Informality and Seguro Popular under Segmented Labor Markets

Robert Duval Hernández

División de Economía, CIDE

Corresponding Author: robert.duval@cide.edu

Ricardo Smith Ramírez

División de Economía, CIDE

5/24/2011

PRELIMINARY DRAFT. COMMENTS WELCOMED

Abstract: This paper analyzes whether a social protection program generates a disincentive to formality. More specifically, we analyze the impact of Mexico's Seguro Popular health protection program on the willingness to search for a formal sector job. Unlike other papers, we don't assume labor markets are perfectly competitive and distinguish between the decision to look for a formal sector job, and the fact of actually getting one. We find that Seguro Popular reduced the probability of looking for a formal sector job, although the negative effect is quite small relative to other determinants of formality.

1. Motivation

In recent years the literature studying informal labor markets in Latin America has focused on analyzing whether there is a relationship between growing informality rates in the region and the introduction of social protection programs that provide benefits without a direct cost over the worker's earnings.

In particular, according to certain authors (see for instance, Baeza and Packard 2006; Perry et. al. 2007, Wagstaff, 2010; Levy, 2008) the introduction of non-contributive programs that provide health benefits to disadvantaged groups generates disincentives to search for employment in the formal sector. The idea underlying this hypothesis is simple: the existence of free social protection programs generates fewer incentives for workers to look for formal sector jobs, where fringe benefits are paid in part by the worker through a payroll tax.

In the case of Mexico, Levy, 2008, has argued that the expansion of such non-contributive social protection programs has generated an increase in the informality rates in the economy, and that this thwarts growth in the long-run.

In spite of the relevance of this topic from both the academic and the public policy perspective, there are to date few empirical studies testing this hypothesis.

The goal of this paper is to study whether the introduction of the health protection program Seguro Popular in Mexico starting in 2002 has generated disincentives to participate in the formal sector. Besides addressing a relatively recent topic in the literature, this study takes a different

methodological approach from others by allowing for a more flexible labor market structure than previous studies.

In particular, in our econometric model we allow for the possibility of rationing of formal sector jobs. This is relevant because if formal sector jobs are rationed then in order to quantify the impact of social protection programs on formality we need to look at the *willingness* to search for formal sector jobs, and not only at the observed formality rates per-se.

While the observed levels of formal employment are the product of market forces that include both workers and employers' decisions, the propensity to apply to the formal sector is a clearer measure to gauge the impact of social protection programs on the incentives workers have to become formal.

We use an econometric specification that separates the worker's application decision to a formal sector job, from the formal sector employer's hiring decision. This allows us to capture the impact of the introduction of the Seguro Popular on the propensity to look for jobs in the formal sector.

Our findings show a negative, yet small, impact of the program on the propensity to look for formal sector jobs. In particular, we estimate that a 10 percentage point increase in the coverage of Seguro Popular reduces the willingness to look for formal sector jobs by about 0.5 percentage points, with larger impacts in rural areas. Relative to the impact of other determinants of the willingness to apply to the formal sector, the Seguro Popular has a negligible effect.

The paper is structured as follows. Section 2 discusses in detail the question to study and the theoretical framework used. Section 3 presents the relevant literature review. The adopted

econometric specification is presented in section 4, while section 5 describes the data used in the paper. Section 6 presents the results, and section 7 concludes.

2. Theoretical Framework

As previously mentioned, the goal of this paper is to analyze whether the introduction of the Seguro Popular health protection program in Mexico generated disincentives to formality in the labor markets.

The main difference between our paper and previous studies analyzing this question is that we propose an econometric specification that does not presume the existence of competitive labor markets, instead it allows for the possibility of rationing of formal sector jobs.

If labor markets are not competitive then studying either aggregate rates of formality or affiliation rates at Social Security institutions does not necessarily capture the disincentives that might be generated by social protection programs like Seguro Popular.

We argue that if labor markets are not competitive, then we should look at the workers' propensity to apply to a formal sector job once the social protection program is introduced. This propensity is a more direct measure of the willingness to enter the formal sector, than formal employment rates per-se, the latter being a measure affected both by the workers' willingness to look for a formal sector job, and by the availability of such jobs in the economy.

Under competitive labor markets this distinction is not relevant given that in equilibrium, frictions aside, the fraction of workers applying to the formal sector is equal to the rate of formal

employment. However, in markets where formal jobs are rationed both rates are not necessarily equal since there might be informal sector workers who would rather be in the formal sector. Furthermore, under such non-competitive setting the introduction of social protection programs might reduce the willingness to apply to formal sector jobs, without a strong impact on the observed formality levels.

To better illustrate this point consider a hypothetical economy with three workers where one of them is employed in the formal sector, another is willingly employed in the informal sector, and the third one is also employed in the informal sector, however, he would rather have a formal sector job.

Now it is possible that after the introduction of a non-contributive social protection program the voluntary informal sector worker remains informal, yet the involuntary informal worker becomes a *voluntary* one, since now the gains from entering the formal sector are smaller. In this imaginary case we have that the creation of the social protection program disincentives formality even if the rates of formal employment remain unchanged in the economy.

Under a competitive market this would not occur since in that case there are no involuntary informal sector workers, and everybody is employed in the sector of their preference. Hence in that case, it is enough to look at the employment share of the formal sector before and after the introduction of the social protection program, after conditioning for other factors.

The methodological implication of the previous discussion is that if we want to test whether social protection programs disincentive formality we need to empirically capture the impact of such

programs on the propensity to apply to formal sector jobs, and not just on the realized formality rates.

3.   Literature Review

So far few empirical studies have directly addressed the impact of social protection programs on the sector composition of the labor market.

Among the studies presenting a theoretical perspective of the labor market impacts of having non-contributive social protection programs in parallel to contributive social security systems in the formal sector we can mention Baeza and Packard, 2006; Perry et. al., 2007; Levy, 2008 and Wagstaff (2010).

The book by Baeza and Packard focuses on Latin America addressing the welfare costs associated with health shocks, especially among poor households with a limited access to a health care system. The authors study the gaps remaining in the social protections systems in the region, and analyze possible alternatives to remedy such gaps.

The authors acknowledge the growing contradiction within a mixed system with both contributive and non-contributive schemes in parallel, and discuss the disincentives this might generate for formality in the labor markets. In particular, the authors mention empirical evidence regarding the negative work impact of payroll taxes, yet not precisely on the degree of formality in the economy.

The report by Perry and coauthors (2007) is a broad study on informality in Latin America analyzing its multiple causes, including the possibility of rationing of formal sector jobs. Among the causes of the rising informality in the region the authors mention the introduction of non-contributive social protection programs as Mexico's Seguro Popular. As Baeza and Packard, the authors discuss the growing mismatch between the formal sector payroll taxes and the benefits packages granted to workers in the formal sector. In particular, if poor workers value more a higher wage today than such benefits package then they will opt to become informal sector workers, and this effect will be reinforced if such workers can find protection in public non-contributive programs.

The paper by Wagstaff (2010) reproduces several of these arguments in its review of the economic impacts of contributive health care systems.

More directly related to the Mexican case, Levy 2008 develops many of the aforementioned arguments to criticize the way social protection is provided today in Mexico. Besides pointing out the disincentives to formality, the author argues that the rise in social protection programs is limiting the growth potential of the economy, since they create a reallocation of labor toward low-productivity sectors.

Evidence on the Impact of Seguro Popular on the Labor Markets

There is a growing literature studying whether the Seguro Popular program disincentives formality in the economy.

The paper by Parker and Scott (2008) uses data from the 2000 Population Census and the 2005 "Conteo de Población" (a mid-decade large survey to estimate the population at the national

level) together with administrative records from participation on the Seguro Popular. The authors complement their analysis with data from the Mexican Family Life Survey (MxFLS) for the years of 2002 and 2005.

The analysis with the Census and Conteo samples is done at the municipality level. In particular, the authors compare the change in the formality rates (measured with the affiliation at the Instituto Mexicano de Segurida Social, IMSS, the private sector contributive social security scheme), between municipalities with and without Seguro Popular in 2005. They also compare how formality is affected by the penetration of this program in terms of number of affiliates and the time a municipality has been participating in the program.

Given that the Difference-in-Difference (DID) approach implemented at the municipal level could be biased if program participation is not random, the authors control for a set of observed characteristics of the municipality to diminish such potential bias.

The authors fail to find any influence of the Seguro Popular in formality rates at the municipal level under any of their econometric specifications. In most of their estimations the impact of the program is not statistically significant, and whenever it is statistically different from zero, the impact has the opposite sign to the one predicted by theory, namely they sometimes find that municipalities enrolled in the Seguro Popular have a higher growth in formality.

The authors recognize that such results could mean either a lack of negative effect from the part of the Seguro Popular or that the data has some hidden problems not controlled for by the empirical approach adopted. One particular problem mentioned is that in rural areas, individuals might be misreporting participation in the formal sector if they confuse enrollment in IMSS with

usage of the health services at IMSS-Oportuidades clinics (i.e. health clinics targeted to the poor, but without any relationship with having a formal sector job). Another potential problem is that there might be substantial variation over time in non-observable characteristics, something that renders invalid the causal interpretation of parameters arising from a DID strategy implemented.

In order to test the robustness of their findings the authors complement their estimations with other ones from the MxFLS surveys in 2002 and 2005. This survey is a multipurpose survey similar to the PSID and is one of the few panel data available for the Mexican economy. With this sample they define an individual to be covered by the Seguro Popular if such individual has a family member affiliated to the program. Given that there might be important differences between participants and non-participants the authors follow a propensity score difference-in-differences approach to capture the causal impact of interest. This method corrects potential selectivity biases that might arise from observable characteristics of individuals, or by time-invariant non-oberservable factors.

In their different econometric exercises with this sample the authors find that beneficiaries of the Seguro Popular have a lower probability of being in the formal sector by about 13 to 15 percent in rural areas, and about 7 percent in urban areas. Hence they conclude that such program leads to strong disincentives to formality in rural areas.

The paper by Campos-Vázquez and Knox (2008) uses the quarterly samples of the Encuesta Nacional de Empleo (ENE) a nationally representative survey designed to study the labor markets, as well as the ENCELURB (Oportunidades), the urban complement of a survey targeting poor households in the poverty-alleviation program Oportunidades. They years they analyze go from 2002 to 2004, and they analyze both aggregate as well as individual level data.

In their DID strategy with aggregate data the authors find no evidence of a negative effect of the Seguro Popular in formality rates at the local level. In fact, all their estimates are statistically non-significant. Similar results are found when they perform an analysis with individual level data. For this reason the authors conclude that if the Seguro Popular has any impact whatsoever on formality rates in the labor markets it must be a small one.

The study by Barros (2009) uses the Encuesta Nacional de Ingreso y Gasto de los Hogares (ENIGH) a nationally representative household survey tracking the income, consumption, and labor status of the population for the years 2000, and 2004-2006. He analyzes the impact of time and regional variations in the supply of the Seguro Popular (as measured by the targeted population to enroll as a fraction of the eligible population) on the decision to become informal, and on informal sector wages following DID and Triple-Difference strategies, after including a set of sociodemographic controls. The author finds no evidence of any impact either on the decision to be informal or on wages in the informal sector.

More recently the study by Azuara and Marinescu (2010) addresses the same question as the previous studies using the temporal and regional variation in the introduction of the program. In particular, the authors estimate the impact of the introduction of the program at the municipal level on the individual probability of becoming informal, after including a set of controls. They estimate their models for different population subsamples using data from the urban subsamples of ENE and it's follow-up survey the Encuesta Nacional de Ocupación y Empleo (ENOE), described below. For the whole population the authors cannot find any impact of Seguro Popular on formality, although they find a relatively small yet statistically significant negative effect among uneducated and older workers, as well as among married workers with children.

In a closely related paper, Bosch and Campos-Vázquez (2010) use administrative records from IMSS on the number of registered employers and employees at the municipal level for the years 2000 to 2009. As explained before, this captures the formal private sector of the economy. The authors exploit the geographical and time variation in the introduction of Seguro Popular at the municipal level to study whether it had any impact on the percent change in formal employment and in formal firms registered at IMSS, by using a DID strategy.

Unlike previous studies, the authors find a negative and statistically significant impact of Seguro Popular on both the number of firms registered and workers covered by IMSS. More precisely, their results indicate that since the program started in 2002 and up to 2009, there would have been 300,000 more formal jobs and 31,000 more formal employers. These represent 2.4 and 3.8% of the stocks of formal employees and employers respectively in 2002 when the program started. These effects are important and they are mainly found among small and medium firms in small municipalities.

The contrasting results between these last two papers may be due to the different dependent variables analyzed. In particular, in the first case the individual-level variable of formal status is analyzed, while the second paper studies the change in enrollment rates at IMSS at the municipal level. This means that the paper by Bosch and Campos-Vázquez cannot distinguish between movements out of formality and movements out of the labor force.

Finally, a study following a different methodological approach is the one by Aguilera (2010). This paper uses information from a randomized experiment conducted in late 2005-early 2006 to assess the impact of Seguro Popular on health related indicators and out-of-pocket health expenditures. Although the experiment was not designed to assess the impact of the program on

informality, the randomization strategy is used to compare enrollment in IMSS in treated and control cluster units, where treatment meant encouraging enrollment in the program. The author uses a DID strategy and fails to find any statistically significant effect of the treatment on formality either at the aggregate or at the individual level in urban areas.

As we previously discussed, an analysis on the impact of the Seguro Popular on the actual formality levels does not necessarily capture the disincentives this program might create in the labor market. In the next section we present the methodology proposed to identify the willingness to apply to a formal sector job and its determinants.

4.   Methodology

In order to estimate the probability of applying to a formal sector job under a model that allows for the possibility of rationing we depart from a sector allocation model similar to Roy 1951, Lee 1979 y 1995, but we now add an additional equation that captures the hiring decisions of formal sector employers.

In particular, consider an unobserved utility V_a (for the workers) from being employed in the formal sector. This utility is a function of both observable characteristics Z, as well as non observed factors captured by ϵ_a.

$$V_a = Z\gamma_a + \epsilon_a \tag{1}$$

This utility function can be interpreted as the net value to the worker of being formal, i.e. as the net value relative to being an informal sector worker.

There is also a propensity V_h for being hired in the formal sector, which is a function of both observable X and non-observable characteristics ϵ_h.

$$V_h = X\gamma_h + \epsilon_h \tag{2}$$

This propensity captures the factors affecting the formal sector employer's decision to hire a worker with observable characteristics X.

Under this framework a worker would be employed in the formal sector if *both* his utility of being formal is positive, and the "utility" of a formal sector employer is positive, i.e. if $V_a > 0$ y $V_h > 0$.[1]

Following this notation we can classify informal sector workers depending on whether they are voluntary, i.e. if $V_a \leq 0$ or if they are involuntary, i.e. if $V_a > 0$ and $V_h \leq 0$.

In other words, the probability of being formal depends both on V_a y V_h, since

$$P(Formal = 1) = P(V_a > 0, V_h > 0)$$

while the informal sector is composed of a mixture of voluntary and involuntary workers.

From an empirical point of view we do not observe either V_a nor V_h, but only the sector where the worker is employed.

---

[1] The normalization of the utility threshold to zero is irrelevant for the analysis.

Given that we can only estimate the determinants of $P(Formal = 1)$, but not of $P(V_a > 0)$ nor of $P(V_h > 0)$ separately we have an identification problem. For this reason we cannot estimate the parameters γ_a and γ_h without additional assumptions.

In order to identify such parameters we present an extension to the procedure proposed by Dickens and Lang (1985) now considering two selection equations (one of the worker and another of the formal sector employer), and three subpopulations instead of two (the formal sector workers, the voluntary informal sector workers, and the involuntary informal sector workers).

In addition to the previous equations we have three earnings equations, one for each subpopulation group, i.e. formal workers, and voluntary and involuntary informal sector workers, respectively.

$$y_1 = X\beta_1 + v_1 \tag{3}$$

$$y_2 = X\beta_2 + v_2 \tag{4}$$

$$y_3 = X\beta_3 + v_3 \tag{5}$$

In order to help us with the identification of the model parameters we assume that some of the observable variables that affect the decision of the worker do not affect the decision of the formal sector employer, i.e. $X \subset Z$. Among the variables excluded from X we have variables that are known to the worker, but non-observable or non relevant for the formal sector employer. These exclusion restrictions are not, however, necessary for the identification of the model, although they make sense from an economic point of view. For the identification of the parameters of the model it

14

suffices to consider different earnings equations for voluntary and involuntary informal sector workers.

Given that the choice of sector is not random we need to consider potential selectivity biases brought by unobservables when estimating equations (3) to (5). [2] For this we assume that the error terms in the previous equations follow a multivariate normal distribution conditional on the observed characteristics, i.e.

$$\zeta = \begin{pmatrix} \varepsilon_a \\ \varepsilon_h \\ \upsilon_1 \\ \upsilon_2 \\ \upsilon_3 \end{pmatrix} \sim N(0,\Sigma) \tag{6}$$

Under this assumption, and with the aforementioned exclusion restrictions we can formulate a likelihood function to estimate the parameters of the model $\gamma_a, \gamma_h, \beta_1, \beta_2, \beta_3$ y Σ. In particular, the log-likelihood function is given by

$$\ln L = \sum_{form=0} \ln\left(L_{2i} + L_{3i}\right) + \sum_{form=1} \ln\left(L_{1i}\right)$$

where *form* is a dummy variable indicating whether the worker is employed in the formal sector and $L_{1,i}$ is the contribution of formal sector workers to the likelihood, i.e.

---

[2] In all the earnings equations the dependent variables are ln earnings.

$$L_{1,i} = \frac{1}{\sigma_1} \phi\left(\frac{y_{1i} - X_i\beta_1}{\sigma_1} \right)$$

$$\Phi_2\left(\frac{Z\gamma_a + \dfrac{\sigma_{a1}}{\sigma_1}\left(\dfrac{y_{1i} - X_i\beta_1}{\sigma_1} \right)}{\sqrt{1 - \left(\dfrac{\sigma_{a1}}{\sigma_1} \right)^2}}, \ \frac{X\gamma_h + \dfrac{\sigma_{h1}}{\sigma_1}\left(\dfrac{y_{1i} - X_i\beta_1}{\sigma_1} \right)}{\sqrt{1 - \left(\dfrac{\sigma_{h1}}{\sigma_1} \right)^2}}, \ \frac{\sigma_{ah} - \dfrac{\sigma_{a1}\sigma_{h1}}{\sigma_1}}{\sqrt{\left(1 - \left(\dfrac{\sigma_{a1}}{\sigma_1} \right)^2 \right)\left(1 - \left(\dfrac{\sigma_{h1}}{\sigma1} \right)^2 \right)}} \right)$$

$L_{2,i}$ is the contribution of voluntary informal sector workers to the likelihood, i.e.

$$L_{2,i} = \frac{1}{\sigma_2} \phi\left(\frac{y_{2i} - X_i\beta_2}{\sigma_2} \right) \Phi\left(\frac{Z\gamma_a + \dfrac{\sigma_{a2}}{\sigma_2}\left(\dfrac{y_{2i} - X_i\beta_2}{\sigma_2} \right)}{\sqrt{1 - \left(\dfrac{\sigma_{a2}}{\sigma_2} \right)^2}} \right)$$

$L_{3,i}$ is the contribution of involuntary informal sector workers to the likelihood, i.e.

$$L_{3,i} = \frac{1}{\sigma_3} \phi\left(\frac{y_{3i} - X_i\beta_3}{\sigma_3} \right)$$

$$\Phi_2\left(\frac{Z\gamma_a + \dfrac{\sigma_{a3}}{\sigma_3}\left(\dfrac{y_{3i} - X_i\beta_3}{\sigma_3} \right)}{\sqrt{1 - \left(\dfrac{\sigma_{a3}}{\sigma_3} \right)^2}}, \ -\frac{X\gamma_h + \dfrac{\sigma_{h3}}{\sigma_3}\left(\dfrac{y_{3i} - X_i\beta_3}{\sigma_3} \right)}{\sqrt{1 - \left(\dfrac{\sigma_{h3}}{\sigma_3} \right)^2}}, \ -\frac{\sigma_{ah} - \dfrac{\sigma_{a3}\sigma_{h3}}{\sigma_3}}{\sqrt{\left(1 - \left(\dfrac{\sigma_{a3}}{\sigma_3} \right)^2 \right)\left(1 - \left(\dfrac{\sigma_{h3}}{\sigma_3} \right)^2 \right)}} \right)$$

and σ_j is the equation j error term variance, while σ_{jk} is the covariance between the errors of

equations j and k. Finally, ϕ, Φ, y Φ_2 are the univariate normal density, the univariate normal cdf,

and the bivariate normal cdf respectively.

In order to capture the impact of the Seguro Popular on the willingness to be formal we include a variable capturing the presence and penetration of this protection scheme at the local level in equation (1), namely in the equation of the propensity of workers to apply for a formal sector job.

The previous model has several advantages over other models used in the literature. First, it allows for the possibility of job rationing in the formal sector of the economy. Without explicitly accounting for this factor in the model one would presume *ipso facto* that labor markets are perfectly competitive.

On this issue it is important to emphasize that while our model allows for the possibility of rationing, it does not impose such rationing, but rather lets the data determine the amount of involuntary workers in the informal sector. For instance, were the slopes in the vector γ_h zero in a joint test of statistical significance, it would mean that the hiring equation is not needed to model sector selection; taking us back to the standard Roy model with free entry.

Secondly, it captures explicitly the possibility that both observables and unobservables affect the sector allocation decisions, as well as earnings. In other words, it relaxes the assumptions of conditional independence used in propensity score models.

Thirdly, it does not assume income maximization as some papers in the literature have done. This assumption is unrealistic given that usually individuals might value the formal sector for reasons other than just wages. Fringe benefits and jobs stability might be valued too at the formal sector, while also freedom, less regulation, and the benefits of social protection programs like Seguro Popular might be valued by workers in the informal sector.

Finally, the model does not rely on sector transitions from a panel data, and hence it does not suffer from potential attrition bias that might arise in this type of data. Although a dynamic version of the present model would be an interesting extension to do, it would require controlling explicitly for the potential biases brought by attrition. Such extension is left for future work.

Among the disadvantages of the model we must mention its reliance on strong distributional assumptions. In particular, both the possibility to account for the impact of unobservables on the selection process and the separation of the informal population in voluntary and involuntary groups rely on the normality assumption (6). If that assumption proves to be false then our estimates may be biased.

It is also important to mention that the present model best describes the labor market decisions of prime-age males and women permanently attached to the labor market. It does not make a good description of women who move in-and-out of the labor force, because none of the equations previously discussed account for entries and exits from the labor force. For this reason all the estimations presented next will be restricted to males.

5.   Data and Specification of the Equations

In order to study the impact of the introduction of the Seguro Popular on the labor markets during the years going from 2002 to 2004 we use data from the Encuesta Nacional de Empleo. For the years 2005 to 2009 we use the Encuesta Nacional de Ocupación (ENOE). We also use administrative data from the program in order to estimate the degree of its presence at the state level.

Both the ENE and the ENOE are nationally representative surveys conducted on Mexican households with the purpose of inquiring about the conditions that prevail in labor markets.

The surveys gather information about socioeconomic characteristics such as age, gender, education, marital status, labor force participation, labor market earnings, sector of employment, occupation, type of fringe benefits, hours worked in the market, type of employment contract, firm size, employment search activity, etc. They are also geographically representative at the state and city level.

The unit of analysis is the individual worker. The subpopulation of study is restricted to employed males between 20 and 60 years of age who report having positive earnings in their work. In other words, we focus on male workers who are more permanently attached to the labor market.

Although these surveys are short-lived rotating panels, the analysis is conducted only on a cross-sectional dimension of the data. In particular, we use data coming from the second quarter of the year.

The definition of informality is based on a social security coverage criterion. In other words, a worker is classified to be in the formal sector if he reports having coverage at IMSS, ISSSTE or a private health insurance provided by his current employment.

The earnings variable is the real monthly eanings measured in 2002 Mexican pesos.

The variables that enter the application equation (1) include individual characteristics like the age and level of education of the respondent, family characteristics like the dependency ratio at the household level, i.e. the ratio of potential dependents to potential earners at the household level, the

fraction of females in the household, as well as a set of controls for local labor market conditions like regional dummies, the share of formal employment at the local level in the year 2001, before Seguro Popular started, and a dummy variable for urban regions. Time dummies and their interactions are also included since we pool data over several years.

In this equation we also include a set of variables that capture the presence of the Seguro Popular at the State where the individual lives.

To measure the presence of Seguro Popular in its initial phase (2002-2004), when not all states had adopted the program, we use a dummy variable indicating the presence of this program at the state level. After the year 2005 all states participate in the program, hence we use the penetration of the program at the state level, as measured by the fraction of the total state population covered by the program.

Given the heterogeneity in the labor markets between urban and rural areas we include an interaction of the Seguro Popular variables previously mentioned with the urban area dummy.

If the parameters of these variables are negative and statistically different from zero then we can claim that the Seguro Popular disincentives participation in the formal sector.

The model is estimated over two subsamples. The first one corresponding to the introductory phase of the program from 2002 to 2004, and the second one over the consolidating phase going from 2005 to 2009.

Since we are pooling several years together we also interact the variables capturing the presence of Seguro Popular at the state level with time dummies, both for urban and rural areas. In

other words we estimate a separate effect of the program for each year in the sample, and for urban and rural areas.

It is important to remark that we do not use individual-level information on whether a worker is affiliated at the Seguro Popular since this information is highly endogenous with the dependent variable of sector of employment. In particular, it is very likely to find a negative association at the individual level between participation in the formal sector and affiliation at Seguro Popular, not necessarily because such scheme disincentives formality, but because informal sector workers will seek affiliation to the program once it becomes available. The use of variables aggregated at the state level minimizes substantially the risk of having such endogeneity bias in our estimations.

Given that workers living in regions with high levels of informality may be less prone to look for jobs at the formal sector, and these regions also have a higher penetration of Seguro Popular, it is important to control by the degree of informal employment at the local level to avoid a potential omitted variable bias. As previously mentioned, we do so by introducing the local fraction of formal employment in 2001, before Seguro Popular started operating.

As previously mentioned, the variables entering the hiring equation (2) include only a subset of the ones entering equation (1). In particular, we assume that the employer only observes or cares about the productive characteristics of the worker, and hence bases his hiring decision on variables like age and education of the worker in addition to the variables controlling for local labor market characteristics. In other words we exclude variables capturing household characteristics, that in principle are not observable to the formal sector employer. Also the variables capturing the presence of Seguro Popular are excluded.

Although it would be desirable to include characteristics of the employer in this equation, we have a problem of observability since we do not know to what type of formal employer the rationed informal sector workers have applied to. Thus, any characteristic of the potential formal employer will have to be captured by the unobserved error term ε_h. The earnings equations (3) to (5) use the same set of regressors of equation (2).

Also, it is important to remark that all the calculations here presented are weighted estimations using the survey factor weights. This is done in order to obtain estimates that are representative of the underlying population of interest.

6.   Results

In this section we present the results of the estimations for the subsamples going from 2002 to 2004 and from 2005 to 2009. However, before that we briefly present some descriptive statistics of the data.

Descriptive Statistics

Tables 1 and 2 present the descriptive statistics for both samples. As previously mentioned we restrict the sample to males with ages between 20 and 60 years of age.

The average age on this subpopulation is of 37 years, while from 2002 to 2004 the educational category with more workers is the Elementary, followed by Intermediate and Higher.

However, due to the increase in educational attainment over time in the second subsample (2005-2009) we have that almost half of our sample has Intermediate schooling.[3]

Regarding the household level variables we have that on average there are 2 active workers for each 1.2 dependents, and women represent around 46 percent of the household population.

Finally, regarding local labor market variables we have that in the year 2001, before the implementation of the Seguro Popular, the local formality rate was about 34 percent of the employed population, with very low levels in rural areas.[4] We also observe that more than 50 percent of the sample lives in the Center and South regions of the country, between 25 and 30 percent live in the North, and the remaining 18 percent does so at Mexico City.

In Figure 1 and Table 3 we present the growing coverage of individuals in the sample by Seguro Popular. In particular we observe that in the consolidation phase of the program (from 2005 onwards) the coverage of the population has grown by about twenty percentage points.

Econometric Estimations

In Table 4 we present the results of the estimations for the initial phase of the program from 2002 to 2004. This table includes the parameters of the model, as well as the variance-covariance matrix.

In these estimations we observe that age has a negative effect on the probability of applying to a formal sector job, but a positive one in the probability of being hired at this sector. This can be

---

[3] For precise definitions of each variable used in the model refer to the Appendix 1
[4] We define a rural area as one with less than 2,500 inhabitants and they comprise about 20 percent of the sample.

explained by the fact that over time workers tend to move into self-employment, while formal firms value the experience accumulated by workers.

Having an intermediate education level increases the probability of being in the formal sector, since both the application and hiring decisions are positively affected by it. However, having a higher education reduces the probability of seeking a formal sector job, but it substantially increases the chances of being hired in this sector.

Living in areas with high rates of formality in 2001 increases the chances of becoming formal, while living in urban areas increases the probability of applying to this sector, yet it diminishes the chances of being hired by a formal sector firm. An opposite effect occurs with the regional dummies whereby living in areas other than Mexico City reduces the chances of applying to the formal sector, although it increases the probability of being hired in this sector.

With respect to the impact of the additional set of exclusion restrictions that affect the probability of applying to the formal sector, we have that workers living in households with a high number of dependents are less likely to search for jobs in the formal sector, probably because the informal sector allows them more flexible schedules to take care of such dependents. Also, workers living in households without females have a smaller probability of applying to the formal sector, since women usually take care of housework chores, allowing men to specialize in full-time jobs like the ones offered at the formal sector. However, a high fraction of females in the household slightly reduces the propensity to look for formal sector jobs.

The presentation of the main results for the variables capturing the impact of the Seguro Popular is postponed until the end of the discussion of the models.

Earnings increase with age (in other words with experience) for workers in the formal sector and for involuntary informal workers. However, we find the opposite effect in the earnings equation of the voluntary informal. Also earnings grow with education, and in particular there are high returns to higher education in the formal sector.

Finally, in terms of the regional variables we find that living in regions with high formality rates and in urban areas is associated with higher earnings. Workers living in the North of the country earn more than their counterparts in Mexico City, while living in the Center and South of the country is associated with lower earnings for the voluntary informal.

The results of the econometric estimations for the years 2005-2009 are presented in Table 5. Most of the parameter estimates are quite similar to the ones obtained for the period 2002-2004. The only differences are that living in urban areas is now associated with a higher probability of being hired in the formal sector. Also the regional dummies have different signs in the application equation and in the earnings equation for the voluntary informal sector workers.

The main goal of this study is to estimate the impact of Seguro Popular on the probability of searching for a formal sector job. Given that the parameter magnitudes in the previous tables are not directly interpretable in term of probabilities, we present instead in Tables 6 and 7 the average marginal effects of these variables on the probability of applying to a formal sector job. These tables contain the main results of the paper.

In Table 6 we see that in 2002 living in a state that provided coverage by the Seguro Popular increased the individual propensity to apply to a formal sector job by 0.8 percentage points, in

contradiction to the predictions of the theory. However, this effect is very small from any practical point of view.

During 2003 living in a state providing coverage by Seguro Popular reduced the probability of applying to formal sector jobs such as predicted by the models previously discussed. However, again these effects are quite small since in urban areas the reduction amounts to 0.7 percentage points, while for rural areas such reduction is about 3 percentage points. For 2004 the effect of this variable was negative and statistically significant only in urban areas where it generated a reduction of 0.4 percentage points.

In other words, from 2003 onwards we start seeing a negative effect of the presence of Seguro Popular on the desire to work at the formal sector. However, the effect is very small in comparison with other factors affecting the decision to become informal. This small magnitude explains why many other studies focusing on the formality rates per-se (and not on the propensity to apply to a formal sector job) and have failed to find any significant impact of the Seguro Popular. In particular, if we find a very small effect of Seguro Popular discouraging application to formal sector jobs, then the impact of this program on the realized formality rates will probably be even smaller, since some of these discouraged workers would not get a job in the formal sector anyway.

In Table 6 we also include the results at a national level for the years 2002-2004. Again there we observe that the overall impact of Seguro Popular on the willingness to search for formal sector jobs is negative. The largest impacts occur in rural areas, but in these areas the jobs covered by Social Security institutions were very rare even before the creation of the program.[5]

---

[5] These results are consistent with the ones found by Parker and Scott 2008.

In Table 7 we present the marginal effects of the variables capturing the coverage rate of Seguro Popular starting on 2005. Given that this penetration variable is calculated as the number of beneficiaries as a fraction of the population at the state level, we have to divide the reported marginal effects by 100 if we want to obtain the impact of increasing the coverage by one percentage point at the state level.

In this table we observe that for all the years and regions a higher penetration of Seguro Popular reduces the probability of applying to the formal sector. However, same as with the previous table this reduction is small.

To illustrate this point consider the year 2005 when the impacts are larger in absolute value. In this case we have than an increment in ten percentage points in the penetration of Seguro Popular at a national level would reduce the probability of applying to formal sector jobs by 0.8 percentage points. In urban areas the effect would be -0.56 percentage points, while in rural areas it would be of -1.8 percentage points.

Over time this impact becomes smaller in absolute value in urban areas, although not necessarily in rural areas. In any case all the effects found in other years are quite small.

The reasons that might explain why the largest negative impacts are found in rural areas include that in such areas the package of benefits offered by the formal sector is less attractive given the limited access to clinics of IMSS, ISSSTE or to private hospitals. Also, the population living in communities with less than 2,500 habitants is in its majority poor and hence they have a high opportunity cost of paying the contributions to social security.

7.   Conclusions

This paper analyzed whether there is a disincentive to formality by the introduction of a social protection program. More specifically we analyze the impact of Mexico's Seguro Popular health program on the probability to look for a formal sector job.

Unlike other papers, we don't assume labor markets are perfectly competitive and distinguish between the decision to look for a formal sector job, and the fact of actually getting one. To make such distinction we propose an econometric model that separates the applying from the hiring decisions in the formal sector. Even if such decisions are not observable in most labor surveys, with the use of exclusion restrictions and distributional assumptions we can estimate the determinants of each one of these decisions. Variables measuring the presence and penetration of Seguro Popular were incorporated as determinants of the propensity to seek formal sector jobs.

Our paper differs from others on the literature on Social Protection and informality in that it allows for the possibility of barriers to entry to the formal sector. This is relevant because if labor markets are not competitive then the impacts of social protection programs on formality rates do not necessarily reflect the impact of such programs on the willingness to look for formal sector jobs.

To avoid potential endogeneity biases we measure the presence and penetration of Seguro Popular at the aggregate level and use this variable in an individual analysis of sector allocation.

The results show that for most of the years studied Seguro Popular generated disincentives to become formal among workers in the economy. However, the negative effects we find are consistently small in comparison with other determinants of the sector allocation decisions. These disincentives are larger in rural areas.

We think of two possible reasons that may lie behind the fact that Seguro Popular has a small impact on the decision to become informal. The first is that health benefits may matter little to workers when choosing which sector to work at. In other words, it is possible that other factors as salary, flexibility in schedules, independence, etc. matter more when choosing whether to apply or not to the formal sector.

The second hypothesis is that workers do not fully value the health services covered by Seguro Popular, and they might not consider them fully comparable to the ones offered by formal sector jobs (Barros 2008 provides some evidence in this direction). This would mean that a higher penetration of Seguro Popular would not imply large increases in informality, but that workers already in the informal sector decide to enroll in this protection program.

To better assess to what extent these hypotheses explain the weak effects of Seguro Popular it is necessary to conduct further research on the reasons underlying the sector allocation decisions of workers and firms.

Bibliography

Aguilera, Nelly (2010) "The Short Term Causal Effects of Seguro Popular on Formal Employment", unpublished manuscript

Azuara, Oliver and Ioana Marinescu (2010) "Informality and the Expansion of Social Protection Programs. Evidence from Mexico" unpublished manuscript

Baeza, Cristian C. and Truman G. Packard (2006) *Beyond Survival. Protecting Households from Health Shocks in Latin America*, Stanford University Press

Barros, Rodrigo (2009) "Wealthier But Not Much Healthier: Effects of a Health Insurance Program for the Poor in Mexico", SIEPR Discussion Paper 09-002.

Bosch, Mariano and Raymundo Campos-Vázquez (2010) "The trade-offs of social assistance programs in the labor market: The case of the Seguro Popular program in Mexico", unpublished manuscript

Campos-Vázquez, Raymundo M. and Melissa Knox (2008) "Social Protection Programs and Employment: The Case of Mexico's Seguro Popular Program" unpublished manuscript

Dickens, W. T. and K. Lang (1985), "A Test of Dual Labor Market Theory", *American Economic Review*, vol. 75(4): 729-805

Lee, Lung-Fei (1979) "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables" *Econometrica*, 47:977-96

Lee, Lung-Fei (1995) "The Computation of Opportunity Costs in Polychotomous Choice Models with Selectivity", *The Review of Economics and Statistics*, vol. 77, issue 3, pages 423-35

Levy, Santiago (2008) *Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico.* Brooking Institution Press, Washington DC

Parker, Susan and John Scott (2008) "An Evaluation of the Mexican Seguro Popular program: Work and Formal sector Participation", unpublished manuscript

Perry, Guillermo E., William F. Maloney, Omar S. Arias, Pablo Fajnzylber, Andrew D.Mason, and Jaime Saavedra-Chanduvi (2007) *Informality. Exit and Exclusion*, World Bank Publications

Roy, A. (1951) "Some thoughts on the Distribution of Earnings", *Oxford Economic Papers*, 3, 135-146

Wagstaff, Adam (2010) "Social Health Insurance Reexamined", *Health Economics,* 19, 503–517

Tables

Table 1. Descriptive Statistics 2002-2004

	Mean	Std. Dev.	Min	Max
Age	37.1	10.82	20	60
Education Level				
Elementary	0.43	0.49	0	1
Middle	0.41	0.49	0	1
Higher	0.17	0.37	0	1
Dependency Ratio	0.62	0.61	0	7
Fraction of Women	0.46	0.19	0	0.91
% Formal Empl. 2001	0.34	0.17	0.09	0.67
Urban	0.77	0.42	0	1
Region				
DF	0.19	0.40	0	1
North	0.26	0.44	0	1
Center South	0.54	0.50	0	1
Log Earnings (2002 Pesos)	7.9	0.97	1.57	12.31

Source: Authors' calculations based on ENE 2002-2004

Table 2. Descriptive Statistics 2005-2009

	Mean	Std. Dev.	Min	Max
Age	37.3	10.78	20	60
Education Level				
Elementary	0.36	0.48	0	1
Middle	0.47	0.50	0	1
Higher	0.17	0.38	0	1
Dependency Ratio	0.58	0.57	0	7
Fraction of Women	0.46	0.19	0	0.90
% Formal Empl. 2001	0.34	0.16	0.09	0.67
Urban	0.81	0.39	0	1
Region				
DF	0.18	0.38	0	1
North	0.30	0.46	0	1
Center South	0.52	0.50	0	1
Log Earnings (2002 Pesos)	8.1	0.80	-0.30	12.59

Source: Authors' calculations based on ENOE 2005-2009

Figure 1 Fraction of the Sample Living in States with Coverage by Seguro Popular, 2002-2004
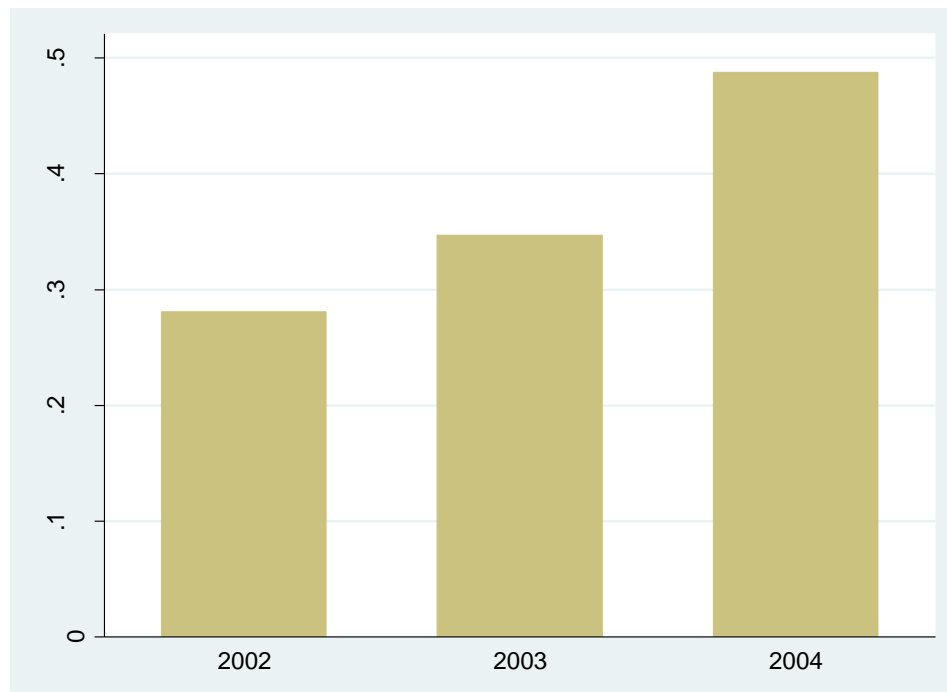
Source: Authors' calculations based on ENE 2002-2004

Table 3 Coverage Rate of Seguro Popular at the State Level, 2005-2009

Year	Mean	1st-Quartile	Median	3rd-Quartile
2005	0.15	0.06	0.11	0.19
2006	0.19	0.09	0.18	0.22
2007	0.26	0.17	0.26	0.32
2008	0.31	0.21	0.30	0.38
2009	0.34	0.24	0.33	0.42
Total	0.25	0.13	0.23	0.33

Source: Authors' calculations with administrative records from Seguro Popular and yearly estimations of population by CONAPO

Table 4. Parameters of the Econometric Model 2002-2004

	Sector Allocation Equations					
	Application			Hiring		
		Standard			Standard	
	Parameter	Error		Parameter	Error	
Constant	0.702	(0.0185)	***	-1.216	(0.0184)	***
Age	-0.028	(0.0003)	***	0.021	(0.0003)	***
Schooling Level						
Elementary (omitted)						
Intermediate	0.214	(0.0062)	***	0.446	(0.0071)	***
Higher	-0.018	(0.0079)	**	1.310	(0.0095)	***
Formality Rate 2001	1.351	(0.0252)	***	0.764	(0.0280)	***
Urban Area	0.547	(0.0104)	***	-0.025	(0.0089)	***
Region						
Mexico City (omitted)						
North	-0.103	(0.0084)	***	0.399	(0.0085)	***
Center South	-0.061	(0.0087)	***	0.207	(0.0085)	***
Year Dummies						
2002 (omitted)						
2003	0.095	(0.0099)	***	-0.073	(0.0073)	***
2004	0.049	(0.0099)	***	-0.002	(0.0072)	
No Dependents	-0.030	(0.0061)	***			
Dependency Ratio	-0.053	(0.0041)	***			
HHold w/out Females	-0.097	(0.0118)	***			
HHold Female Fraction	-0.051	(0.0129)	***			
Seguro Popular Variables						
Seguro Popular	0.022	(0.0101)	**			
SP x 2003	-0.101	(0.0129)	***			
SP x 2004	-0.010	(0.0128)				
SP x Urban Area	0.010	(0.0122)				
SP x Urban Area						
2003	0.041	(0.0122)	***			
2004	-0.038	(0.0120)	***			

(Continues)

Table 4. Parameters of the Econometric Model, 2002-2004 (cont.)

	Voluntary Informal Workers			Involuntary Informal Workers			Formal Workers		
	Parameter	Standard Error		Parameter	Standard Error		Parameter	Standard Error	
Constant	7.937	(0.0295)	***	7.351	(0.0089)	***	6.589	(0.0132)	***
Age	-0.022	(0.0004)	***	0.008	(0.0002)	***	0.018	(0.0002)	***
Schooling Level									
Elementary (omitted)									
Intermediate	0.766	(0.0107)	***	0.010	(0.0036)	***	0.349	(0.0050)	***
Higher	1.002	(0.0135)	***	0.105	(0.0080)	***	1.134	(0.0060)	***
Formality Rate 2001	2.655	(0.0439)	***	0.532	(0.0149)	***	0.659	(0.0187)	***
Urban Area	1.212	(0.0125)	***	0.057	(0.0042)	***	0.041	(0.0065)	***
Region									
Mexico City (omitted)									
North	0.078	(0.0152)	***	0.064	(0.0048)	***	0.141	(0.0055)	***
Center South	-0.034	(0.0157)	**	0.077	(0.0046)	***	0.092	(0.0056)	***
Year Dummies									
2002 (omitted)									
2003	0.067	(0.0106)	***	0.035	(0.0037)	***	-0.022	(0.0049)	***
2004	0.010	(0.0104)		0.042	(0.0037)	***	0.018	(0.0047)	***

Variance Covariance Matrix

cov(e_a, v_1)	1.292	(0.0084)	***
V(v_1)	2.213	(0.0182)	***
cov(e_a, e_h)	-0.608	(0.0051)	***
cov(e_a, v_2)	-0.232	(0.0038)	***
cov(e_h, v_2)	-0.063	(0.0051)	***
V(v_2)	0.183	(0.0012)	***
cov(e_a, v_3)	-0.284	(0.0068)	***
cov(e_h, v_3)	0.565	(0.0039)	***
V(v_3)	0.443	(0.0038)	***
No. Obs.	240,625		

***, **, * H_0: $\beta=0$ rejected at 1, 5, 10 percent of significance

Table 5 Parameters of the Econometric Model, 2005-2009

	Sector Allocation Equations					
	Application			Hiring		
	Parameter	Standard Error		Parameter	Standard Error	
Constant	0.757	(0.0181)	***	-1.725	(0.0176)	***
Age	-0.024	(0.0002)	***	0.017	(0.0003)	***
Schooling Level						
Elementary (omitted)						
Intermediate	0.160	(0.0054)	***	0.660	(0.0061)	***
Higher	-0.131	(0.0068)	***	1.753	(0.0091)	***
Formality Rate 2001	1.156	(0.0193)	***	1.144	(0.0221)	***
Urban Area	0.397	(0.0106)	***	0.192	(0.0085)	***
Region						
Mexico City (omitted)						
North	-0.006	(0.0074)		0.499	(0.0074)	***
Center South	0.057	(0.0077)	***	0.268	(0.0077)	***
Year Dummies						
2005 (omitted)						
2006	0.053	(0.0102)	***	-0.080	(0.0081)	***
2007	0.106	(0.0103)	***	-0.067	(0.0081)	***
2008	0.135	(0.0112)	***	-0.114	(0.0081)	***
2009	0.116	(0.0120)	***	-0.150	(0.0082)	***
No Dependents	-0.026	(0.0057)	***			
Dependency Ratio	-0.065	(0.0043)	***			
HHold w/out Females	-0.100	(0.0105)	***			
HHold Female Fraction	-0.069	(0.0124)	***			
Seguro Popular Variables						
Seguro Popular	-0.474	(0.0654)	***			
SP x 2006	0.087	(0.0740)				
SP x 2007	0.259	(0.0683)	***			
SP x 2008	0.305	(0.0671)	***			
SP x 2009	0.183	(0.0667)	***			
SP x Urban Area	0.289	(0.0716)	***			
SP x Urban Area						
2006	-0.086	(0.0738)				
2007	-0.257	(0.0688)	***			
2008	-0.234	(0.0671)	***			
2009	-0.063	(0.0665)				

(Continues)

Table 5 Parameters of the Econometric Model, 2005-2009 (cont.)

	Salary Equations								
	Volunatry Informal Workers			**Involuntary Informal Workers**			**Formal Workers**		
		Standard			**Standard**			**Standard**	
	Parameter	**Error**		**Parameter**	**Error**		**Parameter**	**Error**	
Constant	7.613	(0.0252)	***	7.416	(0.0080)	***	6.539	(0.0113)	***
Age	-0.011	(0.0004)	***	0.008	(0.0002)	***	0.014	(0.0002)	***
Schooling Level									
Elementary (omitted)									
Intermediate	0.583	(0.0083)	***	0.042	(0.0039)	***	0.402	(0.0040)	***
Higher	0.815	(0.0104)	***	0.113	(0.0111)	***	1.137	(0.0049)	***
Formality Rate 2001	2.354	(0.0312)	***	0.583	(0.0120)	***	0.645	(0.0127)	***
Urban Area	0.861	(0.0094)	***	0.088	(0.0035)	***	0.169	(0.0055)	***
Region									
Mexico City (omitted)									
North	0.050	(0.0113)	***	0.037	(0.0044)	***	0.182	(0.0042)	***
Center South	0.026	(0.0120)	**	0.060	(0.0040)	***	0.117	(0.0045)	***
Year Dummies									
2005 (omitted)									
2006	0.061	(0.0106)	***	0.045	(0.0038)	***	0.021	(0.0047)	***
2007	0.114	(0.0107)	***	0.055	(0.0038)	***	0.042	(0.0046)	***
2008	0.121	(0.0108)	***	0.067	(0.0038)	***	0.017	(0.0046)	***
2009	0.008	(0.0109)		0.001	(0.0038)		-0.047	(0.0047)	***
Variance Covariance Matrix									
cov(e_a, v_1)	1.044	(0.0096)	***						
V(v_1)	1.675	(0.0165)	***						
cov(e_a, e_h)	-0.108	(0.0112)	***						
cov(e_a, v_2)	-0.168	(0.0053)	***						
cov(e_h, v_2)	-0.021	(0.0077)	***						
V(v_2)	0.158	(0.0011)	***						
cov(e_a, v_3)	-0.039	(0.0072)	***						
cov(e_h, v_3)	0.500	(0.0024)	***						
V(v_3)	0.380	(0.0019)	***						
No. Obs.		355,646							

*** , ** , * H_0: $\beta=0$ rejected at 1, 5, 10 percent of significance

Table 6. Average Marginal Effects of Presence of Seguro Popular at the State Level on the Probability of Searching Formal Sector Jobs

	Marginal Effect	Asy. Std. Err.	
2002			
Global Effect	0.0086	0.0024	***
Urban Areas	0.0089	0.0026	***
Rural Areas	0.0080	0.0039	**
2003			
Global Effect	-0.0154	0.0027	***
Urban Areas	-0.0077	0.0027	***
Rural Areas	-0.0293	0.0042	***
2004			
Global Effect	-0.0014	0.0026	
Urban Areas	-0.0044	0.0026	*
Rural Areas	0.0045	0.0043	
2002-2004			
Global Effect	-0.0028	0.0015	*
Urban Areas	-0.0013	0.0017	
Rural Areas	-0.0057	0.0030	*

***, **, * H$_0$: β=0 rejected at 1, 5, 10 percent of significance

Table 7. Average Marginal Effects of Penetration of Seguro Popular at the State Level on the Probability of Searching Formal Sector Jobs

	Marginal Effect	Asy. Std. Err.	
2005			
Global Effect	-0.0802	0.0126	***
Urban Areas	-0.0565	0.0134	***
Rural Areas	-0.1797	0.0247	***
2006			
Global Effect	-0.0725	0.0108	***
Urban Areas	-0.0547	0.0111	***
Rural Areas	-0.1468	0.0215	***
2007			
Global Effect	-0.0583	0.0075	***
Urban Areas	-0.0530	0.0073	***
Rural Areas	-0.0811	0.0171	***
2008			
Global Effect	-0.0383	0.0078	***
Urban Areas	-0.0323	0.0076	***
Rural Areas	-0.0636	0.0156	***
2009			
Global Effect	-0.0363	0.0081	***
Urban Areas	-0.0185	0.0080	**
Rural Areas	-0.1098	0.0147	***
2005-2009			
Global Effect	-0.0572	0.0045	***
Urban Areas	-0.0432	0.0048	***
Rural Areas	-0.1163	0.0124	***

***, **, * H_0: $\beta=0$ rejected at 1, 5, 10 percent of significance

Appendix 1

This appendix includes a description of the variables used in the estimation of the model.

Dependent Variables

Form – Dummy Variable with value 1 for workers covered by the social security schemes of IMSS, ISSSTE, or by a private medical insurance provided by the employer

Lnrearn – ln of real earnings in 2002 Mx Pesos

Independent Variables

Age – Years of age of the worker

Education Level (dummy variables)

       Elementary – 6 yrs. or less (omitted category)

       Intermediate – 6 to 12 yrs of education

       Higher – 13 yrs. or more

Region (dummy variables)

       DF – Mexico City

       North – Northern states : Baja California, Baja California Sur, Coahuila de Zaragoza, Chihuahua, Durango, Nayarit, Nuevo León, San Luis Potosí, Sinaloa, Sonora, Tamaulipas, Zacatecas

       Center South – States in the Center and South región of the country: Aguascalientes, Campeche, Colima, Chiapas, Guanajuato, Guerrero, Hidalgo, Jalisco, México (without Mexico City suburbs), Michoacán de Ocampo, Morelos, Oaxaca, Puebla, Querétaro Arteaga, Quintana Roo, Tabasco, Tlaxcala, Veracruz de Ignacio de la Llave, Yucatán

Dependency Ratio –

       Household dependency ratio, i.e.

$$\frac{Younger\ 15\ yrs + Older\ 65\ yrs + Disabled}{Individuals\ in\ Working\ Age}$$

       No Dependents – dummy variable if the previous ratio is zero

HHold Female Fraction – Fraction of women in the household

HHold w/out Females – dummy variable if previous fraction is zero

Formality Rate 2001 – Fraction of workers employed in the formal sector (see definition supra) at a local level. For metropolitan areas self-represented in the survey, this fraction is estimated at the city level. For urban areas not self-represented in the survey and for rural areas this is approximated by the formality rate at the state level excluding the self-represented cities.

Urban Area – Cities with more than 2,500 inhabitants

Presence of Seguro Popular

Period 2002-2004: dummy variable taking value 1 if there is coverage by Seguro Popular at the state level

Period 2005-2009 – Coverage rate of Seguro Popular at the state level, measured as the No. de Beneficiaries as a fraction of the total population of the state.

Table A1 – No. of Observations in the sample by year

Year	No. Obs.
2002	83,819
2003	80,323
2004	76,483
2005	72,856
2006	73,723
2007	72,778
2008	70,720
2009	65,569