

**IS FORMAL EMPLOYMENT SECTOR  
HEREDITARY? DETERMINANTS OF  
FORMAL/INFORMAL SECTOR CHOICE  
FOR MALE MEXICAN WORKERS**  
**¿ES HEREDITARIO EL EMPLEO EN EL SECTOR  
FORMAL? DETERMINANTES DE LA ELECCIÓN  
DEL SECTOR FORMAL/INFORMAL PARA  
LOS TRABAJADORES MEXICANOS HOMBRES**

**Adelaido García-Andrés**

**Ernesto Aguayo-Téllez**

*Universidad Autónoma de Nuevo León*

**Jose N. Martínez**

*California State University, Dominguez Hills*

*Resumen:* Entender la relación entre el trabajo formal e informal de padres e hijos es esencial para promover la movilidad social en México. Con base en la encuesta de movilidad social 2011 (EMOVI), este artículo contribuye a estudiar la movilidad intergeneracional del empleo. Se encuentra una fuerte conexión entre las decisiones intergeneracionales de empleo y una selección positiva de los trabajadores. Individuos cuyos padres trabajaron en el sector formal tienen mayores probabilidades de trabajar en el sector formal y viceversa. Además, después de controlar por el sector de empleo, la escolaridad permanece como un vehículo importante para acceder al sector formal.

*Abstract:* Understanding the relationship between parents' and sons' formal employment is essential for promoting social mobility in Mexico. Using the 2011 Survey of Social Mobility in Mexico (EMOVI), this paper contributes to the literature by addressing the intergenerational mobility of employment. Findings show a strong connection between intergenerational employment choices and suggest a positive selection for workers. Individuals with parents who worked in the formal sector are more likely to be enrolled in formal work and vice versa. Also, after controlling for parent's employment sector, schooling remains as a significant vehicle to transit to the formal sector.

*Clasificación JEL/JEL Classification:* J62, J24, J31, J46

*Palabras clave/keywords:* job; occupational, and intergenerational mobility; occupational choice; wage differentials; informal labor markets, movilidad laboral, ocupacional e intergeneracional; elección ocupacional; diferenciales salariales; mercados laborales informales

*Fecha de recepción:* 30 VII 2017

*Fecha de aceptación:* 24 IV 2018

*Estudios Económicos, vol. 34 núm. 1, enero-junio 2019, páginas 91-121*

## 1. Introduction

The intergenerational mobility (IGM) model is concerned with the relationship between the socio-economic status of parents and the socio-economic outcomes of their children as adults (Blanden, 2013). According to Torche (2013), this socioeconomic standing is captured by different measures, but the most common ones are individual earnings, family income, social class, and employment status. In particular, employment status provides a good proxy for long-term economic standing because it remains relatively stable over the individual's employment career, so that a single measure can yield adequate long-run information (Fox, Torche and Waldfogel, 2016).

Recent evidence shows that family background has a significant impact on the sector of employment of their children. In particular, evidence shows that children of self-employed or informal workers are significantly more likely to be self-employed or work in the informal sector (Hout and Rosen, 2000; Colombier and Masclat, 2008; Pasquier-Doumer, 2012). Some researchers argue that informal employment is the result of an optimal choice. As such, informal workers choose informality because they expect higher welfare there than in the formal sector (Maloney and Ribeiro, 1999; Maloney, 2004; Packard, 2007).

Understanding the relationship between parents and sons' work choices is essential for assessing the opportunities for social mobility. In fact, the intergenerational transmission of self-employed or informal sector status is frequently connected with higher expected earnings (Fairlie and Robb, 2006; Colombier and Masclat, 2008). Although there is no consensus on the voluntary nature of working in the informal sector, evidence points to a strong intergenerational transmission of employment status (Pasquier-Doumer, 2012).

In Mexico, the way that social mobility has been measured depends on the specific aspects of social mobility being investigated, as well as on the available data. Until now, research regarding intergenerational transmissions of formal/informal employment sector has focused only on sons and parents living in the same household (Valero and Tijerina, 2003; Castillo and Vela, 2013), mainly because there is no adequate data from longitudinal surveys.

This paper uses a retrospective data collection from the *ESRU survey of social mobility in Mexico*, EMOVI 2011, CEEY (2011), which allows us to explore how a worker's formal/informal employment sector is affected by the employment sector of her/his parents. Our study differs from previous studies in Mexico in two ways: (1) the micro-econometric framework is derived from a structural model with

expected wages explicitly determining employment sector decisions, (2) selectivity bias is controlled for using Lee's (1982) procedure.

The outline of the paper is as follows. Section 2 briefly examines some of the associated literature. Section 3 introduces the model and discusses its identification. Section 4 highlights a few features from the data and presents results. Section 5 concludes.

## 2. Literature review

A large number of economists suggest that labor markets are segmented. Lewis (1954) and Harris and Todaro (1970) propose that developing countries are characterized by a dual labor market, consisting of a modern formal sector and a large, less-efficient, informal sector. These two sectors seem to be operating different labor markets. One offers better paid and more attractive jobs, while the secondary, informal sector, is characterized by rather low pay, unqualified workers, and short term jobs (Eichhorst and Kendzia, 2014). Secondary sector jobs are characterized by low (or even zero) marginal productivity which justifies in part the low wages paid. Hence, informal employment appears as a constrained choice (Rauch, 1991; Magnac, 1991; Pasquier-Doumer, 2012; Mboutchouang, Kenneck and Mbenga, 2013).

However, recent literature for some countries have proposed a different perspective: informal sector is the result of an optimal choice where individuals expect a higher welfare there than they would obtain in the formal sector (Maloney and Ribeiro, 1999; Packard, 2007). This literature views the informal sector as an active and voluntary entrepreneurial small-firm sector, where individuals choose to work because they expect a greater welfare than if they were wage-earners or formal-sector entrepreneurs (Maloney, 2004). In this sense, a proportion of informal employment may reflect an efficient allocation of labor (PasquierDoumer, 2012).

In addition, the possibility that both types of informal employment exist simultaneously has also been proposed (Fields, 1990; Perry *et al.*, 2007). Alcaraz, Chiquiar and Salcedo (2015) estimate that 10 to 20 percent of informal workers in Mexico are voluntarily self-selected into the informal sector.

Although there is no consensus on the -at least partially- voluntary nature of informal employment, some facts lead to certain reflections on individuals self-selecting to work in the informal sector and on the intergenerational transmission of informal entrepreneurship status (De Paul, Massil, and Modeste, 2013). In this sense,

some authors have suggested that segmented labor markets must be characterized by low mobility between formal and informal jobs must be observed; while self-selected informal employment markets should have the opposite characteristic (Fiess, Fugazza and Maloney, 2010).

In addition, in the economic literature, the intergenerational transmission of formal or informal employment is sometimes connected with higher expected earnings, which is consistent with the non-segmented labor market view (Colombier and Masclet, 2008; Fairlie and Robb, 2006). In Mexico, nearly 60 percent of workers have informal jobs (ILO, 2014), operating small establishments with low capital labor ratios (Leal-Ordoñez, 2014). Moreno (2007) estimates the average conditional difference between both sectors using Mincer equations. After controlling for selectivity among workers, he finds that workers with higher levels of education earn more in the formal sector; and workers with high school or less receive higher wages in the informal sector. These results are similar for women with basic and secondary schooling.

With regard to the choice of becoming employed in the formal or in the informal sector in Mexico, Duval and Smith (2011) modify Dickens and Lang (1985) methodology to allow for voluntary and involuntary informal workers. The authors find that free social protection programs (such as Mexico's "Seguro Popular") reduce the incentives for employment in formal sector jobs. In addition, Huesca and Padilla (2012) use a counterfactual technique to estimate wages if workers were employed in both sectors. The authors estimate that workers with higher schooling are more likely to work in the formal sector and women with lower education levels have a higher probability of working in the informal sector. However, he found that these probabilities change with age.<sup>1</sup>

Few studies have addressed individual wages from an intergenerational approach in Mexico. In an early paper, Valero and Tijerina (2003) estimate Mincer equations for workers controlling for their parents' characteristics, including wages, schooling, and employee status. The authors find that sons of employers and self-employed parents have higher wages than sons of employees. The authors suggest that this is caused by the transmission of skills, training, and entrepreneurial abilities.

---

<sup>1</sup> Evidence of wage differentials is not necessarily a proof that the labor market is segmented. Relatively higher wages in the formal sector may reflect other non-observed benefits for working in the informal sector, such as labor flexibility, which could compensate informal workers for such differentials (Maloney, 2004).

As evidence for the intergenerational transmission of sectors of employment in Mexico, Castillo and Vela (2013) use a probit model to estimate the probability that sons choose the same sector than their parents. Their results suggest that the social-domestic context has an influence in their labor decisions, suggesting that self-employed parents transmit informal human capital to their sons during childhood.

However, the main limitation of these last two studies is that they only obtain information for sons and parents living in the same household at the moment of the interview. This fact not only reduces the sample size considerably (by almost 90%) but also may introduce important biases given that sons may choose to stay at parent's home for non-random reasons.

This research overcomes this problem using EMOVI 2011. Unlike previous surveys in Mexico, EMOVI 2011 provides retrospective data that allows us to connect current respondents' information and their comparable retrospective data from parent's and family conditions when the interviewee was 14 years old. This allows us to recover data for both cohabiting and non-cohabiting children.

### 3. Empirical strategy

#### 3.1. Identification

A key issue for examining the determinants of labor payments is the preferences of workers. In a narrow sense, we might be interested in asking why a worker chooses to work in a particular sector. The model used here is based on a binary representation of the sector of employment decision. It assumes that there are two employment options available to each individual: formal and informal sectors. In addition, two important features should be considered. First, the informal sector regime is associated with more flexibility and independence than the formal sector. And second, each sector has different working conditions and different market institutions.

Similar to Rees and Shan (1986), we assume that the sector of employment and wages are determined simultaneously. This requires a 3SLS estimation procedure. In the first stage, we estimate a reduced-form probit model of sector of employment decision. This is used to construct a sample selection correction term. In the second stage, we estimate an OLS standard Mincer equation to estimate the earnings function. This is used to compare the differences in wages between both sectors, and to correct the bias in sample selection. And finally,

in the third stage, the differences in wages are used to estimate a probit model of the decision to be employed in the formal or informal sector.<sup>2</sup>

Assuming that workers choose between working in the formal or informal sector, we specify a linear utility function,  $U_k$  that represents the utility derived by individual  $i$  working in either sector: Individual  $i$  decides to work in the formal sector if  $I_i^* = (U_{formal} - U_{informal}) > 0$ . We cannot observe differences in each individual's utility level but we can observe whether the individual works in the formal or in the informal sector, where  $I_i = 1$  (working in the formal sector) if  $I_i^* = (U_{formal} - U_{informal}) > 0$  and zero otherwise.

Equation (1) below defines the probability that individual  $i$  works in the formal sector as a function of his own earnings differences ( $\log Y_{formal} - \log Y_{informal}$ ) and other worker characteristics ( $X_i$ ) as follows:

$$I_i = Pr(U_{formal} - U_{informal} > 0) = \Pr(\alpha + \beta(\log Y_{formal} - \log Y_{informal}) + \gamma X_i + \varepsilon_i > 0) \quad (1)$$

Equation (1) can be estimated as a probit model. However, a worker's earnings are only observed for the sector in which the worker is actually working. In order to run equation (1) we need an estimate of the potential earnings in the other sector for each worker. Since worker's earnings in a particular sector may depend on the potentially self-selected characteristics of the workers, the Lee (1982) two-stage procedure must be used to construct such predicted earnings (Koumenta, 2011).

To address this issue, we estimate two Mincer's (1974) semi-logarithmic wage equations, one for the formal workers ( $\log \hat{Y}_{formal}$ ) and other for the informal workers ( $\log \hat{Y}_{informal}$ ). As independent variables, we use worker characteristics ( $X_i$ ) and their parent's characteristics ( $Z_i$ ).

Mincer's semi-logarithmic wage equations are defined as follows:

$$\ln(Y_{formal}) = \alpha_{formal} + \phi_{formal} X_i + \delta_{formal} Z_i + e_{formal} \quad (2)$$

---

<sup>2</sup> The self-selection problem has been analyzed in different context. For example, Lee (1978) investigates the joint determination of the extent of unionism and the effect of unions on wage rates. Adamchik and Bedi (1983) examine differences in the public and private sector.

$$\ln(Y_{informal}) = \alpha_{informal} + \phi_{informal}X_i + \delta_{informal}Z_i + e_{informal} \quad (3)$$

The model is identified by the exclusion in equations (2) and (3) of at least one element of  $X_i$  in equation (1). Equation (1) can then be estimated using the standard maximum likelihood procedure. Estimating income equations (2) and (3) by Ordinary Least Square (OLS) might be inappropriate because it fails to reflect the possibility of self-selection in the decision to choose a sector of employment: workers might have some unobserved characteristics that affect their income generating capabilities in each sector.

To deal with possible self-selection bias, Lee's (1982) methodology recommends substituting income equations (2) and (3) into (1) and obtaining the reduced form of the sector of employment decision equation (1).

$$I_i = \beta_0 + \beta_1X_i + \beta_2Z_i + \varepsilon_i^* \quad (4)$$

Assuming that the error term ( $\varepsilon_i^*$ ) is normally distributed with unit variance, equation (4) can be estimated by a maximum likelihood probit and the fitted values ( $\hat{\psi}_i$ ) can then be used to calculate the selectivity correction variables or inverse Mill's ratios as follows:  $H1_i = \left(-f(\hat{\psi}_i)/F(\hat{\psi}_i)\right)$  and  $H0_i = \left(f(\hat{\psi}_i)/(1 - F(\hat{\psi}_i))\right)$ , where  $F(\hat{\psi}_i)$  is the standard cumulative distribution function and  $f(\hat{\psi}_i)$  is the standard normal density function.<sup>3</sup>

The selectivity correction variables,  $H1_i$  and  $H0_i$ , measure the truncation effect associated with sample selectivity (see, *e.g.* Lee, 1982) and are included in the two income equations (2) and (3) to control for self-selection as follows:

$$\ln(Y_{formal}) = \alpha_{formal} + \phi_{formal}X_i + \delta_{formal}Z_i + \sigma_{formal}^*H1 + \eta_{formal} \quad (5)$$

$$\ln(Y_{informal}) = \alpha_{informal} + \phi_{informal}X_i + \delta_{informal}Z_i + \sigma_{informal}^*H0 + \eta_{informal} \quad (6)$$

---

<sup>3</sup> If a univariate probit estimation is provided, the variables  $IMR1(H1)$  and  $IMR0(H0)$  are the Inverse Mill's Ratios to correct for a sample selection bias of  $y = 1$  and  $y = 0$ , respectively.

Thus,  $(\sigma_{formal}^*)$  and  $(\sigma_{informal}^*)$  estimate the selectivity bias in the observations. For example, if  $(\hat{\sigma}_{formal}^* > 0)$  and  $(\hat{\sigma}_{informal}^* < 0)$  each group has a comparative advantage in its chosen employment sector. This is called positive self-selection and typically occurs when  $\hat{\sigma}_{informal}^*$  is very large relative to  $\hat{\sigma}_{formal}^*$ . Hence, individuals with higher skills enter professions with the greater variance in earnings (see, for example, Maddala, 1983).<sup>4</sup>

Finally, the OLS predicted values of earnings for individuals in both formal  $\ln(\hat{Y}_{formal})$  and informal  $\ln(\hat{Y}_{informal})$  sectors from equations (5) and (6) are substituted into the structural form of the employment sector decision, equation (1), to obtain consistent parameter estimates.<sup>5</sup>

$$I_i = \Pr(\alpha + \beta(\log \hat{y}_{formal} - \log \hat{y}_{informal}) + \gamma X_i + \varepsilon_i > 0) \quad (7)$$

## 4. Data

### 4.1. Informal employment definition

Measuring the size of the informal economy and the incidence of informal employment is a difficult task. Also, different definitions have been put forward to make the concept operational. The appropriate methodology for the statistical measurement of informal employment depends on the users' requirements, measurement objectives, and the organization of the information.

Several definitions of the dividing line between formal and informal employment exist. For statistical purposes, the International Labor Organization (ILO) in The Fifteen International Conference of Labor Statisticians Characterized (15th ICLS resolution) uses the following criteria: informal employment "...encompasses a person in employment who, by law or in practice, are not subject to national

---

<sup>4</sup> The selectivity bias can be positive or negative and four different cases are possible (see, for example, Maddala, 1983).

<sup>5</sup> The model presented above is referred to as an endogenous switching regression model. It is used to address issues of self-selection and the estimation of treatment effects when there is no random allocation of subjects to treatment and control groups as is generally the case with observational (as opposed to experimental) data.



labor legislation and income tax or entitled to social protection and employment benefits...” (ILO, 2013: 4).<sup>6</sup>

Following one of the above criteria, in this paper, informal employment is defined as employees without access to public or private health insurance. This criterion is especially useful in countries where the registration of workers entails the registration of the enterprises employing them with social security institutions (most notably, in Mexico, through its social security agencies, such as IMSS and ISSSTE).

Although the employment relationships of workers in informal employment are heterogeneous, they share a basic vulnerability, which is that they need to be self-supporting and to rely on informal arrangements with respect to social protection needs.<sup>7,8</sup>

#### 4.2. *Social mobility survey*

The individual information used in this study is taken from the 2011 ESRU Survey of Social Mobility in Mexico, EMOVI 2011, that covers both rural and urban areas. EMOVI provides information about individuals between 25 and 64 years old (both household heads and non-household heads) from 11 001 households over two generations. This survey tracks the socioeconomic variables of a given household; each household member is asked detailed questions about age, gender, marital status, educational level, labor market participation, working hours, employment status and other work-related variables, as well as household size and other family features. In addition, a retrospective survey is applied, with each adult member of the household

---

<sup>6</sup> For more details see, ILO (2013) Statistics of employment in the informal sector, Fifteenth International Conference of Labor Statisticians (15th ICLS resolution), January 1993.

<sup>7</sup> The ILO has reached an international consensus with respect to the concept of informal employment in two dimensions. First, the nature of the economic unit, in other words, whether the unit uses household resources and keeps basic accounts records. The second dimension refers to whether employment is subject to labor law or an institutional framework, regardless of whether the economic unit employing the workers is an unregistered enterprise or a formal enterprise.

<sup>8</sup> The definition of informal employment used here may not be equivalent in Mexico at the time of the survey, 2011, as it was when the interviewee was 14 years old. Worker in the informal sector in 2011, have access to a larger set of free social benefits compared to those available 20 or 30 years earlier. Hence, a worker in 2011 may not have the same exogenous reasons for working in the informal sector as his father. Results should be taken with caution.

asked detailed information about characteristics of their parents and family conditions when the interviewee was 14 years old. Such information includes parents' labor market participation and parents' working conditions.<sup>9,10</sup>

#### 4.3. *Sample and descriptive statistics*

Since our goal is to study the interaction between sons and parents' sector of employment, we define our estimation sample according to the criteria that emphasizes coverage by the social security system. In the EMOVI 2011 survey, informal work regime is identified by the question: *As part of this job do you receive health care benefits?* This question is applicable for sons and their parents.

As in many labor research projects, we exclude female workers to prevent the results from being affected by sample selection bias, due to women's low rates of labor participation in Mexico (Caamal, 2007; Campos and Velez, 2014). This restriction is justified by the aim of forming a relatively homogeneous sample of employment occupations and wages.

We restrict the sample to full-time male workers (defined for our purpose as those who only have one job and work 30 or more hours per week) who provide information on their earnings and sector of employment. Part-time workers usually have more unstable work behavior, affecting the precision and significance of the estimators.<sup>11</sup> We keep only those who report being wage or salaried workers and

---

<sup>9</sup> This survey is designed by the *Espinosa Rugarcia Foundation* and the *Center for Studies Espinosa Yglesias* (CEEY) and was conducted in 2006 and 2011.

<sup>10</sup> A weakness of retrospective surveys such as EMOVI 2011 is that respondents may not remember or may not know their parents' labor conditions. In addition, the quality of the answers may decline with the age of the respondent. The more time passes, the harder it is to remember. An option would be to use a longitudinal survey. Unfortunately, the only longitudinal survey in Mexico, the MXFLS, ran only from 2002 to 2012, not enough time to follow an individual time span. <http://www.ennvih-mxfls.org>.

<sup>11</sup> This more unstable behavior is related to short-run labor decisions such as holidays or seasonal jobs or student jobs. Hence, the choice of part-time *vs* full-time jobs may be more related to job characteristics such as industry and occupation than to individual characteristics (Barrett and Doiron, 2001). If present in the Mexican labor market, these occupation or industry conditions may generate some bias in the estimators.

exclude all others, because salaried workers are those to whom health care obligations apply. In the case of earnings, we exclude observations with values smaller than the 1st percentile or larger than the 99th percentile. This cutoff point is of course arbitrary, but it is frequently used in related studies.<sup>12</sup>

The final sample consists of 2 851 full-time male workers between 25 and 64 years old. Table A1 in the appendix contains a summary of the statistics of the sample, including maximum, minimum, mean and standard deviation of each variable.

As a starting point, table 1 reports the intergenerational mobility regime. Each row of the table shows the sector of employment of the fathers while columns indicate the sector of employment of the sons. It shows that 62 percent of respondents work in the same sector as their parents: 39 percent of formal-sector workers reported that their parents also worked in the formal sector, and 75 percent of informal-sector workers report that their parents were also employed in the informal sector. Clearly, the percentage of sons that stay in the same sector as their parents is considerably larger in the informal sector.

**Table 1**  
*Intergenerational mobility between sectors of employment*

<i>Father's employment sector</i>	<i>Son's employment sector</i>		<i>Total</i>
	<i>Formal worker</i>	<i>Informal worker</i>	
Formal worker	403	628	1 031
(%)	39	61	100
Informal worker	452	1 368	1 820
(%)	25	75	100
Total	855	1 996	2 851
(%)	30	70	100

Source: Author's calculation based on 2011 EMOVI.

<sup>12</sup> Sample definition excludes: 1) unemployed, unpaid workers, pensioners, students, as well as people with a disability (4 037); 2) missing values in wage variable (2 544); 3) female population (1 420); 5) cutoff  $\leq$  1st and  $\geq$  99th (68); missing values in working hours (81); final sample consists of (N = 2 851 workers).

Table 2 provides descriptive statistics by employment status. Although there are significant differences between individuals working in the formal and informal sector in all respects, the informality wage penalty cannot be calculated simply looking at differences in the logarithm of the hourly wage. Regression analysis is required to find the *ceteris paribus* effect of the sector of employment upon earnings.

Our data shows that the standard deviation in the logarithm of the informal-sector hourly wage exceeds the standard deviation of the formal sector wage, and it seems plausible that informal sector incomes may in fact be more vulnerable to shocks than formal sector incomes. In addition, significant differences are observed in age, potential labor experience, and schooling levels. Also, we observed that formal workers were more likely to have parents with more years of schooling.<sup>13</sup>

**Table 2**  
*Descriptive statistics by sector of employment*

Variable	Formal workers (Obs = 855)		Informal workers (Obs = 1 996)		Mean difference
	Mean	SD	Mean	SD	
Log hourly wage	3.244	0.655	2.876	0.749	0.368***
Age	35.724	9.852	37.511	11.667	-1.787***
Labor experience	18.957	11.156	23.177	13.572	-4.220***
Less than primary	0.053	0.223	0.161	0.368	-0.109***
Primary completed	0.143	0.35	0.24	0.427	-0.097***
Secondary completed	0.636	0.481	0.556	0.497	0.080***
University completed	0.168	0.374	0.043	0.202	0.126***
Father's years of schooling	5.172	4.428	3.532	3.905	1.640***
Mother's years of schooling	4.988	4.074	3.472	3.762	1.516***

Note: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Author's calculation based on 2011 EMOVI.

<sup>13</sup> The Mexican education system is characterized by 6 years of primary education, followed by 6 years of secondary education. Secondary education is divided in 3 years of lower secondary education (*secundaria*) and 3 years of upper secondary education (*preparatoria*). Hence, we grouped education into the following categories: Less than primary (less than 6 years of schooling), Primary completed (6 to 11 years of schooling), Secondary completed (12 to 15 years of schooling), College (16 years or more completed).

## 5. Results

### 5.1. *Correcting for the self-selection bias*

Following the empirical strategy defined in section 3, the first step in correcting for self-selection bias is to estimate the probability of working in the formal sector as in equation (4), including as independent variables: marital status, labor experience, labor experience squared, parents' employment sector, community size, and dummies for region. The estimated coefficients of these variables are reported in table A2 in the appendix. With the fitted values of equation (4) at hand, we compute both inverse Mill's ratios  $H1_i = \left[ -f(\hat{\psi}_i)/F(\hat{\psi}_i) \right]$  and  $H0_i = \left[ f(\hat{\psi}_i)/1 - F(\hat{\psi}_i) \right]$  for the whole sample. Then we estimate the corrected wage estimates for both formal and informal workers as in equations (5) and (6). This involves regressing the natural logarithm of the hourly wage upon levels of work experience, marital status, having children and other demographic characteristics and parents' characteristics such as schooling and formal or informal employment sector employment.

The formal workers wage equation (5) and informal workers wage equation (6) also include, as regressors, the selection coefficient or inverse Mills ratio,  $H1$  and  $H0$ , respectively. The exercise is performed for the entire sample and for two groups of workers according to their level of schooling: middle school or less and high school or more. Estimated coefficients are displayed in table 3. The coefficients in the wage equations report the estimates for selectivity correction terms  $H1$  and  $H0$ . In both cases they are statistically significant.<sup>14</sup>

It can be shown that there is a positive selection into formal and into informal sectors. That is, those who choose the formal sector have unobserved abilities that make them relatively more productive (able to earn relatively higher wages) in the formal sector than in the informal sector; while, those who choose the informal sector have unobserved abilities that make them relatively more productive (able to earn relatively higher wages) in the informal sector than in the formal sector. The finding of positive selection bias for both kinds of

---

<sup>14</sup> The reduced-form equation contains the selectivity coefficient which is estimated rather than observed. It is well known that although OLS produces consistent estimates of the parameters, this is not true for the standard error. Then, we use the correction to the variance covariance matrix following Lee (1978) procedure.

workers is consistent with the hypothesis that those who have chosen the employee status possess comparative advantages in it.<sup>15</sup>

Dividing the sample, positive selection remains for informal sector workers, while for formal sector workers positive selection remains only for the highly-skilled jobs. Thus, the wages of formal sector workers with low schooling are not significantly different from what their wages would have been had they chose to work in the informal sector.

Table 4 reports the corrected predicted average wages for workers working in the formal and the informal sector. Working in the formal sector, formal workers would receive in average higher wages than similar informal sector workers. This gap is statistically significant except for the *less than primary* education. On the other hand, informal workers working in the informal sector would receive in average higher wages than formal workers with the same observed and unobserved characteristics working in the informal sector. This gap is statistically significant at all levels of education. Also notice that the difference is larger than when working in the formal sector.

**Table 3**  
*Coefficients of selectivity variables in wage regressions*

	<i>Equation (5)</i>			<i>Equation (6)</i>		
	<i>Formal wage equation</i>			<i>Formal wage equation</i>		
	<i>All sample</i>	<i>Less-skilled workers</i>	<i>More-skilled workers</i>	<i>All sample</i>	<i>Less-skilled workers</i>	<i>More-skilled workers</i>
Married	-.088 (.050)	-.021 (.066)	-.086 (.054)	-.041 (.038)	-.031 (.046)	-.079 (.049)
Age	.012*** (.003)	.003 (.004)	.013*** (.003)	.005** (.002)	.005* (.002)	.010*** (.003)

<sup>15</sup> A tighter interpretation is as follows. Consider the subsample of formal and informal workers with the same measured characteristics. Then, the earnings distribution actually observed for formal sector workers is higher than the distribution that would be observed for the average individual in the subsample had he chosen to work in the informal sector (see, for example, Rees and Shan, 1986). In other words, the average earnings of individuals with given measured characteristics who have chosen to be in the formal sector is greater than what formal earnings would be for those with the same measured characteristics who chose the informal sector. See graph A1 in the appendix.

**Table 3**  
(continued)

	<i>Equation (5)</i>			<i>Equation (6)</i>		
	<i>Formal wage equation</i>			<i>Formal wage equation</i>		
	<i>All sample</i>	<i>Less-skilled workers</i>	<i>More-skilled workers</i>	<i>All sample</i>	<i>Less-skilled workers</i>	<i>More-skilled workers</i>
H1	.563*** (.102)	.144 (.146)	.491*** (.116)			
HO				-1.092*** (.120)	-1.029*** (.172)	-1.017*** (.143)
Const.	3.628*** (.148)	3.276*** (.196)	3.612*** (.173)	2.350*** (.119)	2.303*** (.152)	2.297*** (.147)
Obs.	852	403	686	1 985	1 416	1 189
$R^2$	.107	.048	.118	.056	.035	.057

Notes: Controls: Dummies for region: North, Center North, Center, Capital, Gulf, South and Pacific. Dummies for occupation: Employer, self-employed, government employee, private sector employee, public company employee, domestic service. Less-skilled jobs: Workers with 9 or less years of schooling completed (middle school or less). More-skilled jobs: Workers with 10 or more years of schooling completed (high school or more). Standard error in parentheses, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Author's calculation based on 2011 EMOVI.

It can also be observed that age has a nonlinear effect on wages for informal workers, increasing relatively slowly at first, and then decreasing after about age 35. This effect is opposite for workers in the formal sector where a gradually increasing effect is observed.

After controlling for self-selection, on average, if formal-sector workers were artificially allocated to the informal sector, they would have higher wages than similar informal-sector workers in the same informal sector. Intuitively, that is consistent with the idea that formal workers enjoy a significant comparative earnings advantage over informal workers in both sectors, given a particular set of other observed characteristics.

However, this difference is smaller in the informal sector, corroborating the existence of positive self-selection of both formal and informal workers.

**Table 4**  
*Predicted wage corrected by employment status*

<i>Characteristics</i>	<i>Working in the formal sector</i>			<i>Working in the informal sector</i>		
	<i>Formal-sector workers</i>	<i>Informal-sector workers</i>	<i>Mean diff.</i>	<i>Formal-sector workers</i>	<i>Informal-sector workers</i>	<i>Mean diff.</i>
Less than primary	3.085	3.079	0.006	2.779	2.758	0.021
Primary complete	3.195	3.117	0.077***	2.886	2.832	0.054***
Secondary complete	3.230	3.147	0.083***	2.965	2.917	0.048***
University complete	3.390	3.264	0.126***	3.105	3.040	0.064**
Age = 25 to 34 years	3.141	3.068	0.072***	2.944	2.882	0.063***
35 to 44 years	3.196	3.099	0.097***	2.983	2.898	0.085***
45 to 54 years	3.219	3.119	0.100***	2.979	2.897	0.082***
55 to 64 years	3.393	3.170	0.223***	2.888	2.773	0.115***

Note: \* p< 0.1, \*\* p<0.05, \*\*\* p<0.01. Source: Author's calculation based on 2011 EMOVI.



### 5.2. *The structural probit equation*

Finally, using the corrected estimated coefficients of the two wage equations, we compute the differences in earnings between formal- and informal-sector employment for each worker in the sample  $(\log \hat{Y}_{formal} - \log \hat{Y}_{informal})$ . This term is included in the structural probit equation (1) to obtain consistent parameters.<sup>16</sup> Equation (1) is estimated for the entire sample and for two groups of workers according to their schooling level. Results by schooling level are presented in the tables A3, and A4 in the appendix.

Table 5 shows, for the entire sample, the marginal effects at mean of the probability of working in the formal sector. Bootstrapped robust standard errors are reported in parentheses. The coefficient of the wage difference is positive and strongly significant. A larger wage difference between the formal and the informal sector increases the probability of working in the formal sector, after controlling for other individual and family background characteristics.

In other words, a one percent increases in the ratio of wages (i.e. the wages that a worker would get in the formal sector, relative to the wage that the same worker would get in the informal sector) measured by  $\log(\frac{\hat{Y}_{formal}}{\hat{Y}_{informal}})$  increases the probability of working in the formal sector. More specifically, the probability of formal workers depends positively on the earning difference term and therefore provides support for rational occupational choice.

The coefficient of the schooling variable is positive and significant. This is due to rewarding job opportunities being more readily available to the highly educated workers in the formal sector. The coefficient of marital status confirms that being married increases the probability of working in the formal sector. This result may be a consequence of the workers' need to provide social security to their families. The number of children at home has no effect.<sup>17</sup>

---

<sup>16</sup> Given that the third stage regression includes as a regressor the selection-corrected wage differential variable computed from the second stage, standard errors for the third stage regression are calculated using bootstrapping procedures.

<sup>17</sup> The variable age is included in the regressions to control for the moment in the workers life cycle. However, this variable is perfectly correlated with the year when the worker was 14 years old. This correlation implies that the variable age implicitly represents the economic conditions faced by the father when the worker was 14 years old. Hence, the coefficient of age must be taken with caution because it does not only report the effect of the age of the respondent; it also reports a

Regarding retrospective data conditions when the interviewee was 14 years old, as we expected, having a father who worked in the formal sector has a positive effect on the probability of choosing formal sector. This coefficient is statistically significant in all models. For the entire sample, having a father who worked in the formal sector increases the probability of working in the formal sector by almost 20 percentage points. In addition, we found that a higher level of education for a parent increased the probability of being employed in the formal sector.

**Table 5**  
*Probability of working in the formal sector for the whole sample*

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Earning dif- ference	.0691*** (7.227)	.110*** (9.602)	.0850*** (8.411)	.112*** (10.004)
Age (in years)	.0184** (2.627)	.0193** (2.703)	.0205** (2.862)	.0197** (2.598)
Age squared	-.000304*** (-3.637)	-.000338*** (-3.964)	-.000330*** (-3.849)	-.000343*** (-3.811)
Less than pri- mary completed	Reference	Reference	Reference	Reference
Primary com- pleted	.0468 (1.730)	.0223 (.739)	.0329 (1.181)	.0172 (.573)
Secondary com- pleted	.166*** (6.752)	.116*** (3.880)	.130*** (4.698)	.110*** (3.768)
University com- pleted	.513*** (12.976)	.386*** (8.274)	.443*** (9.752)	.380*** (8.192)
1 if married (d)	.0659* (2.520)	.0883*** (3.330)	.0735** (2.793)	.0866*** (3.316)
1 if has children (d)	-.0251 (-.813)	-.0316 (-.991)	-.0237 (-.761)	-.0300 (-.994)

fixed effect for the phase of the business cycle when the respondent was 14 years old.

**Table 5**  
(continued)

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Number of children	.00808 (.922)	.0125 (1.370)	.0117 (1.305)	.0137 (1.503)
1 if father in formal work (d)		.175*** (8.649)		.178*** (8.492)
Father's years of schooling		.0165*** (6.362)		.0120*** (3.626)
1 if father encouraged his children to study (d)		.0510** (2.691)		.0491* (2.464)
1 if father is indigenous (d)		-.0567* (-2.017)		-.00753 (-.170)
1 if mother in formal work (d)			.0911* (2.575)	.0767* (1.991)
Mother's years of schooling			.0152*** (5.801)	.00681 (1.944)
1 if mother encouraged her children to study (d)			.0186 (.579)	-.00387 (-.112)
1 if mother is indigenous (d)			-.0749** (-2.650)	-.0625 (-1.356)
Community size dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Observations	2 837	2 837	2 837	2 837
Log likelihood	-1524.3	-1458.5	-1499.4	-1452.4
McFadden's <i>R</i> squared	.121	.159	.135	.162
Count <i>R</i> squared (%)	73.70	74.62	74.55	74.97

Log hourly wage difference =  $\ln \hat{Y}_{formal} - \ln \hat{Y}_{informal}$ . Controls: Community size dummies: Less than 2 500; 2 500 to 14 999; 15 000 to 99 000; 100 000 or more. Region dummies: North, Center North, Center, Capital, Gulf, South and Pacific. Marginal effects at mean; Bootstrapped robust standard errors in parentheses; (d) discrete change of dummy variable from 0 to 1. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . Source: Author's calculation based on 2011 EMOVI.

## 6. Conclusions

This paper uses EMOVI 2011 retrospective database for Mexico. Unlike previous surveys in Mexico, this database provides information about family background characteristics in addition to individual and family characteristics. This database allows us to study the effects of a parents formal or informal sector of employment on the son's formal or informal sector of employment and, in some sense, to contribute to the understanding of intergenerational social mobility in Mexico.

To do this, we use a micro econometric framework based on Lee (1982) to control for the presence of self-selection in the sector of employment decisions. Our results highlight the following aspects. Results suggest positive selection for workers into formal and into informal sectors; significant coefficients show that not correcting for self-selection bias might skew results. This evidence is consistent with the hypothesis that those who have selected a sector of employment possess a comparative advantage for it. A larger wage difference between both sectors increases the probability that workers will choose to work in the formal sector, which provides support for the idea that occupational-sector choice is rational.

With respect to social mobility variables, we find that parents' schooling is an important vehicle of upward social mobility. This paper corroborates previous literature in Mexico that found that sons tend to follow the same employment sector as their parents, probably through the intergenerational transmission of formal- or informal-sector employment skills.

These findings support the view that investing in better schools and better educational quality, as well as in programs to increase workers human capital such as labor training, will have a positive effect on the acquisition of formal-sector employment skills, and with it, better chances to access formal-sector labor conditions; improving the labor conditions of Mexican workers and the quality of life of their families.

Future research could use life-long longitudinal surveys to fully control for differences in economic conditions facing parents and sons at the time they choose to work. Such differences should have an effect on the decision to work in the formal or in the informal sector.

### *Acknowledgments*

This research project was supported by the Graduate Research Fellowship Program sponsored by ESRU Foundation and the *Espinosa Yglesias Research Center (Centro de Estudios Espinosa Yglesias, CEEY)*. The views expressed here are those of the authors' and

should not be attributed to the CEEY. The authors acknowledge valuable comments from two anonymous referees. adelaido.garciaan@uanl.edu.mx, ernesto.aguayotl@uanl.edu.mx, jnmartinez@csudh.edu

## References

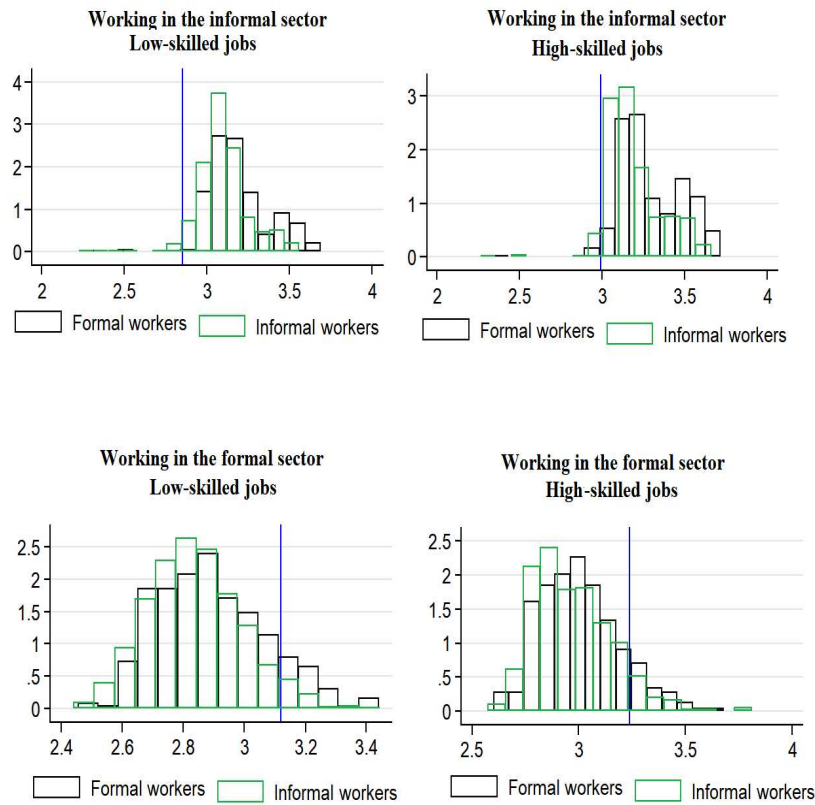
- Adamchik, V., and V. Bedi. 1983. Wage differentials between the public and the private sectors: Evidence from an economy in transition, *Labor Economics*, 7(1): 203-224.
- Alcaraz, C., D. Chiquiar, and A. Salcedo. 2015. Informality and segmentation in the Mexican labor market, Working Papers, no. 2015-25, Banco de México.
- Barret G., and D. Doiron. 2001. Working part time: By choice or by constraint, *The Canadian Journal of Economics*, 34(4): 1042-1065.
- Blanden, J. 2013. Cross-national ranking of intergenerational mobility: A comparison of approaches from economics and sociology, *Journal of Economic Surveys*, 27(1): 38-73.
- Caamal, C. 2007. Oferta laboral en México: un enfoque de variables instrumentales, *Ensayos Revista de Economía*, 0(1): 115-154.
- Campos, R. and R. Velez. 2014. Female labour supply and intergenerational preference formation: Evidence for Mexico, *Oxford Development Studies*, 42(4): 553-569.
- Castillo, D., and F. Vela. 2013. Labor mobility and intergenerational transmission of self-employment in Mexico, *Iberoamericana*, 1(6): 193-235.
- Centro de Estudios Espinoza Iglesias (CEEY). 2011. *Encuesta ESRU de movilidad social en México 2011* or ESRU Survey of Social Mobility in Mexico 2011, EMOVI 2011, Fundación ESRU, <https://ceey.org.mx>.
- Colombier, N., and D. Masclat. 2008. Intergenerational correlation in self-employment: some further evidence from French ECHP data, *Small Business Economics*, 30(1): 423-437.
- De Paul, V., J. Massil, and K. Modeste. 2013. Intergenerational transmission of self-employed status and informal production units performance in Cameroon, MPRA Paper, no. 50133.
- Dickens, W., and K. Lang. 1985. A test of dual labor market theory, *American Economic Review*, 75(4): 729-805.
- Duval, R., and R. Smith. 2011. Informality and *Seguro Popular* under segmented labor markets, CIDE, preliminary draft. [http://conference.iza.org/conference\\_files/worldb2011/3229.pdf](http://conference.iza.org/conference_files/worldb2011/3229.pdf)
- Eichhorst, W., and M. Kendzia. 2014. Workforce segmentation in Germany: From the founding era to the present time, IZA, Discussion Paper, no. 8648.
- Fairlie, R., and A. Robb. 2006. Families, human capital, and small business: Evidence from the characteristics of business owner's survey, *Industrial and Labor Relations Review*, 60(2): 225-245.
- Fields, G. 1990. Labour market modeling and the urban informal sector: Theory and evidence, in D. Turnham, and A. Schwarz (eds.), *The Informal Sector Revisited*, Paris, Development Centre of the OECD.

- Fiess, N., M. Fugazza, and W. Maloney. 2010. Informal self-employment and macroeconomic fluctuations, *Journal of Development Economics*, 91(2): 211-226.
- Fox, L., F. Torche, and J. Waldfogel. 2016. Intergenerational mobility, in D. Brady, and L. Burton (eds.), *The Oxford Handbook of the Social Science of Poverty*, Oxford University Press.
- Harris, J., and M. Todaro. 1970. Migration, unemployment and development: A two sector analysis, *American Economic Review*, 60(1): 126-142.
- Hout, M., and H. Rosen. 2000. Self-employment, family background, and race, *Journal of Human Resources*, 32(4): 670-692.
- Huesca, L., and A. Padilla. 2012. Empleo, escolaridad y sector informal en la frontera norte de México y Chihuahua: expectativas de ocupación en la crisis, *Ensayos Revista de Economía*, 31(2): 57-86.
- ILO. 2013. Measuring informality: A statistical manual on the informal sector and informal employment, Geneva, International Labor Organization, <[http://www.ilo.ch/global/publications/ilo-bookstore/order-online/books/WCMS\\_222979/lang-en/index.htm](http://www.ilo.ch/global/publications/ilo-bookstore/order-online/books/WCMS_222979/lang-en/index.htm)>.
- . 2014. Informal employment in Mexico: Current situation, policies and challenges, FORLAC Programme.
- Koumenta, M. 2011. Modernization, privatization, and the public services ethos in the United Kingdom, in D. Marsen (ed.), *Employment in the Lean Years. Policy and Prospects for the Next Decade*, Oxford University Press.
- Leal-Ordoñez, J. 2014. Tax collection, the informal sector, and productivity, *Review of Economic Dynamics*, 17(1): 262-286.
- Lee, L. 1978. Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables, *International Economic Review*, 19(1): 415-433.
- Lee, L. 1982. Some approaches to the correction of selectivity bias, *Review of Economic Studies*, 49(3): 355-372.
- Lewis, W. 1954. Economic development with unlimited supply of labor, *Manchester School of Economics and Social Studies*, 22(1): 139-191.
- Maddala, G. 1983. *Limited-dependent and Qualitative Variables in Economics*, New York, Cambridge University Press.
- Magnac, T. 1991. Segmented or competitive labor markets, *Econometrica*, 59(1): 165-187.
- Maloney, W. 2004. Informality revisited, *World Development*, 32(7): 1159-1178.
- , and E. Ribeiro. 1999. Efficiency wage and union effects in labor demand and wage structure in Mexico, World Bank, Working Paper, no. 2131.
- Mboutchouang, V., J. Kenneck, and K. Mbenga. 2013. Intergenerational transmission of self-employed status and informal production units performance in Cameroon, MPRA Munich Personal RePEc Archive.
- Mincer, J. 1974. *Schooling, Experience and Earnings*, New York, Columbia University Press.
- Moreno, J. 2007. Los salarios del sector formal e informal en México: análisis de ganancias y pérdidas por formalización, *Ensayos Revista de Economía*, 31(1): 57-86.
- Packard, T. 2007. Do workers in Chile choose informal employment? A dynamic analysis of sector choice, World Bank, Working Paper, no. 4232.

- Pasquier-Doumer, L. 2012. Intergenerational transmission of self-employed status in the informal sector: A constrained choice or better income prospects? Evidence from seven West African Countries, *Journal of African Economies*, 22(1): 73-111.
- Perry, G., W. Maloney, O. Arias, P. Fajnzylber, A. Mason, and J. Saavedra. 2007. *Informality: Exit and exclusion*, Latin American and Caribbean Studies, no. 40008, Washington, DC., World Bank.
- Rauch, J. 1991. Modeling the informal sector formality, *Journal of Development Economics*, 35(1): 33-47.
- Rees, H., and A. Shan. 1986. An empirical analysis of self-employment in the U.K., *Journal of Applied Econometrics*, 1(1): 95-108.
- Torche, F. 2013. How do we characteristically measure and analyze intergenerational mobility? Stanford Center on Poverty and Inequality (unpublished).
- Valero, J., and A. Tijerina. 2003. Transmisión intergeneracional del ingreso, empresariado y educación en México, *Ensayos Revista de Economía*, 22(2): 19-34.

Appendix

**Graph A1**  
*Predicted log wage distributions for  
 all sample by occupation status*



Note: The y-axis and x-axis measure, respectively, the density and linear prediction of log hourly wage. Source: Author's calculation based on 2011 EMOVI.



**Table A1**  
*Summary statistics*

<i>Variable</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>Std. dev.</i>	<i>N</i>
1 if formal worker(d)	0	1	0.30	0.46	2 851
Log hourly wage	0.85	6.44	2.99	0.74	2 851
Age (in years)	25	64	36.98	11.18	2 851
Age square	625	4096	1492.16	936.12	2 851
Labor experience	1	58	21.91	13.04	2 851
Labor experience squared	1	3364	650.03	733.96	2 851
Completed years of schooling	0	26	9.06	4.04	2 851
1 if married (d)	0	1	0.64	0.48	2 851
1 if has children (d)	0	1	0.61	0.49	2 851
Number of children	0	13	1.36	1.55	2 851
1 if father in formal work(d)	0	1	0.36	0.48	2 851
Father's years of schooling	0	23	4.02	4.14	2 851
1 if father is indigenous (d)	0	1	0.16	0.37	2 851
1 if father encouraged his children to study (d)	0	1	0.50	0.5	2 851
1 if mother in formal work(d)	0	1	0.12	0.32	2 851
Mother's years of schooling	0	22	3.93	3.92	2 851
1 if mother is indigenous (d)	0	1	0.15	0.36	2 851
1 if mother encouraged her children to study (d)	0	1	0.13	0.33	2 851

Note: (d) indicates dummy variable. Source: Author's calculation based on 2011 EMOVI.

**Table A2**  
*Probability of working in the formal sector (reduced form)*

<i>Variables</i>	<i>dy/dx</i>	<i>SE</i>
1 if married (d)	0.0899***	(4.758)
Labor experience	0.0011***	(0.391)
Labor experience squared	-0.0001*	(-2.030)
1 if father in formal work (d)	0.1280***	(7.197)
Father's years of schooling	0.0098**	(3.229)
1 if mother in formal work (d)	0.0446	(1.607)
Mother's years of schooling	0.0074*	(2.246)
Community size dummies	Yes	
Region dummies	Yes	
Observations	2,851	
Log likelihood	-1639.6	
McFadden's R squared	0.0584	
Count R squared	70.96	

Notes: controls: Community size dummies: Less than 2 500; 2 500 to 14 999; 15 000 to 99 000; 100 000 or more. Region dummies: North, Center North, Center, Capital, Gulf, South and Pacific. Marginal effects at means; standard errors in parentheses; (d) discrete change of dummy variable from 0 to 1. \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01. Source: Author's calculation based on 2011 EMOVI.

**Table A3**  
*Probability of working in the formal sector for more-skilled workers*

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Earning difference	0.0768*** (6.714)	0.129*** (9.377)	0.0934*** (8.138)	0.133*** (9.367)
Age (in years)	0.0245* (2.255)	0.0251* (2.181)	0.0257* (2.479)	0.0246* (2.241)
Age squared	-0.000381** (-2.816)	-0.000423** (-2.889)	-0.000396** (-3.051)	-0.000420** (-3.042)
Years of schooling	0.0445*** (9.811)	0.0334*** (7.118)	0.0394*** (8.475)	0.0331*** (7.203)
1 if married (d)	0.0906** (2.662)	0.123*** (3.506)	0.0986** (2.806)	0.121*** (3.558)
1 if has children (d)	-0.0809 (-1.886)	-0.0932* (-2.176)	-0.0785 (-1.868)	-0.0915* (-2.108)
Number of children	0.0278 (1.781)	0.0351* (2.329)	0.0314* (2.155)	0.0373* (2.437)
1 if father in formal work (d)		0.215*** (7.780)		0.218*** (7.662)
Father's years of schooling		0.0174*** (5.540)		0.0144*** (3.460)
1 if father encouraged his children to study (d)		0.0544* (1.781)		0.0500 (1.625)

**Table A3**  
(continued)

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
1 if father is indigenous (d)		(2.040) -0.0747 (-1.874)		(1.905) -0.0214 (-0.347)
1 if mother in formal work (d)			0.0953* (2.053)	0.0826 (1.680)
Mother's years of schooling			0.0137*** (3.975)	0.00485 (1.116)
1 if mother encouraged her children to study (d)			0.0394 (0.925)	0.0138 (0.305)
1 if mother is indigenous (d)			-0.0907* (-2.366)	-0.0711 (-1.157)
Community size dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Observations	1 875	1 875	1 875	1 875
Log likelihood	-1107.5	-1057.9	-1093.6	-1054.9
McFadden's <i>R</i> squared	0.101	0.141	0.112	0.143
Count <i>R</i> squared (%)	68.59	69.87	69.71	70.19

Notes: Log hourly wage difference =  $\ln \hat{Y}_{formal} - \ln \hat{Y}_{informal}$ , controls: Community size dummies: Less than 2 500; 2 500 to 14 999; 15 000 to 99 000; 100 000 or more. Region dummies: North, Center North, Center, Capital, Gulf, South and Pacific. Marginal

effects at means; bootstrapped robust standard errors in parentheses; (d) discrete change of dummy variable from 0 to 1. \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01. Source: Author's calculation based on 2011 EMOVI.

**Table A4**  
*Probability of working in the formal sector for less-skilled workers*

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Earning difference	0.0547*** (4.490)	0.0797*** (5.896)	0.0619*** (4.835)	0.0801*** (5.944)
Age (in years)	0.0141 (1.876)	0.0149* (2.017)	0.0163* (2.100)	0.0157* (2.029)
Age squared	-0.000234** (-2.665)	-0.000253** (-2.954)	-0.000256** (-2.860)	-0.000261** (-2.912)
Years of schooling	0.0158*** (3.533)	0.0103* (2.256)	0.0119** (2.622)	0.00972* (2.080)
1 if married (d)	0.0335 (1.064)	0.0410 (1.369)	0.0368 (1.184)	0.0403 (1.399)
1 if has children (d)	-0.0000824 (-0.002)	-0.00803 (-0.243)	-0.00374 (-0.107)	-0.00858 (-0.258)
Number of children	0.00447 (0.535)	0.00767 (0.938)	0.00654 (0.777)	0.00753 (0.879)

**Table A4**  
(continued)

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
1 if father in formal work (d)		0.142*** (6.766)		0.143*** (6.469)
Father's years of schooling		0.0125*** (3.883)		0.0108** (2.633)
1 if father encouraged his children to study (d)		0.0550** (2.654)		0.0550* (2.552)
1 if father is indigenous (d)		-0.0280 (-0.989)		0.0195 (0.430)
1 if mother in formal work (d)			0.0218 (0.579)	0.0121 (0.289)
Mother's years of schooling			0.0108** (3.268)	0.00257 (0.605)
1 if mother encouraged her children to study (d)			0.0283 (0.750)	0.0107 (0.266)
1 if mother is indigenous (d)			-0.0469 (-1.603)	-0.0588 (-1.299)
Community size dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Observations	1 819	1 819	1 819	1 819

**Table A4**  
(continued)

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Log likelihood	-867.4	-834.7	-859.5	-833.5
McFadden's <i>R</i> squared	0.0984	0.132	0.107	0.133
Count <i>R</i> squared (%)	78.61	79.44	78.56	79.49

Notes: Log hourly wage difference =  $\ln \hat{Y}_{formal} - \ln \hat{Y}_{informal}$ , controls: Community size dummies: Less than 2 500; 2 500 to 14 999; 15 000 to 99 000; 100 000 or more. Region dummies: North, Center North, Center, Capital, Gulf, South and Pacific. Marginal effects at means; bootstrapped robust standard errors in parentheses; (d) discrete change of dummy variable from 0 to 1. \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01. Source: Author's calculation based on 2011 EMOVI.