Choices and Constraints: The Nature of Informal Employment in Urban Mexico

ROBERT DUVAL-HERNÁNDEZ

Faculty of Economics and Management, Open University of Cyprus, Latsia, Cyprus, Institute of Labor Economics (IZA), Bonn, Germany, CIDE, Mexico City, Mexico

ABSTRACT Using a special module of the 2015 Mexican Labour Force Survey with information on workers’ preferences for jobs with social security coverage, it is found that 80 per cent of informal workers in large urban areas would prefer to work in a job that provides them with such coverage. The estimation of a discrete choice econometric model which distinguishes between wanting a formal job and the probability of getting one shows that schooling increases the chances of being hired in formal employment and of having higher earnings in it. Women with greater responsibilities at home are less likely to want formal employment, and they also face a lower probability of being hired in such jobs. The findings indicate the segmentation of Mexican labour markets and the rationing of formal jobs, together with the existence of workers who voluntarily participate in informal employment. However, the estimated fraction of involuntary informal workers is quite high.

KEYWORDS: Employment; wages; workers; Latin America

1. Introduction

A central question in the study of labour markets in the developing world is whether its large share of informal employment is mainly composed of workers who cannot find a better employment option in the formal economy, or whether this sector is formed by workers who voluntarily seek this type of employment given the incentives in the market.

Broadly speaking, traditional theories of dualistic labour markets conceive informal employment as a second-best option for workers who cannot find a formal job, and who cannot afford to remain unemployed while searching for a good employment option (see, for instance, the seminal model of Fields, 1975). In contrast, a different strand of the literature has argued that workers optimally choose informal employment because, given their individual characteristics, they obtain a higher utility/payoff in that type of work (see, for instance, Maloney, 1999). This
can arise when workers prefer to be informally employed because they will pay lower taxes, face less regulations, or have more flexible work arrangements than if they were formally employed.

Despite the above opposing views, the literature recognises that, in practice, informal employment is heterogeneous. It is composed of workers who participate in it voluntarily and of those who are there because they cannot find a formal job (Fields, 1990; Maloney, 2004). While there is agreement about the heterogeneous nature of informal employment, there is less agreement about how many informal workers are voluntarily in informal employment and how many simply cannot find a better job elsewhere. Furthermore, more needs to be known about the characteristics of these different types of workers, as this information is crucial for the proper design of policy based on an understanding of the functioning and structure of the labour market.1

Traditionally, researchers have devoted a substantial amount of energy to test the existence of rationing in labour markets through: (i) the estimation of complex structural econometric models of sector allocation (e.g. Günther & Launov, 2012; Magnac, 1991); (ii) the analysis of sector transitions over the business cycle (e.g. Bosch & Maloney, 2010; Fiess, Fugazza, & Maloney, 2010); and iii) the use of ancillary information about the characteristics of informal employment in order to identify workers who perform successfully in it and, thus, are more likely to participate in it voluntarily (e.g. Gindling & Newhouse, 2014).

The contribution of this paper is to use a special supplement of the 2015 Mexican Labour Force Survey (Instituto Nacional de Estadística y Geografía [INEGI], 2015a) which directly asks urban workers about their valuation of jobs with social security coverage. Since, in Mexico, having social security coverage in a job is the defining characteristic of formal wage employment, this piece of information gives us a proxy measure of workers’ preferences for formal jobs.

This variable is used to address three key questions: What fraction of informal workers would rather be formal employees? What individual characteristics increase the probability of applying to and of being hired in a formal job? And How earnings are related to individual characteristics, after accounting for the non-random selection into each type of employment?

To answer these questions the paper presents a combination of descriptive statistics, a discrete choice econometric model that distinguishes the worker’s decision to apply for a formal job from the formal sector employer’s hiring decision, as well as selectivity-corrected earnings equations for three different types of workers: formal, voluntary informal, and other informal workers.

Overall, the empirical findings confirm the view that informal employment is formed of a heterogeneous group of workers, some of whom participate voluntarily in it while others do so because of a lack of better options. However, contrary to previous findings in the literature, the fraction of involuntary informal workers is quite high. In particular, the analysis indicates that around 80 per cent of the respondents who lack social security coverage would prefer to have a job with such benefits, even if that entailed having to pay the corresponding contributions for them.

Broadly speaking, the findings of this paper highlight two main factors on the supply side of the labour market that limit the number of workers employed in formal jobs. The first is related to household demographics and the division of housework. In particular, women who have a higher burden of work at home are less likely to seek formal employment and are less likely to be hired in formal jobs. The second is related to human capital, as higher levels of schooling increase the chances of being hired in formal jobs and of obtaining higher earnings in them.

The paper is organised as follows. Section 2 presents the data on the valuation of jobs with social security benefits, including its main descriptive statistics. Section 3 discusses the econometric methodology for the ceteris paribus analysis of job allocation and earnings, while the results of these estimations are presented in Section 4. Section 5 concludes.
2. Voluntary versus involuntary informal employment

For the case of Mexico, a significant literature has studied segmentation in labour markets through the analysis of the patterns of sector transitions over the business cycle (see, for instance, Bosch & Maloney, 2010; Maloney, 1999, 2004) and it concludes that the informal sector is mainly composed of voluntary workers. While interesting on their own, sector transitions over the business cycle only provide indirect evidence about the preferred employment of workers changing jobs and, by construction, cannot inform us about the preferred jobs of stayers, since they don’t reveal if stayers remain informal by choice or by lack of better options.

Other studies have tried to estimate the proportion of involuntary informal workers in Mexico using structural econometric methods. Two examples are the papers by Duval-Hernández and Smith (2010) and Alcaraz, Chiquiar and Saucedo (2015). Both use discrete choice models that allow for rationing of formal jobs under a context of partial observability because they lack information on the preferred sector of workers. The predicted proportion of involuntary informal workers vary widely between these studies, and sometimes even within a given paper, depending on the sample analysed.

In general, the literature has done only a limited use of information on stated preferences for formal employment. For instance, Maloney (1999, 2004) analyses a sample of microentrepreneurs and presents statistics on their reasons for opening a business and for quitting their last job. Based on this evidence the author estimates the fraction voluntary informal workers to be around 60 per cent.

A related literature analyses job satisfaction by type of employment (see for instance Perry et al. 2007; Falco, Maloney, Rijkers, & Sarrias, 2015). While these studies often find greater job satisfaction in formal employment, there is substantial heterogeneity in job satisfaction among informal workers, and the results also vary substantially across countries.

Few studies in other Latin-American countries have analysed surveys where workers are explicitly asked whether they would prefer to have a formal job, similar to the survey question used in this paper. Using a special module of a household survey in Brazil in 1990, Soares (2004) estimated that around 70 per cent of informal wage workers and 30 per cent of the self-employed would rather be employed as formal wage workers. Also, using data for Chile in 2009, Contreras, Gillmore, and Puentes (2017) estimated that around 35 per cent of the self-employed would rather work as wage employees.

Probably the most promising use of information on preferences for formal employment is the one arising from discrete choice experiments that randomize hypothetical job attributes among respondents in order to elicit their willingness to pay for such attributes. The works of Assy et al. (2019) for youth workers in Kenya and Mahmud et al. (2021) for Bangladesh indicate a strong preference (i.e. a high willingness to pay) for jobs that have characteristics similar to those found in formal employment. The preferred characteristics include greater job stability, health insurance, and good working conditions. While the survey instrument analysed in this paper does not permit such a detailed analysis, it still provides interesting insights regarding the preferences of informal workers in Mexico.

2.1. Descriptive statistics

In Mexico, having social security coverage is the defining characteristic of formal wage employment (see for instance Levy, 2008 and INEGI, 2014). For this reason, the analysis in this paper equates ‘formal employment’ with having a job that is covered by social security benefits.

In the second quarter of 2015, the Mexican Labour Force Survey was supplemented by a module inquiring about the employment trajectories of workers and their contribution to and valuation of social security protection. The MOTRAL module (after its acronym in Spanish) was applied to a representative sample of workers aged 18 to 54, living in large urban centres who were either employed or had previous labour market experience. This target population
represented around 90 per cent of the labour force in large urban centres and 60 per cent of the overall urban labour force in 2015.\textsuperscript{7}

This module included the following key question: Do you think it is better to have a job with social security, even if you have to make payments to be eligible for it?\textsuperscript{8} This question is central to the study of informality because most labour surveys contain information on the sector of employment, but they do not collect information on the types of jobs workers value.

To the extent that social security coverage is the defining characteristic of formal wage employment in Mexico, the answer to this question can be used as a proxy for the workers’ preference for formal employment. Linking this variable to information on the actual type of job can help us to approximate the fraction of involuntary informal workers.\textsuperscript{9}

Table 1 presents the key sociodemographic characteristics of employed workers in the MOTRAL module depending on whether they have social security coverage in their job and on whether they would prefer to have a job with coverage.\textsuperscript{10} The table shows that the age and gender composition is more or less homogeneous across groups except for workers with coverage who do not value their social security benefits (in column 2), who are predominantly male and slightly younger. Workers with coverage are more educated than those without coverage, and within each market segment (with coverage or without coverage) respondents who do not value social security have higher levels of schooling than those who value it. In addition, workers without coverage are more likely to still be enrolled in school. Also, workers without coverage who do not value social security benefits are less likely to be married, have a higher number of dependents at home, and have the lowest earnings of all groups. In contrast, the group that exhibits higher average earnings is the workers with coverage who do not value social security coverage.

Based on the numbers reported at the bottom of Table 1, around 80 per cent of informal workers would rather have a covered (i.e. formal) job. This contrasts with the view put forward by a strand of the literature which considers Mexican informal employment to be mainly composed of workers who voluntarily avoid formal employment (e.g. Maloney, 1999, 2004).

To better understand the previous result, Table 2 provides information on the job separations over the five years prior to the survey interview for workers who are currently in informal employment. In particular, the table distinguishes the job separations of currently uncovered

---

Table 1. Characteristics of employed population

| Has social security | Does not have social security |
|--------------------|-----------------------------|
|                    | Wants social security | Does not want social security | Wants social security | Does not want social security |
| Years of age 37.1  | 35.4                  | 38.4                         | 37.1                  |
| Male (%) 48.2     | 63.0                  | 48.4                         | 49.7                  |
| Years of schooling 11.8 | 12.4          | 9.9                         | 10.7                  |
| Enrolled in school (%) 3.9 | 2.9           | 6.1                         | 6.3                  |
| Married (%) 45.1  | 42.5                  | 44.7                         | 32.7                  |
| Household composition |               |                             |                      |
| Dependency ratio 0.61 | 0.59           | 0.61                         | 0.72                  |
| # Adult females 1.06 | 1.55           | 1.06                         | 1.00                  |
| Earnings (monthly Mx Pesos) 7456 | 9167        | 4461                         | 3652                  |
| Share of employment 43.5 | 8.9            | 38.3                         | 9.3                  |
| # Obs. (unweighted) 1906 | 393           | 1759                         | 439                  |

Notes: Unless otherwise stated, all numbers are averages across the different employment groups. All estimates use sampling weights. 
Source: Author’s calculations based on MOTRAL 2015 (INEGI, 2015a) and ENOE 2nd Quarter 2015 (INEGI, 2015b).
workers according to sector of origin of the terminated job, and for those who separated from a formal (covered) job, it distinguishes between several reasons for the separation. This table shows that half of informal workers in the sample did not experience a job separation over the last five years, and almost forty per cent had separations from informal jobs. About seven per cent of informal workers had separations from formal jobs due to involuntary ‘market’ reasons (such as firings, firm closures, etc.), while ten per cent had separations due to voluntary reasons such as looking for greater income, bad job conditions, or other personal motives (pregnancies, sickness, etc.).11

The last column of the table shows the percentage that states preferring a formal job, for each separation category. It is interesting to note, that these reported preferences for formal employment are very high among the workers that experienced separations from formal jobs, irrespective if their separation was voluntary or not.12

To further compare these numbers with another source of information, the Online Appendix presents a similar exercise for a sample of microentrepreneurs from the 2012 Encuesta Nacional de Micro-Negocios (ENAMIN, or National Survey on Micro-enterprises) (INEGI, 2012). That exercise shows that almost sixty per cent of the informal microentrepreneurs express they would prefer having a formal wage job if it were to pay an income equivalent to the one they currently obtain.13

In both the MOTRAL and ENAMIN surveys around ten per cent of informal workers quit formal jobs voluntarily and yet many of them report preferring formal jobs. This indicates that a stated preference for formal jobs does not imply that these workers will accept any job offer coming from the formal sector. In particular, if a formal job offer pays too little or has poor working conditions, such worker might rather remain informal. In light of this, the question on preferences for formal jobs in the MOTRAL survey should be interpreted as indicating the potential pool of workers who would be interested in formal employment, provided that the job offer they receive from the formal sector is attractive enough.

In this sense, the discussion on whether to consider an informal worker as involuntary is similar to the one that arises in macroeconomics regarding involuntary unemployment. For instance, Robert Solow characterizes someone as involuntarily unemployed if “he or she would be prepared to work at a job that he or she knows how to do, at the going wage, and cannot find employment. […] by this definition a person can be involuntary unemployed and employed at the same time. If you take an inferior job you can still be an involuntary

| Table 2. Job separations of informal workers in last 5 years and their preference for covered jobs |
|---------------------------------------------------------------|
| Percent | Number of observations | Wants social security (%) |
|-------------------|-----------------|---------------------------|
| No separation in last 5 years | 52.0 | 1046 | 77.5 |
| Separation from uncovered jobs | 37.9 | 882 | 82.0 |
| Separation from covered jobs due to: |
| Market reasons | 6.8 | 217 | 87.9 |
| Voluntary reasons: |
| Insufficient income | 2.7 | 77 | 84.6 |
| Poor job conditions | 2.7 | 83 | 92.8 |
| Personal reasons | 5.4 | 138 | 84.5 |
| Other | 1.5 | 27 | 91.5 |
| Total | | 2470 | 80.5 |

Notes: The ‘Percent’ column reports the % of informal workers with a specified type of separation. The second column reports the unweighted number of workers with a specified type of separation. Workers can have more than one type of separation. ‘Market reasons’ include a contract ending, a firm closing down, firings, strikes, end of seasonal work, etc. ‘Poor job conditions’ include unsuitable schedules, poor hygiene, poor safety, etc. Sum of percentages exceed 100 because some workers experience more than one job separation over the period. All percentage estimates use sampling weights.

Source: Author’s calculations based on MOTRAL 2015 (INEGI, 2015a)
unemployed skilled worker.” (Snowdon & Vane, 1999, p.287). By the same token, one can consider informal workers as involuntary employed if they prefer, yet cannot find, formal employment that they know how to do, at the going wage, without this implying that they will accept any formal job offer that comes their way.

While it would be preferable to have more detailed information to disentangle the individual valuation of each component associated with formal employment, (e.g. health benefits, pensions, greater job security, higher wages, etc.), it is still informative to proceed by analysing the valuation question in hand and use it as a proxy that helps us identify the potential pool of candidates interested in good quality formal employment.

Based on the above considerations, I classify informal workers as involuntary if they state preferring a covered job. Furthermore, I assume that all formal workers prefer their employment to an uncovered job, irrespective of their answer to the valuation question. In other words, I assume they remain in formal employment because of other job qualities such as higher wages, greater job security, and in general better working conditions.14

The next section presents the econometric model used to exploit the information on the preferred type of job, with the goal of identifying the sociodemographic factors associated with applying for and being hired in formal employment.

3. Econometric methodology

As previously emphasised, this paper’s main contribution is its analysis of a dataset that contains information about the preferred type of job for a representative sample of urban workers. By having information that distinguishes between the desired and the actual types of job, one can estimate econometric models of sector assignment that disentangle application from hiring decisions for formal jobs.

To set the notation, \( V^a_i \) denotes the utility of worker \( i \) of applying and being employed in a formal job, and \( V^h\) denotes the corresponding propensity of a formal sector employer of hiring this worker \( i \). Assume that such propensities depend on vectors of observable characteristics of the worker, \((Z_i, X_i)\), as well as on a set of unobservables, \((u_{ai}, u_{hi})\).15 In particular, assume that these components are related by the following system of equations:

\[
V^a = Z \gamma_a + u_a \\
V^h = X \gamma_h + u_h
\]

(1.1)

(1.2)

where \((\gamma_a, \gamma_h)\) is a pair of vectors of unknown parameters.16 To estimate the parameters from this model, we assume the unobservables \((u_{ai}, u_{hi})\) follow a standard bivariate normal distribution with (unknown) correlation parameter \(\rho\).

The vectors of individual observable characteristics, \((Z, X)\) need not be the same across equations, and they will only include characteristics of the worker. Ideally, one would like to include characteristics of the various potential employers with whom a worker might be matched, but this information is not available in the data analysed. Another piece of missing information is whether individuals who do not apply for a formal job would be hired in that position if they were to apply.17

In practice, for any given worker, there are three possible scenarios that can be distinguished in the data: i) being a formal worker, which occurs with probability \(P(V^a > 0, V^h > 0 | Z, X)\); ii) being an involuntary informal worker, which occurs with probability \(P(V^a > 0, V^h \leq 0 | Z, X)\); and iii) being a voluntary informal worker, which occurs with probability \(P(V^a \leq 0 | Z, X)\). These three scenarios are incorporated into a discrete choice model that captures the joint decisions of workers and formal sector employers.18 The likelihood function of the discrete choice problem is
\[
\mathcal{L} = \prod_{i \in \mathcal{V}} [1 - F(Z_{i}^{a})] \prod_{i \in \mathcal{I}} G(Z_{i}^{a}, -X_{i}^{h}; -\rho) \prod_{i \in \mathcal{F}} G(Z_{i}^{a}, X_{i}^{h}; \rho)
\]

where \(F(\cdot)\) and \(G(\cdot, \cdot, \cdot)\) are the standardised normal and bivariate normal distributions, respectively, \(\mathcal{V}\) is the set of voluntary informal workers, \(\mathcal{I}\) is the set of involuntary informal workers, and \(\mathcal{F}\) is the set of formal workers. This model can be estimated via maximum likelihood. The estimates of the model provide a reduced form *ceteris paribus* answer to the question of which individual characteristics are associated with applying for and being hired in a formal job.\(^{19}\)

The discrete choice model can also be extended to estimate, in a second stage, selectivity-corrected earnings equations for different types of workers through a switching-regression model that allows for sector allocation based on both the application and hiring decisions modelled in equations (1.1) and (1.2). Under the appropriate assumptions the selectivity-adjusted log-earnings functions will be

\[
\log y_{s} = X_{s}\beta_{s} + \theta_{sa}\hat{\lambda}_{a}(Z, X) + \theta_{sh}\hat{\lambda}_{h}(Z, X) + \epsilon_{s},
\]

where the subscript \(s\) refers to the three different groups of workers characterised previously, i.e. formal, involuntary informal, and voluntary informal workers.\(^{20}\) The terms \(\hat{\lambda}_{a}(Z, X)\) and \(\hat{\lambda}_{h}(Z, X)\) are selectivity correction terms that adjust for the fact that individuals are not randomly assigned across sectors.\(^{21}\)

Key demographic characteristics are included among the observables \(Z\) that affect the probability of applying for formal employment. In particular, \(Z\) includes age, gender, schooling variables, marital status, as well as the household dependency ratio, and the interaction of this ratio with the number of adult females in the household. As previously mentioned, more dependents will need more hours of care and a higher income to sustain them. Since females typically do most of the housework, this variable is expected to negatively affect their probability of applying for a formal job as these jobs are less flexible in their schedules. However, as the number of adult females increases within a household, the housework load per woman will be smaller, hence the need to interact these two variables. The effects of these demographic variables are estimated separately by gender.\(^{22}\) In addition, a set of dummy variables at the city level are included to control for varying conditions in local labour markets.

The variables included in vector \(X\), which enters the hiring equation (1.2), are a subset of \(Z\), and only include variables that would be observable by a formal sector employer, such as age, gender, marital status, the aforementioned schooling variables, and city-level dummies. In other words, the demographic characteristics of the household (i.e. the dependency ratio and its interaction with the number of adult females) are excluded, as these characteristics are typically not observable by employers and, thus are less likely to affect their hiring decisions. Finally, all three log-earnings equations (3) include age, years of tenure in the current job, gender and marital status (interacted), schooling variables, and city-level dummies.

It is important to note that neither in vector \(Z\) nor in \(X\) I include information about firm characteristics or the current occupation, as these characteristics occur after the sector selection process has taken place, and thus are already an endogenous outcome of the job allocation process.

4. Results

This section presents the results of the estimation of the models described above, beginning with the results pertaining to the discrete choice model.
4.1 Discrete choice model

The results of the parameter estimates of equations (1.1) and (1.2) are included in the Online Appendix. Instead, Table 3 presents the average partial effects of the model as these are easier to interpret, i.e. I report the average derivatives of a probability of interest (e.g. the probability of applying for a formal job) with respect to one explanatory variable, holding all other observable variables constant.

The first two columns of Table 3 report the average partial effects (and their standard errors) of the ‘apply’ equation (1.1), while the last pair of columns presents the partial effects of the ‘hiring’ decision conditional on having applied for a formal job position, i.e. the average partial effects of covariates on the conditional probability:

$$P(V^a > 0, V^h > 0| Z, X) \quad P(V^a > 0| Z, X)$$.

Analysing first the partial effects corresponding to the decision to apply to a formal job, one observes a concave relation between the propensity to apply and age for females. For males, however, the partial effect of age is negative among young workers and statistically insignificant otherwise. Married workers are about 4 percentage points more likely to apply for formal jobs, although the effect is only statistically significant for males. While several of the partial effects differ by gender, on average males are equally likely as females to apply to formal jobs. Also, the education variables are not significantly associated with the applying decision.

1356  R. Duval-Hernández
A marginal increase in the dependency ratio decreases for women the probability of applying for formal employment by almost 5 percentage points, but this negative effect disappears if there is another adult female present in the household, presumably because that other female will help with the care of the dependents. In contrast, the opposite effect is found for males, i.e. there is a positive association between this ratio and the propensity to apply. These findings support the idea that the division of labour at home has an important influence on whether workers want formal jobs, but the effects differ depending on the gender of the worker. More specifically, having more dependents and no extra help leads women to search for more flexible informal jobs, as these will allow them to care for dependents. For men, in contrast, having more dependents increases the propensity to apply for formal jobs. This finding is consistent with a more traditional role of men as the main breadwinners in the household.

Regarding the probability of being hired in a formal job, conditional on having applied, this probability presents a concave pattern with respect to age for females and a convex pattern association for males. Married men are almost 12 percentage points more likely to be hired in formal employment than single males, while the opposite is the case for married females. This hiring penalty for married females likely reflects that formal sector employers perceive them as less attractive employees, a finding that is consistent with recent experimental evidence for Mexico (Arceo-Gomez & Campos-Vazquez, 2014). These findings, together with those pertaining to the applying decision, highlight the importance of the division of labour at the household level in shaping both applying and hiring decisions in the formal sector.

Finally, education is a key factor affecting the probability of being hired. In particular, an extra year of schooling is associated with a higher probability of being hired by 3 percentage points, while being enrolled in school decreases this hiring probability by almost 25 percentage points.23

---

Table 4. Selectivity-adjusted log-earnings OLS equations.

| Has social security | Wants social security | Does not want social security |
|---------------------|-----------------------|-------------------------------|
| Age                 | 0.00189 (0.0205)      | 0.0295 (0.0341)               | 0.0543 (0.0421) |
| Age sq.             | 0.0000378 (0.000267)  | -0.000230 (0.000324)          | -0.000651 (0.000567) |
| Single female       | -0.118 (0.105)        | -0.117 (0.155)                | -0.410** (0.155) |
| Married female      | 0.187 (0.0424)**      | 0.486 (0.144)**               | 0.389 (0.104)** |
| Married male        | 0.484 (0.0898)**      | 0.363 (0.131)**               | 0.449 (0.101)** |
| Tenure              | 0.0223 (0.0186)       | 0.0206 (0.00643)**            | 0.0133 (0.0376) |
| Tenure sq.          | -0.000827 (0.000748)  | -0.00105 (0.000426)**         | -0.000110 (0.00121) |
| Years schooling     | 0.148 (0.0347)**      | 0.0445 (0.0335)               | 0.0744 (0.0193)** |
| Enrolled in school  | -0.486 (0.221)**      | 0.0535 (0.466)                | -0.317 (0.188)  |
| λ(apply)            | -0.868 (0.620)        | -0.0633 (0.702)               | -0.411 (0.179)** |
| λ(hire)             | 1.059 (0.612)*        | -0.265 (0.776)                |                  |
| Constant            | 6.087 (0.498)**       | 6.466 (1.222)**               | 4.834 (0.677)** |
| R²                  | 0.393                 | 0.186                        | 0.501           |
| N                   | 2139                  | 1553                         | 368             |

Notes: Standard errors robust to clustering at the city level and adjusted for generated regressors in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01. Inference done with a t-distribution with 32 clusters – 1 as degrees of freedom. Formal workers are those in a job with social security coverage. City fixed effects included. Sample includes only individuals with positive earnings. All estimates use sampling weights. Source: Author’s calculations based on MOTRAL 2015 (INEGI, 2015a) and ENOE 2nd Quarter 2015 (INEGI, 2015b).
4.2 Selectivity-adjusted earnings equations

If the above discrete choice model is complemented with log-earnings equations, one can obtain parameters for the latter which are adjusted for potential sample-selectivity biases. As not all individuals report their earnings or have positive earnings, the previous discrete choice model is re-estimated for a subsample of workers with positive earnings and is used to estimate the selectivity correction terms in equation (3).24

The results of these earnings estimations are presented in Table 4 which shows that the age profiles are statistically insignificant. In general, males obtain substantially higher earnings relative to otherwise comparable single females. The earnings premiums range between 20 and 60 per cent.25 In contrast, among voluntary informal workers, married women display a \textit{ceteris paribus} earnings penalty of more than 30 per cent relative to their single female counterparts. These numbers indicate that demographic factors at the household level affect the job allocation process as well as the earnings obtained in the market.

The number of years of tenure at a firm have increasing concave returns. However, this shape is statistically significant only among involuntary informal workers, with an inflection point around 10 years of tenure. An extra year of schooling is associated with an increase in earnings of about 15 per cent in formal jobs and 7 per cent among those who do not want a job with coverage and has no significant effect on the earnings of involuntary informal workers. Being enrolled in school leads to \textit{ceteris paribus} earnings losses of almost 50 per cent among formal workers only.

Finally, from the selectivity correction terms we observe that the unobserved factors that affect the probability of being hired in formal jobs are positively correlated with the unobservable term in the earnings equation of formal workers, while the unobserved factors that affect the probability of applying for formal jobs are negatively correlated with the unobservables in the earnings equation of voluntary informal workers.

It is worth mentioning that the fit of the earnings equations is higher for workers employed in their preferred sector, i.e. formal and voluntary informal workers. This indicates that there is a greater degree of unexplained heterogeneity in the earnings of involuntary informal workers.

Taking stock of the above results, the econometric analysis performed indicates that the division of labour at home across gender lines plays an important role in determining who applies for and who is hired in formal jobs, as well as the earnings therein gained. Also, the levels of schooling are a crucial factor affecting both the probability of being hired in a formal job and the earnings obtained in it. Before concluding some robustness checks are discussed in the next section.

4.3 Robustness checks

In addition to the baseline specification above, two additional models are estimated to test the robustness of the previous results. The results of these estimations are presented in the Online Appendix.26

The first robustness check adds to the baseline specification a regressor measuring whether an adult member in the household has a formal job and examines whether this influences the respondent’s decision to apply to a formal job.

The results from this exercise show that having a member of the household employed in a formal job \textit{increases} the respondent’s probability of applying to formal employment by about 7 percentage points, while the coefficients of other variables are very similar to the ones reported in Table 3.27 This means that rather than strategically dividing who is formal and who informal, respondents with a formally employed household member are more likely to seek such type of employment as well. This probably reflects positive assortative mating among workers, whereby individuals with a higher valuation of formal jobs form households together.28
In a second robustness check the definition of ‘formal employment’ is modified to incorporate in this group the formal self-employed, as defined by the Mexican Statistical Agency (INEGI, 2014). In this alternative sample, the interpretation of the hiring equation (1.2) changes, as it now describes factors that affect hiring by a formal sector employer, as well as the costs and benefits for a firm operating formally. In general, the average partial effects of this specification are very similar to the base specification and the differences are mainly in the magnitudes and statistical significance of some effects. One of the main differences observed in this exercise is the fact that the years of schooling positively affect the probability of applying for a formal job, while in the previous sample this effect was statistically insignificant. This last finding indicates that more educated workers are more likely to seek establishing a formal business.

5. Conclusions

This paper exploits a unique dataset containing information about the preferred type of jobs of workers in large urban centres in Mexico. Comparing this information with the actual jobs they have, it finds that 80 per cent of informal workers consider it preferable to be employed in a formal job, even if this entails paying the corresponding taxes associated with such job. Rather than a preference for social security coverage per se, this number likely reflects a preference for jobs with a good set of characteristics (e.g. high pay, job stability, fringe benefits, etc.) such as those often found in the formal sector.

The empirical analysis conducted highlights factors that hinder the workers’ access to formal employment from the supply side of the market. In particular, the traditional division of labour at home is a likely culprit for limiting the willingness of females to apply for formal wage jobs and the probability of being hired in such jobs. In addition, having a higher level of education plays a significant role in increasing the chances of being hired in a formal job and of earning a higher income from it.

The creation of more good quality jobs is a complex task that requires policies operating at many different levels. It requires policies that affect both the demand as well as the supply side of the labour market, in addition to other policies in other markets (e.g. credit) that complement the labour policies.

The results from this research indicate that in addition to any efforts to foster the creation of good formal jobs from the demand side of the labour market, it is important for policymakers to consider how they can affect the two aforementioned dimensions in order to encourage the successful transition of workers into formal employment.

Finally, the analysis conducted here shows that there is much to be gained by studying the stated preferences of workers about potential jobs and their characteristics. In future work, it would be profitable to perform this type of analysis with more detailed survey instruments such as those used in discrete choice experiments in order to disentangle what characteristics of formal jobs are attractive to informal workers.

Notes

1. For instance, Levy (2008, 2018) has argued that the design of social protection policy in Mexico has led to the misallocation of resources and contributes to the low growth performance of the economy.
2. Other studies in this tradition include Gong and van Soest (2002) and Gong, van Soest, and Villagomez (2004).
3. Other papers have tried to test segmentation by comparing formal and informal wage equations after correcting for self-selection into each sector. However, the methods used to correct for sample selectivity are often based on sector choice models that assume free entry into formal employment. This is problematic as free choice among sectors is precisely the issue these papers try to test (see, for instance, Marcouiller, Ruiz de Castilla, & Woodruff, 1997 for the case of Mexico).
4. For instance, Perry et al. (2004) find greater satisfaction associated with formal jobs in Colombia, while the sectoral differences in satisfaction are milder in the Dominican Republic. Falco et al. (2015) infer that about
half of the informal workers in Ghana would prefer formal employment, yet their results show substantial heterogeneity in job satisfaction depending on the type of informal worker under consideration.

5. Alternatively, one could define the formality status for the self-employed based on whether a business is registered with tax authorities and has fixed work premises (INEGI, 2014). A robustness check that considers this criterion of formality for the self-employed is discussed in section 4.3.

6. The supplementary module is the Módulo de Trayectorias Laborales, 2015 (MOTRAL) and its data can be publicly accessed online (see INEGI, 2015a). A similar module was also applied in 2012, but in that edition the key variable used in the analysis that follows was not included.

7. Information on these individuals is also included in the Labour Force Survey (INEGI, 2015b), and the two datasets can therefore be linked, as is done here, in order to have a richer set of variables.

8. The original question reads: ‘¿Considera que es mejor tener un empleo con seguridad social, aunque tenga que realizar pagos para tener derecho a ella?’

9. If respondents were to answer the valuation question based on their preference over social security benefits per se, and not over formal jobs in general, then one could potentially misclassify them as involuntary informal. However, Levy (2008) argues that for many workers the net benefits package offered by Social Security is not necessarily superior to the benefits available through non-contributory Social Protection schemes. Furthermore, very few self-employed voluntarily enrol for social security coverage (as permitted by the law), indicating that few of them value social security benefits per se. Therefore, it seems unlikely that informal workers declare valuing covered jobs based on their preference over social security benefits alone, and not because of a more general preference for formal employment.

10. The Online Appendix that accompanies this paper presents an additional set of descriptive statistics for the samples used in the various econometric estimations.

11. The percentages in the first column of the table do not sum up to 100 because a worker might have more than one separation over the previous 5 years.

12. In fact, the lowest preference for covered jobs is found among those who did not experience separations over the past 5 years.

13. The fact that the preference for a formal wage job is 20 percentage points lower in the ENAMIN survey relative to the MOTRAL sample suggests that some of the respondents in the MOTRAL survey state their preference while envisioning moving to a better paying formal job.

14. This entails assuming that the 8.9 per cent of the employed population who has a covered job but do not value their social security benefits (see column 2 in Table 1), still prefer formal to informal employment, as otherwise they would voluntarily move to an uncovered informal job.

15. In this context, ‘unobservable’ means characteristics not available to the econometrician.

16. In the above equations and the ones that follow, I drop the individual subscript \( i \).

17. It is conceptually possible that voluntary informal workers could be offered a formal job if they were to apply for that job. However, I do not allow for that possibility in the econometric model estimated.

18. This model is sometimes called a bivariate probit with ‘sample selection’ (see for instance Van de Ven & Van Praag, 1981).

19. The model is a reduced form one because it does not explicitly incorporate wages in the choice decisions equations.

20. A detailed presentation of the model is included in the Online Appendix. A more detailed exposition, including the formulas for the standard errors of the parameters, can be found in Tunali (1986).

21. In the case of the earnings equation of voluntary informal workers, the term \( \theta_{0,1} \beta_{1} \Gamma \) is zero.

22. Talamas (2021) estimates that in Mexico the deaths of grandmothers (the main childcare providers) reduce mothers’ employment by almost 30 per cent and have no effect on fathers.

23. Online Appendix Table OA.8 reports the correlation parameter between the unobservables (\( u_{g}, u_{h} \)), which is -0.53 (and statistically significant at the 1 per cent level), indicating that the unobservable factors that lead to a greater probability of applying for formal employment are negatively correlated with the unobservables that affect the probability of being hired.

24. The parameter estimates of the bivariate probit over this slightly more restricted sample are available upon request from the author.

25. These premia are obtained by noting that for a given gender \( \times \) marital group \( g \), the parameters estimated are \( \beta_{g} = \ln(y_{g}) - \ln(y_{g}^*) \), where \( \ln(y_{g}) \) and \( \ln(y_{g}^*) \) are the predicted log earnings of group \( g \) and single females, respectively. Hence \( \exp \left( \beta_{g} \right) - 1 \) approximates the percentage premium for group \( g \) relative to single females.

26. In addition, the Online Appendix presents predicted counterfactual earnings of informal workers if they worked in formal employment.

27. See Table OA.4 in the Online Appendix.

28. An interesting extension of this exercise would be to jointly model the sectoral labour supply of couples, rather than treating the employment status of other family members as exogenously given. This exercise is however, beyond the scope of this paper.

29. The second sample is smaller than the first one because I import information from the Labour Force Survey (INEGI, 2015b) to determine who among the self-employed is formal, and in the process of merging both
samples. I discard observations with discrepancies at the job characteristics level (e.g. occupation, industry, etc.). Additional descriptive statistics for this sample are found in Table OA.1 in the Online Appendix.

30. The evidence discussed in Ulyssea (2020) indicates that the costs that matter most for formal firms are not the entry costs to the formal sector (e.g. registration) but rather the costs of the ongoing operation of a formal business, such as taxes.

Acknowledgements:
This paper is part of the project ‘Transforming informal work and livelihoods’ of the United Nations University World Institute for Development Economics Research (UNU-WIDER). I would like to thank Gary S. Fields, Tim Gindling, Kunal Sen, Simone Schotte, Michael Danquah, and seminar participants at the workshops organised by the project, as well as four anonymous referees for their valuable comments. The opinions and views here expressed are my own and do not represent those of my employer or of UNU-WIDER. All the data used in this paper is freely available to download from INEGI (see bibliography). All the codes used to estimate the results in the paper are available from the author upon request.

Disclosure statement
No potential conflict of interest was reported by the author(s).

References
Alcaraz, C., Chiquiar, D., & Saucedo, A. (2015). Informality and segmentation in the Mexican Labor Market (Banco de México Working Papers 2015-25). Mexico City: Mexico. doi:10.36095/banxico/di.2015.25
Arceo-Gomez, E.O., & Campos-Vazquez, R. M. (2014). Race and marriage in the labor market: A discrimination correspondence study in a developing country. American Economic Review, 104(5): 376–380. doi:10.1257/aer.104.5.376
Assy, A. E., Ribeiro, T., Robalino, D., Rosati, F., Sanchez Puerta, M. L., & Weber, M. (2019). The jobs that youth want and the support they need to get them: Evidence from a discrete choice experiment in Kenya (IZA Discussion Paper Series, No. 12864).
Bosch, M., & Maloney, W. F. (2010). Comparative analysis of labor market dynamics using Markov processes: An application to informality. Labour Economics, 17(4): 621–31. doi:10.1016/j.labeco.2010.01.005
Contreras, D., Gillmore, R., & Puentes, E. (2017). Self-employment and queues for wage work: Evidence from Chile. Journal of International Development, 24(4): 473–99. doi:10.1002/jid.3074
Duval-Hernández, R., & Smith, R. (2010). Counterfactual wages and segmentation between formal and informal sectors in Mexico (Unpublished Manuscript). División de Economía, CIDE, Mexico City, Mexico.
Falco, P., Maloney, W. F., Rijkers, B., & Sarrias, M. (2015). Heterogeneity in subjective wellbeing: An application to occupational allocation in Africa. Journal of Economic Behavior and Organization, 111, 137–153.
Fields, G. S. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. Journal of Development Economics, 2(2): 165–87. doi:10.1016/0304-3878(75)90014-0
Fields, G. S. (1990). Labour market modelling and the urban informal sector: Theory and evidence. In D. Turnham, B. Salome, & A. Schwarz (Eds.), The informal sector revisited. Paris: Organisation for Economic Co-operation and Development.
Fiess, N. M., Fugazza, M., & Maloney, W. F. (2010). Informal self-employment and macroeconomic fluctuations. Journal of Development Economics, 91(2): 211–26. doi:10.1016/j.jdeveco.2009.09.009
Gindling, T. H., & Newhouse, D. (2014). Self-employment in the developing world. World Development, 56: 313–31.
Gong, X., & Van Soest, A. (2002). Wage differentials and mobility in the urban labour market: A panel data analysis for Mexico. Labour Economics, 9, 513–529.
Gong, X., van Soest, A., & Villagomez, E. (2004). Mobility in the urban labor market: A panel data analysis for Mexico. Economic Development and Cultural Change, 53(1), 1–36.
Günther, I., & Launov, A. (2012). Informal employment in developing countries: Opportunity or last resort? Journal of Development Economics, 97, 88–98.
INEGI (Instituto Nacional de Estadística y Geografía) (2012). Encuesta Nacional de Micronegocios [National Survey on Micro-entreprises]. Retrieved from https://www.inegi.org.mx/programas/enamin/2012/
INEGI (Instituto Nacional de Estadística y Geografía) (2014). La informalidad laboral: Encuesta Nacional de Ocupación y Empleo: marco conceptual y metodológico [Labour informality. National survey of occupation and employment: Conceptual and methodological framework]. México: INEGI.
INEGI (Instituto Nacional de Estadística y Geografía) (2015a). *Módulo de Trayectorias Laborales (MOTRAL) 2015* [Module of Labour Trajectories (MOTRAL) 2015]. Retrieved from https://www.inegi.org.mx/programas/motral/2015/

INEGI (Instituto Nacional de Estadística y Geografía) (2015b). *Encuesta Nacional de Ocupación y Empleo, población de 15 años y más de edad* [National Survey of Occupation and Employment, population aged 15 years and more]. Retrieved from https://www.inegi.org.mx/programas/enoe/15ymas/

Levy Algazi, S. (2018). *Esfuerzos mal recompensados: La elusiva búsqueda de la prosperidad en México*. Washington, DC: Banco Interamericano de Desarrollo.

Levy, S. (2008). *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico*. Washington, DC: Brookings Institution Press.

Magnac, T. (1991). Segmented or competitive labor markets. *Econometrica, 59*(1): 165–87. doi:10.2307/2938245

Mahmud, M., Gutierrez, I. A., Kumar, K. B., & Nataraj, S. (2021). What aspects of formality do workers value? Evidence from a choice experiment in Bangladesh. *The World Bank Economic Review, 35*(2), 303–327.

Maloney, W. F. (1999). Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *World Bank Economic Review, 13*(2): 275–302. doi:10.1093/wber/13.2.275

Maloney, W. F. (2004). Informality revisited. *World Development, 32*(7): 1159–78. doi:10.1016/j.worlddev.2004.01.008

Marcouiller, D., Ruiz de Castilla, V., & Woodruff, C. (1997). Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural Change, 45*(2): 367–92. doi:10.1086/452278

Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and exclusion*. Washington, D.C.: The World Bank.

Snowdon, B., & Vane, H. R. (1999). *Conversations with leading economists. Interpreting modern macroeconomics*. Chentelham, UK.: Edward Elgar.

Soares, F. V. (2004). *Do informal workers queue for formal jobs in Brazil?* (Texto para Discussão #1021). Rio de Janeiro: IPEA.

Talama, M. (2021). *Grandmothers and the gender gap in the Mexican labor market* (Unpublished manuscript).

Tunali, I. (1986). A general structure for models of double selection and an application to a joint migration/earnings process with remigration. *Research in Labor Economics, 8*, 235–82.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics, 12*(1): 525–46. doi:10.1146/annurev-economics-082119-121914

Van de Ven, W. P. M. M., & Van Praag, B. M. S. (1981). The demand for deductibles in private health insurance. A probit model with sample selection. *Journal of Econometrics, 17*(2), 229–52. doi:10.1016/0304-4076(81)90028-2