
Observations (unweighted)	439,653	229,247	210,406
Observations (weighted)	77,511,936	36,115,122	41,396,814
<hr/>			
Labor variables		mean	
Informality		52.14	
Population 12-65 years		72.68	
Economically active population		62.96	
Employment		85.33	
Unemployment		14.68	
Observations (unweighted)		439,653	
Observations (weighted)		77,511,936	
<hr/>			

Note: ENH-ECH Household Survey 1992-2004. Authors' calculations. Results are in % unless otherwise noted. See Section 3 for details.

Table 3: Characteristics on the Eve of Census of the Poor Interviews, 1992

Adopters:	Earlier (1994)	Later (1996)
Labor variables (%)		
Informal	50.63	50.78
Working age population	70.61	70.81
Active population	64.36	64.20
Employed	90.54	86.69
Economic sectors*		
Group 1 (66-79 informality)	25.79	30.13
Group 2 (47-60 informality)	38.71	40.16
Group 3 (35-43 informality)	13.35	11.68
Group 4 (16-33 informality)	13.71	12.02
Group 5 (1-13 informality)	8.45	6.02
Demographic variables		
Average Population	4,184,331	1,392,317
Average Age	33.85	34.20
Proportion male	58.48	58.85
Education variables (%)		
No education	2.02	2.24
Some primary	10.12	14.05
Completed primary	18.00	19.28
Some secondary	27.37	28.06
Completed secondary	21.66	19.54
Some college	8.36	7.28
Completed College	12.47	9.55

Note: ENH-ECH Household Survey 1992. Authors' calculations. Using data from 1992, prior to the onset of the Census of the Poor interviews. *See appendix Table A1 for a detailed listing of the sectors in each group.

Table 4: Characteristics on the Eve of Census of the Poor (COP) Interviews

	1st COP 1994 - 2003	1st COP matched	2nd COP 2003-2005	2nd COP matched
Informality	81.48	82.58	70.53	75.09
Gender				
Male	67.07	73.06	63.89	73.06
Female	32.93	26.94	36.11	26.94
Age				
12 to 17	0.03	0.02	0.01	0.00
18 to 24	13.11	12.22	14.28	3.21
25 to 34	32.29	34.37	29.38	26.27
35 to 44	29.39	32.20	29.69	36.31
45 plus	25.18	21.19	26.63	34.22
Mean (years)	37.10	36.24	37.31	40.74
Education				
None	6.55	5.67	4.3	4.98
Some primary	25.88	26.35	18.42	23.37
Completed primary	26.08	27.45	21.44	26.35
Some secondary	22.76	23.32	21.64	22.68
Completed secondary	15.10	14.21	24.96	17.61
Some university	2.06	1.76	4.67	2.6
Completed university	1.35	1.04	3.93	1.96
Marital status				
Cohabit	34.53	37.88	33.51	37.59
Married	28.67	29.59	26.02	31.94
Divorced/Widow(er)	13.68	11.42	13.14	13.13
Single	23.13	21.11	27.33	17.34
Household composition				
Prop. Children	5.68	6.97	8.38	8.00
Prop. elderly in household	1.96	1.66	2.57	2.61
Prop. Potential beneficiaries to SR	4.32	4.28	4.73	4.25
Observations	3,073,025	1,065,855	5,725,028	1,065,855

Note: Information for people in the labor force who appear in the first and second Census of the Poor.

Table 5: Changes in Informality after the Onset of Census of the Poor Interviews

Dependent variable:	Informal employment					
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.040** [0.020]	0.050*** [0.019]	0.042** [0.020]	0.039** [0.020]	0.040** [0.020]	0.041** [0.021]
Observations	439,653	439,653	437,510	437,510	414,647	414,647
Pseudo R-squared	0.025	0.025	0.165	0.171	0.175	0.241
Municipality effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Labor controls		Yes	Yes	Yes	Yes	Yes
Individual controls			Yes	Yes	Yes	Yes
Household controls				Yes	Yes	Yes
SES controls					Yes	Yes
Sector controls						Yes

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Results are clustered at the municipal 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 or not working in activities which are likely to be seasonal. Household controls include: proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the Subsidized Regime, proportion employed, average education, and average age. Employment sector controls correspond to the economic sectors as defined in Table A1.

Table 6: Changes in Informality Using the Proportion of Census of the Poor Interviews

Dependent variable:	Informal employment					
	(1)	(2)	(3)	(4)	(5)	(6)
Prop. Interviews	0.015	0.023	0.023**	0.020**	0.022**	0.010
	[0.013]	[0.015]	[0.010]	[0.010]	[0.009]	[0.011]
Observations	439,653	439,653	437,510	437,510	414,647	414,647
Pseudo R-squared	0.025	0.025	0.165	0.171	0.175	0.241
Municipality effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Labor controls		Yes	Yes	Yes	Yes	Yes
Individual controls			Yes	Yes	Yes	Yes
Household controls				Yes	Yes	Yes
SES controls					Yes	Yes
Sector controls						Yes

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Results are clustered at the municipal 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 or not working in activities which are likely to be seasonal. Household controls include: proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the Subsidized Regime, proportion employed, average education, and average age. Employment sector controls correspond to the economic sectors as defined in Table A1.

Table 7: Informality After the Onset of Interviews by Proportion of Vulnerable Members in the Household

Dependent variable:	Informal employment					
	(1)	(2)	(3)	(4)	(5)	(6)
No vulnerable members	0.037* [0.022]	0.048** [0.019]	0.043** [0.021]	0.041** [0.020]	0.042** [0.021]	0.045** [0.022]
Observations	314,838	314,838	313,345	313,345	296,953	296,953
Pseudo R-squared	0.024	0.025	0.163	0.168	0.172	0.240
≤ median vulnerable	0.042 [0.026]	0.049* [0.028]	0.034 [0.029]	0.031 [0.027]	0.034 [0.026]	0.027 [0.025]
Observations	67,397	67,397	67,039	67,039	63,289	63,289
Pseudo R-squared	0.025	0.025	0.165	0.175	0.179	0.244
> median vulnerable	0.052* [0.029]	0.063** [0.030]	0.038 [0.030]	0.033 [0.029]	0.030 [0.026]	0.031 [0.027]
Observations	57,418	57,418	57,126	57,126	54,405	54,405
Pseudo R-squared	0.029	0.029	0.176	0.182	0.188	0.246
Municipality effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Labor controls		Yes	Yes	Yes	Yes	Yes
Individual controls			Yes	Yes	Yes	Yes
Household controls				Yes	Yes	Yes
SES controls					Yes	Yes
Sector controls						Yes

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%. Vulnerable members are elderly and children less than 1 year old. Results are clustered at the municipal 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 or not working in activities which are likely to be seasonal. Household controls include: proportion employed, average education and average age. SES controls include controls for geographic socio-economics status. Employment sector controls correspond to the economic sectors as defined in Table A1.

Table 8: Informality After the Onset of Interviews by Proportion of Potential Beneficiaries in the Household

Dependent variable:	Informal employment					
	(1)	(2)	(3)	(4)	(5)	(6)
No potential beneficiaries	0.033*	0.048***	0.038**	0.036*	0.034*	0.035
	[0.020]	[0.017]	[0.019]	[0.018]	[0.020]	[0.022]
Observations	262,202	262,202	261,119	261,119	246,994	246,994
Pseudo R-squared	0.028	0.028	0.171	0.176	0.181	0.246
≤ median potential beneficiaries	0.045*	0.047*	0.045	0.044	0.046*	0.051*
	[0.026]	[0.025]	[0.029]	[0.028]	[0.027]	[0.029]
Observations	92,157	92,157	91,596	91,596	87,590	87,590
Pseudo R-squared	0.023	0.023	0.156	0.163	0.167	0.233
> median potential beneficiaries	0.052***	0.055***	0.046**	0.044**	0.050***	0.049***
	[0.019]	[0.021]	[0.020]	[0.019]	[0.019]	[0.017]
Observations	85,294	85,294	84,795	84,795	80,063	80,063
Pseudo R-squared	0.023	0.023	0.156	0.162	0.166	0.233
Municipality effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Labor controls		Yes	Yes	Yes	Yes	Yes
Individual controls			Yes	Yes	Yes	Yes
Household controls				Yes	Yes	Yes
SES controls					Yes	Yes
Sector controls						Yes

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Potential Beneficiaries is the number of household members other than a spouse or unmarried children. Results are clustered at the municipal 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 or not working in activities which are likely to be seasonal. Household controls include: proportion employed, average education and average age. SES controls include controls for geographic socio-economics status. Employment sector controls correspond to the economic sectors as defined in Table A1.

Table 9: Informality After the Onset of Interviews by Eligibility

Dependent variable:	Informal employment				
	(1)	(2)	(3)	(4)	(5)
Post	0.017 [0.020]	0.034* [0.019]	0.021 [0.020]	0.017 [0.019]	0.016 [0.019]
Eligible	0.162*** [0.017]	0.162*** [0.017]	0.010 [0.010]	-0.007 [0.008]	-0.009 [0.008]
Post*Eligible	0.071*** [0.012]	0.070*** [0.012]	0.073*** [0.012]	0.075*** [0.011]	0.082*** [0.011]
Observations	439,653	439,653	437,510	437,510	437,510
Pseudo R-squared	0.05	0.05	0.17	0.17	0.24
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 controls					Yes

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Results are clustered at the municipal level. Results only use information for strata levels less than 4, to make it comparable to the interviews targeted by the Census of the Poor. 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 or not working in activities which are likely to be seasonal. Household controls include: proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the Subsidized Regime, proportion employed, average education, and average age. Employment sector controls correspond to the economic sectors as defined in Table A1.

Table 10: Informality Using Proportion of Interviews by Eligibility

Dependent variable:	Informal employment				
	(1)	(2)	(3)	(4)	(5)
Proportion of interviews	-0.004	-0.007	-0.009	-0.012	-0.029***
	[0.009]	[0.010]	[0.010]	[0.010]	[0.009]
Eligible	0.180***	0.180***	0.026**	0.009	0.007
	[0.020]	[0.020]	[0.012]	[0.010]	[0.008]
Proportion of interviews*Eligible	0.084***	0.084***	0.094***	0.095***	0.109***
	[0.015]	[0.016]	[0.016]	[0.016]	[0.016]
Observations	439,653	439,653	437,510	437,510	437,510
Pseudo R-squared	0.052	0.052	0.168	0.173	0.238
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 controls					Yes

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; + significant at 1%. Results are clustered at the municipal level. Results only use information for strata levels less than 4, to make it comparable to the interviews targeted by the Census of the Poor. 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 or not working in activities which are likely to be seasonal. Household controls include: proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the Subsidized Regime, proportion employed, average education, and average age. Employment sector controls correspond to the economic sectors as defined in Table A1.

Table 11: Informality Using a Panel Dataset of Individuals

Dependent variable: Informal Employment	(1)	(2)	(3)	(4)
Eligible	0.017*** [0.004]	0.016*** [0.003]	0.015*** [0.003]	0.015*** [0.003]
Observations	1,055,271	1,055,271	413,772	413,772
Pseudo R-squared	0.26	0.27	0.07	0.07
Municipality effects	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Household controls		Yes		Yes
Quintic terms	Yes	Yes		
LLR			5.6	5.6

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; + significant at 1%. Parametric regressions are clustered at the score level. Local linear regression results use an optimal bandwidth selection algorithm. Individual controls include age group, schooling, marital status, sex and whether or not working in activities which are likely to be seasonal. Household controls include proportion of children in the household, proportion of elderly, proportion of potential beneficiaries to the Subsidized Regime, proportion employed, average education and average age.

Figure 1: First Census of the Poor Informality Rates Using a Panel Dataset of Individuals

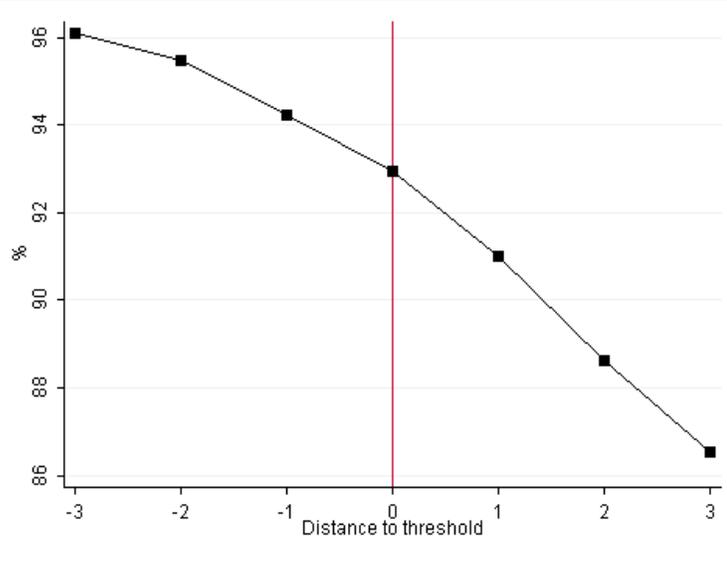


Figure 2: Second Census of the Poor Informality Rates Informality Using a Panel Dataset of Individuals

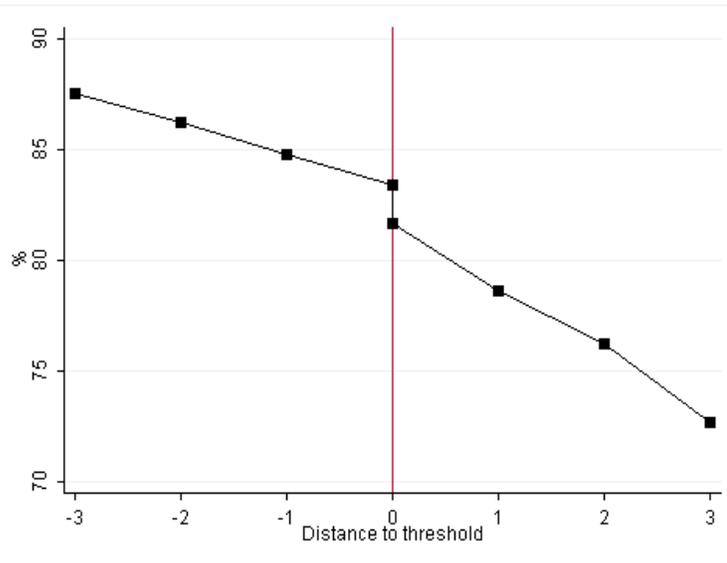


Table A1: Appendix: Average Informality Rate by Economic Sector in 1992

	Informal 1992 (2)
<hr/>	
Group 1- Informality rate 79-66	
Personal and household services	79.05
Other mining	71.85
Construction	69.47
Manuf. of wood products, furniture	66.50
Restaurants and hotels	66.27
<hr/>	
Group 2- Informality rate 60-47	
Wholesale and retail trade	59.99
Transport and storage	57.56
Textile, apparel and leather industries	52.11
Recreational and cultural services	50.19
Metal ore mining	48.11
Agriculture and hunting	46.91
<hr/>	
Group 3- Informality rate 43-35	
Manuf. of non-metallic mineral products	42.94
International and extra-territorial bodies	42.07
Manuf. of food, beverages and tobacco	40.00
Manuf. of metal, machinery and equipment	39.84
Real estate and business services	35.09
Basic metal industries	34.80
<hr/>	
Group 4- Informality rate 33-16	
Manuf. of paper, printing and publishing	32.96
Fishing	30.84
Coal mining	29.34
Activities not adequately defined	26.10
Manuf. of chemicals, oil, coal and plastic	21.94
Forestry and logging	19.02
Social and related community services	18.38
Sanitary and similar services	16.20
<hr/>	
Group 5- Informality rate 13-1	
Electricity, gas and steam	12.55
Insurance	9.11
Communication	6.86
Crude oil and natural gas production	6.57
Financial institutions	4.86
Water works and supply	2.86
Public administration and defense	1.08
Observations (unweighted)	26,590
Observations (weighted)	30,430

Note: ENH Household Survey, 1992. Authors' calculations.