The Role of the Informal Sector in the Early Careers of Less-Educated Workers

Javier Cano Urbina*
University of Western Ontario

November 4, 2011

Abstract

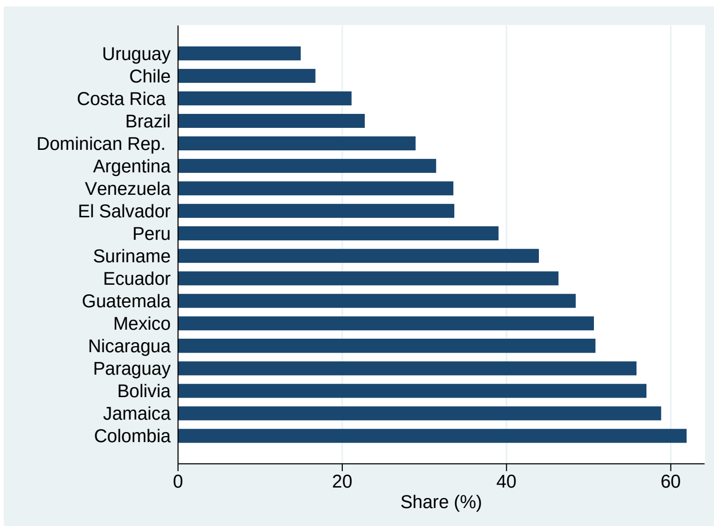
Does work experience gained in the informal sector affect the career prospects of less-educated workers? This paper examines two roles that informal sector jobs play in the early stages of a worker's career: informal jobs may (*i*) provide the opportunity to accumulate skills, and (*ii*) act as a screening device that enables employers to learn a worker's ability. This paper develops a matching model of the informal and formal sectors that can accommodate both roles. Implied hazard rates from informal to formal sectors as a function of tenure are shown to differ depending on whether the dominant role is human capital accumulation or screening. Using the ENOE, a longitudinal employment survey from Mexico, hazard functions are estimated for less-educated workers. The estimated hazard functions suggest that screening plays a more important role in the informal sector than does skill formation in the early stages of a worker's career. The estimation results also imply that employers would only learn the ability of 14% of their workers after one month of employment. This finding suggests that employers' capacity to select workers is limited in government employment programs requiring employers to provide permanent positions to a predetermined fraction of workers after a short period of time.

*Financial support from CONACYT was greatly appreciated. I thank Audra Bowlus and Lance Lochner for their support, guidance, and suggestions throughout this project, and I thank Youngki Shin for his suggestions in the empirical implementation. I also thank Carlo Alcaraz, Daniel Montanera and Deanna Walker for useful comments, and Aldo Colussi for inspiring talks early on in this project. All remaining errors are mine.

1 Introduction

The informal sector is an important feature of labor markets in developing countries. This sector, composed of all jobs not complying with labor regulations, occupies a significant portion of these countries' labor markets. In Latin America and the Caribbean, the fraction of workers employed in the informal sector ranges from 15% to 62% (see Figure 1). Jobs in this sector employ the majority of young unskilled workers usually paying very low wages, not to mention the lack of health and employment insurance enjoyed by workers holding formal sector jobs.

Figure 1: Share of Salaried Workers in Informal Jobs in Latin America and the Caribbean



Source: Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank). Data obtained in the fall of 2010. Males and females ages 25-64 in urban areas. Varying years. A worker is considered informal if (s)he does not have the right to a pension when retired.

The presence of large informal sectors has typically been a concern for researchers and policymakers. Some are concerned that the informal sector could be the disadvantaged sector in a segmented labor market market (Magnac, 1991; Maloney, 1999; Farrell, 2004; Amaral and Quintin, 2006; Arias and Khamis, 2008). Others are concerned that the informal sector might adversely affect productivity and growth (Loayza, 1996; Schneider and Enste, 2000; Levy, 2007; Fajnzylber, 2007). Whether these concerns are supported by the evidence is still unresolved. However, they have induced policymakers to introduce tighter regulations to reduce or control the size of the informal sector.

Before attempting to restrict the informal sector, it is important to investigate the potential benefits that workers obtain during informal sector employment. Previous studies have found that less-educated workers start their working careers in salaried jobs in the informal sector and move into formal jobs as they grow older (Maloney, 1999; Arias and Maloney, 2007). We would like to know if informal sector jobs provide some value above and beyond make-shift low-paying work while people wait to find a "good" formal sector job: do these jobs also provide skills or help screen workers to facilitate a transition to higher paying formal sector jobs? If rules designed to reduce the informal sector are implemented, would we lose some valuable worker training or screening? If so, restrictions on informal sector employment should be accompanied by policies that replace the productive functions of these jobs.

We investigate two potential roles that informal sector jobs could play in the early stages of a worker's career. First, these jobs may provide the opportunity to accumulate skills, making workers more productive and more attractive to formal sector employers. While more-educated workers tend to access greater training opportunities in formal sector employment, less-educated workers may turn to the informal sector to gain work skills.[1] Second, informal sector jobs may serve as a screening device that enables employers to learn a worker's ability. The lack of compliance with labor regulations, especially firing costs and severance payments, suggests that informal sector employers may be more prone to hire young unskilled workers entering the labor market than are formal sector employers. Hence, an informal sector worker who reveals that he is productive may increase his likelihood of finding a formal sector job.

The role of the informal sector as a provider of training opportunities was first suggested by Hemmer and Mannel (1989) and has been advocated by Maloney (1999) and Arias and Maloney (2007). The role of the informal sector as a screening device is rarely discussed. One exception is Arias and Maloney (2007) who argue that labor regulations and information asymmetries "impede young workers' entry into the formal sector."[2] The study presented here contributes to this literature by providing an analytical framework and empirical evidence about these roles of the informal sector.

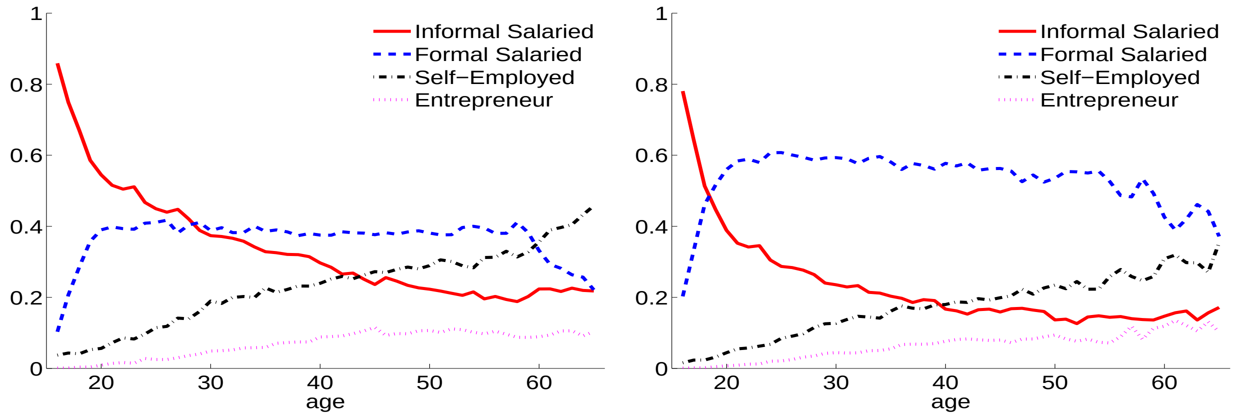
To determine the relative importance of the training or screening roles of the informal sector, we develop a two-sector matching model to study worker movements from the informal to the formal sector. The model is designed to better understand the labor market dynamics in Mexico, a country with a significant informal labor market. In Mexico, the informal sector is a port of entry to the labor market for less-educated workers. These workers are concentrated in the informal sector in the early stages of their working careers, moving to the formal sector as they age (see Figure 2). Figure 3 shows that the probability of moving

[1]Barron et al. (1997) find that more educated workers in the U.S. have grater access to on-the-job training (see Table 4.2). We believe that formal sector employers have a closer resemblance to employers in the U.S. than informal sector employers.

[2]Bosch (2006) and Bosch et al. (2007) present evidence that labor regulations affect the patterns of job creation in the formal sector in economies with large informal sectors. Some argue that these regulations disproportionately affect the youth (World Bank, 2007, chap. 4).

Figure 2: Distribution of Workers by Employment Sector in Mexico

(a) Years of Education: [0,9) (b) Years of Education: [9,12)



Source: Author's calculations using ENOE I:2005 - IV:2010. A worker is considered informal if he is not enrolled in government health care program. Males not attending school.

from the informal to the formal sector increases during the early stages of workers' careers.

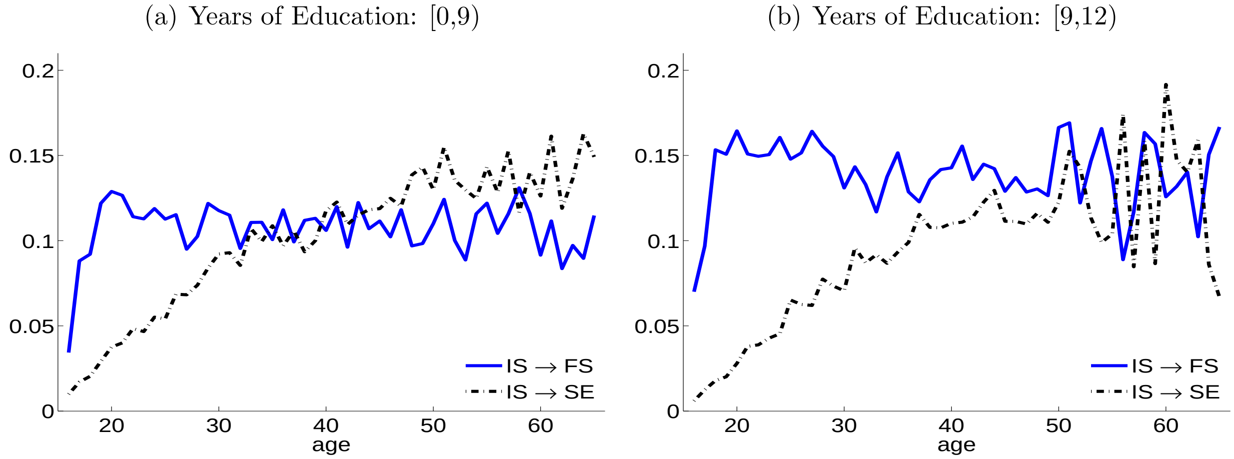
The empirical analysis is based on the analytical implications for hazard rates from the informal to the formal sectors derived from the model. It is shown that hazard rates from informal to formal sectors as a function of tenure differ depending on the relative importance of human capital accumulation or screening. On the one hand, if workers accumulate human capital while working in the informal sector, the likelihood of moving into the formal sector increases with informal sector tenure. On the other hand, if workers' productivities are screened while working in the informal sector, those discovered as highly productive move faster to the formal sector, leaving behind those with low productivity who have difficulties to access formal sector jobs. Thus, the likelihood of moving into the formal sector decreases with informal sector tenure.

Using an employment survey from Mexico to obtain measures of duration of employment in the informal sector, we estimate the hazard functions and test the two hypotheses. The results indicate that in the early stages of a worker's career, screening plays a more important role in the informal sector than does skill formation.

Our results give us the means to evaluate one stream of the *Bécate* training program for the unemployed in Mexico, which is targeted at less-educated youth.[3] One of the streams of *Bécate* is a mixture of skill formation and worker placement. In this stream, training takes place at the workplace, and the hosting firm must have empty vacancies that it is looking to fill. The training program lasts for one to three months. At the end of the training program,

[3] *Bécate* was launched in 1984 and was designed to assist individuals with less than 9 years of education between the ages of 16 and 30. Currently, the program has more streams to assist a broader set of workers and needs. Delajara et al. (2006) provides a comprehensive evaluation of the program.

Figure 3: Transitions Out of the Informal Sector in Mexico



Source: Author's calculations using ENOE I:2005 - IV:2010. Number of transitions relative to the size of the informal sector. A worker is considered informal if he is not enrolled in government health care program. Males not attending school. IS = Informal Salaried, FS = Formal Salaried, SE = Self-Employed.

the firm is committed to hire at least 70% of the participants.[4] Given this short amount of time, it seems likely that the program works more as a screening device than a source of significant skill formation.

Based on the estimated hazard, we can deduce the rate at which an employer learns about a worker's ability. For workers with less than 12 years of education, the estimates indicate that an employer learns about a worker's ability at a rate of 14% per month. Consequently, if an employer commits to hire 70% of the program participants, a one or two month program requires the employer to take a gamble on a considerable portion of the program participants, since the employer must bear the firing costs of terminating any unsuitable workers. This highlights the importance of better understanding the role of the informal sector in the design of policy.

The study is organized as follows. In Section 2, we present the baseline model and its implications for hazard rates from the informal to the formal sector. In Section 3, we present models with human capital accumulation and with employer learning, deriving their implications for hazard rates. Once the theoretical implications are described, in Section 4 we describe the data used in the empirical analysis. The details of the estimation follow in Section 5. Section 6 summarizes the empirical results, and Section 7 concludes with some remarks on the results and suggestions for future research.

[4]In this stream of the program, the firm can participate in the selection and recruitment of workers participating in the program.

2 Baseline Model

The labor market is composed of two sectors, a formal sector and an informal sector. Formal sector firms comply with labor regulations represented by a firing cost incurred by firms when jobs are destroyed. The firing cost is assumed to be a wasteful tax as in Mortensen and Pissarides (2003) and Dolado et al. (2005), so no transfer to the worker takes place. Informal sector firms do not comply with labor regulations.

We follow Albrecht et al. (2006, 2009) by assuming that workers differ in their productivity in the formal sector, but they are equally productive in the informal sector. Workers in the formal sector produce px units per period, where $p \in \{p_L, p_H\}$, with $p_H > p_L$, and x is a measure of *match quality*. Match quality is a random draw from a known distribution $G(x)$ with support on $[0, 1]$ that is made when the worker and firm meet; match quality stays constant until the job is destroyed. A fraction ϕ of the workers have the innate productivity p_L in the formal sector; we refer to these workers as L-skilled and the others as H-skilled. Innate productivity is perfectly observable. All workers in the informal sector produce p_I units per period. It is assumed that $p_I \geq z$, where z is the flow utility in unemployment.

Job destruction in both sectors follows from an idiosyncratic shock that arrives to occupied jobs at Poisson rate δ. If the job is destroyed in the formal sector, the firm incurs a firing cost D. Jobs are also destroyed due to worker's death. A worker dies with probability τ regardless of the worker's employment status. Every dead worker is replaced by a new unemployed worker who is L-skilled with probability ϕ. Job destructions due to death do not generate firing costs.

Unemployed workers search for jobs in both sectors, and all informal sector workers search for jobs in the formal sector.[5] The number of meetings between workers and firms in the informal sector is $m(u, v_I)$ and $m(u + e_I, v_F)$ in the formal sector, where u and e_I are the number of workers in unemployment and in informal sector jobs, respectively, v_j is the number of open vacancies in sector $j \in \{F, I\}$, and $m(\cdot, \cdot)$ is the meeting function. The meeting function is homogeneous of degree one, concave and increasing in both its arguments. As a result, a job seeker meets a firm in sector $j \in \{F, I\}$ with probability $m(\theta_j) = m(1, \theta_j)$, and a firm in sector j meets a job seeker with probability $m(\theta_j)/\theta_j$, where $\theta_I = v_I/u$ and $\theta_F = v_F/(u + e_I)$ are the measures of market tightness in the informal and the formal labor markets, respectively.

Given the assumptions on productivity in the informal sector, all meetings between an informal sector firm and an unemployed worker lead to job creation. Due to firing costs and

[5] To focus on flows from the informal to the formal sector, we abstract from on-the-job search in the opposite direction and from on-the-job search within each sector.

to the assumptions on productivity in the formal sector, a job in this sector is created if and only if the match quality is higher than a reservation match quality. The reservation match quality is endogenous and depends on both the skill level and the current employment status of the worker.

The payoffs for workers are:

$$(1) \quad \tilde{r}U(p) = z + m(\theta_I)\big[W_I(p) - U(p)\big] + m(\theta_F)\int_{C(p)}^{1}\big[W_F(s,p) - U(p)\big]dG(s)$$

$$(2) \quad \tilde{r}W_F(x,p) = w_F(x,p) + \delta\big[U(p) - W_F(x,p)\big]$$

$$(3) \quad \tilde{r}W_I(p) = w_I(p) + \delta\big[U(p) - W_I(p)\big] + m(\theta_F)\int_{Q(p)}^{1}\big[W_F(s,p) - W_I(p)\big]dG(s)$$

where $\tilde{r} \equiv r + \tau$ and r is the discount rate. $U(p)$, $W_F(x,p)$, and $W_I(p)$ denote the present discounted value of the expected income stream of an unemployed worker, a worker employed in the formal sector, and a worker employed in the informal sector, respectively. Employed workers earn wage $w_I(p)$ or $w_F(x,p)$ when they work in the informal or the formal sector, respectively. The reservation match quality for the unemployed is $C(p)$ and for informal sector workers is $Q(p)$.

The payoffs for firms are:

$$(4) \quad \tilde{r}J_F(x,p) = px - w_F(x,p) + \delta\Big[V_F - D - J_F(x,p)\Big] + \tau V_F$$

$$(5) \quad \tilde{r}J_I(p) = p_I - w_I(p) + \big[\delta + \mu(p)\big]\Big[V_I - J_I(p)\Big] + \tau V_I$$

$$(6) \quad rV_F = -k_F + \frac{m(\theta_F)}{\theta_F}\Big(E_{X,P}\big[J_F(x,p)|\phi_U,\phi_I\big] - V_F\Big)$$

$$(7) \quad rV_I = -k_I + \frac{m(\theta_I)}{\theta_I}\Big(E_P\big[J_I(p)|\phi_U\big] - V_I\Big)$$

where $\mu(p) \equiv m(\theta_F)\big[1 - G\big(Q(p)\big)\big]$, $J_F(x,p)$, and $J_I(p)$ denote the present discounted value of the expected profit from an occupied job in the formal and the informal sector, respectively, and V_j denotes the present discounted value of expected profit from a vacant job in sector $j \in \{F, I\}$. Note that (4) incorporates firing costs, (5) incorporates the possibility that the worker moves to the formal sector, and that the value of an open vacancy depends on the recruitment costs, k_j, and on the fraction of low-skilled job seekers, given by ϕ_U in unemployment and ϕ_I in the informal sector.

Wages in both sectors are determined according to a surplus sharing rule that entitles workers to a fraction β of the match surplus. The match surplus in the informal sector is $S_I(p) = W_I(p) - U(p) + J_I(p) - V_I$, and in the formal sector is given by $S_F(x,p) =$

$W_F(x,p) - U(p) + J_F(x,p) - V_F$. The resulting wages are presented in Appendix A.

The decision to create a job in the formal sector depends on the match quality drawn when the worker and the firm meet. If the firm meets with an unemployed worker, both the firm and the worker require $x \geq C(p)$ to match, where $C(p)$ is such that $S_F(C(p), p) = 0$ for $p \in \{p_L, p_H\}$. If the firm meets with a worker in the informal sector, they require $x \geq Q(p)$, where $Q(p)$ is such that $S_F(Q(p), p) = S_I(p)$ for $p \in \{p_L, p_H\}$. Using the payoffs and wages, these cut-offs are given by:

$$(8) \quad C(p) = \frac{\tilde{r}U(p)}{p} + \frac{\delta D}{p}$$

$$(9) \quad Q(p) = C(p) + \frac{(\tilde{r} + \delta)S_I(p)}{p}$$

where $p \in \{p_H, p_L\}$. Note that from (8) and (9) we cannot determine if $C(p_H) < C(p_L)$ and $Q(p_H) < Q(p_L)$ without some assumptions on productivity levels in the formal and informal sectors. Lemma 1 provides a sufficient condition that enables us to determine the relative size of the cut-offs.

Lemma 1. *If $(1 - \beta)(p_I - z) < (r + \tau + \delta)p_L$, then $C(p_H) < C(p_L)$ and $Q(p_H) < Q(p_L)$.*

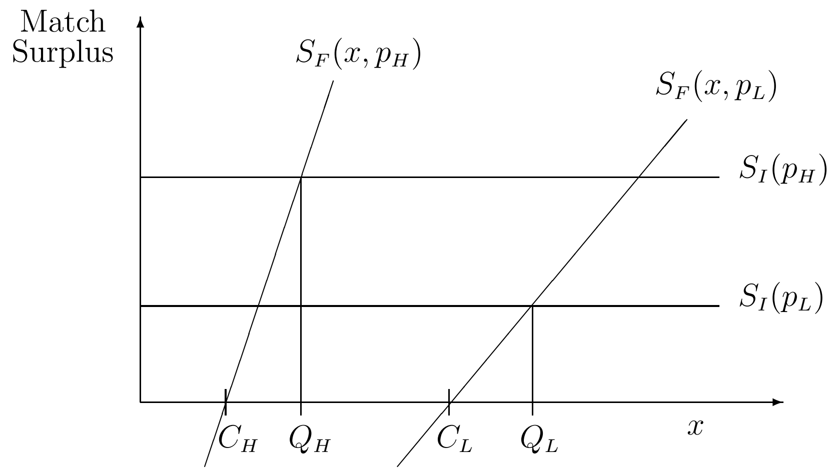
Appendix B.1 presents the proof of Lemma 1.[6] Lemma 1 provides a lower bound on the productivity of formal sector jobs that ensures that $C(p_H) < C(p_L)$ and $Q(p_H) < Q(p_L)$. The lower bound depends on the discounted gain in the flow surplus generated when an unemployed worker finds a job in the informal sector. Since formal sector jobs are harder to find than informal sector jobs, the productivity level in the formal sector must be high enough to make it worth while.

After substituting wages and cut-offs in the match surplus in the formal sector, we find that $S_F(x,p) = \frac{p}{\tilde{r} + \delta}(x - C(p))$. Then, given the result in Lemma 1 and that $p_H > p_L$, it follows that $\forall x \in [0,1]$, $S_F(x, p_H) > S_F(x, p_L)$ and $\partial S_F(x, p_H)/\partial x > \partial S_F(x, p_L)/\partial x$. Figure 4 illustrates this result, and the fact that $C(p_H) < C(p_L)$ and $Q(p_H) < Q(p_L)$. Note that $S_I(p) > 0$ implies that $Q(p) > C(p)$ for $p \in \{p_L, p_H\}$, as a consequence informal sector workers are more selective than unemployed workers when it comes to matching with a formal sector firm.

The baseline model produces implications for the hazard rate from the informal to the formal sector. We distinguish between the hazard rate conditional on worker skill level,

[6] A closer look into the proof of Lemma 1 reveals that these inequalities hold under milder conditions, but these conditions involve the measures of labor market tightness which are endogenous variables in the model. Numerical exercises suggest that for a wide range of parameter values $C'(p) < 0$ and $Q'(p) < 0$ even when the condition in Lemma 1 is violated.

Figure 4: Reservation Match Quality for Employed and Unemployed Workers



NOTE: $C_H = C(p_H)$, $C_L = C(p_L)$, and $Q_H = Q(p_H)$, $Q_L = Q(p_L)$

denoted $\lambda(t|p)$, and the unconditional (or average) hazard rate, denoted $\lambda(t)$; where t is the realization of a random variable $T \geq 0$ measuring duration in the informal sector and $p \in \{p_L, p_H\}$. These results are summarized in Propositions 2 and 3.

Proposition 2. *Suppose that the condition in Lemma 1 holds. Then, in the baseline model, the hazard rate from the informal to the formal sector conditional on the worker skill level, $\lambda(t|p)$, is constant for each $p \in \{p_L, p_H\}$, and it is higher for H-skilled workers than for L-skilled workers.*

Proof. In the baseline model, the hazard rate conditional on worker skill is given by $\lambda(t|p) = \mu(p) = m(\theta_F)\big[1 - G(Q(p))\big]$, so that $\partial\lambda(t|p)/\partial t = 0$. By Lemma 1, $Q(p_H) < Q(p_L)$, which implies that $\lambda(t|p_H) > \lambda(t|p_L)$. $\square$

Proposition 3. *Suppose that the condition in Lemma 1 holds. Then, in the baseline model, the unconditional hazard rate, $\lambda(t)$, is decreasing in duration.*

The proof of Proposition 3 follows the arguments of Lancaster (1990) and is presented in Appendix B.2. In this model, the fraction of L-skilled workers in the *risk set* (i.e. those that have not left the informal sector yet) increases with duration, pushing down the average hazard rate. This fraction increases with duration because H-skilled workers move from the informal to the formal sector at a faster rate than L-skilled workers. Lancaster (1990) calls this a "selection effect."

3 Extensions to the Baseline Model

The baseline model provides an analytical framework that helps us understand the key factors underlying the transitions from the informal to the formal sector. However, this model predicts that the transition rates from the informal to the formal sector remain constant as workers age. Yet, as shown in Figures 2 and 3, this is not the case in the data. Instead, we observe that transition rates increase as workers age (during early stages of the workers' careers).

We consider two extensions to the baseline model intended to explain this feature in the data. First, we assume that workers can accumulate human capital while working, which increases the chance of finding a formal sector job. Second, we assume that employers gradually learn about workers' skills. As a result, workers who are found to be H-skilled increase their chances of finding a formal sector job. We implement each extension separately because, as shown below, each mechanism generates opposing implications that would be hard to disentangle in a model with both mechanisms.

We focus on the implications for the hazard rate from the informal to the formal sector. On the one hand, when we assume that a worker can become more productive while in the informal sector, the longer such a worker stays in this sector, the more likely he is to make a transition into the formal sector. On the other hand, when we assume that a worker's productivity is gradually learned, those discovered as highly productive move to the formal sector faster, leaving behind those with low productivity levels and hence greater difficulties to access formal sector jobs. Thus, the longer a worker stays in the informal sector, the lower the likelihood that he makes a transition to the formal sector.

3.1 Human Capital Accumulation

First, we extend the baseline model by adding the possibility that workers accumulate skills through learning-by-doing. We follow Rebière (2008) and assume that a L-skilled worker can accumulate skills and become H-skilled with probability κ.[7] The accumulation of skills can only take place on the job, so the unemployed L-skilled workers cannot become H-skilled. Human capital does not depreciate, but since workers die and are replaced, the model does not converge to a degenerate distribution of skills.

The payoffs for unemployed workers and for vacancies have the same formulation as in the baseline model. The payoffs for employed workers and for filled vacancies now incorporate

[7] In Rebière (2008), workers start as beginners and become experienced while working in the beginners' sub-market; once they are experienced they search for jobs in the experienced sub-market. The labor market is segmented, so only beginners search for jobs in the beginners' sub-market, and only experienced search for jobs in the experienced sub-market.

the possibility of accumulating skills. These are given by:

$$(10)\quad \tilde{r}W_F(x,p) = w_F(x,p) + \delta\big[U(p) - W_F(x,p)\big] + \kappa\big[W_F(x,p_H) - W_F(x,p)\big]$$

$$(11)\quad \tilde{r}W_I(p) = w_I(p) + \delta\big[U(p) - W_I(p)\big] + \kappa\big[W_I(p_H) - W_I(p)\big] + m(\theta_F)\int_{Q(p)}^{1}\big[W_F(s,p) - W_I(p)\big]dG(s)$$

$$(12)\quad \tilde{r}J_F(x,p) = px - w_F(x,p) + \delta\big[V_F - D - J_F(x,p)\big] + \kappa\big[J_F(x,p_H) - J_F(x,p)\big] + \tau V_F$$

$$(13)\quad \tilde{r}J_I(p) = p_I - w_I(p) + \big[\delta + \mu(p)\big]\Big[V_I - J_I(p)\Big] + \kappa\Big[J_I(p_H) - J_I(p)\Big] + \tau V_I.$$

The terms that account for the accumulation of skills disappear when $p = p_H$, so the value functions for H-skilled workers have the same formulation as in the baseline model. Wages are determined by the surplus sharing rule. The resulting wages for this model are presented in Appendix A. The reservation match qualities for unemployed and employed workers are determined in terms of the match surplus in the formal sector. That is, $S_F(C(p),p) = 0$ and $S_F(Q(p),p) = S_I(p)$. In this model the cut-offs are given by:

$$(14)\quad C(p) = \frac{\tilde{r}U(p)}{p} + \frac{\delta D}{p} - \kappa\left(\frac{U(p_H) - U(p)}{p}\right) - \kappa\left(\frac{S_F(C(p),p_H)}{p}\right)$$

$$(15)\quad Q(p) = C(p) + \frac{(\tilde{r}+\delta)S_I(p)}{p} - \kappa\left(\frac{S_F(Q(p),p_H) - S_I(p) - S_F(C(p),p_H)}{p}\right)$$

where the terms that account for the accumulation of skills disappear when $p = p_H$. Note that the direct effect of human capital accumulation is to reduce the cut-offs for L-skilled workers; this effect is picked up by the negative terms in both (14) and (15). An indirect effect of human capital accumulation increases the cut-offs for L-skilled, because both the value of unemployment and the match surplus in the informal sector increase.

Obtaining results similar to those in Lemma 1 is much more complicated with the inclusion of human capital accumulation. Consider the following assumptions.

Assumption 1. $\forall x \in [0,1]$, $S_F(x,p_H) > S_F(x,p_L)$ and $\partial S_F(x,p_H)/\partial x > \partial S_F(x,p_L)/\partial x$.

Assumption 2. $\forall x \in [0,1]$, $S_F(x,p_H) - S_F(x,p_L) > S_I(p_H) - S_I(p_L)$.

These two assumptions impose complementarities between the production technology in the formal sector and worker skills. Assumption 1 implies that formal sector firms have a strict preference for H-skilled workers. If satisfied, then $C(p_H) < C(p_L)$. Assumption 2 is similar to assuming that the marginal value of skills is higher in the formal sector than in the informal sector. If satisfied, then $Q(p_H) < Q(p_L)$. These two implications can be easily verified in Figure 4.

If Assumptions 1 and 2 are satisfied, the human capital model preserves the same ranking in cut-offs as in the baseline model. With this, we can derive similar implications for the conditional and unconditional hazard rates. These results are summarized in Propositions 4 and 5.

Proposition 4. *Suppose that Assumptions 1 and 2 are satisfied. Then, in the model with human capital accumulation, the hazard rate from the informal to the formal sector conditional on worker's initial skill level, $\lambda(t|p)$, is constant for H-skilled workers and increasing for L-skilled workers.*

Proof. The conditional hazard rate for H-skilled workers is $\lambda(t|p_H) = \mu(p_H)$, which is constant with respect to duration, t. Next, for L-skilled workers, the conditional hazard rate is given by $\lambda(t|p_L) = (1-\kappa)^t \mu(p_L) + [1-(1-\kappa)^t]\mu(p_H)$. Then: $\partial\lambda(t|p)/\partial t = (1-\kappa)^t \ln(1-\kappa)\big[\mu(p_L) - \mu(p_H)\big] > 0$, which is positive because $\mu(p_L) < \mu(p_H)$ and $\kappa \in (0,1)$. $\square$

When workers accumulate skills while working in the informal sector, the increase in productivity derived from the accumulation of skills facilitates access to job opportunities in the formal sector. Consequently, the likelihood of moving from the informal to the formal sector for L-skilled workers increases with tenure in the informal sector, resulting in an increasing hazard for L-skilled workers.

Proposition 5. *Suppose that Assumptions 1 and 2 are satisfied. Let ϕ_I be the probability that $p = p_L$ in the informal sector. Then, in the model with human capital accumulation, the unconditional hazard rate, $\lambda(t)$, is:*

(i) increasing if: $-\ln(1-\kappa) > (1-\phi_I)\big[\mu(p_H) - \mu(p_L)\big]$

(ii) U-shaped otherwise.

The proof of Proposition 5 follows the arguments of Lancaster (1990) and is presented in Appendix B.2. This Proposition states that when κ is large, the higher transition rate to the formal sector of H-skilled workers does not increase the fraction of L-skilled in the risk set, because L-skilled workers accumulate skills at a faster rate. As such, the hazard rate is increasing in duration. In contrast, if κ is not very large, it takes some time for the L-skilled to accumulate skills, and the higher transition rate of the H-skilled results in a higher fraction of L-skilled in the risk set. In this case, the hazard rate is initially decreasing; however, eventually L-skilled workers accumulate skills, so the fraction of L-skilled in the risk set decreases, resulting in an increasing hazard for higher durations.

3.2 Employer Learning (Screening)

In this extension of the baseline model, we abstract from human capital accumulation. Instead, we assume that when workers enter the labor market, their skill level (or type) is not known, but it is eventually revealed while they are working. We refer to these workers as "newcomers". We assume that neither the worker nor the employer knows the newcomer's type, and that once the type is revealed, everybody can observe the worker's skill level, as in Farber and Gibbons (1996). The revelation process is a stochastic process such that the worker's skill is revealed with probability σ.

All newcomers start unemployed, and it is common knowledge that a fraction ϕ of them are L-skilled. Newcomers also follow a reservation match quality strategy when facing formal sector job opportunities, taking informal sector opportunities as they arrive. When the worker's type is revealed in a formal sector job, the job could be destroyed if the current match quality is below the reservation match quality for that worker's type.

Let C be the reservation match quality for unemployed newcomers, and Q be the reservation match quality for newcomers holding an informal sector job. In the current study we focus on cases that satisfy the following condition.

Condition 1. $C(p_H) < C < C(p_L)$ and $Q(p_H) < Q < Q(p_L)$.

If Condition 1 holds, then all formal sector workers found to be H-skilled keep their job. On the contrary, a formal sector worker found to be L-skilled with match quality $x < C(p_L)$ loses his job, in which case the firm incurs firing costs. If the worker is found to be L-skilled but match quality is $x > C(p_L)$, then the worker keeps his job.

The payoffs and the reservation match quality for L-skilled and H-skilled workers have the same formulation as that in the baseline model. Let $\bar{p} \equiv \phi p_L + (1-\phi)p_H$ reflect the expected formal sector productivity for newcomers. Given Condition 1 holds, the payoffs for newcomers are given by:

$$(16)\quad \tilde{r}U = z + m(\theta_I)[W_I - U] + m(\theta_F)\int_C^1 [W_F(s) - U]dG(s)$$

$$(17)\quad \tilde{r}W_F(x) = w_F(x) + \delta[U - W_F(x)] + \sigma(1-\phi)W_F(x, p_H) + \sigma\phi\Big[\Gamma_L(x)U(p_L) + (1-\Gamma_L(x))W_F(x, p_L)\Big] - \sigma W_F(x)$$

$$(18)\quad \tilde{r}W_I = w_I + \delta[U - W_I] + m(\theta_F)\int_Q^1 [W_F(s) - W_I]dG(s) + \sigma\phi W_I(p_L) + \sigma(1-\phi)W_I(p_H) - \sigma W_I$$

$$(19)\quad \tilde{r}J_F(x) = \bar{p}x - w_F(x) + \delta[V_F - D - J_F(x)] + \sigma(1-\phi)J_F(x,p_H) + \sigma\phi\Big(\Gamma_L(x)[V_F - D] + (1-\Gamma_L(x))J_F(x,p_L)\Big) - \sigma J_F(x) + \tau V_F$$

$$(20)\quad rJ_I = p_I - w_I + [\delta + \bar{\mu}][V_I - J_I] + \sigma\phi J_I(p_L) + \sigma(1-\phi)J_I(p_H) - \sigma J_I + \tau V_I$$

where $\bar{\mu} \equiv m(\theta_F)[1 - G(Q)]$, and $\Gamma_L(x) = \mathbf{1}\{x < C(p_L)\}$.

Wages for this model are presented in Appendix A. Given Condition 1, reservation match qualities for newcomers are:

$$(21)\quad C = \frac{\tilde{r}U}{\bar{p}} + \frac{\delta D}{\bar{p}} + \frac{\sigma\phi D}{\bar{p}} - \frac{\sigma[\phi U(p_L) + (1-\phi)U(p_H) - U]}{\bar{p}} - \frac{\sigma(1-\phi)S_F(C,p_H)}{\bar{p}}$$

$$(22)\quad Q = \frac{\tilde{r}U}{\bar{p}} + \frac{\delta D}{\bar{p}} + \frac{\Gamma_L(Q)\sigma\phi D}{\bar{p}} - \frac{\sigma[\phi U(p_L) + (1-\phi)U(p_H) - U]}{\bar{p}} - \frac{[1-\Gamma_L(Q)]\sigma\phi S_F(C,p_L)}{\bar{p}} - \frac{\sigma(1-\phi)S_F(Q,p_H)}{\bar{p}} + \frac{(\tilde{r}+\delta+\sigma)S_I}{\bar{p}}.$$

Note that if $\Gamma_L(Q) = 1$, then $Q \approx C + \frac{(\tilde{r}+\delta+\sigma)S_I}{\bar{p}}$. Again, these hiring standards give us some implications in terms of the hazard rates from the informal to the formal sector, which are summarized in Propositions 6 and 7.

Proposition 6. *Suppose that the condition in Lemma 1 and Condition 1 hold. Then, in the model with employer learning, the hazard rate from the informal to the formal sector conditional on the worker skill level,* $\lambda(t|p)$*, is increasing for H-skilled workers and decreasing for L-skilled workers.*

Proof. The conditional hazard rate is given by $\lambda(t|p) = (1-\sigma)^t\bar{\mu} + [1-(1-\sigma)^t]\mu(p)$, for each $p \in \{p_H, p_L\}$. Let $\partial\lambda(t|p)/\partial t = \lambda'(t|p)$, then $\lambda'(t|p) = (1-\sigma)^t \ln(1-\sigma)\big[\bar{\mu} - \mu(p)\big]$, which is positive for $p = p_H$ because $\bar{\mu} < \mu(p_H)$ and $\sigma \in (0,1)$, and negative for $p = p_L$ because $\bar{\mu} > \mu(p_L)$ and $\sigma \in (0,1)$. □

In this model, employers can distinguish three different groups of workers. However, everyone knows that newcomers are either L-skilled or H-skilled. H-skilled workers face an increasing hazard in their informal sector career because once they are revealed as H-skilled, the likelihood of finding a formal sector job increases. On the contrary, L-skilled workers face a decreasing hazard.

Proposition 7. *Suppose that the condition in Lemma 1 and Condition 1 hold. Let* ϕ *be the probability that* $p = p_L$ *in the labor market. Then, in the model with employer learning, the unconditional hazard rate,* $\lambda(t)$*, is:*

(i) decreasing if $\bar{\mu} > \phi\mu(p_L) + (1-\phi)\mu(p_H)$

(ii) hump-shaped if $\bar{\mu} < \phi\mu(p_L) + (1-\phi)\mu(p_H)$

(iii) flat and then decreasing if $\bar{\mu} = \phi\mu(p_L) + (1-\phi)\mu(p_H)$.

The proof of Proposition 7 follows the arguments of Lancaster (1990) and is presented in Appendix B.2. Proposition 7 states that the shape of the unconditional hazard function initially depends on whether the hazard rate of newcomers is higher than, lower than, or equal to the average hazard rate of workers with revealed types. Cases (*i*) to (*iii*) compare these two hazard rates. Eventually, as more worker types are revealed, the hazard function decreases with duration due to selection, as in the baseline model.

Whether case (*i*), (*ii*), or (*iii*) arises depends on: *a*) the mixture of H-skilled and L-skilled workers in the population, summarized by ϕ; *b*) the location of Q with respect to $Q(p_L)$ and $Q(p_H)$; and *c*) the properties of the distribution of match quality, $G(x)$. Note that Q is not determined by $Q\big(\phi p_L + (1-\phi)p_H\big)$, and so it is hard to raise conclusions in terms of the properties of $Q(\cdot)$, defined in equation (9). Even so, case (*i*) is more likely to occur if $\overline{G}(x) \equiv [1 - G(x)]$ is concave (or $G(x)$ convex), so that the convex combination $\phi\overline{G}\big(Q(p_L)\big) + (1-\phi)\overline{G}\big(Q(p_H)\big)$ is lower than $\overline{G}(Q)$. In contrast, case (*ii*) is more likely to arise if $\overline{G}(x)$ is convex (or $G(x)$ concave), so that the convex combination of $\overline{G}\big(Q(p_L)\big)$ and $\overline{G}\big(Q(p_H)\big)$ is higher than $\overline{G}(Q)$. In addition, case (*ii*) is more likely to come up if formal sector employers proceed with extreme caution when hiring newcomers, so that Q is located close to $Q(p_L)$.

3.3 Understanding the Role of the Informal Sector in the Early Careers of Less-educated Workers

We are now in a position to assess the role of the informal sector in the early stages of the careers of less-educated workers. The results from the previous sections indicate that we can use the estimated hazard functions to determine which mechanism, the accumulation of skills or the screening of workers' abilities, plays a more important role in explaining the dynamics of transitions from the informal to the formal sector in the data. We estimate these hazard functions using data from an employment survey from Mexico. In the next section, we describe the data, the sample, and some details of the variables used in estimation.

4 Data: The ENOE

We use a household survey from Mexico called the Occupation and Employment Survey, ENOE (its acronym in Spanish). The ENOE is a rotating panel where households are visited five times during 12 months, one visit every three months. Every three months, 20% of the sample is replaced. Although information from each family member is recorded, this information is provided by only one member; the respondent is not necessarily the same individual on each visit.

The ENOE records the demographics of each family member (e.g. education, age, marital status), and information on the main and secondary jobs of family members older than 12 years of age. Job information includes working hours, earnings, fringe benefits, job position, firm size, industry, occupation and job tenure. The job tenure information is only recorded in the long form of the ENOE, which is answered at least once during the five visits to the household. For further details about the ENOE see INEGI (2005, 2007).

4.1 Sample

To focus on less-educated workers, we restrict the sample to individuals not currently attending school and with less than 12 years of education. To focus on young workers, our sample only includes workers between the ages of 16 and 25. Age 16 is the minimum age at which a worker can be hired according to Mexican Labor Law (see Congress, 1970), and age 25 is the age at which transitions from the informal to the formal sector plateau (see Figures 2 and 3). Our sample only includes male workers because women may have different reasons for joining the informal sector, e.g. job flexibility to balance work and child rearing (Arias and Maloney, 2007).

We divide our sample of less-educated workers into two groups based on completion of the mandatory level of education in Mexico, which is 9 years. In one group, we include less-educated workers who failed to complete the mandatory level of education, and in the other those who completed the mandatory level of education but who failed to complete high school (i.e. 12 years). Since the mandatory level of education in Mexico could be compared to junior high school in the U.S., we refer to the first group as junior high school dropouts, and the second as junior high school graduates.[8]

Table 1 presents the sample summary statistics. Note that the group of junior high school dropouts is further divided in two groups. Junior high school graduates represent

[8]In Mexico, compulsory education comprises primary school (grades 1 to 6) and junior high school (grades 7 to 9). In terms of our labeling, note that some of the individuals in the junior high school dropout group may not have even started junior high school.

63% of the sample. Workers in all three groups are mainly concentrated in small firms, but the junior high school graduates have the highest percentage in large firms. Also, note that the two groups of junior high school dropouts are mainly concentrated in the construction industry, while graduates are mainly concentrated in the services industry. Finally, note that graduates are more likely to have a parent working in a formal sector job. Firm size, industry, and family head employment status could be important determinants of the probability of moving from the informal to the formal sector.

4.2 Identification of Informal Salaried Workers

When a worker is hired in Mexico, it is the employer's responsibility to register the worker in the IMSS or the ISSSTE.[9] These institutions provide a bundle of benefits to their affiliates. For example, the bundle offered by IMSS includes: health insurance, day-care services for children, life insurance, disability pensions, work-risk pensions, sports and cultural facilities, retirement pensions, and housing loans (Levy, 2007). Both the worker and the employer must pay fees to fund these institutions, but the portion paid by the employer is much higher than that paid by the worker. If the firm is caught not complying with these regulations, it incurs a penalty.

Once a worker is registered in the IMSS or the ISSSTE the work relationship must abide by the labor regulation in Mexico. This means that the employer will incur firing costs if the work relationship is terminated.

The questionnaire of the ENOE does not ask the individual whether he is a formal or an informal worker. Instead, the survey asks the individual if he has access to medical services provided by the IMSS or the ISSSTE. We consider a worker to belong to the formal sector if he is salaried and has access to the IMSS or the ISSSTE, and to belong to the informal sector if he is salaried and does not have access to these services. Note that the self-employed are not included in our definition of the informal sector.

4.3 Measuring Duration in the Informal Sector

To obtain the measure of duration of employment in the informal sector, we employ two sampling schemes: flow sampling and stock sampling. In the flow sample, we only include individuals that made the transition into the informal sector during the period of observation. Given the design of the ENOE, we include in this sample individuals that moved into the

[9] IMSS is the acronym in Spanish for the Mexican Institute of Social Security and ISSSTE is the acronym in Spanish for the Institute of Security and Social Services for the State's Workers.

informal sector after the first but before the fourth visits.[10] Hence only those who made a transition into the informal sector recorded in the second, third, or fourth visits are included in the flow sample.

In the stock sample, we draw individuals from the stock of informal sector workers at the time that the long form of the ENOE was used, since the long form enables us to identify some information about the start date of the current job. We include in this sample informal sector workers whose long form interview took place in the first, second, third, or fourth visits.[11] The level of detail of the start date recorded by the ENOE depends on the current tenure of the worker. If the individual's job started in the current or the previous calendar year, then the survey records the month and the year of the start date. If the job started before the previous calendar year, then it only records the year.

The stock sample is constructed using two pieces of information. The first concerns the measure of *elapsed duration*, which is the time between the job start and the long form interview. The second concerns the *residual duration*, which is the time between the long form interview and the moment of transition to the formal sector.[12] Depending on the tenure of the individual, the elapsed duration will be stored in months (if the job started in the current or previous calendar year) or in a 12-month interval (if the job started before the previous calendar year). Similarly, if the individual changed employers when moving from the informal to the formal sectors, and the month of this event is recorded, then the residual duration will be stored in months; otherwise it will be stored in a 3-month interval.

The complete duration in the stock sample is the sum of the elapsed and the residual duration. Let T_i denote duration in the informal sector of individual i. Hence, in the stock sample we have three possible situations: (*i*) T_i is the exact number of months, (*ii*) T_i is only known to be contained in a 12-month interval, or (*iii*) T_i is only known to be contained in a 15-month interval.

All duration measures from the flow sample are interval-censored. For these individuals, we only know that between two consecutive visits the individual made a transition into the informal sector, but the exact moment of the transition is only known within a 3-month interval. When the worker makes a transition to the formal sector, we may observe the month of the transition to the formal sector if the individual changed employers and recorded this

[10]Those who made a transition between the fourth and fifth visits are not included, because we are not able to follow them after the fifth visit.

[11]We rely on five interviews for each individual, one for each visit. We refer to the long form interview as the one in which the long form of the ENOE is answered. If an individual has more than one long form interview we use the first one as the reference interview.

[12]There are instances in which an individual does not make a transition to the formal sector during the period of observation. Right-censoring is discussed below, but for ease of exposition the duration measures are described first abstracting from it.

information; otherwise we only know that the transition occurred within a 3-month interval. Hence, in the flow sample we have two possible cases: (i) T_i is only known to be contained in a 3-month interval, or (ii) T_i is only known to be contained in a 6-month interval.

Summing up, we may observe T_i up to the exact number of months, or we may only observe an interval $(L_i, R_i]$ such that $T_i \in (L_i, R_i]$. Note that $(L_i, R_i]$ may overlap for different individuals, hence we cannot use the techniques of discrete time duration analysis (e.g. Prentice and Gloeckler, 1978; Meyer, 1990; Han and Hausman, 1990). We must instead work with interval-censored data. In the final sample, one fifth of the duration data is not intervallic and the most frequent type of interval is the 3-month interval (see Table 2). Two-thirds of the sample comes from the stock sample, and one-third of the stock sample has observations with interval-censored elapsed durations.

A spell is only completed when the informal sector worker makes a transition to the formal sector.[13] This event is known as *failure* (Kalbfleisch and Prentice, 1980). However, if the individual is still working in the informal sector at the time of the last visit, his spell will be right-censored because we do not observe his failure. We also treat transitions from the informal sector into a state different from the formal sector (e.g. self-employment) as right-censored. We treat these spells as right-censored because, for individuals in this situation, we do not observe their failure. In the next section, we discuss the assumptions that the censoring mechanism must satisfy in order to make inference.

Finally, some of the spells in the sample have starting points on a date before the individual reaches age 16. Individuals that started their informal sector jobs before age 16 may delay their transition to the formal sector owing to legislative restrictions, and not for the reasons stipulated in the model. We adjust the duration measure of these individuals by subtracting from their duration the number of months worked before age 16, and create an indicator variable for them, which is included in the covariates. In this way, all spells measure the time that the individuals were "at risk" of making a transition to the formal sector.

Table 3 summarizes the duration data generated from the ENOE. For this table, we impute interval-censored duration measures with the midpoint in the interval. Note that the mean duration of employment in the informal sector ("time to exit" in Table 3) is lower for junior high school graduates. In fact, the distribution of duration for junior high school graduates first-order-stochastically dominates that of the dropouts, suggesting that graduates move to the formal sector at a faster rate than the dropouts. Finally, note that in all three groups we have a high degree of censoring reflected in the small number of failures. The number of spells adjusted because of their pre-age 16 starting point ranges from 7 to

[13] A spell is the measure of duration of employment in the informal sector.

12% of the observations.

5 Estimation

5.1 Likelihood Function

The likelihood function to be estimated depends on the hazard, the survivor, and the density functions. Once the hazard function is defined, we can formulate the survivor function as $S(t) = \exp\{-\int_0^t \lambda(s)ds\}$, and the density of T as $f(t) = \lambda(t)S(t)$. For estimation, we use a set of covariates, x_i, as described below. Hence, we work with $\lambda(t|x_i)$, $S(t|x_i)$, and $f(t|x_i)$.

In order to make inference in the presence of right-censoring, the censoring mechanism must satisfy the assumption of *independent censoring*. Kalbfleisch and Prentice (1980, page 13) define a censoring scheme as independent if "the probability of censoring at time t depends only on the covariate x, the observed pattern of failures and censoring up to time t in the trial, or on random processes that are independent of the failure times in the trial."[14] Let T_i^* be the completed spell of individual i in the absence of censoring, and let C_i be the censoring time for i. Independent censoring implies that, given the covariates, C_i are independent of each other and of the uncensored duration T_i^*.

In the ENOE, given that the individual is only visited for a fixed number of times, censoring as a result of the individual working in the informal sector during the last visit is independent. But we must be cautious with the duration of employment of individuals that did not fail, because they moved to another state (e.g. self-employment). The duration of employment of these individuals is right-censored, but the assumption of independent censoring could be violated if they were systematically more (or less) likely to make a transition to the formal sector.

To that end, in our covariates we include variables that also provide information on why these individuals move to another state *before* moving to the formal sector. The covariates include industry, firm size, educational attainment, degree of government support to self-employment by state of residence, marital status, and employment status of the family head. In the final sample, about 30% of the observations are censored because the individual moved to unemployment or to "another risk" (see Table 4). The state "another risk" is mainly composed of self-employment, but also includes unpaid family work, entrepreneurship, and out of the labor force.

Interval-censoring also imposes a requirement in order to make inference, which is very

[14] As the authors emphasize "mechanisms in which the failure times of individuals are censored because the individuals appear to be at unusually high (or low) risk of failure are not independent."

similar to the one for right-censoring. Kalbfleisch and Prentice define this requirement as *independent interval censoring.* Let $0 < C_{i1} < C_{i2} < \cdots < C_{im_i} < \infty$ be the inspection times for individual i, then independent interval censoring requires that: "having observed that the individual is alive at time $C_{i,j-1}$, the timing of the next inspection is distributed independently of the time of the failure"(Kalbfleisch and Prentice, 1980, page 79). Since the household visits are scheduled every three months, this assumption is also satisfied in the ENOE. The assumption would be violated if the next visit is determined to be sooner (or later) depending on the probability that the individual moves from the informal to the formal sector.

Given that both censoring mechanisms are independent, the contribution of a right-censored observation to the likelihood is $\Pr(T_i^* > C_i|x) = S(C_i|x)$, and the contribution of an interval-censored observation is $S(L_i|x_i) - S(R_i|x_i)$. Define the indicator functions $\Upsilon_i = \mathbf{1}\{T_i \text{ Not Interval}\}$ and $d_i = \mathbf{1}\{T_i \text{ Not Right-censored}\}$, and let E_i denote the elapsed duration. The likelihood function is given by:

$$(23) \quad L(\theta|x_i) = \prod_{\{i|\Upsilon_i=1\}} \frac{f(t_i|x_i)^{d_i} S(t_i|x_i)^{(1-d_i)}}{S(E_i|x_i)} \prod_{\{i|\Upsilon_i=0\}} \frac{S(L_i|x_i) - S(R_i|x_i)}{S(E_i|x_i)}$$

where the elapsed duration, E_i, is used to weight the likelihood because of the sample selection caused by stock sampling (see Kiefer, 1988; Wooldrige, 2002, chap. 20).

In our sample, we have the additional problem that for some observations the elapsed duration is interval-censored, i.e. $E_i \in [E_i^L, E_i^R)$. We investigated several alternatives for overcoming coarseness in the start date by imputing the interval-censored starting times and performing a Monte Carlo analysis (see Appendix C). The results indicate that imputed interval midpoints outperform the alternatives. Therefore, in the estimation we use this imputed measure of elapsed duration.

5.2 Hazard Function

To estimate the hazard, instead of imposing the functional form implied by each model, we estimate a flexible hazard function. Widely used parametric models such as the Weibull or the Log-logistic impose restrictions on the shape of the hazard (see Wooldrige, 2002, chap. 20). For this reason, our main results rely on the estimation of a piecewise constant hazard, which allows more flexibility in the shape of the hazard function. We assume a proportional hazards model $\lambda(t|x_i) = \exp(x_i'\beta)\lambda_0(t)$, where:

$$(24) \quad \lambda_0(t) = \lambda_m, \quad a_{m-1} \leq t < a_m, \quad \lambda_m > 0, \quad m = 1, 2, \ldots, M$$

and $\{a_0, a_1, \ldots, a_M\}$ are known break points that define $M+1$ intervals $[a_0, a_1), [a_1, a_2), \ldots,$ $[a_{M-1}, a_M), [a_M, \infty)$ that may contain t. We set $a_0 = 0$, and choose the other break points using the distribution of T. The distribution of T is divided into six quantiles, so that $M = 6$, with break points determined by the quantiles.

The survivor function is given by:

$$(25) \quad S(t|x_i) = \exp\left\{-\exp(x_i'\beta)\left[\sum_{k=1}^{I(t)-1} \lambda_k\Big(a_k - a_{k-1}\Big) + \lambda_{I(t)}\Big(t - a_{I(t)-1}\Big)\right]\right\}$$

where $I(t)$ is such that $a_{I(t)-1} \leq t < a_{I(t)}$, i.e. t is contained in the $I(t)^{th}$ interval.

We estimate the hazard function for the whole sample and for two mutually exclusive education groups. The break points for each of these samples are:

Education Group	a_1	a_2	a_3	a_4	a_5	a_6
$[0, 12)$	3	4.5	7.5	13	30	128
$[0, 9)$	3	4.5	7.5	14	33	118
$[9, 12)$	3	4.5	6.5	12.5	28.5	128

6 Results

6.1 Piecewise Constant Hazard Function

We maximize the likelihood function in equation (23) using all of the elements discussed in the previous section. The estimation results for the whole sample and for junior high school dropouts and graduates are summarized in Table 5. Figure 5 depicts the estimated baseline hazard with the 95% pointwise confidence intervals. The plot of the baseline hazard in Figure 5 depicts the hump-shaped pattern predicted by the model with employer learning. Note that this pattern holds for the whole sample, and for the junior high school dropouts and graduates.

Even though both junior high school dropouts and graduates show signs of employer learning, those who completed the mandatory level of education have a higher hazard rate at all times. In terms of Proposition 7, this result indicates that the proportion of L-skilled workers is higher among dropouts than among graduates as one might expect.[15]

Estimated effects of the covariates in Table 5 are fairly similar for the whole sample and for junior high school graduates and dropouts. The estimation results for the whole sample show that graduation from primary school (grade 6) has little effect on the hazard rates, but

[15] Alternatively, there could be more than two worker skill levels, with some of them concentrated in one education group, e.g. the highest concentrated in group of graduates and the lowest concentrated in the group of dropouts. Note that we could extend the models to a continuum of worker types, as in Albrecht et al. (2006, 2009). This would yield similar results to those derived above.

graduation from secondary school (grade 9) has a significant effect. This is consistent with Arias and Maloney (2007) who claim that "graduation to formal salaried work is unlikely for youth who drop out of school before completing at least a full course of secondary education" (Arias and Maloney, 2007, page 62) .

Not surprisingly, one of the most important covariates is the size of the firm. The higher the firm size, the higher the hazard rate from the informal to the formal sector. There are two potential explanations for this result. On the one hand, many of the transitions could be happening within the same employer. Alternatively, it could be that larger firms have a larger network and as a result expose workers' skills to other employers more than small firms do.

Industry does not play a big role in explaining the hazard rate from the informal to the formal sector. Married workers have higher hazard rates than single workers, consistent with the incremental demand for health services when individuals form their own families. And when the family head works in the formal sector, the individual also has a higher hazard rate, which could also be the result of the individual having access to a larger network of formal sector employers.[16]

Finally, note that a hump-shaped hazard rules out the baseline model. The baseline model predicts constant hazard rates conditional on worker skill level, which in turn implies that the unconditional survivor function is a mixture of exponential distributions. Based on comments made by Chamberlain (1980), Heckman, Robb, and Walker (1990) argue that "all mixtures of exponentials models have nonincreasing hazards." The pointwise confidence intervals for our estimated hazard imply that the hazard is increasing for short spells, thereby ruling out the baseline model (with any arbitrary number of worker types).[17]

6.2 Parametric Hazard Functions

As a robustness check, we estimated two widely used parametric hazards, the Weibull and the Log-logistic hazard models. We are mainly interested in the estimation result from the Log-logistic model. The Weibull is characterized by the hazard function:

$$(26) \quad \lambda(t|x) = \varphi\alpha t^{\alpha-1}$$

[16]In Mexico, dependents of workers registered in the IMSS can only use the medical services of this institution up to age 18. The coverage can be extended if the dependent is attending school, which is not the case in our sample.

[17]Using the estimated hazard function, and following the procedure suggested by Chamberlain, we conclude that the survivor function for the data in this study cannot be generated by a mixture of exponentials. For a description of the rejection criterion and the procedure see Chamberlain (1980) or Heckman et al. (1990).

and the Log-logistic by:

$$(27) \quad \lambda(t|x) = \frac{\varphi \alpha t^{\alpha-1}}{1+\varphi t^{\alpha}},$$

where $\varphi = \exp(x'\beta)$ is the most common choice in empirical applications. The shape of the hazard function in each case is determined by the parameter α, as summarized in the following table:

	Weibull	Log-logistic
$\alpha < 1$	Decreasing	Decreasing from ∞ at $t = 0$, to 0 as $t \to \infty$
$\alpha = 1$	Constant	Decreasing from φ at $t = 0$, to 0 as $t \to \infty$
$\alpha > 1$	Increasing	Increasing from 0 at $t = 0$, to a single maximum, and then approaches 0 as $t \to \infty$

When $\alpha > 1$ in the Log-logistic, the maximum occurs at $T^* = [(\alpha-1)/\varphi(x)]^{1/\alpha}$ (see Lancaster, 1990, chap. 3).

The estimated hazards for these two models are presented in Table 6. The estimated coefficients for the covariates in the Weibull hazard are very similar to the ones in the piecewise constant hazard, since both of these are proportional hazards models. For the Log-logistic model, they are not identical but have the same pattern across the groups of dropouts and graduates from junior high school. Given the restrictions of the Weibull hazard, the estimates suggest a monotonically decreasing hazard, but the Log-logistic suggests a hump-shaped hazard. Most importantly, the predicted maximum in the Log-logistic hazard function is very similar to the maximum we have in the piecewise constant hazard in Figure 5; where T^* was computed using $x = \bar{x}$.

6.3 Screening in *Bécate* Training Program

In this section we use the estimated piecewise constant hazard to infer the parameters governing the employer learning process. Knowledge of these parameters gives us the means to evaluate *Bécate*'s screening program introduced in Section 1. In terms of the employer learning model, we want to know how fast employers learn about their workers' abilities. This information is obtained using the model-generated hazard and the estimated hazard. The unconditional hazard in the employer learning model is a function of five parameters, $\left(\bar{\mu}, \mu(p_L), \mu(p_H), \sigma, \phi\right)$; while the piecewise constant hazard is a function of seven parameters, $(\lambda_1, \ldots, \lambda_6, \beta)$. We use the estimated parameters $\left(\hat{\lambda}_1, \ldots, \hat{\lambda}_6, \hat{\beta}\right)$ to infer the value of the parameters of the employer learning model.

Let $\nu(t) \equiv \lambda_M\left(t; \bar{\mu}, \mu(p_L), \mu(p_H), \sigma, \phi\right) - \lambda_{PW}\left(t; \hat{\lambda}_1, \ldots, \hat{\lambda}_6, \hat{\beta}\right)$, for $t = 0, 1, \ldots, \overline{T}$, denote the residual between the model generated hazard, $\lambda_M(\cdot)$, and the estimated piecewise constant hazard, $\lambda_{PW}(\cdot)$. To get the parameters governing the employers' learning process, we look for the vector $(\bar{\mu}, \mu(p_L), \mu(p_H), \sigma, \phi)$ that minimizes the sum of squared residuals. The details of the optimization algorithm are explained in Appendix D.

The estimated and the model-generated hazards are shown in Figure 6. The resulting parameters indicate that employers learn their workers' abilities at a rate of $\sigma = 0.1437$ per month, and that the proportion of L-skilled workers in the population is $\phi = 0.5416$. Then, at the end of a three-month *Bécate* program, employers know the skill level of about 37% of the recruited workers, where 54% of these workers are expected to be L-skilled. The firm will be happy to hire those identified as H-skilled, but must also fulfill its promise to take 70% of the workers recruited for the program. This implies that the firm must take a gamble in hiring 53% of the original number of workers whose skill level is still unknown. However, since 54% of these workers are expected to to be L-skilled, the firm will end up hiring 29% of the original number of workers that are L-skilled. If the firm does not have a good match quality with these L-skilled workers, it will incur firing costs.

7 Final Remarks

The present study asks whether work experience in the informal sector can affect the career prospects of less-educated workers. The analysis focuses on two potential roles of informal sector jobs: accumulation of skills and screening of workers' ability. In the traditional queuing model of the informal sector with heterogenous workers' abilities, the hazard rate from the informal into the formal sector decreases with duration of informal sector employment. This study shows that, when informal sector jobs also enable workers to accumulate skills or employers to screen workers' abilities, the shape of the hazard function can be different from that predicted by the traditional queuing model. Human capital accumulation implies an increasing or U-shaped hazard due to the accumulation of skills (and the fact that more skilled workers leave the informal sector faster). Screening can generate a hump-shaped hazard if workers with observable ability leave (on average) faster than informal sector entrants, resulting in an increasing hazard; eventually, as more skilled workers leave faster, the hazard decreases with duration. These differences in the predicted hazard suggests a procedure to decide which role of informal sector jobs is more important.

The hazard function was estimated using an employment survey from Mexico. The estimated hazard reflects the hump-shaped pattern predicted by the screening model, indicating that informal sector jobs play mostly the role of a screening device that enables employers to distinguish the best workers from the worst. Furthermore, the estimation results reject the traditional queuing model with heterogenous workers' abilities, indicating that informal sector jobs provide some value above and beyond make-shift work while waiting to find a formal sector job.

The employment survey used in this study is a rotating panel with a periodic follow-up, and so a significant fraction of the duration measures are interval-censored. In addition, for a good share of the spells the starting time is only known to fall within a twelve-month interval. These features of the data required the application of techniques for interval-censored failure time data, and a Monte Carlo study to investigate several alternatives for overcoming coarseness of the starting time of spells. The latter is part of a larger research project in progress.

The parameters characterizing the employer learning process were inferred to determine how fast employers learn about their workers' abilities. The exercise suggests that employers learn about their workers' abilities at a much slower rate than that required by a government employment program, *Bécate*. This finding highlights the importance of a firm's involvement in the recruitment of workers participating in the program. In this way firms can minimize expected firing costs by recruiting candidates with a good match quality. Firm participation in the selection of candidates is allowed in the current format of *Bécate*.

Finally, the results in this study suggest a limited role for human capital accumulation in the informal sector. On the one hand, if low skill formation is due to the absence of opportunities to produce human capital in informal sector jobs, then policies that control these sort of jobs might be necessary, taking into consideration the loss of the screening services that these jobs provide. On the other hand, if these workers lack the ability to produce more human capital, then the problem needs a deeper solution that may involve a reassessment of education policies. Perhaps the basic education system in Mexico is not building the base to engage in the production of more human capital for these individuals.

References

Albrecht, J., L. Navarro, and S. Vroman (2006): "The Effects of Labor Market Policies in an Economy with an Informal Sector," Discussion Paper IZA DP No. 2141, The Institute for the Study of Labor (IZA).

——— (2009): "The Effects of Labour Market Policies in an Economy with an Informal Sector," *The Economic Journal*, 119, 1105–1129.

Amaral, P. S. and E. Quintin (2006): "A competitive model of the informal sector," *Journal of Monetary Economics*, 53, 1541–1553.

Arias, O. and M. Khamis (2008): "Comparative Advantage, Segmentation And Informal Earnings: A Marginal Treatment Effects Approach," Mimeo.

Arias, O. and W. F. Maloney (2007): "The *Razón de Ser* of the Informal Worker," in *Informality: Exit and Exclusion*, The World Bank, chap. 2, 43–78.

Barron, J. M., M. C. Berger, and D. A. Black (1997): *On-the-Job Training*, Kalamazoo, Michigan: Upjohn Institute for Employment Research.

Bosch, M. (2006): "Job Creation and Job Destruction in the Presence of Informal Labour Markets," Working Paper, London School of Economics.

Bosch, M., E. Goni, and W. F. Maloney (2007): "The Determinants of Rising Informality in Brazil: Evidence from Gross Worker Flows," Working Paper 4375, The World Bank.

Chamberlain, G. (1980): "Comment on 'The Analysis of Re-Employment Probabilities for the Unemployed' by T. Lancaster and S. Nickell," *Journal of the Royal Statistical Society. Series A (General)*, 143, 160.

Congress (1970): *Federal Labor Law*, Mexico, D.F.

Delajara, M., S. Freije, and I. Soloaga (2006): "An Evaluation of Training for the Unemployed in Mexico," Working Paper OVE/WP-09/06, Inter-American Development Bank.

Dolado, J. J., M. Jansen, and J. F. Jimeno (2005): "Dual Employment Protection Legislation: A Framework for Analysis," Discussion Paper IZA DP No. 1564, The Institute for the Study of Labor (IZA).

Fajnzylber, P. (2007): "Informality, Productivity and the Firm," in *Informality: Exit and Exclusion*, The World Bank, chap. 6, 157–178.

Farber, H. S. and R. Gibbons (1996): "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111, 1007–1047.

Farrell, D. (2004): "The hidden dangers of the informal economy," *The McKinsey Quarterly*, 27–37.

Han, A. and J. A. Hausman (1990): "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5, 1–28.

Heckman, J. J., R. Robb, and J. R. Walker (1990): "Testing the Mixture of Exponentials Hypothesis and Estimating the Mixing Distribution by the Methods of Moments," *Journal of the American Statistical Association*, 85, 582–589.

Hemmer, H.-R. and C. Mannel (1989): "On the Economic Analysis of the Urban Informal Sector," *World Development*, 17, 1543–1552.

INEGI (2005): *Encuesta Nacional de Ocupación y Empleo 2005. Una Nueva Encuesta para México. ENOE*, Aguascalientes: Instituto Nacional de Estadística, Geografía e Informática, http://www.inegi.org.mx.

——— (2007): *Conociendo la base de datos de la ENOE*, Aguascalientes: Instituto Nacional de Estadística, Geografía e Informática, http://www.inegi.org.mx.

Kalbfleisch, J. D. and R. L. Prentice (1980): *The statistical analysis of failure time data*, New York: Wiley.

Kiefer, N. M. (1988): "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26, 646–679.

Lancaster, T. (1990): *The econometric analysis of transition data*, New York: Cambridge University Press.

Levy, S. (2007): "Can Social Programs Reduce Productivity and Growth? A Hypothesis for Mexico," Mimeo.

Loayza, N. V. (1996): "The economics of the informal sector: a simple model and some empirical evidence from Latin America," *Carnegie-Rochester Conference Series on Public Policy*, 45, 129–162.

Magnac, T. (1991): "Segmented or Competitive Labor Markets," *Econometrica*, 59, 165–187.

Maloney, W. F. (1999): "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico," *The World Bank Economic Review*, 13, 275–302.

Meyer, B. D. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58, 757–782.

Mortensen, D. T. and C. A. Pissarides (2003): "Taxes, subsidies and equilibrium labor market outcomes," in *Designing inclusion: tools to raise low-end pay and employment in private enterprise*, ed. by E. S. Phelps, Cambridge; New York: Cambridge University Press, 44 – 73.

Prentice, R. L. and L. A. Gloeckler (1978): "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data," *Biometrics*, 34, 57–67.

Rebière, T. (2008): "Young workers' professional experience and access to high-skill jobs," Working paper, CERENE, University of Le Havre, France.

Schneider, F. and D. H. Enste (2000): "Shadow Economies: Size, Causes, and Consequences," *Journal of Economic Literature*, 38, 77–114.

Wooldrige, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Mass.: The MIT Press.

World Bank (2007): *World Development Report 2007: Development and the Next Generation*, Washington, DC: The World Bank.

Table 1: Summary Statistics by Education Group

	Years of Education		
	[0 , 6)	[6 , 9)	[9 , 12)
Age	20.96	20.31	20.60
Married	0.13	0.16	0.14
Monthly Earnings†	3498.44	3641.89	3428.78
Minimum Wage‡			
Zone A	0.13	0.13	0.15
Zone B	0.10	0.14	0.15
Zone C	0.76	0.73	0.70
Firm Size			
1-5	0.66	0.62	0.60
6-20	0.24	0.27	0.25
21+	0.10	0.11	0.15
Industry			
Construction	0.43	0.32	0.23
Manufactures	0.20	0.20	0.19
Commerce	0.12	0.15	0.21
Services	0.24	0.33	0.37
Family Head Status§			
Formal Sector Job	0.08	0.16	0.21
Self-employed	0.12	0.13	0.12
Unemployed	0.01	0.01	0.02
Entrepreneur	0.05	0.09	0.10
Out of Labor Force	0.13	0.08	0.08
Obs.	402	1,508	3,207

†Average monthly earnings in Mexican Pesos as of the 2nd half of December 2010. ‡Minimum wage by zone: A > B > C. §Employment status of the family head, when the family head is different from the individual in the sample.

Table 2: Distribution of Duration Data in the Sample

	Type of Sample			
Type of Interval	Flow	Stock 1†	Stock 2§	Total
Non-interval	0	1,129	0	1,129
3-month	839	958	0	1,797
6-month	1,075	0	0	1,075
12-month	0	0	527	527
15-month	0	0	589	589
Total	1,914	2,087	1,116	5,117

†Workers with job start in the current or previous calendar year. §Workers with job start before the previous calendar year.

Table 3: Summary Statistics of Duration Data

	Years of Education		
	[0 , 6)	[6 , 9)	[9 , 12)
Time to exit			
Mean	19.6	17.6	15.7
25th pctile	3.5	3.5	3
50th pctile	9	7.5	6.5
75th pctile	23	21	18
Mean elapsed duration	12.62	10.98	9.60
Number of failures	113	513	1,321
Obs. w/adjusted duration	43	187	226
Obs.	402	1,508	3,207

Note: For the purposes of getting these summary statistics, we imputed the interval-censored duration data using the middle point in the interval.

Table 4: Censoring in the Sample

	Freq.	Percent
Uncensored	1,947	38.05
Unemployed	615	12.02
Another risk§	939	18.35
Right-censored	1,616	31.58
Total	5,117	100

§ Mainly composed by self-employment, but also includes unpaid family work, entrepreneurship, and out of the labor force.

Table 5: Estimated Piecewise Constant Hazard

	Years or Education		
	[0 , 12)	[0 , 9)	[9 , 12)
Firm size 6-20	0.4082	0.3707	0.4331
	(0.0548)	(0.0946)	(0.0676)
Firm size 21+	0.7887	0.9469	0.7384
	(0.0639)	(0.1179)	(0.0765)
Manufactures Ind	-0.1000	-0.0425	-0.1549
	(0.0712)	(0.1196)	(0.0892)
Commerce Ind	0.1214	0.0080	0.1239
	(0.0714)	(0.1337)	(0.0860)
Services Ind	-0.0418	0.1685	-0.1507
	(0.0618)	(0.1028)	(0.0772)
Graduate Grade 6	0.1446	0.1124	
	(0.1058)	(0.1074)	
Graduate Grade 9	0.2098		
	(0.0538)		
Married	0.2797	0.4823	0.1666
	(0.0657)	(0.1072)	(0.0840)
Family Head FS	0.2604	0.2966	0.2479
	(0.0588)	(0.1100)	(0.0700)
Gov. Support to SE	-0.0054	-0.0061	-0.0047
	(0.0019)	(0.0034)	(0.0023)
λ_1	0.0235	0.0226	0.0345
	(0.0034)	(0.0049)	(0.0049)
λ_2	0.1817	0.1590	0.2710
	(0.0213)	(0.0248)	(0.0273)
λ_3	0.0507	0.0454	0.0948
	(0.0073)	(0.0099)	(0.0163)
λ_4	0.0335	0.0307	0.0536
	(0.0044)	(0.0055)	(0.0061)
λ_5	0.0381	0.0431	0.0525
	(0.0047)	(0.0070)	(0.0058)
λ_6	0.0415	0.0392	0.0637
	(0.0053)	(0.0070)	(0.0073)
Log likelihood	-3,444.20	-1,190.92	-2,236.59
Number of obs	5,117	1,910	3,207

Omitted industry is Construction, omitted firm size is 1-5. The covariates also include indicators for other family head employment status (see Table 1), and an indicator for adjusted duration measures. The "Gov. Support to SE" variable corresponds to the number of supported applications for self-employment scholarships in the state of residence, relative to the size of the local labor market. Standard errors in parenthesis.

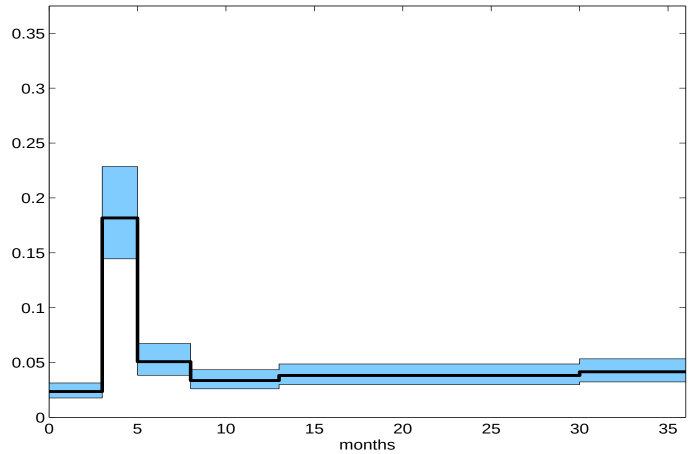
Table 6: Estimated Weibull and Log-logistic Hazards

	Weibull Years or Education			Log-logistic Years or Education		
	[0 , 12)	[0 , 9)	[9 , 12)	[0 , 12)	[0 , 9)	[9 , 12)
Firm size 6-20	0.4211	0.4001	0.4320	0.6269	0.5505	0.6742
	(0.0546)	(0.0943)	(0.0671)	(0.0952)	(0.1585)	(0.1197)
Firm size 21+	0.8250	0.9838	0.7725	1.1266	1.1477	1.1147
	(0.0633)	(0.1170)	(0.0756)	(0.1172)	(0.2031)	(0.1438)
Manufactures Ind	-0.1078	-0.0616	-0.1520	-0.1960	0.0776	-0.4176
	(0.0709)	(0.1195)	(0.0886)	(0.1220)	(0.1939)	(0.1582)
Commerce Ind	0.1289	0.0057	0.1419	0.1925	0.2457	0.1078
	(0.0710)	(0.1333)	(0.0854)	(0.1226)	(0.2219)	(0.1500)
Services Ind	-0.0493	0.1771	-0.1627	-0.1657	0.1940	-0.3910
	(0.0616)	(0.1027)	(0.0767)	(0.1051)	(0.1698)	(0.1354)
Graduate Grade 6	0.1229	0.0942		0.2592	0.1895	
	(0.1056)	(0.1071)		(0.1724)	(0.1758)	
Graduate Grade 9	0.2052			0.3964		
	(0.0535)			(0.0913)		
Married	0.2950	0.4882	0.1877	0.5511	0.7917	0.3802
	(0.0652)	(0.1061)	(0.0834)	(0.1177)	(0.1890)	(0.1507)
Family Head FS	0.2876	0.3061	0.2779	0.3363	0.4352	0.3114
	(0.0585)	(0.1097)	(0.0695)	(0.1047)	(0.1906)	(0.1259)
Gov. Support to SE	-0.0055	-0.0067	-0.0046	-0.0121	-0.0123	-0.0117
	(0.0019)	(0.0035)	(0.0024)	(0.0033)	(0.0054)	(0.0042)
α	0.8845	0.9177	0.8762	1.6362	1.5710	1.6885
	(0.0192)	(0.0346)	(0.0233)	(0.0410)	(0.0675)	(0.0521)
Log likelihood	-3,597.05	-1,234.41	-2,352.53	-3,621.42	-1,253.86	-2,354.12
Number of obs	5,117	1,910	3,207	5,117	1,910	3,207
Log-log. T^*				5.8	6.7	5.5

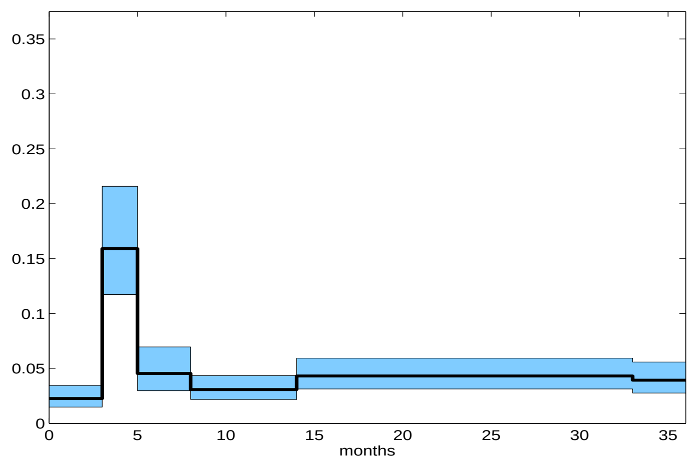
Omitted industry is Construction, omitted firm size is 1-5. The covariates also include indicators for other family head employment status (see Table 1), an indicator for adjusted duration measures, and an intercept. The "Gov. Support to SE" variable corresponds to the number of supported applications for self-employment scholarships in the state of residence, relative to the size of the local labor market. Standard errors in parenthesis.

Figure 5: Piecewise Constant Baseline Hazard with 95% Pointwise Confidence Interval

(a) Years of Education: [0,12)



(b) Years of Education: [0,9)



(c) Years of Education: [9,12)

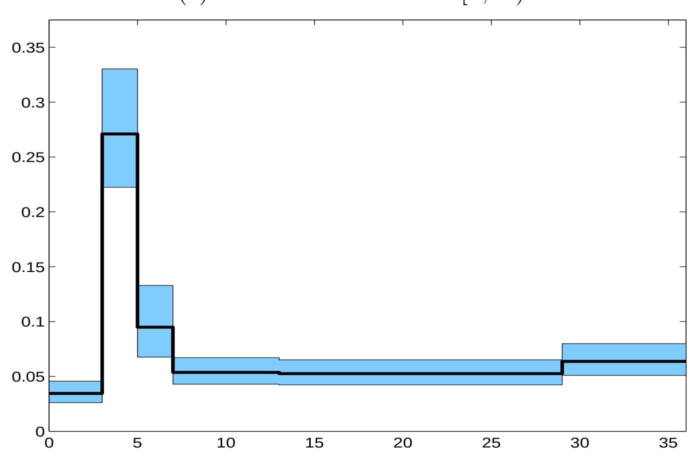
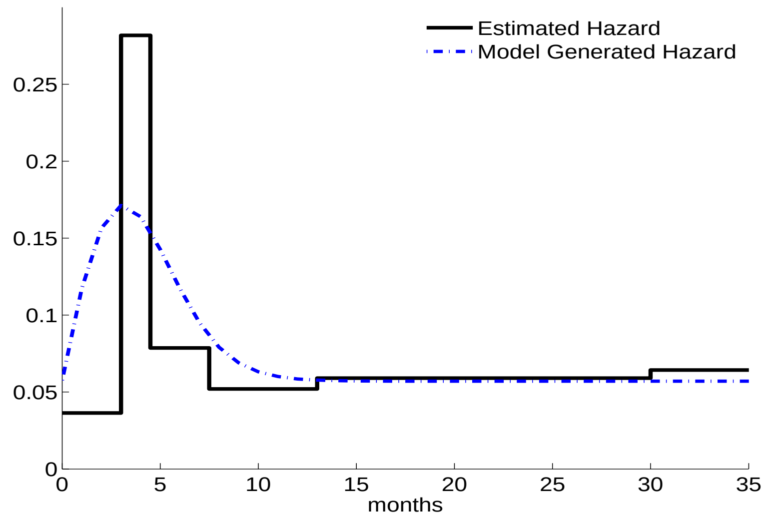


Figure 6: Estimated and Model-Generated Hazards



NOTE: The model-generated hazard uses $\bar{\mu} = 0.0576$, $\mu(p_L) = 0.057024$, $\mu(p_H) = 1.0$, $\phi = 0.5416$, and $\sigma = 0.1437$. The estimated hazard uses $(\hat{\lambda}_1, \ldots, \hat{\lambda}_6)$ from Table 5 and $\exp(\bar{x}'\hat{\beta}) = 1.55$.

APPENDIX

A Wages

The surplus sharing rule implies that:

$$w_F(x,p) \quad \text{is such that:} \qquad W_F(x,p) - U(p) = \frac{\beta}{1-\beta}\big[J_F(x,p) - V_F\big]$$

$$w_I(p) \quad \text{is such that:} \qquad W_I(p) - U(p) = \frac{\beta}{1-\beta}\big[J_I(p) - V_I\big]$$

where in equilibrium, free entry implies that $V_F = 0$ and $V_I = 0$.

Wages in the Baseline Model:

$$w_F(x,p) = \beta\big(px - \delta D\big) + (1-\beta)\tilde{r}U(p)$$
$$w_I(p) = \beta p_I + (1-\beta)\left(\tilde{r}U(p) - \beta m(\theta_F)\int_{Q(p)}^{1} S_F(s,p)dG(s)\right)$$

Wages in the Human Capital Model:

$$w_F(x,p) = \beta\big(px - \delta D\big) + (1-\beta)\Big(\tilde{r}U(p) - \kappa\big[U(p_H) - U(p)\big]\Big)$$

$$w_I(p) = \beta p_I + (1-\beta)\left(\tilde{r}U(p) - \kappa\big[U(p_H) - U(p)\big] - \beta m(\theta_F)\int_{Q(p)}^{1} S_F(s,p)dG(s)\right)$$

Wages in the Learning Model:

$$w_F(x) = \beta\Big(\bar{p}x - \delta D - \sigma\phi\Gamma_L(x)D\Big) + (1-\beta)\Big(\tilde{r}U - \sigma[\phi U(p_L) + (1-\phi)U(p_H) - U]\Big)$$

$$w_I = \beta p_I + (1-\beta)\left(\tilde{r}U - \sigma[\phi U(p_L) + (1-\phi)U(p_H) - U] - \beta m(\theta_F)\int_{Q}^{1} S_F(x)dG(x)\right)$$

B Proofs

B.1 Proof of Lemma 1

Note that $S_I(p) = \dfrac{J_I(p)}{1-\beta}$. In the proof we replace $S_I(p)$ with $J_I(p)/(1-\beta)$ in (9). Consider the following result which proves to be useful in the proof of Lemma 1.

Lemma 8. *An upper bound for* $\big[Q(p) - C(p)\big]$ *is* $\dfrac{\tilde{r}+\delta}{\tilde{r}+\delta+\mu(p)+m(\theta_I)\beta}\left(\dfrac{p_I - z}{p}\right)$.

Proof. First, note that:

$$\begin{aligned}
J_I(p) &= \frac{1-\beta}{\tilde{r}+\delta+\mu(p)}\left(p_I - \tilde{r}U(p) + m(\theta_F)\int_{Q(p)}^{1}[W_F(x,p) - U(p)]dG(x)\right) \\
&< \frac{1-\beta}{\tilde{r}+\delta+\mu(p)}\left(p_I - \tilde{r}U(p) + m(\theta_F)\int_{C(p)}^{1}[W_F(x,p) - U(p)]dG(x)\right) \\
&= \frac{1-\beta}{\tilde{r}+\delta+\mu(p)}\Big(p_I - z - m(\theta_I)[W_I(p) - U(p)]\Big) \\
&= \frac{1-\beta}{\tilde{r}+\delta+\mu(p)}\Big(p_I - z - m(\theta_I)\frac{\beta}{1-\beta}J_I(p)\Big)
\end{aligned}$$

And so, $J_I(p) < \dfrac{1-\beta}{\tilde{r}+\delta+\mu(p)+\beta m(\theta_I)}(p_I - z)$. Since $Q(p) - C(p) = \left(\dfrac{\tilde{r}+\delta}{p}\right)\dfrac{J_I(p)}{1-\beta}$, the result follows. □

Next, we proceed to prove Lemma 1.

Proof. Note that once we substitute equilibrium wage equations and use the surplus sharing rules to substitute for unknown value functions, equations (1), (8), and (9) represent a system of three equations with three unknowns and one parameter:

(28) $F(\tilde{r}U, C, Q; p) = 0.$

Note that we treat $\Omega = (\delta, p_I, m(\theta_I), m(\theta_F), r, z, D, \beta)$ as given because Ω does not change when the parameter p changes. Linearizing $F(\cdot)$ we get:

$$\begin{aligned}
(1-A)d(\tilde{r}U) + (N+AK)dC + (A(L-E))dQ - (AH+M)dp &= 0 \\
-\frac{1}{p}d(\tilde{r}U) + dC + \frac{C}{p}dp &= 0 \\
Bd(\tilde{r}U) + (BK-1)dC + (1+B(L-E))dQ - \left(BH - \frac{Q-C}{p}\right)dp &= 0
\end{aligned}$$

where:

$$\begin{aligned}
A &= \frac{m_I\beta}{\tilde{r}+\delta+\mu(p)}, & K &= \frac{\mu(p)\beta}{\tilde{r}+\delta}p, \\
B &= \frac{\tilde{r}+\delta}{p(\tilde{r}+\delta+\mu(p))}, & L &= \frac{m_F\beta}{\tilde{r}+\delta}p(Q-C)g(Q), \\
E &= m_F g(Q)\frac{J_I}{1-\beta}, & M &= \frac{m_F\beta}{\tilde{r}+\delta}\int_C^1 (x-C)dG(x), \\
H &= \frac{m_F\beta}{\tilde{r}+\delta}\int_Q^1 (x-C)dG(x), & N &= \frac{m_F\beta}{\tilde{r}+\delta}p[1-G(C)]
\end{aligned}$$

Note that for ease of exposition we denoted $m(\theta_j) = m_j$ for $j \in \{F, I\}$. By the Implicit Function Theorem and using Cramer's rule, we can derive dC/dp, which is given by:

$$\frac{dC}{dp} = \frac{A(L-E)(Q-C) + pB(L-E)(M-C) + pA(H-C) + p(M-C)}{p\big[(L-E)(BN+A+pB) + AK + Ap + N + p\big]}.$$

It is straightforward to show that all the terms in the numerator, except for the second one, are negative. Adding the first, second, and fourth terms in the numerator, and after some algebra, we get:

$$p\left(\frac{\delta D + z}{p}\right)\left[\frac{m_F(1-\beta)}{\tilde{r}+\delta+\mu(p)}(Q-C)g(Q) - 1\right] - \frac{m_I\beta}{\tilde{r}+\delta}p(Q-C)$$

which is negative if the term in square brackets is negative. Using Lemma 8 to bound this term from above, and the fact that that m_F, $g(Q) \in (0,1)$, and that $\mu(p)$, $m_I > 0$, we find:

$$\left[\frac{m_F(1-\beta)}{\tilde{r}+\delta+\mu(p)}(Q-C)g(Q) - 1\right] < \left(\frac{1-\beta}{\tilde{r}+\delta}\right)\left(\frac{p_I - z}{p}\right) - 1$$

hence, a sufficient condition for the numerator to be negative is that:

(CDN 1) $\dfrac{1-\beta}{\tilde{r}+\delta}(p_I - z) < p.$

Now, we focus on the denominator of dC/dp. It is straightforward to show that $(L-E)(BN+A+pB) < 0$, and using Lemma 8 again, we can bound from above the absolute value of the this term:

$$(29)\quad \Big|(L-E)(BN+A+pB)\Big| < g(Q)\frac{m_F(1-\beta)(p_I-z)}{\tilde{r}+\delta+\mu(p)+m_I\beta}\left[\frac{m_I\beta+m_F\beta[1-G(Q)]+\tilde{r}+\delta}{\tilde{r}+\delta+\mu(p)}\right].$$

The other term in the denominator is positive and it is given by:

$$(30)\quad (AK+Ap+N+p) = p\left[\left(\frac{m_I\beta}{\tilde{r}+\delta+\mu(p)}\right)\left(\frac{\mu(p)\beta+\tilde{r}+\delta}{\tilde{r}+\delta}\right)+\frac{m_F\beta[1-G(C)]}{\tilde{r}+\delta}+1\right].$$

Next, we compare (29) and (30). Using the sufficient condition (CDN 1) we can show that $p > g(Q)\dfrac{m_F(1-\beta)(p_I-z)}{\tilde{r}+\delta+\mu(p)+m_I\beta}$, so the outer term is higher for (30). Finally, it is straightforward to show that the term in square brackets is also higher in (30) than the term in square brackets in (29), so that the denominator is positive. As a result, $dC/dp < 0$.

Now, we apply the Implicit Function Theorem and use Cramer's rule again to derive dQ/dp, which is given by:

$$\frac{dQ}{dp} = \frac{pB\big[H(N+p)-M(K+p)+C(K-N)\big]+p\big[(M-Q)+A(H-Q)\big]+(N+AK)(C-Q)}{p\Big[(L-E)(BN+A+pB)+AK+Ap+N+p\Big]}.$$

We already proved that under certain parameter conditions the denominator of dQ/dp is positive. Then it just remain to show that the numerator is negative. It is straightforward to show that the second and third terms of the numerator are negative. To show that the first term is positive, note that $M > H$ so:

$$\begin{aligned} pB[H(N+p)-M(K+p)+C(K-N)] &< pB[M(N+p)-M(K+p)+C(K-N)] \\ &= pB[MN-MK+C(K-N)] \\ &= pB(N-K)[M-C] \\ &< 0 \end{aligned}$$

where the last inequality from the fact that $C > M$. As a result, $dQ/dp < 0$. And this completes the proof.

□

B.2 Proofs of the Shape of the Unconditional Hazard Rates

Before proving Propositions 3, 5, and 7, consider the following result about the unconditional hazard rate. The proof of Lemma 9 follows the arguments of Lancaster (1990, chap. 4).

Lemma 9. *Let $\lambda(t|p)$ be the hazard rate conditional on worker skill level, and $\lambda'(t|p) = \partial\lambda(t|p)/\partial t$. Let ϕ_I be the probability that $p = p_L$ in the informal sector. Then, the unconditional hazard rate and its derivative are given by:*

$$\lambda(t) = \gamma(t)\lambda(t|p_L) + [1-\gamma(t)]\lambda(t|p_H)$$

$$\lambda'(t) = \gamma'(t)\big[\lambda(t|p_L) - \lambda(t|p_H)\big] + \gamma(t)\lambda'(t|p_L) + [1-\gamma(t)]\lambda'(t|p_H)$$

where $\gamma(t) = \frac{1}{1+\eta(t)}$, $\eta(t) = \left(\frac{1-\phi_I}{\phi_I}\right)e^{-[\Lambda(t|p_H)-\Lambda(t|p_L)]}$, $\Lambda(t|p) = \int_0^t \lambda(s|p)ds$, *and* $\eta'(t) = \eta(t)\big[\lambda(t|p_L) - \lambda(t|p_H)\big]$.

Proof. The conditional survivor function is given by $S(t|p) = e^{-\Lambda(t|p)}$. Then, the unconditional survivor function is given by $S(t) = \phi_I e^{-\Lambda(t|p_L)} + (1-\phi_I)e^{-\Lambda(t|p_H)}$, and the unconditional hazard is given by $\lambda(t) = -d\ln S(t)/dt$, then by the First Fundamental Theorem of Calculus:

$$\lambda(t) = \frac{\phi_I\lambda(t|p_L)e^{-\Lambda(t|p_L)} + (1-\phi_I)\lambda(t|p_H)e^{-\Lambda(t|p_H)}}{\phi_I e^{-\Lambda(t|p_L)} + (1-\phi_I)e^{-\Lambda(t|p_H)}} = \gamma(t)\lambda(t|p_L) + [1-\gamma(t)]\lambda(t|p_H)$$

and

$$\gamma(t) = \frac{\phi_I e^{-\Lambda(t|p_L)}}{\phi_I e^{-\Lambda(t|p_L)} + (1-\phi_I)e^{-\Lambda(t|p_H)}} = \frac{1}{1+\eta(t)}$$
$$\eta(t) = \left(\frac{1-\phi_I}{\phi_I}\right)e^{-[\Lambda(t|p_H)-\Lambda(t|p_L)]},$$

so that $\eta(t) > 0$. $\lambda'(t)$ is straightforward and applying the First Fundamental Theorem of Calculus again we have:

$$\eta'(t) = \eta(t)[\lambda(t|p_L) - \lambda(t|p_H)].$$

□

B.2.1 Proof of Proposition 3

Proof. From Lemma 9 and Proposition 2 we have that:

$$\eta'(t) = \eta(t)[\mu(p_L) - \mu(p_H)] < 0,$$
$$\gamma'(t) = -\gamma(t)^2\eta(t)[\mu(p_L) - \mu(p_H)] > 0, \text{ and}$$
$$\lambda'(t) = \gamma'(t)[\mu(p_L) - \mu(p_H)] < 0.$$

□

B.2.2 Proof of Proposition 5

Proof. From Lemma 9 and Proposition 4 we have that:

$$\eta'(t) = \eta(t)(1-\kappa)^t\big[\mu(p_L) - \mu(p_H)\big] < 0,$$
$$\gamma'(t) = -\gamma(t)^2\eta(t)(1-\kappa)^t\big[\mu(p_L) - \mu(p_H)\big] > 0, \text{ and}$$
$$\lambda'(t) = \gamma(t)(1-\kappa)^t[\mu(p_L) - \mu(p_H)]^2\left[\frac{\ln(1-\kappa)}{\mu(p_L) - \mu(p_H)} - \gamma(t)\eta(t)(1-\kappa)^t\right],$$

where each term in the square brackets is positive. However, the first term is constant while the second one decreases with time. To see this, define $\Phi(t) = \gamma(t)\eta(t)(1-\kappa)^t$, then it is easy to check that

$$\Phi'(t) = \frac{\eta'(t)}{[1+\eta(t)]^2}(1-\kappa)^t + \frac{\eta(t)}{1+\eta(t)}(1-\kappa)^t \ln(1-\kappa) < 0$$

where negativity follows from $\eta'(t) < 0$ and $\kappa \in (0,1)$. Evaluating $\Phi(t)$ at $t = 0$ we find that $\Phi(0) = 1 - \phi_I$, therefore:

(i) if $\ln(1-\kappa)/[\mu(p_L) - \mu(p_H)] > (1 - \phi_I)$, then the term in square brackets is always positive, and

(ii) if $\ln(1-\kappa)/[\mu(p_L) - \mu(p_H)] < (1-\phi_I)$, then the term in square brackets is initially negative, but becomes eventually positive, so that $\lambda(t)$ decreases initially, but eventually increases.

□

B.2.3 Proof of Proposition 7

Proof. From Lemma 9 and Proposition 6 we have that

$$\begin{aligned}
&\eta'(t) = \eta(t)[1-(1-\sigma)^t]\big[\mu(p_L) - \mu(p_H)\big] < 0 \\
&\gamma'(t) = -\gamma(t)^2\eta(t)[1-(1-\sigma)^t]\big[\mu(p_L) - \mu(p_H)\big] > 0, \text{ and} \\
&\lambda'(t) = -\gamma(t)^2\eta(t)[1-(1-\sigma)^t]^2\big[\mu(p_L) - \mu(p_H)\big]^2 \\
&\qquad\qquad + (1-\sigma)^t \ln(1-\sigma)\big[\bar{\mu} - \phi\mu(p_L) - (1-\phi)\mu(p_H)\big].
\end{aligned}$$

Note that in the definition of $\eta(t)$ in Lemma 9, ϕ replaces ϕ_I. Inspection of $\lambda'(t)$ reveals that for low values of t, the second term dominates but it is eventually overtaken by the first term, much more faster the higher σ is. Next, evaluating $\lambda'(t)$ at $t = 0$, we find

$$\lambda'(0) = \ln(1-\sigma)\big[\bar{\mu} - \phi\mu(p_L) - (1-\phi)\mu(p_H)\big].$$

Therefore, if the term in square brackets is:

(i) Positive, then the hazard is monotonically decreasing.

(ii) Negative, then the hazard increases initially, but eventually decreases.

(iii) Zero, then the hazard is initially flat, but eventually decreases.

□

C Imputing Interval-censored Elapsed Duration

In the Monte Carlo experiment, the duration data is generated taking into account the features of the ENOE, primarily: (i) only individuals with job start before the previous calendar year have interval elapsed duration, (ii) the complete duration from the stock sample is interval-censored.

To construct one stock sample, we made repeated draws from spells until all individuals accumulated enough duration to reach certain point in time, which represents the moment at which the stock sample is taken. The spell in which an individual reaches the stock sampling point is taken as that individual's duration of employment. We assumed that the individuals' duration has a Weibull-Gamma distribution, and tried six different parameter sets to cover cases with positive, negative and no duration dependence, as well as cases with and without unobserved heterogeneity. We only simulate stock sample data.

We tried three imputation methods. Let $\tilde{E}_i$ be the imputed duration, then we tried: (i) $\tilde{E}_i = E_i^L$, (ii) $\tilde{E}_i = E_i^R$, and (iii) $\tilde{E}_i = (E_i^L + E_i^R)/2$. The simulation exercise shows that using either E_i^L or E_i^R yields very poor results, and that using the midpoint in the interval yields the best results. We also tried a random draw from the interval using the uniform distribution in the interval, but the results are very similar to those using the midpoint. Thus, we use this midpoint imputation method for estimation.

D Minimization Algorithm to Find Parameters of the Employer Learning Model

The estimated hazard suggest starting values for $\big(\bar{\mu}, \mu(p_L), \mu(p_H)\big)$. In particular, by Condition 1, $Q(p_H) < Q < Q(p_L)$, and so $\mu(p_L) < \bar{\mu} < \mu(p_H)$. However, the estimated hazard in Figure 5 suggests that $Q \approx Q(p_L)$. This is because at $t = 0$ the hazard must equal $\bar{\mu}$ and for longer durations the hazard must equal $\mu(p_L)$. Then, we set $\mu(p_L) = 0.99 \cdot \bar{\mu}$, so that $\mu(p_L)$ is arbitrarily close to, but below $\bar{\mu}$, and use $\exp(\bar{x}'\hat{\beta})\hat{\lambda}_1 = 0.036$ as a starting value for $\bar{\mu}$. Similarly, we know that $\mu(p_H)$ must be higher than the maximum of the hazard function, then we use $\exp(\bar{x}'\hat{\beta})\hat{\lambda}_2 = 0.281$ as a starting value for $\mu(p_H)$. The estimated hazard does not provide much information to select starting values for (σ, ϕ). Hence we use different starting values given by $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ for each parameter. This gives a total of 25 different starting values, in all cases we use $\overline{T} = 50$. For all starting values, the resulting vector of parameters is: $\bar{\mu} = 0.0576$, $\mu(p_H) = 1.0$, $\phi = 0.5416$, and $\sigma = 0.1437$.