Labor Market Frictions, Search and Informality with Endogenous Schooling Investments*

Matteo Bobba   Luca Flabbi   Santiago Levy

Preliminary draft

May 27, 2015

Abstract

Many labor markets, typically in medium- and low-income countries, are characterized by high levels of informality. Informality is associated with particularly weak indicators of labor market performance but may also constitute an important margin to increase labor market flexibility and improve workers allocation. Even if informality can be an optimal response given the labor market environment, it is important to consider its long-term costs. In this paper we focus on one margin of the workers side long-term cost: the possible under-investment in education. We propose a search-matching-bargaining model of a frictional labor market in which potential employees are randomly matched with firms that could offer legal and illegal wage contracts and individuals choose (i) whether becoming self-employed or remaining unemployed prior to searching for a job as employees and (ii) whether acquiring productivity-enhancing schooling prior to labor market entry. Structural estimation of the model's parameters using Mexican micro-data enables us to quantify the sensitivity of educational investments with respect to the labor market environment and to evaluate the equilibrium impacts of alternative labor market policies.

Keywords: Labor market search, Nash bargaining, Hold-up, Informality; Schooling decisions.
JEL Codes: J24, J3, J64, O17

*Inter-American Development Bank, 1300 New York Avenue 20577 Washington DC (e-mail: matteob@iadb.org; lucaf@iadb.org; slevy@iadb.org). We thank the participants at the Workshop *Labor Market Opportunities and Schooling Investments* held at the IDB on September 28th, 2012 for very useful comments. Marco Pariguana, Jose Mauricio Salazar and especially Matias Morales provided excellent research assistance.

1 Introduction

Many labor markets, typically in medium- and low-income countries, are characterized by high levels of informality. At the most general level, informality is defined as any deviation from the labor contract as designed by law, including absence of workers protections, no or significantly lower payroll contributions and lack of conformity to other labor law statutes. The issue is particularly acute in Latin America where even large middle-income economies with well developed labor market institutions see level of informality larger than 50% of the labor force. In Mexico, the country where we will perform our empirical analysis, about 63% of employed workers is informal. The population average for the region is about 55%[1]

Informality is associated with particularly weak indicators of labor market performance. Informal workers tend to earn lower wages, to experience more earnings instability and to accumulate less human capital. Informal firms tend to be less productivity, to innovate less and to fail at higher rates.On the other hand, informality may constitute an important margin to increase labor market flexibility and improve workers allocation. The latest claim is supported by the relative fluidity of the informality status. Both on the demand and the supply side and for a level of skills in the middle range, it is rare to find workers always informal and firms fully informal. More and more workers tend to move in and out the informality status over their labor market career while more and more firms tend to hire a mix of formal and informal workers.

Even if informality can be an optimal response given the labor market environment, it is important to consider its long-term costs. On the firms side, informality may lead to sub-optimal size and to poor access to financial markets. On the workers side, informality may lead to sub-optimal human capital investment and to excessive job instability.In this paper we focus on one margin of the workers side long-term cost: the possible under-investment in education. We focus on a particularly active schooling level: the secondary level. If primary schooling completion is by now universal in all the middle-income countries in LAC and if tertiary schooling level is still beyond the means of large part of the population, the secondary schooling level is really the next step to address in order to reduce education inequality and increase the skill composition of the population.

Given the trade-offs involved in the choice of formality status and level of education, we propose an equilibrium model of the labor market with a detailed modeling of the institutional setting, a careful definition of informality and an endogenous schooling process. We focus on the Mexican labor market and on the completion of High School. Mexico is a major economy in the Region with a high and persistent level of informality. The completion of High School (Secondary Educa-

[1] See Levy and Shady (2013). Our definition of informality is based on presence of absence of social security contributions.

tion) is a major milestone in middle income economies and it is the education level where policy actions and individual decisions are concentrated.

We propose a search-matching-bargaining model of a frictional labor market in which individuals choose (i) whether becoming self-employed or remaining unemployed prior to searching for a job as employees and (ii) whether acquiring productivity-enhancing schooling prior to labor market entry. Potential employees are randomly matched with firms that could offer legal and illegal wage contracts. Upon observing match-specific productivity, workers and firms engage in bargaining over both wages and legality status of the contract. Workers are allowed to value differently the extra-wage benefits embedded in the two contract types. Since the potential costs of an illegal contract are increasing in productivity, only relatively lower productivity matches will be associated with an illegal job status.

However, the outside options in self-employment are heterogenous among workers, generating substantial overlap in the wages and productivity distributions of legal and illegal jobs in equilibrium. This is a crucial result to match the data of typical high-informality labor markets. Estimation of the model using micro-data from the Mexican labor force survey enables us to quantify the sensitivity of educational investment with respect to the labor market parameters and evaluate the impacts of alternative labor market policies on the returns to schooling. Preliminary results show highly non-linear impacts on education levels and informality levels of the two main policy parameters available in this context: payroll contributions and non-contributory benefits.

The informality literature in an equilibrium context is scarce. Bosch and Esteban-Pretel (2011) calibrate a two sector model for Brazil where firms have a choice of hiring workers formally or informally. Albrecht, Navarro and Vroman (2009) is an equilibrium search model studying the distributional implications of labor market policy in a labor market with an informal sector Meghir, Narita and Robin (2015) develop an equilibrium search model with wage posting endogenous firms decisions about locating in the formal or informal sector. They estimate the structural parameters of the model on Brazilian labor force data.

Our contribution is different with respect to the previous literature along several dimensions. First, we include an endogenous schooling decisions and therefore policy experiments are allowed to impact returns to schooling investment and the equilibrium impact of the resulting schooling choices. Second, we propose a better and more nuanced definition of informality and we incorporate both in the theoretical model and in the empirical implementations. Third, we implement a more careful treatment of the insitutional context which allows for counterfactual experiments which are more policy-relevant. Finally as Meghir, Narita and Robin (2015), we allow for endogenous choice of the formality status, no restrictions of labor market mobility between states and for a structural estimation using representative labor market data. However, our bargaining structure allows for a joint decision of the formality status between worker and firm.

Finally, our paper relates to the labor market search literature with endogenous schooling decisions. Schooling decisions in labor market search model are usually taken as given due to the theoretical complexity of embedding endogenous schooling decisions in a framework already comprising a fairly rich labor market decision process. Adding this ingredient, however, is essential to satisfy the objective of our research agenda since we want to study the impact of labor market conditions on schooling decisions and not only the returns to skills once these decisions have been taken.[2] Moreover, a quantitative assessment of the extent of the hold up problem in human capital investments requires a framework with endogenous schooling decisions. Since workers must invest in schooling *before* finding a job and since, in a non-competitive labor market, part of the payoff of the schooling investment can be appropriated by firms, agents may *held up* their schooling investment. The resulting inefficiencies are larger the larger the frictions in the labor market.[3] Flinn and Mullins (2015) extend the standard search and matching framework to allow for endogenous schooling decisions. They propose a labor market characterized by search frictions and by firms and workers deciding not only about the job relation but also about investment decisions prior to market entry. Interestingly, after appropriate distributional assumptions are imposed, the model is tractable enough to be estimated using standard labor market dynamics data. We will use results from their environment to embed endogenous schooling decisions in our model.

2 Model

A model of labor market search with matching and bargaining in the presence of formal and informal labor market opportunities is presented. When a potential employee and a firm meet, the productive value of the match is immediately observed by both the applicant and the firm. At this point a division of the match value and the legal status of the job is proposed using a Nash-bargaining framework. We first characterize the equilibrium conditions under which firms and workers agree on a match and whether a legal or illegal job is formed. We then extend the basic framework by allowing individuals to permanently choose (i) whether to become self-employed or remaining unemployed prior to searching for a job as employees and (ii) to acquire productivity-enhancing schooling prior to labor market entry.

[2] Eckstein and Wolpin (1995) is a classic reference in this respect, showing how returns to schooling change when we are looking at offered wages as opposed to (incorrectly) accepted wages. Flabbi and Tejada (2012) is a more recent contribution in this vein, evaluating the equilibrium impact of discrimination on wage offers distributions by gender and education.

[3] Acemoglu and Shimer (1999) study under what conditions these inefficiencies can be alleviated or eliminated. **?** estimates that, in the context of the US labor market, the extent of the hold up inefficiency is very sensitive to the workers' bargaining power parameter.

2.1 Environment

The model is formulated in continuos time and assumes stationarity of the labor market environment. Unemployed individuals search for a job as employees and they receive an instantaneous utility (or disutility) flow ξ which is assumed constant in the population. The arrival of job offers follows a Poisson process at instantaneous rates λ. Once an employer and a worker meet, they observe a match-specific productivity value x, modeled as a draw from an exogenous distribution denoted by the cdf G. Once a match is formed, it can be terminated following a Poisson process at instantaneous rate η. There are no offers while the individual is working as an employee, and the instantaneous common discount rate is ρ. Workers' utility when employed is

$$v(x, f) = w(x, f) + f\beta_1 B_1 + (1 - f)\beta_0 B_0, \tag{1}$$

where w is the wage, B_1 represents the monetary value of the bundle of contributory social-security insurance (CSI) of working as a legal employee $(f = 1)$, B_0 is the monetary value of the lump-sum transfer that both illegal employees $(f = 0)$ and unemployed individuals receive through non-contributory social-security insurance (NCSI). In this framework, β_1 and β_0 define the marginal willingness to pay for CSI and NCSI, both of which are assumed constant in the population. Given the diverse nature of CSI, we parametrize B_1 as following:

$$B_1 = \gamma\tau[w(x, f)] + b_1, \tag{2}$$

where γ denotes the relative share of the bundle which is assumed to be returned one-to-one to the workers through proportional extra-wage benefits (e.g. pensions) and b_1 is the lump-sum transfer that formal workers receive in equilibrium once all firms' withdrawals have been collected (e.g. health insurance). The function $\tau(\cdot)$ possibly captures both the withdrawal at the source operated by firms on their formal employees' wages and income taxes. In most countries, the former tends to be regressive while the latter progressive, so for simplicity we assume a linear function in wages: $\tau[w(x, f)] = \tau w(x, f)$.

The corresponding value functions for the potential worker in the states of employment and unemployment is:

$$V(x, f) = \frac{w(x, f)(1 + f\beta_1\gamma\tau) + f\beta_1 b_1 + (1 - f)\beta_0 B_0 + \eta Z}{\rho + \eta}, \tag{3}$$

where Z is the value of unemployment, which is defined by the following expression:

$$Z = \frac{\xi + \beta_0 B_0 + \lambda \int \max\{V(x, f), Z\}\, dG(x)}{\rho + \lambda}. \tag{4}$$

Firms' instantaneous profits from a filled job are:

$$p(x,f) = x - w(x,f) - f\tau w(x,f) - (1-f)c(x), \tag{5}$$

where the function $c(\cdot)$ is the expected value of the per-worker penalty the firm has to pay if discovered employing an illegal worker, which is assumed step-wise linear in the match-specific productivity value:

$$c(x) = \begin{cases} 0 & \text{if } x \le c_0, \\ c_1 x & \text{if } x > c_0. \end{cases} \tag{6}$$

This specification captures the notion that the legal authorities are more likely to inspect more productive firms, since fines are actually increasing with illegal workers' wage (that is a positive function of the match-specific productivity). According to this specification, only sufficiently high-productive firms (above c_0) interiorize the expected penalty from evading the social security law.

The corresponding value functions for a job filled at the firm are:

$$P(x,f) = \frac{[1-(1-f)c_1]x - w(x,f)(1+f\tau)}{\rho+\eta}, \tag{7}$$

where the value of the alternative state, an unfilled vacancy, is zero because we assume free-entry of firms in the market (or alternatively that posting a vacancy costs nothing).

2.2 Equilibrium

Potential employees and firms bargain over the surplus. We assume that the outcome of the bargaining game is the pair (w,f) described by the axiomatic Generalized Nash Bargaining Solution, using the value of unemployment (4) and zero, respectively, as threat points, and $(\alpha, 1-\alpha)$ as the workers and the firm's bargaining power parameters. We define the surplus of the match as the weighted product of the worker's and firm's net return from the match: $\Sigma(x,f) \equiv [V(x,f)-Z]^{\alpha} \times [P(x,f)]^{1-\alpha}$. Nash Bargaining implies the following equilibrium wage schedules:

$$w(x,f) = \frac{\alpha\,[1-(1-f)c_1]}{1+f\tau}x + \frac{1-\alpha}{1+f\beta_1\gamma\tau}\,[\rho Z - f\beta_1 b_1 - (1-f)\beta_0 B_0] \tag{8}$$

Equations (3) and (7) show that the values of being employed are monotone increasing in the match-specific productivity values and wages. Equation (4) shows that the value of unemployment is constant with respect to match-specific productivity values and wages. As a result and under mild regularity conditions, the optimal decisions rules are characterized by the following reservation match-specific productivity value:

$$x^*(f) = \frac{1+f\tau}{(1+f\beta_1\gamma\tau)\left[1-(1-f)c_1\right]}\left[\rho Z - f\beta_1 b_1 - (1-f)\beta_0 B_0\right]. \tag{9}$$

Potential employees only accept matches with productivity higher than $x^*(f)$. An analogous decision rule holds for firms. Because of Nash Bargaining, the reservation value at which the worker is indifferent between accepting the firm's offer or keep searching for a better match is equal to the reservation value at which the firm is indifferent to the option of holding a vacancy or hiring the worker. Equation (9) states that job legal status $f \in \{0,1\}$ has two opposite effects on the reservation productivity value at which the match is formed. It decreases the reservation value because employees receive a valuable job amenity B_1 (B_0), but it also increases the reservation value because the firm pays costs τ (c) to provide the amenity.

By substituting the equilibrium wage schedule (8) into the value function (3), we obtain the match-specific productivity value such that both employees and firms will be indifferent between a legal and an illegal job:

$$\tilde{x} = \frac{\beta_0 B_0 - \beta_1 b_1}{c_1 - \frac{(1-\beta_1\gamma)\tau}{1+\tau}}. \tag{10}$$

So long as workers do not fully value CSI benefits ($\beta_1 < 1$) and some firms expect to pay a non-zero penalty in case they decide to evade the social security law ($c_1 > 0$), the value function (3) for legal jobs is steeper than the one for illegal jobs at any match-specific productivity value. Because the cost of providing the legal status is lower for highly productive matches, only relatively lower productivity matches will be associated with an illegal job status. The range of productivities over which illegal and legal jobs are accepted is directly related to the model parameters though the reservation values defined in (9) and (10). The optimal decision rule is to reject the match if $x < x^*(0)$, accept the match as an illegal worker if $x^*(0) < x < \tilde{x}$ and accept the match as a legal worker if $x > \tilde{x}$.

Figure 1 depicts the value functions and the resulting decision rules. The value of illegal employment $V(0)$ is represented by a broken line with a change in slope at the match-specific productivity threshold at which firms expect to pay a penalty for evading the social security law. Accordingly, potential illegal matches with a productivity below c_0 will be accepted only if they do not entail any additional cost associated to the illegal status of the job, whereas in those above c_0 a fraction $c_1 \in (0,1)$ of the total surplus will be retained in order to face the potential penalty costs.

Using these decision rules and plugging the equilibrium wage schedule (8) onto the value functions (3), we can express the value of unemployment (4) as a function of the primitive parameters of the model

$$\rho Z = \xi + \beta_0 B_0 + \frac{\lambda\alpha}{\rho+\eta}\left\{\int_{x^*(0)}^{c_0}\left[x - (\rho Z - \beta_0 B_0)\right] dG(x) + \int_{c_0}^{\tilde{x}}\left[x - \frac{\rho Z - \beta_0 B_0}{1-c_1}\right] dG(x)\right\}$$
$$+ \frac{\lambda\alpha}{\rho+\eta}\left\{\int_{\tilde{x}}\left[x - \frac{(1+\tau)(\rho Z - \beta_1 b_1)}{1+\beta_1\gamma\tau} dG(x)\right]\right\}. \quad (11)$$

Given that $G(x)$ is an increasing function, this implicit function in ρZ has a unique solution.[4] We can therefore propose the following

Definition 1 *Given a vector of parameters $\Gamma = (\xi, \lambda, \eta, \rho, \alpha, \beta_0, \beta_1, \tau, \gamma, c_0, c_1, B_0)$ and one probability distribution function for the productivity of match values $G(x)$ a **labor market equilibrium** is a value of unemployment Z that solves equation (11). The equilibrium value Z determines all the reservation values that constitute each agent's decision rules.*

2.3 Self-employment Decision

Self-employment decisions are assumed to be taken once and for all before searching for a job as an employee. In particular, we allow individuals to be heterogeneous in the potential earnings derived from their own business activity. The latter variable is modeled as a draw from an exogenous distribution denoted by the cdf Y which summarizes heterogeneity with respect to the monetary costs, ability-related returns and preference-related determinants of being self-employed. We further assume that when a potential employee and a firm bargain over the wage and the job status, the realized value (if any) of self-employment profits y is immediately observed by both the applicant and the firm. In this setting, the value of searching for a job as an employee is

$$\rho Z(ys) = \xi(1-s) + \beta_0 B_0 + ys + \frac{\lambda_s\alpha}{\rho+\eta}\left\{\int_{x^*(0;ys)}^{c_0}\left[x - (\rho Z(ys) - \beta_0 B_0)\right] dG(x)\right\}$$
$$+ \frac{\lambda_s\alpha}{\rho+\eta}\left\{\int_{c_0}^{\tilde{x}}\left[x - \frac{\rho Z(ys) - \beta_0 B_0}{1-c_1}\right] dG(x) + \int_{\tilde{x}}\left[x - \frac{(1+\tau)(\rho Z(ys) - \beta_1 b_1)}{1+\beta_1\gamma\tau} dG(x)\right]\right\}, \quad (12)$$

where we have allowed the instantaneous rate of arrival job offers to vary (exogenously) according to the two regimes $s = 0$ and $s = 1$ of the search status.

[4] Analogously, it can be shown that the equilibrium value of the lump-sum component of the CSI benefits can be expressed as a function of the primitive parameters of the model and the value of unemployment (11).

$$b_1 = \left[\frac{(1-\gamma)\tau}{1+\tau\beta_1(1-\alpha+\alpha\gamma)}\right]\left\{\frac{\alpha(1+\beta_1\gamma\tau)}{[1-G(\tilde{x})](1+\tau)}\int_{\tilde{x}} x dG(x) + (1-\alpha)\rho Z\right\}.$$

This expression together with (11) defines a system of two implicit functions that can be solved iteratively.

Self-employment decisions are taken by comparing the value functions in expression (12) in the two regimes $s = 0$ and $s = 1$. The wage schedule, together with the value functions, implies that the optimal decision rule has a reservation value property. The value of unemployment $Z(0)$ is independent of the potential earnings accruing from starting-up a business, whereas the value of self-employment $Z(y)$ is increasing with respect to self-employment earnings. Hence, there exists a reservation value such that $Z(0) = Z(y^*)$[5], so that we have

$$y^* = \xi + \frac{\alpha(\lambda_0 - \lambda_1)}{\rho + \eta}\left\{\int_{x^*(0;0)}^{c_0} [x - (\rho Z(0) - \beta_0 B_0)]\, dG(x) + \int_{c_0}^{\tilde{x}} \left[x - \frac{\rho Z(0) - \beta_0 B_0}{1 - c_1}\right] dG(x)\right\}$$
$$+ \frac{\alpha(\lambda_0 - \lambda_1)}{\rho + \eta}\left\{\int_{\tilde{x}} \left[x - \frac{(1+\tau)(\rho Z(0) - \beta_1 b_1)}{1 + \beta_1 \gamma \tau} dG(x)\right]\right\}. \tag{13}$$

Given that $G(x)$ is an increasing function, this equation has a unique solution which positively depends on the difference in the (exogenous) arrival rates of job offers between unemployed and self-employed searchers.

Figure 2 displays the change in the value functions and in the reservation values which is induced by the presence of self-employed searchers in our model. The outside options in self-employment shift upward the intercepts without affecting the slopes, thereby increasing the match-specific productivity reservation value at which self-employed individuals accept an illegal job offer ($x^*(0;0) < x^*(0;y), \forall y > y^*$) without altering the reservation value for accepting a legal job offer ($\tilde{x}$).

2.4 Schooling Decision

We now introduce an additional dimension of heterogeneity across individuals: the stock of human capital they have accumulated prior to any occupational choice decision. For simplicity, we consider two levels of schooling: $h = 0$ and $h = 1$, and we allow all the parameters of the model to vary along this dimension. In terms of the labor market equilibrium described above, this implies

[5] According to expression (9), this implies that the match-specific reservation productivity values at which a potential employee is indifferent between accepting or not a given job (legal or illegal) are equal for the individual who is indifferent between becoming self-employed or remaining unemployed: $x^*(f; y^*) = x^*(f; 0)$.

that the value of searching can be expressed as

$$
\begin{aligned}
\rho Z(ys,h) = \xi_h(1-s) + \beta_{0,h}B_0 + ys + \frac{\lambda_{s,h}\alpha}{\rho+\eta_h}\left\{\int_{x^*(0;ys,h)}^{c_{0,h}} \left[x - (\rho Z(ys,h) - \beta_{0,h}B_0)\right] dG_h(x)\right\} \\
+ \frac{\lambda_{s,h}\alpha}{\rho+\eta_h}\left\{\int_{c_{0,h}}^{\tilde{x}} \left[x - \frac{\rho Z(ys,h) - \beta_{0,h}B_0}{1-c_{1,h}}\right] dG_h(x)\right\} \\
+ \frac{\lambda_{s,h}\alpha}{\rho+\eta_h}\left\{\int_{\tilde{x}} \left[x - \frac{(1+\tau)(\rho Z(ys,h) - \beta_{1,h}b_1)}{1+\beta_{1,h}\gamma\tau} dG_h(x)\right]\right\}.
\end{aligned}
\tag{14}
$$

Analogously, the reservation value of potential self-employment earnings in the presence of different schooling sub-markets becomes:

$$
\begin{aligned}
y^*(h) = \xi_h + \frac{\alpha(\lambda_{0,h}-\lambda_{1,h})}{\rho+\eta_h}\left\{\int_{x^*(0;0,h)}^{c_{0,h}} \left[x - (\rho Z(0,h) - \beta_{0,h}B_0)\right] dG_h(x)\right\} \\
+ \frac{\alpha(\lambda_{0,h}-\lambda_{1,h})}{\rho+\eta_h}\left\{\int_{c_{0,h}}^{\tilde{x}} \left[x - \frac{\rho Z(0,h) - \beta_{0,h}B_0}{1-c_{1,h}}\right] dG_h(x)\right\} \\
+ \frac{\alpha(\lambda_{0,h}-\lambda_{1,h})}{\rho+\eta_h}\left\{\int_{\tilde{x}} \left[x - \frac{(1+\tau)(\rho Z(0,h) - \beta_{1,h}b_1)}{1+\beta_1\gamma\tau} dG_h(x)\right]\right\}.
\end{aligned}
\tag{15}
$$

Individuals are characterized by an individual-specific cost $\kappa \sim K(\kappa)$ which summarizes any monetary cost, ability-related cost and preference-related cost in acquiring a schooling level $h = 1$ with respect to a schooling level $h = 0$. The discounted equilibrium value of unemployment corresponds by definition to the present discounted value of participating in a labor market characterized by $\langle \Gamma_h, Y_h(\cdot), G_h(\cdot)\rangle$. This is the value individuals will look at ex-ante when deciding the optimal level of schooling. If $Z(ys,0) > Z(ys,1)$ the decision is trivial and no one will acquire education level $h = 1$. If $Z(ys,0) < Z(ys,1)$ then there will exist a reservation cost value

$$
\kappa^* = \left[Z(0,1)Y_1(y^*(1)) + \int_{y^*(1)} Z(y,1)dY_1(y)\right] - \left[Z(0,0)Y_0(y^*(0)) + \int_{y^*(0)} Z(y,0)dY_0(y)\right]
\tag{16}
$$

such that a fraction $K(\kappa^*)$ of individuals will acquire education level $h = 1$.

3 Identification

We describe the identification of the model's parameters based on data containing the following information: accepted wages, self-employed earnings, unemployment and self-employment durations and an indicator of the legal status of the job.

A first set of parameters is fixed to assumed or calibrated values. This is the case for three parameters that we calibrate to fit the specific institutional setting of the Mexican labor market: $\gamma = 0.55, B_0 = 1.27, \tau = 0.33$. But this is also the case for a parameter we are not able to identify since we do not have demand side information:[6] the bargaining parameter α. In our benchmark specification we assume symmetric Nash bargaining for both schooling levels, fixing $\alpha_0 = \alpha_1 = \frac{1}{2}$. In robustness exercises, we show the sensitivity of the estimates to this assumption.

The identification of a second set of parameters is standard in the search literature, dating back to the contribution of Flinn and Heckman (1982). They show that a parametric assumption is necessary on the exogenous match specific productivity distribution $G_h(x)$ which is responsible for the ex-post wage dispersion. Following previous literature, we assume it to belong to a two parameters lognormal distribution. We denote the location and scale parameters as (μ_h, σ_h). The lognormal distribution possesses the recoverability condition necessary for identification and guarantees a good fit of accepted wages data. Dispersion and location of observed accepted data together with the knowledge of the truncation point allows to identify wage offers distributions. Wage offers distributions are then mapped into the match-specific distributions by inverting the one-to-one mapping implied by Nash bargaining (see equation 8).

Knowledge of $G_h(x)$ and the truncation point allow to separate the probability of accepting an employee job offer in the exogenous arrival rate component and in the acceptance probability component. The additional information provided by the unemployment and self-employment durations is then enough to identify the arrival rates $\lambda_{0,h}$ and $\lambda_{1,h}$. Termination rates η_h are identified by exploiting the equilibrium rate of unemployment implied by the model and the unemployment rate observed in the data.

The parameters ξ_h and ρ can only be jointly identified using the equilibrium equation (14). Our strategy will be to fix a value for the discount rate ρ and to use equation (14) to recover ξ_h. Notice that we will not allow ρ to be schooling-specific.

The identification of a third set of parameters is discussed in more detail because it is original to our model and exploits specific features of the ordering and overlapping of accepted wage under a legal or illegal labor contract. The set of parameters we are referring to are the preference parameters β_0 and β_1 and the cost of signing illegal job contracts described by the parameters c_0 and c_1. The specific empirical features we are exploiting are: (i) the first order stochastic dominance of the legal contracts accepted wage distribution with respect to the illegal contracts accepted wage distribution; and (ii) the substantial overlap between the two distributions. Since employers and job seekers bargain over both the wage and the legal status of the job, workers' valuations of the relative extra-wage benefits not only affect equilibrium wages (8) but also determine the level of the match-specific productivity value at which both employees and firms are indifferent between a legal and an illegal job (10). In equilibrium, the cost of providing a legal job is lower for highly productive

[6] See Flinn (2006) for a formal discussion.

matches and hence only relatively lower productivity matches will be associated with illegal jobs. Hence, the resulting reservation wage of working as a legal employee embeds a positive premium which compensates the worker for not having an illegal job that depends on workers' valuation over the extra-wage benefits implied by the two job contracts and the sensitivity of the expected penalty for evading with respect to the match-specific productivity value. This optimal reservation rule sorts accepted wages with legal contracts in the right tail of the productivity distribution and accepted wages with illegal contracts in the left tail, generating the first order stochastic dominance observed in the data.

At the same time, the wage at the reservation productivity value generating indifference between the two contracts is higher if the worker chooses an illegal contract than if she chooses a legal one. This fact generates an overlap in the two accepted wage distributions which is larger or smaller depending on the set of parameters β_0, β_1, c_0 and c_1. Notice that without the presence of this set of parameters, there would be no overlap between the two distributions of legal and illegal contracts accepted wages.[7] If the data feature guaranteeing the joint identification of the four parameters is clear, it is more difficult to tease out empirical features generating the separate identification of each of them. Setting all of them to zero, would generate perfect sorting but first order stochastic dominance of the illegal distribution over the legal. The different sorting in the data is then a decisive feature suggesting that at least some of the four parameters are different from zero. If only the preference parameters were set close to zero or only the cost of the marginal contract were set close to zero, we could still generate significant overlapping between the two distributions. Figure 3 illustrates this point. The top panel displays simulation results for the equilibrium wages resulting from an exogenous increase in the marginal willingness to pay for CSI (β_1). As a result, illegal job matches are accepted at lower productivity values and the associated wage distribution shrinks dramatically around its mean. In the limit in which individuals fully value CSI benefits ($\beta_0 = \beta_1 = 1$), no illegal matches are accepted in equilibrium. A very similar effect is produced by an increase in the proportional cost for evading the social security law (c_1), as shown in the bottom panel of Figure 3. Given the limit of this identification strategy, the only feature allowing for the separate identification of the four parameters is the shape and the extent of the overlap of the wage distributions associated to legal and illegal jobs. However, shape and extent of the overlap are also closely a function of our specific functional form assumptions. As a result, we are in the process of acquiring additional empirical information to reach a more convincing separate identification of preference parameters and cost of illegality parameters.

[7] to clarify the discussion, assume for the moment that there is no self-employment and therefore there is a unique reservation wage for each schooling level. The argument is generalizable to the case with self-employment: the difference is that the same argument is repeated for each level of self-employment income. The presence of self-employment may increase or decrease the overall overlap of the two distributions depending on parameters. Since we observe self-employment income, this result does not invalidate the identification argument.

Finally, the identification of a fourth set of parameters is again standard even if the features they are referring to are not. The primitive distribution of self-employment potential earnings $Y_h(y)$ can be identified from the earnings of the currently self-employed in the same way we have identified $G_h(x)$ from the wages of the currently employed: by assuming a recoverable distribution. We assume again a lognormal distribution and denotes its parameters with (μ_h^y, σ_h^y). The K distribution for the cost of acquiring schooling is assumed negative exponential with parameter δ, which is identified by the proportion of individuals in our sample who have completed secondary education. We are forced to use a one-parameter distribution for heterogeneity in schooling costs because we do not observe direct information on them. We only know, based on the model, that they generate a threshold-crossing impact: above a certain threshold the individual will not acquire additional education, below she will. The resulting equilibrium distribution can therefore identify only one parameter. The choice of a negative exponential distribution is arbitrary but has the advantage of providing analytical convenience.

4 Data and Estimation

For identification purposes, we need a data set reporting a vector of accepted wages (w) and an indicator vector for the legal status of the job (f - defined according to whether employees are registered or not in the social-security payrolls), self-employment earnings (y), unemployment and self-employment durations (t_0 and t_1) and information regarding the latest level of schooling attained (h).

The data is extracted from Mexico's national labor force survey (ENOE, by its spanish acronym) for the first quarter of the year 2013. We restrict the sample to nonagricultural male workers between the ages of 35 and 55 who reside in urban areas. We further drop those who did not complete primary schooling and those who completed college or a higher educational degree and split the resulting sample in two groups according to whether the worker has completed high school ($h = 1$) or not ($h = 0$). In order to obtain a more homogenous population of self-employed individuals, we drop those who report having paid employees (roughly 30%). Table 1 reports descriptive statistics for the final sample we use in the empirical analysis.

We estimate the model's parameters using a Simulated Method of Moments (SMM) procedure in which, for a given parameters vector, we simulate moments that we compare with the corresponding moments obtained from the data sample. We estimate $\lambda_{s,h}$ and η_h by matching two moments exactly: the mean durations of unemployment and self-employment spells as equal to the corresponding hazard rate, which, together with the unemployment rate, defines a system of three linear equations in three unknowns for each schooling sub-market.

In the preliminary version of the estimates, we normalize β_0 due to the weak separate identification of the preference parameters. We also estimate δ directly from the proportion of individuals acquiring the high schooling level. The remaining vector of parameters is estimated by Simulated Method of Moments (SMM) since standard regularity condition to perform maximum likelihood estimation do not hold. Formally, the set of parameter $\theta_h \equiv \{\mu_h, \sigma_h, \mu_h^y, \sigma_h^y, \beta_{1,h}, c_{0,h}, c_{1,h}, \xi_h, \lambda_{0,h}, \eta_h, \lambda_{1,h}\}$ and is estimated as

$$\theta = \underset{\theta}{\operatorname{argmin}}\ \Psi(\theta, t_0, t_1 w, y, f)' W \Psi(\theta, t_0, t_1, w, y, f) \tag{17}$$

$$\text{such that } \Psi(\theta, t_0, t_1, w, y, f) = \left[\Gamma_R(\theta | x^\star\widehat{(0; ys}, h), \widehat{c_{0,h}}, \widehat{\tilde{x}(h)}) - \Gamma_N(t_0, t_1, w, y, f)\right],$$

where Γ_N is the vector of the sample moments obtained by our sample of dimension N, and $\Gamma_R(\theta | x^\star\widehat{(0; ys}, h), \widehat{c_{0,h}}, \widehat{\tilde{x}(h)})$ is the vector of the corresponding moments obtained from a simulated sample of size R conditional on the estimated productivity reservation values of $x^\star(0; ys, h)$, $\widehat{c_{0,h}}$ and $\tilde{x}(h)$. The weighting matrix W is a diagonal matrix with elements equal to the inverse of the bootstrapped variances of the sample moments.

The moments we match are extracted from the unemployment and self-employment durations, from the accepted wages distributions at legal and illegal jobs and from the accepted self-employed earning distributions. For the durations, we simply compute the mean and the proportion of individuals in unemployment and self-employment. For the wage distributions, we want to exploit - on top of the first two moments of the accepted wages by schooling and legality status - that the distribution of legal and illegale employed workers overlap. We attain that by computing means and standard deviations of wages at legal and illegal jobs over various percentile ranges defined by accepted wages at illegal jobs. We use percentiles 20, 40, 60, 80 and 100 of the illegal workers' accepted wage distribution to define 5 intervals. Within these 5 intervals, we compute the proportion of workers holding legal jobs and the mean and standard deviations of wages with legal and illegal jobs.

5 Results (preliminary)

The model is estimated separately for individuals with a complete high-school degree and for individuals with incomplete high-school. The implicit assumption is that the labor market is segmented along observable workers' characteristics so that the two education groups do not compete for the same jobs.

Estimated parameters are reported in Table 2. The parameters governing the rates of job arrival and termination differ across schooling sub-markets, with higher arrival and lower termination rates for individuals with a complete high-school degree. Important differences between the two schooling groups are also observed in the parameters of the match specific productivity distributions. The preference parameters and the cost of illegality parameters, instead, do not differ by schooling group.

Table 3 reports some useful statistics implied by the estimated structural parameters. Two main results emerge. First, we can look at returns to formality as measured by wage offers received. Focusing on expected values, the estimates imply a much larger return to formality for workers with the higher education level completed. Second, we can look at returns to schooling corrected for selection in employment versus unemployment and self-employment. A first measure of this is given by average wage offers received by the two education groups. We find modest returns to schooling in the illegal labor contract but substantial returns to schooling in the legal labor contract. These preliminary results suggest that policies affecting the incentive to work legally or illegaly can potentially have big impacts on schooling decisions.

References

Abramitzky, R. and Lavy, V. (2011), How responsive is investment in schooling to changes in redistribution policies and in returns, Nber working papers, National Bureau of Economic Research, Inc.

Acemoglu, D. and Shimer, R. (1999), 'Holdups and efficiency with search frictions', *International Economic Review* **40**(4), 827–49.

Albrecht, J., Navarro, L. and Vroman, S. (2009), 'The effects of labour market policies in an economy with an informal sector', *Economic Journal* **119**(539), 1105–1129.

Banerjee, A. V., Cole, S., Duflo, E. and Linden, L. (2007), 'Remedying Education: Evidence from Two Randomized Experiments in India', *The Quarterly Journal of Economics* **122**(3), 1235–1264.

Bosch, M. and Esteban-Pretel, J. (2012), 'Job creation and job destruction in the presence of informal markets', *Journal of Development Economics* **98**(2), 270–286.

Burke, K. and Beegle, K. (2004), 'Why Children Aren't Attending School: The Case of Northwestern Tanzania', *Journal of African Economies* **13**(2), 333–355.

Duflo, E. (2001), 'Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment', *American Economic Review* **91**(4), 795–813.

Eckstein, Z. and van den Berg, G. J. (2007), 'Empirical labor search: A survey', *Journal of Econometrics* **136**(2), 531–564.

Eckstein, Z. and Wolpin, K. I. (1995), 'Duration to first job and the return to schooling: Estimates from a search-matching model', *Review of Economic Studies* **62**(2), 263–86.

Filmer, D. (2007), 'If you build it, will they come? School availability and school enrollment in 21 poor countries', *Journal of Development Studies* **43**(5), 901–928.

Flabbi, L. and Tejada, M. (2012), Gender gaps in education and labor market outcomes in the u.s.: The impact of employers' prejudice, IDB Working Paper Series 357, IDB and Georgetown University.

Flinn, C. (2006), 'Minimum wage effects on labor market outcomes under search, bargaining and endogenous contact rates', *Econometrica* **73**, 1013–1062.

Flinn, C. and Heckman, J. (1982), 'New methods for analyzing structural models of labor force dynamics', *Journal of Econometrics* **18**(1), 115 – 168.

Flinn, C. and Mullins, J. (2015), 'Labor Market Search and Schooling Investment', *International Economic Review* **56**, 359–398.

Jensen, R. (2012), 'Do labor market opportunities affect young women's work and family decisions? experimental evidence from india', *The Quarterly Journal of Economics*

127(2), 753–792.

Kazianga, H., Levy, D., Linden, L. L. and Sloan, M. (2013), 'The Effects of Girl-Friendly; Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso', *American Economic Journal: Applied Economics* **5**(3), 41–62.

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

Mortensen, D. (2003), *Wage Dispersion: Why are similar workers paid differently?*, Zeuthen Lecture Book Series, Cambridge, MA, MIT Press.

Pissarides, C. A. (2000), *Equilibrium Unemployment Theory*, Cambridge, MA, MIT Press.

Robalino, D. A., Zylberstajn, E. and Robalino, J. D. (2011), Incentive effects of risk pooling, redistributive and savings arrangements in unemployment benefit systems: Evidence from a job-search model for brazil, IZA Discussion Papers 5476, Institute for the Study of Labor (IZA).

Rogerson, R., Shimer, R. and Wright, R. (2005), 'Search-theoretic models of the labor market: A survey', *Journal of Economic Literature* **43**(4), 959–988.

Tejada (2012), Dual labor markets and labor protection in an estimated search model with matching and bargaining, mimeo, Georgetown University.

Figures and Tables

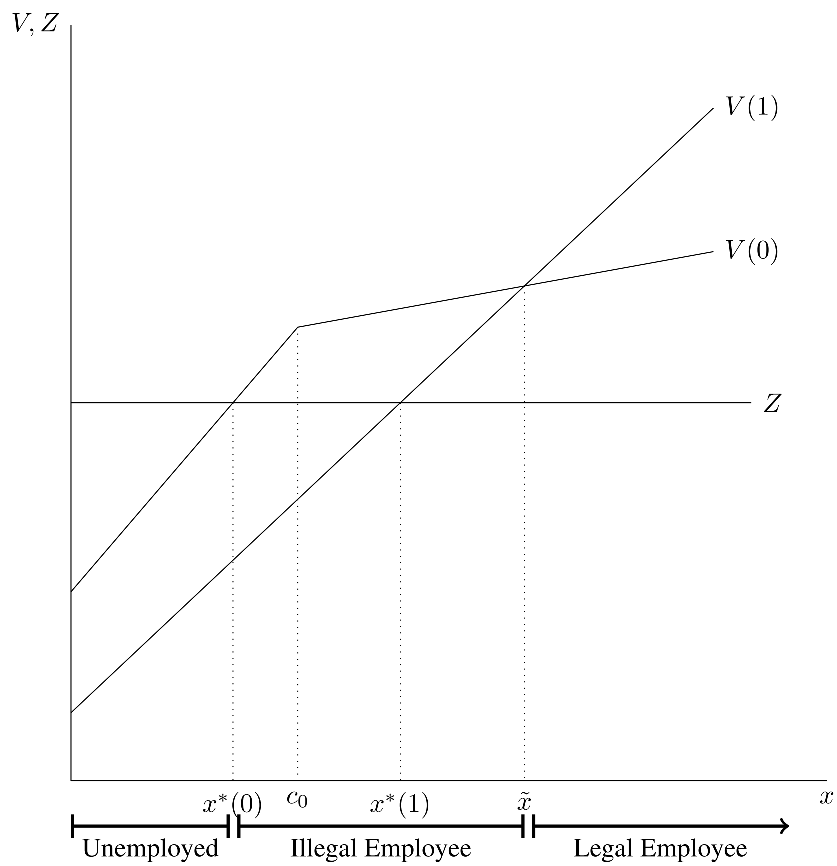


Figure 1. Equilibrium Without Self Employment

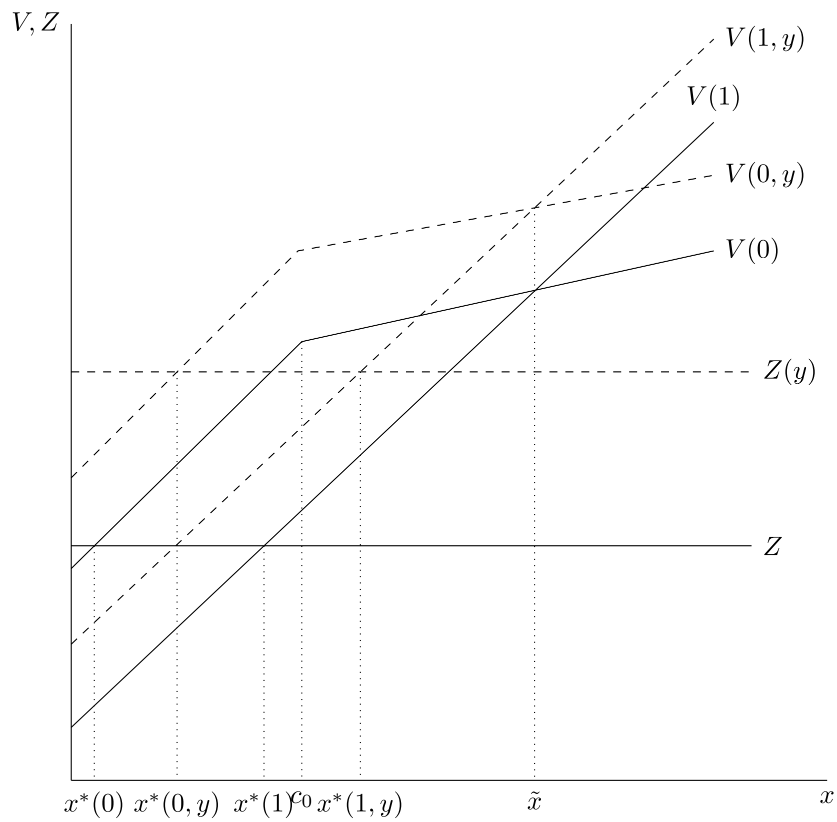
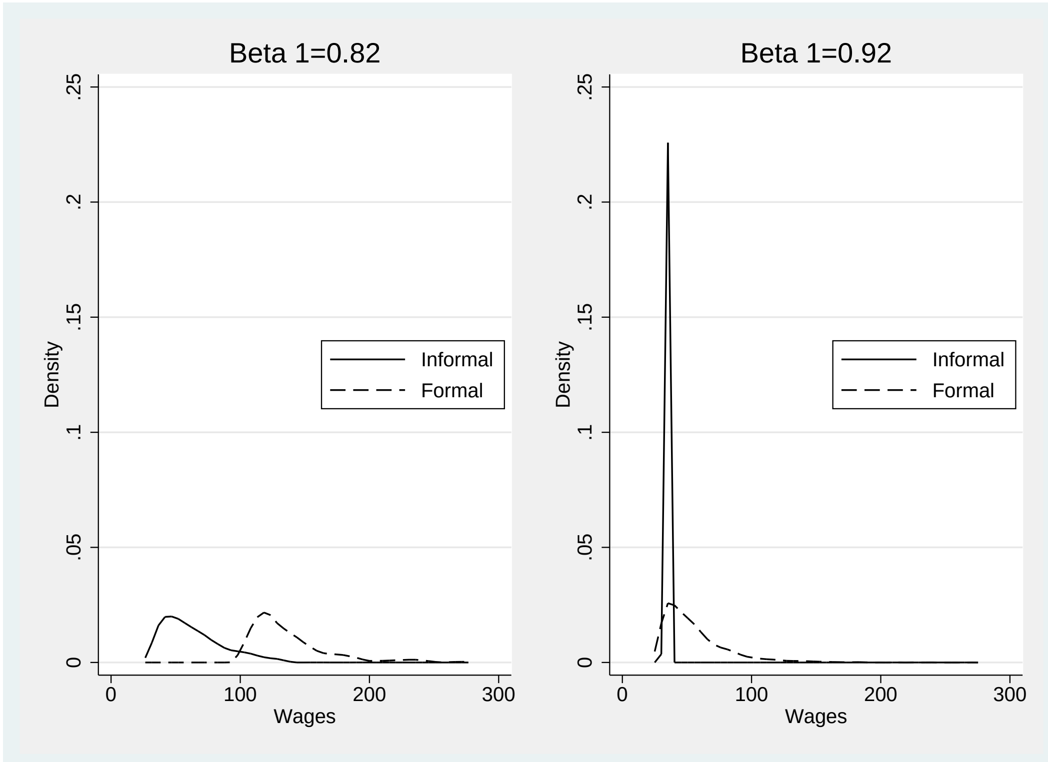
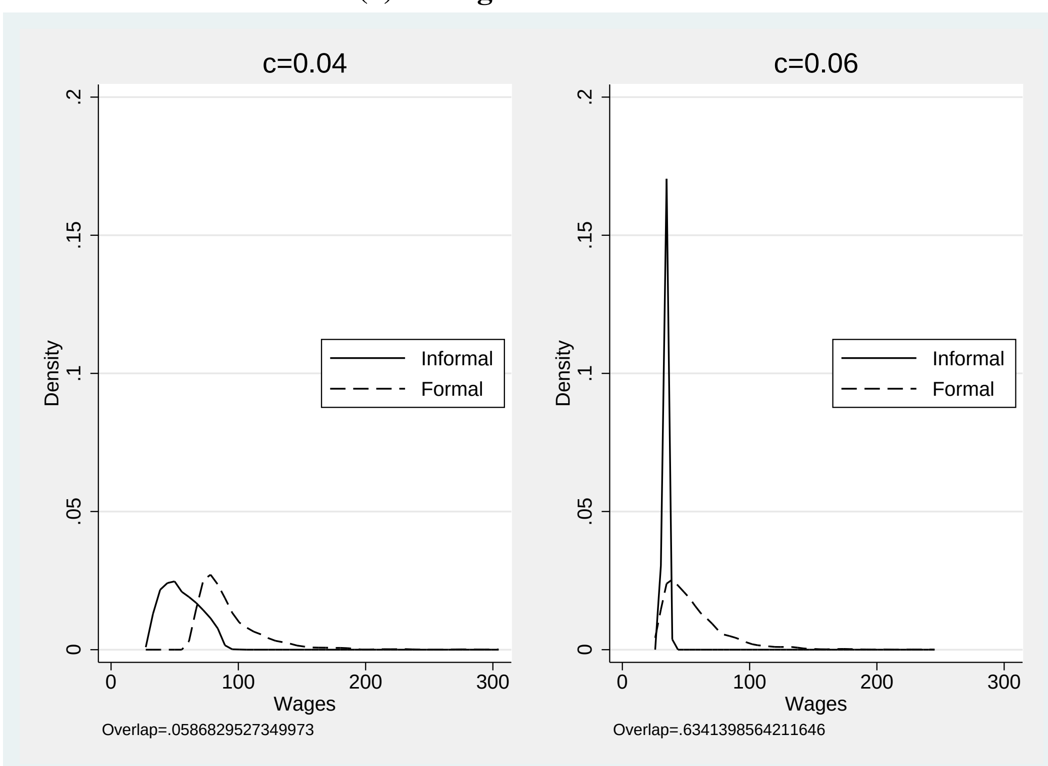


Figure 2. Equilibrium With Self Employment



(a) Changes in Beta 1



(b) Changes in c

Figure 3. Simulated Wage Distributions

Table 1. Descriptive Statistics - ENOE 2013:Q1

Education group	Incomplete Secondary			Complete Secondary		
Labor market state	Formal	Informal	Self-Empl.	Formal	Informal	Self-Empl.
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Hourly Wage	32.42	24.39	31.10	42.68	29.11	38.40
SD Hourly Wage	16.89	13.47	19.46	26.25	21.68	30.67
Employment Rates	0.46	0.29	0.20	0.59	0.17	0.18
Mean Unempl. Duration (months)		2.23			2.72	
Mean Self-Empl. Duration (years)		12.46			11.16	
Workers with Complete Secondary	29%					
Number of Observations	5390	3400	2436	2956	866	895

NOTE: The definition of earnings in the publicly available version of the ENOE refers to monthly ”equivalent” earnings from the main job after taxes and Social Security contributions, including overtime premia and bonuses. For those paid by the week, the survey transforms weekly earnings into monthly earnings by multiplying the former by 4.3. Similar adjustments are used for workers paid by the day or every two weeks. Wage and self-employed earning distributions are trimmed at the top and bottom 0.5 percentile.

Table 2. Estimated Parameters

	High Schooling	**Low Schooling**
λ_1	0.318	0.188
	(0.049)	(0.038)
λ_0	0.303	0.296
	(0.075)	(0.060)
η	0.040	0.061
	(0.014)	(0.020)
μ	3.916	3.516
	(0.214)	(0.197)
μ_y	2.396	2.235
	(0.252)	(0.248)
σ	1.346	1.354
	(0.096)	(0.090)
σ_y	1.387	1.321
	(0.052)	(0.074)
β_0	1.000	1.000
	(-)	(-)
β_1	0.827	0.823
	(0.037)	(0.037)
c_0	37.266	27.793
	(8.780)	(10.995)
c_1	0.060	0.061
	(0.009)	(0.009)
ξ	-2.711	-2.780
	(0.517)	(0.658)

Bootstraped Mean and Standard Errors from 100 samples in parentheses

Table 3. Reservation and Implied Values

	Complete Secondary	Incomplete Secondary
Panel A	*Reservation Values*	
$x^*(0)$	25.71	15.69
$x^*(1)$	28.19	17.73
$\hat{x}$	37.26	27.42
$\tilde{x}$	73.18	73.35
y^*	10.42	9.62
κ^*	4.51	
Panel B	*Implied Values*	
Productivity, Average	124.14	84.18
Productivity, SD	280.84	193.01
Offered Wage Informal, Average	23.98	19.49
Offered Wage Informal, SD	150.09	102.08
Accepted Wages Informal, Average	35.51	34.07
Accepted Wages Informal, SD	6.94	6.36
Offered Wage Formal, Average	50.58	31.41
Offered Wage Formal, SD	114.29	79.09
Accepted Wages Formal, Average	69.44	62.09
Accepted Wages Formal, SD	42.15	30.10
Wages Self Employed, Average	51.30	39.04
Wages Self Employed, SD	38.01	23.03
Unemployment Duration(months)	10.84	15.47
Self Employment Duration(months)	106.77	152.69
% Unemployed	0.12	0.20
% Self Employed	0.12	0.13
% Formal	0.33	0.23
% Informal	0.43	0.44
% Informal Below $\hat{x}$	0.28	0.37
% Informal Above $\hat{x}$	0.72	0.63