Firing Costs and Stigma: A Theoretical Analysis and Evidence from Micro Data.

Patrizia Canziani
J. P. Morgan Securities Ltd (London)[*]

Barbara Petrongolo
Universidad Carlos III (Madrid)
and CEP (LSE)

November 1999[†]

Abstract

This paper uses a general equilibrium search model to study the effects of firing costs in the presence of imperfect information about workers' ability. Firing costs change the way firms form expectations about workers' abilities from their employment history. The model exhibits the standard implication that firing costs lower the option value of hiring workers of uncertain productivity, thus youth unemployment is higher. More importantly, firing costs increase the stigma associated with a bad employment history which may lead to higher long-term unemployment.

Using micro data on labor market transitions, we test and confirm the model's prediction that firing costs increase the stigma of poor employment histories.

Keywords: firing costs, stigma, unemployment.

JEL classification: D83, J64, J65.

---

[*]Patrizia Canziani worked on this paper while at the Massachusetts Institute of Technology. The views expressed in the paper are the personal views of the author and do not necessarily represent the views and opinions of J.P. Morgan.

[†]We have received valuable insights from Daron Acemoglu, Olivier Blanchard and Peter Diamond. Thanks also to Pierre-Olivier Gourinchas, Jonathan Parker, Robert Shimer, Rodrigo Valdes and two anonymous referees for many helpful discussions and comments, and to Sergi Jiménez for providing the data. B. Petrongolo gratefully acknowledges financial support from the Spanish Ministry of Education (Grant No. DGES PB97-0091).

1 Introduction

This paper analyzes …rms' labor demand and the relative incidence of youth and long-term unemployment in economies with imperfect information and …ring costs.

There are several motivations for this work. First, unemployment is not only higher in Europe than in the US, but is also relatively more concentrated on young workers. More speci…cally, countries with stricter employment protection legislations tend to have a higher ratio of youth to adult unemployment rate (see OECD 1994a, Table 1.13 and OECD 1994b, Table 6.7). This evidence points out a possible relationship between a …rm's cost to adjust its employment level and the limited information problem of learning about the quality of workers new to the labor market.

Second, because we observe …rms devoting time and resources to learn about the quality of job applicants, the inclusion of imperfect information to a model of …ring costs is important. The e¤ect of imperfect information about workers' abilities on labor market outcomes has long been recognized (see Akerlof 1970, Spence 1973 and, more closely related to this work, Greenwald 1986, Gibbons and Katz 1991, and Riordan and Staiger 1993). This literature, however, generally neglects to account for …ring costs.

Finally, exit rates out of unemployment seem lower in countries with high …ring costs (see Machin and Manning, 1999). This paper argues that, as …ring costs increase, so does the stigma attached to bad employment histories. This helps to explain why high-…ring-cost economies have higher unemployment duration.

This work includes imperfect information about workers' abilities in a labor market equilibrium model with …ring costs, and argues that policies that make it costly for …rms to adjust their employment levels have an impact on the distribution of unemployment. This occurs because …ring costs simultaneously decrease the option value of hiring a worker of uncertain productivity, while increasing the stigma associated with bad employment histories. In economies with lower …ring costs being …red is relatively more common, so there is less stigma attached to it, making it easier for a …red person to …nd a job. Conversely, in high-…ring-cost economies, because few workers are …red, those who are become stigmatized as low-productivity individuals and have greater di¢culty …nding another job.

Four main results are obtained using a simple overlapping generation model, in which heterogeneous workers stay in the labor market for two periods. First, when …ring costs are high the option value of hiring a worker new to the market is lower, resulting in higher youth unemployment. Second, workers who have been …red su¤er from a stigma because perspective employers assume they have lower productivity than workers new to the market. This results in a novel implication: the stigma from bad employment histories is increased by …ring costs. The more di¢cult it is for …rms to dismiss low-productivity workers, the greater the stigma associated with those who are …red. This implies that it is harder for a worker to …nd a new job after being …red, and there may be higher long term unemployment as stigmatized workers are not hired.

Third, in high-…ring-cost economies the quality of those who were never hired would improve as …rms raise their hiring standards. On its own, this would increase hiring and reduce unemployment of those that are not new to the market. In our model, this only slightly mitigates againts the stigma e¤ect that …ring costs produce on those who have been …red.

Finally, …ring costs may inhibit the posting of job vacancies, through their e¤ect on the expected pro…tability of a job-worker pair.

The empirical analysis of the paper tests the prediction that …ring costs increase the stigma from being …red and therefore reduce re-employment probabilities of displaced workers. We focus on Spain, that in the mid 1980's relaxed employment protection legislation through the introduction of …xed-term contracts. Such typology of contracts gives employers the opportunity to hire a worker and learn better about her ability. When the contract expires, the …rm can choose to keep the worker o¤ering her a regular contract of undetermined duration. Alternatively, the worker can be easily dismissed. In this way, the adoption of …xed term contracts corresponds to a decrease in …ring costs.

To the extent that lower …ring costs indeed lower the stigma attached to bad employment histories, we would expect unemployment spells to be shorter after the termination of a …xed-term contract, with respect to the case of layo¤ within a permanent contract. Such prediction is tested using micro data, by estimating a duration model for unemployment. Our …ndings are that those who terminated a …xed-term contract or quit voluntarily their previous job experience signi…cantly shorter unemployment spells than those who were dismissed through costly …ring procedures.

By building a model containing both imperfect information and …ring costs, this paper brings together two strands of literature.

The consequences of …ring costs on unemployment have long been analyzed by economists. In a widely cited paper, Bentolila and Bertola (1990) propose a partial equilibrium model of labor demand in the presence of …ring costs. They suggest two opposing e¤ects of turnover costs on unemployment. On the one hand, …rms may be less likely to dismiss workers in response to adverse shocks, possibly waiting for their situation to improve. On the other hand, given that dismissals are costly, …rms may be less likely to hire workers in response to positive shocks, possibly waiting to see whether the situation persists before committing to hiring additional workers. The overall e¤ect on aggregate employment is therefore ambiguous. Bertola (1990) also provides a model with a dynamic labor demand and linear costs of employment adjustment. He shows that job security provisions neither bias labor demand toward lower average unemployment at given wages, nor do they bias labor demand toward higher wages. In both Bentolila and Bertola (1990) and Bertola (1990) …rms are forward-looking and take into account the dynamic behavior of wages, productivity, and demand conditions. Hopenhayn and Rogerson (1993) concentrate instead on the welfare impact of government policies that make it costly for …rms to adjust their employment levels. They build a general equilibrium model and …nd that a tax on job destruction has a signi…cant negative impact on total employment, lowers average productivity, and reduces utility in terms of consumption.

The novel contribution of the present analysis consists in introducing imperfect information on workers' ability in a labor market with …ring costs. The e¤ect of imperfect information on job tenure has been recognized by Jovanovic (1979, 1984). Similarly as in his models, here we treat job matches as "experience goods", i.e. instruments that can be used by employers to update their beliefs about workers's abilities. Stigma e¤ects have also been studies in the context of the asymmetric information literature. In particular, Gibbons and Katz (1991) assume that a worker's current employer is better informed about the worker's ability than prospective employers. If …rms have discretion over whom to …re, then the decision to dismiss somebody signals the market that the worker is probably of low ability, so that past employment histories are important determinants of wages and unemployment durations. We embody this kind of imperfect information problem in an environment with …ring restrictions in order to draw conclusions on the distribution of unemployment, and explain why high-…ring-cost economies tend to experience a higher incidence of youth and long-term unemployment.

The organization of the paper is as follows: Section 2 introduces our theoretical framework with imperfect information on workers' ability and …ring costs. The main results concerning the level and composition of unemployment are derived in Section 3. Section 4 closes our model by endogenizing …rms' vacancy posting decisions. Section 5 empirically tests that, in a high-…ring-cost environment, those who were …red have worse reemployment prospects than other categories of job losers. Finally, Section 6 summarizes and interprets our …ndings.

2 Problem Setup

We consider a labor market where firms and workers search for matches and produce output in discrete time. Firms are homogeneous, but workers are heterogeneous.

2.1 Workers

At the beginning of each period a new generation of mass 1 enters the labor market, and stays there for two periods. Let us call "young" and "old" the people in their first and second period respectively.

Workers have different abilities depending on observable and unobservable characteristics. We assume individual ability is exogenous and constant over time. Let a_i be the measure of the worker's productivity; the subscript "i" denotes individuals. Firms have some prior beliefs about a_i. Specifically, assume that this prior is normally distributed with mean $\bar{a}$ and precision h_a (equal to the inverse of the variance):[1]

$$\text{prior about } a_i \sim N\left(\bar{a}; \frac{1}{h_a}\right) \tag{1}$$

Figure 1 shows the signalling structure of the model, under the assumption that unemployed workers always meet with a firm in the beginning of a period. Firms do not know a worker's ability or productivity, nor workers have a better knowledge of their productivity than firms. However, when they meet an individual, firms receive an imperfect signal of her productivity, s_i, through interviews and the curriculum vitae, and decide whether to hire her or not. The signal is defined as

$$s_i \equiv a_i + \varepsilon_i; \tag{2}$$

where ε_i is independently distributed as a normal:

$$\varepsilon_i \sim N\left(0; \frac{1}{h_\varepsilon}\right): \tag{3}$$

So that the distribution for one signal s_i is:

$$s_i \sim N\left(\bar{a}; \sigma_s^2\right); \qquad \sigma_s^2 = \frac{1}{h_a} + \frac{1}{h_\varepsilon}: \tag{4}$$

[1] According to (1), workers are all equally "risky" from employers' perspective. Lazear (1995) provides instead a model in which workers of uncertain ability compete for jobs against those whose ability is known with certainty. Interestingly, it is shown that risky workers tend to be prefereed to safe ones, because firms are willing to pay more in order to hire a worker with upside potential. If risky and safe workers coexisted in our model, firing and stigma would clearly apply to risky ones.

As time progresses, the firm that hired a worker has a better, though still imperfect, knowledge of her productivity. We assume that the firm observes a second signal of the worker's productivity s_i' at the end of the first period and on the basis of s_i and s_i' it decides whether to keep or fire the worker.

The distribution for the average of two independent signals s_i and s_i' is:

$$\bar{s}_i \equiv \frac{s_i + s_i'}{2} \sim N\left(a, \sigma_{\bar{s}}^2\right); \qquad \sigma_{\bar{s}}^2 = \frac{1}{h_a} + \frac{1}{2h_\varepsilon} \tag{5}$$

Were the process to continue, the distribution of the mean of n independent signals would become more and more concentrated as the number of signals increases.

If an individual worked and was not fired when young, she keeps working for the same firm in the second period of her working life. Otherwise, as an unemployed old worker she searches for a new job. When a firm meets an old unemployed person, it does not observe her productivity, but it knows her past employment history and observes a new signal of her productivity, s''. Conditional on this information a firm decides whether to hire the worker or not.

2.2 Firms

There is an infinite number of identical, infinitely lived firms. Each firm can meet only one unemployed person per period. The cost to the firm of maintaining a job vacancy is $c > 0$ per period. The cost to fire a worker is $f > 0$. The production function is

$$y = \sum_i a_i \tag{6}$$

Firms make three decisions: how many vacancies to post, whether to hire a worker, and whether to fire a worker.

2.3 Matches

Firms and workers search for matches. A firm and an unemployed worker meet at the beginning of a period with a probability that is a function of total unemployment, U, and the number of vacant jobs, V.

When meeting with a worker, the firm does not know the worker's type, but it can observe the worker's (i) age, (ii) employment history, and (iii) current signal. If a worker is not offered a job when young, she remains unemployed during one period, and resumes job search when old. If she is hired, she starts working with the firm, and at the end of the period she is either kept or …red. In the latter case, she looks for another job at the beginning of the following period.

Once a job-worker pair has been agreed, the worker receives a …xed, positive wage, equal to her reservation value. Reservation wages are therefore the same for all workers, and all job o¤ers are accepted. In this set-up, we can think of a_i as individual productivity net of wages.[2]

3 Computing Equilibrium Unemployment

We start with the simplest possible model of labor demand with …ring costs and imperfect information.

For the moment, let us assume that the number of vacancies is exogenous and that all workers meet a …rm in the beginning of each period. This assumption eliminates one possible reason of unemployment – that some workers do not meet with any …rm. While this increases the simplicity of the model, it does not change its predictions.

With this simple structure, it is shown that in high-…ring-costs economies the following occurs: (i) …rms tend to hire fewer young workers, leading to high youth unemployment; (ii) …rms …re less; (iii) the stigma for being …red is higher and (iv) that of not being hired is lower; (v) old-age, total and long-term unemployment are higher if the e¤ects in (i) and (iii) overcome the e¤ects in (ii) and (iv).

3.1 Young workers

3.1.1 Updating beliefs

Before receiving any signal, …rms expect a worker to have ability $\hat{a}$, the expected value for a whole cohort of labor market entrants. This expectation is updated any time a new signal arrives.

[2] The assumption of …xed wages might seem problematic. However, on the one hand …ring costs have hardly any e¤ect if wages are perfectly ‡exible (see - among others - Lazear, 1990). On the other hand, the assumption that wages are somewhat rigid is generally well accepted, and it may interestingly be argued that, in a political economy perspective, wage compression and …ring costs tend to arise together (see Bertola and Rogerson, 1995). Our results are obtained for …xed wages because, without loss of generality, it simpli…es proofs greatly. However, sluggish wages would lead to similar results. The most intuitive way to check it is to assume that wages are proportional to the expected productivity of workers, conditional on the available information, i.e. w_i = ® E(a_ij¢); where 0 < ® < 1. In this case, net productivity would still be a monotonically increasing function of gross productivity, and all the results would follow.

Given that a_i and ϵ_i have independent normal distributions, the posterior after signal s_i has a normal distribution:

$$E(a_i|s_i) \sim N\left(\frac{h_a}{h_a + h_\epsilon}\bar{a} + \frac{h_\epsilon}{h_a + h_\epsilon}s_i; \frac{1}{h_a + h_\epsilon}\right) \tag{7}$$

Similarly, the posterior after two signals s_i and s_i^0 has a normal distribution:

$$E(a_i|s_i; s_i^0) \sim N\left(\frac{h_a}{h_a + 2h_\epsilon}\bar{a} + \frac{2h_\epsilon}{h_a + 2h_\epsilon}\bar{s}_i; \frac{1}{h_a + 2h_\epsilon}\right) \tag{8}$$

The mean of the posterior distribution is a weighted average of $\bar{a}$ and s_i (or $\bar{s}_i$), where $\bar{a}$ is the mean of the prior distribution and s_i (or $\bar{s}_i$) is the signal(s) received, as defined in equations (2)-(5). The weights are proportional to the precision of the conditional distribution of the signals of a_i for any given value of a_i, and the precision of the prior distribution of a_i. The larger the number of signals and the higher their precision, the greater their weight.

The precision of the posterior distribution increases by the amount h_ϵ with each signal that is received, regardless of the observed values. Therefore, as the number of signals increases, the distribution of the posterior becomes more and more concentrated around its mean. Moreover, the concentration must increase in a fixed and predetermined way, while the values of the mean will depend on the observed signals.

3.1.2 Hiring decisions

A firm hires a worker when the expected profits from that worker are non-negative. Let us denote π_i the profitability from hiring worker i. In general, the hiring condition is:

$$E[\pi_i|s_i; \text{age, employment history}] \geq 0 \tag{9}$$

When a firm meets a young unemployed worker, it forms expectations about her ability on the basis of s_i. For a young worker, the expected profitability is equal to the sum of the expected profitability in the current and the next period, allowing for the possibility that the worker may be fired at the end of the first period. Assuming for simplicity a discount rate equal to zero, the expected profitability is equal to:

$$\begin{aligned} E[\pi_i|s_i] = \; & \frac{h_a}{h_a + h_\epsilon}\bar{a} + \frac{h_\epsilon}{h_a + h_\epsilon}s_i + \\ & - \Pr[\text{fire}|s_i]\, f + \Pr[\text{keep}|s_i]\, E[a_i|\bar{s}_i \geq \bar{s}^0; s_i] \end{aligned} \tag{10}$$

where

$$\Pr[\text{fire}|s_i] = \Pr(\tilde{s}_i \leq \bar{s}'|s_i) = \Phi\left(\frac{\bar{s}' - \frac{1}{2}\left[\frac{h_a}{h_a+h_\varepsilon}\bar{a} + \frac{h_a+2h_\varepsilon}{h_a+h_\varepsilon}s_i\right]}{\sigma_{\tilde{s}|s}}\right) \equiv \Phi(\bar{s}'; s_i; \cdot);$$

$$\Pr[\text{keep}|s_i] = 1 - \Phi(\bar{s}'; s_i; \cdot);$$

$\bar{s}'$ is a cutoff value for $\tilde{s}_i$ above which the worker is retained, $\sigma_{\tilde{s}|s}$ is the standard deviation of $\tilde{s}$ conditional on s, and Φ is the c.d.f. of a standard normal distribution. $E(a_i|\tilde{s}_i \geq \bar{s}'; s_i)$ is computed in Appendix A and is equal to:

$$E(a_i|\tilde{s}_i \geq \bar{s}'; s_i) = \bar{a}\frac{h_a}{h_a+h_\varepsilon} + s_i\frac{h_\varepsilon}{h_a+h_\varepsilon} + \sigma_{\tilde{s}|s}\frac{2h_\varepsilon}{h_a+2h_\varepsilon}\frac{\phi(\bar{s}'; s_i; \cdot)}{1-\Phi(\bar{s}'; s_i; \cdot)};$$

where ϕ is the p.d.f. of a standard normal distribution.

Equation (10) becomes:

$$\begin{aligned} E[\pi_i|s_i] &= \frac{h_a}{h_a+h_\varepsilon}\bar{a} + \frac{h_\varepsilon}{h_a+h_\varepsilon}s_i - f\,\Phi(\bar{s}'; s_i; \cdot) + \\ &+ [1-\Phi(\bar{s}'; s_i; \cdot)]\left[\frac{h_a}{h_a+h_\varepsilon}\bar{a} + \frac{h_\varepsilon}{h_a+h_\varepsilon}s_i + \sigma_{\tilde{s}|s}\frac{2h_\varepsilon}{h_a+h_\varepsilon}\frac{\phi(\bar{s}'; \hat{s}; \cdot)}{1-\Phi(\bar{s}'; \hat{s}; \cdot)}\right] \\ &= [2-\Phi]\left[\frac{h_a}{h_a+h_\varepsilon}\bar{a} + \frac{h_\varepsilon}{h_a+h_\varepsilon}s_i\right] - f\,\Phi + \sigma_{\tilde{s}|s}\frac{2h_\varepsilon}{h_a+2h_\varepsilon}\phi. \end{aligned} \quad (11)$$

Expected profits are a continuous and monotonically increasing function of the signal s_i (proof in Appendix B), so that a cutoff level $\hat{s}$ exists such that any young worker with a signal $s_i \geq \hat{s}$ is hired, while young workers with lower signals remain unemployed.

The hiring cutoff level, $\hat{s}$, is implicitly defined by equating the expected profits to zero:

$$[2-\Phi]\left[\frac{h_a}{h_a+h_\varepsilon}\bar{a} + \frac{h_\varepsilon}{h_a+h_\varepsilon}\hat{s}\right] - f\Phi + \sigma_{\tilde{s}|s}\frac{2h_\varepsilon}{h_a+2h_\varepsilon}\phi = 0 \quad (12)$$

The hiring cutoff $\hat{s}$ is positively related to f (proof in Appendix B):

$$\frac{\partial \hat{s}}{\partial f} > 0. \quad (13)$$

When firing costs increase, the option value of hiring a new worker decreases, the hiring cutoff increase, and firms hire less.

3.1.3 Firing decisions

In the end of each period firms decide whether to keep a young worker or fire her. Given that in equilibrium the option value of a filled vacancy is zero, firms optimally fire a young worker when her expected productivity is negative and larger than the firing costs f. According to Equation (8), this corresponds to:

$$\frac{h_a}{h_a + 2h_\varepsilon}\bar{a} + \frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\bar{s}_i \quad < -f. \tag{14}$$

Because this condition is monotonic in $\bar{s}_i$, there exists a cutoff level for the average of the two signals, $\bar{s}^{*\prime}$, such that firms will keep workers for whom the mean signal is higher than $\bar{s}^{*\prime}$, and will fire others. The firing cutoff must satisfy

$$\bar{s}^{*\prime} = -\frac{h_a}{2h_\varepsilon}\bar{a} - \frac{h_a + 2h_\varepsilon}{2h_\varepsilon} f. \tag{15}$$

When the variance of the signal is high, firms fire less because it is more difficult to infer the productivity of a worker from the signals she gives. Conversely, when the variance of the distribution of ability is high, firms fire more because workers tend to have different abilities and it is easier to separate those with high and low productivity. When the average ability increases, firms fire less because the option value of keeping a worker is higher. Moreover, when firing costs increase, firms fire less:

$$\frac{\partial \bar{s}^{*\prime}}{\partial f} = -\frac{h_a + 2h_\varepsilon}{2h_\varepsilon} < 0 \tag{16}$$

Firms fire workers whose posterior expected productivity after one period of employment is negative and larger than the cost of firing them. The higher f, the lower the expected productivity has to be for an individual to be laid off (i.e. the lower the firing cutoff $\bar{s}^{*\prime}$), and the fewer people are fired.

3.1.4 Youth Unemployment

In this model the only reason why young people are unemployed is that they sent a bad signal and were not hired.

$$\begin{aligned} YU &= \text{Youth Unemployment} = \Pr[s_i < \bar{s}^*] \\ &= \Phi\left(\frac{\bar{s}^* - \bar{a}}{\sigma_s}\right) \end{aligned} \tag{17}$$

The level of youth unemployment is related to the distribution of ability in the population: the more numerous "good" workers (i.e. the higher $\bar{a}$), the better the average quality of the unemployed, and the more young people are hired by firms. Furthermore, firing costs increase youth unemployment by raising $\hat{s}$.

3.2 Old Workers

3.2.1 Hiring Decisions

At the beginning of each period an old worker might be unemployed either because she was fired or because she was not hired when young. When a firm meets an old worker at the beginning of a period, it observes her age, her current signal, and her employment history. On the basis of this information it forms expectations about the worker's productivity and hires her if the expected profitability from hiring is non-negative. Given that old workers only have one period ahead, their expected profitability is simply given by their expected current productivity.

Let s_i'' denote the signal an old unemployed worker sends when she meets a firm. The expected productivity of an old worker who did not work when young and currently gives signal s_i'' is equal to (proof in Appendix C):

$$\begin{aligned} E[a_i \mid \text{not hired when young}; s_i''] &= E[a_i \mid s_i < \hat{s}; s_i''] \\ &= \bar{a}\frac{h_a}{h_a + h_\varepsilon} + s_i''\frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''}\frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\frac{\phi(\hat{s}; s_i''; \cdot)}{\Phi(\hat{s}; s_i''; \cdot)} \end{aligned} \tag{18}$$

where $\sigma_{s|s''}$ is the standard deviation of s conditional on s''. The posterior for the productivity of a worker who was not hired in the past is increasing in the level of the current signal and the mean of the prior: if the average quality of workers is high, or if the current signal is very good, the prospective employer tends to think that the low signal in the previous period was not indicative of low productivity, but came from noise.

As we would expect, the one-period expected productivity is lower for a worker that was not hired in the past than for a worker new to the market (proof in Appendix D)

$$E[a_i \mid \text{not hired when young}; s_i''] < E[a_i \mid s_i'']. \tag{19}$$

The fact that the worker sent a bad signal when young affects firm's expectations about her productivity when old. However, the stigma for not being hired decreases with firing costs. This can be seen in equation (18): when firing costs go up, the hiring cutoff $\hat{s}$ rises, so that the whole ratio $\frac{\phi(\cdot)}{\Phi(\cdot)}$ decreases. This is also intuitive: the higher the firing costs, the harder it is for firms to fire even non-productive workers, resulting in firms being less keen to hire new workers. When turnover costs are high, the average quality of the workers that are not hired increases.

Similarly, the expected productivity of an old worker who was fired in the previous period is (proof in Appendix E):

$$
\begin{aligned}
E[a_i | \text{fired}; s_i''] &= E[a_i | \bar{s}_i < \hat{s}'; s_i''] \\
&= \bar{a} \frac{h_a}{h_a + h_\varepsilon} + s_i'' \frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{\bar{s}|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\phi(\hat{s}'; s_i''; \cdot)}{\Phi(\hat{s}'; s_i''; \cdot)} \qquad (20)
\end{aligned}
$$

where $\sigma_{\bar{s}|s''}$ is the standard deviation of $\bar{s}$ conditional on s''. We can see that (proof in Appendix F)

$$E[a_i | \text{fired}; s_i''] < E[a_i | s_i'']. \qquad (21)$$

Moreover,

$$E[a_i | \text{fired}; s_i''] < E[a_i | \text{not hired when young}; s_i''] \qquad (22)$$

if

$$\frac{\sigma_{\bar{s}|s''}}{\sigma_{s|s''}} > \frac{\phi(\hat{s}'; s_i''; \cdot)/\Phi(\hat{s}'; s_i''; \cdot)}{\phi(\hat{s}; s_i''; \cdot)/\Phi(\hat{s}; s_i''; \cdot)}. \qquad (23)$$

The LHS is always greater than 1. The RHS is continuous and increasing in f. Firing costs affect the numerator and the denominator in opposite ways. In particular, when f increases, $\frac{\phi(\hat{s}';s_i'';\cdot)}{\Phi(\hat{s}';s_i'';\cdot)}$ increases and $\frac{\phi(\hat{s};s_i'';\cdot)}{\Phi(\hat{s};s_i'';\cdot)}$ decreases. So that when $f \to 0$ the whole ratio in the RHS of (23) converges to a positive limit smaller than 1; while when $f \to \infty$ the ratio goes to $+\infty$.

In other words, the condition in (23) is satisfied if firing costs are high enough. The economic intuition for these effects is the following. When firing costs are very low firms are more willing to hire even people with low signals. Then, the stigma for not being hired when young is very high. On the other hand, since firing costs are low, firms are ready to adjust their employment levels by firing workers in case they do not prove to be as good as expected. Relatively many workers are fired when firing costs are significantly low, and the stigma for being fired is low. When firing costs are sufficiently high, the stigma for being fired is stronger than that for not being hired. A worker who has sent a low signal when young and was not hired has a higher expected ability than a worker who was employed but was fired on the basis of more than one signal about her ability.

We can also see these effects by looking at the hiring cutoffs for old workers. Firms hire old workers if their expected productivity is non-negative. Given the monotonicity of the expected productivity in the signal, firms will hire old workers whose signal is above a certain cutoff. Let $\hat{s}^{nh}$ and $\hat{s}^{f}$ denote the hiring cutoffs for workers that were not hired or were fired when young, respectively. The hiring cutoffs are implicitly defined by equating to zero the relevant profitability function.

In particular, $\hat{s}^{nh}$ is implicitly defined as the level of the signal s'' such that the profitability in Equation (18) is equal to zero:

$$\bar{a}\,\frac{h_a}{h_a + h_\varepsilon} + \hat{s}^{nh}\,\frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''}\,\frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\,\frac{\phi(\hat{s}; \hat{s}^{nh}; \cdot)}{\Phi(\hat{s}; \hat{s}^{nh}; \cdot)} = 0 \tag{24}$$

Similarly, $\hat{s}^f$ is implicitly defined as the level of the signal s'' such that the profitability in Equation (20) is equal to zero:

$$\bar{a}\,\frac{h_a}{h_a + h_\varepsilon} + \hat{s}^f\,\frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''}\,\frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\,\frac{\phi(\hat{s}_0; \hat{s}^f; \cdot)}{\Phi(\hat{s}_0; \hat{s}^f; \cdot)} = 0 \tag{25}$$

The two hiring cutoffs for old workers, $\hat{s}^{nh}$ and $\hat{s}^f$ react in opposite ways to an increase in firing costs (proofs in Appendixes G and H):

$$\frac{\partial \hat{s}^{nh}}{\partial f} < 0 \tag{26}$$

and

$$\frac{\partial \hat{s}^{f}}{\partial f} > 0. \tag{27}$$

When firing costs increase, fewer young workers are hired as firms set higher hiring cutoffs, $\hat{s}$. Thus, as f goes up the expected ability of non-hired workers increases, and $\hat{s}^{nh}$ decreases. The higher firing costs, the lower the stigma for not being hired. Conversely, when f is higher, less young people are fired, and $\hat{s}^f$ tends to be large; when it is more expensive to fire workers, being fired carries a stronger stigma.

3.2.2 Old-Age Unemployment

Old unemployed are those workers who were never hired or were fired when young, and were not employed when old because they did not send a high enough signal to compensate for their negative past employment history. Therefore:

$$\begin{aligned} OU &= \text{Old-Age Unemployment} = \\ &= \Pr[\text{not hired when young}]\Pr[s'' < \hat{s}^{nh} | \text{not hired when young}] \\ &\quad + \Pr[\text{fired}]\Pr[s'' < \hat{s}^f | \text{fired}] \\ &= \Phi(\hat{s}; \cdot)\,\Phi\left(\hat{s}^{nh}; \cdot\right) + [1 - \Phi(\hat{s}; \cdot)]\,\Phi(\hat{s}_0; \cdot)\,\Phi\left(\hat{s}^f; \cdot\right) \end{aligned} \tag{28}$$

The higher the prior about the ability in the population, the lower is the level of the unemployment of the "old" workers. Firing costs have ambiguous effects on the old-age unemployment: firing costs increase the probability that a young individual is not hired $\Phi(\hat{s}; ¢)$, and that an old worker that was fired will not be re-employed $\Phi(\hat{s}^{f}; ¢)$, but they also decrease the probability of firing somebody $\Phi(\hat{s}^{10}; ¢)$ and the probability that those who are not hired when young find a job when old $\Phi(\hat{s}^{nh}; ¢)$.

3.2.3 Long Term Unemployment

In this simple economy the long term unemployed are all those workers who never worked because they were not hired in either period. Therefore:

$$\begin{aligned} LTU &= \text{Long-Term Unemployment} = \\ &= \Pr[\text{not hired when young}] \Pr[\text{not hired when old} | \text{not hired when young}] = \\ &= \Phi(\hat{s}; ¢)\Phi(\hat{s}^{nh}; ¢). \end{aligned}$$

LTU increases with firing costs if:

$$\frac{\partial LTU}{\partial f} = \frac{\partial \Phi(\hat{s})}{\partial f}\Phi(\hat{s}^{nh}) + \frac{\partial \Phi(\hat{s}^{nh})}{\partial f}\Phi(\hat{s}) > 0 \qquad (29)$$

The first term in the RHS of (29) is positive. It represents the negative effect of firing costs on firms' hiring decisions for young workers. The higher f, the fewer young job applicants will be employed, and the more young workers will be unemployed. When this cohort is ageing, they have a lower probability of being hired than new entries, thus increasing the incidence of long-term unemployment. The second term is negative. It represents the beneficial effect of firing costs on the stigma of not being hired when young. When f is high many workers are not hired when young. When this cohort is ageing, they have a higher probability of being hired than dismissed workers, potentially reducing long-term unemployment.

4 Endogenizing vacancies

In this section we drop the assumption that all unemployed workers meet a firm in the beginning of a period. Suppose that young workers firms with a positive probability μ. In the case there is a contact between the two parties, the model works as in the previous section. The main difference is that young workers can now be unemployed not only because they met a firm and were not hired, but also because they did not meet a firm.

When a firm meets an old unemployed worker it forms expectations about her productivity. Firms observe the past employment history of old job applicants. In particular, not only can firms distinguish between workers that were fired and workers who did not have a job; they also observe whether the worker did not have a job because she did not meet a firm, or because she met a firm but was not hired. In case the worker did not meet a firm in the previous period, her expected productivity is simply equal to the one-period expected productivity updated for the current signal. However, if the worker is unemployed because she was not hired or was fired when young, the prospective employer takes her negative employment history into consideration when forming expectations about her productivity.

The main point of this section is that firing costs may affect the posting of new vacancies.

How many new jobs are being posted depends on the cost of maintaining a job vacancy, the probability of filling it, and the expected profitability of the job-worker pair. Let us suppose that the cost of maintaining an open vacancy c is fixed. The probability of filling a vacancy depends on an aggregate matching function, which gives the number of job matches formed in terms of the inputs of firms and workers into the search process. With U unemployed workers and V vacant jobs at the beginning of each period, job matches are given by the function:

$$x = AV^{\alpha}U^{1-\alpha} \tag{30}$$

with $0 \leq \alpha \leq 1$ and $A \geq 0$.[3] Unemployed workers move into employment according to a Poisson process with rate $\mu = x/U = A(V/U)^{\alpha}$, while the process that changes the state of vacant jobs is Poisson with rate $x/V = A(U/V)^{1-\alpha}$:

Let us call ω the fraction of unemployed that are old in the beginning of each period. Then, a firm will profitably post a vacancy if the following condition holds:

$$c \leq A(U/V)^{1-\alpha}\left[(1-\omega)E(a_i|young; s_i) + \omega E(a_i|old; s_i'')\right] \tag{31}$$

In equilibrium all profit opportunities from new jobs are exploited. Therefore, the vacancy posting condition (31) holds with equality.

The expected productivity of a young worker was already computed in Equation (10). When firing costs are higher the expected productivity of young workers decreases as it is not possible to fire those workers whose posterior about productivity is negative but smaller than firing costs.

The expected productivity of an old worker can be computed as

$$\begin{aligned} E[a_i|old; s_i''] &= \frac{U^{nm}}{\omega}E(a_i|did\ not\ meet; s_i'') + \\ &+ \frac{U^{nh}}{\omega}E(a_i|not\ hired\ when\ young; s_i'') + \frac{U^{f}}{\omega}E(a_i|fired; s_i'') \end{aligned} \tag{32}$$

[3] See, among others, Pissarides (1990), chapter 1, for the underlying motivation.

where U^{nm}; U^{nh} and U^{f} denote the proportion, among the old unemployed, of those who never met a firm, met a firm but were not hired, and were hired and subsequently fired, respectively. $E(a_i|\text{did not meet}; s_i'')$ is simply equal to $-\frac{h_a}{h_a+h_2}\bar{a} + \frac{h_2}{h_a+h_2}s_i''$ and unaffected by firing costs. $E(a_i|\text{not hired}; s_i'')$ is computed in equation (18), and increases with firing costs. Finally, $E(a_i|\text{fired}; s_i'')$ is given by (20) and decreases with firing costs.

Similarly as for old-age unemployment, for old-age productivity to decrease with firing costs the tendency to hire fewer displaced workers has to be stronger than the tendency to hire more among those that were not hired upon entry . This is a sufficient - although not necessary - condition for firms to post fewer vacancies, as shown in the free entry condition (31).

5 Empirical Analysis

This section tests the main prediction of the preceding analysis by empirically studying the effects of poor labor market histories and stigma on workers' re-employment prospects. In doing this, we focus on one country, Spain. This is firstly because Spain provides an interesting case study of a two-tier system where firing regulations only apply to the termination of some of existing labor contracts. This allows to verify whether workers who lost a job protected by firing costs suffer from higher stigma and therefore lower job-finding probabilities than others. Secondly, the panel version of the Spanish Labor Force Survey contains abundant information on workers' transitions across labor market states and on the type of contract held, providing the adequate data set for our analysis.

5.1 Fixed-term labor contracts

Until mid-1980s, the Spanish labor market was one of the most heavily regulated among the OECD set, especially as far as severance payments and firing restrictions were concerned (see OECD 1994a,b). This situation, combined with a 20% unemployment rate, has probably triggered the experiment of "flexibility at the margin", launched in 1984 with the introduction of a new typology of labor contract, characterized by limited duration and negligible firing costs. Such contracts have been massively used: soon after their introduction, as much as 98% of newly registered contracts were of this type[4] and a decade later one third of Spanish employees was holding a fixed-term contract. At the same time, all

[4] See Bentolila and Saint-Paul (1992).

…ring restrictions on permanent contracts were left unchanged.[5]

Fixed-term contracts give employers the opportunity to hire a worker and learn better about her ability. Upon expiry, the …rm can choose to keep the worker o¤ering her a regular contract of undetermined duration. Alternatively, the worker can be easily dismissed. In this way, the adoption of …xed term contracts corresponds to a decrease in …ring costs. To the extent that lower …ring costs indeed lower the stigma attached to poor employment histories, we should expect workers who terminated a …xed-term contract to face shorter unemployment spells than those that were (costly) dismissed from a permanent position.

To carry out this analysis, we estimate a duration model of unemployment, including an individual control for the reason of job loss. Previous work that analyzed unemployment exit rates in Spain (Ahn and Ugidos 1995, García Pérez 1997, and Bover et al. 1998) did not focus on such an issue. An exception is Alba (1998), who estimated a logit model for unemployment termination and found a non-signi…cant e¤ect of layo¤s on youth exit rates. In our analysis we consider workers of all ages, and adopt a duration model of exit rates. Duration models should adequately describe the dynamics of the transition from unemployment to employment by exploiting the potential strength of cohort panel studies. Such studies allow in fact to track individuals over time, and observe exactly how long they take to make a transition.

Below we describe our data, and set-up an econometric model that would …t the structure of our data set. Finally, we provide the estimation results.

5.2 The data

The data used come from the Spanish Labor Force Survey (Encuesta de Población Activa), carried out every quarter on a sample of some 60,000 households. It is designed to be representative of the total Spanish population, and contains very detailed information about labor force status of individuals. Each household remains in the survey for a maximum of six consecutive quarters: each quarter a new cohort is selected, and one sixth of households leave the sample. Labor force transitions are analyzed exploiting the panel structure of the survey (EPA enlazada), available for the period 1987 onwards.

Our sample includes individuals belonging to all cohorts who entered the survey between 1987:2 and 1996:3 and completed six quarterly interviews. This allows us to monitor re-employment probabilities over (more than) a full cycle of the Spanish economy. We select workers who reported to be jobless and looking for a job, excluding those who had never been employed before, those who entered the military service, those who retired, and those

[5] This was true until the 1997 reform, that created a new type of permanent contract with lower …ring costs, targeted at speci…c categories of workers.

with missing information on the cause of termination of their last employment spell. This restricts us to a …nal sample of 91,664 individuals. More speci…cally, we concentrate on the transition out of the …rst unemployment spell that is observed during the survey period.

Roughly 60% of our observations are unemployment entrants, so we observe the start of their spell. The remaining 40% of spells started before the worker was selected for the survey, so that we condition the hazards on elapsed duration of search, using the information on the uncompleted duration of the current spell that is reported at the …rst interview. This duration is reported in months if it is lower than two years, and in years otherwise. Such data bunching problem could be eliminated by focusing only on entrants into unemployment, as in Bover et al. (1998) and Alba (1998). However, this procedure has the disadvantage of removing from the sample all the long-term unemployed. We therefore prefer to use information on both the unemployment in‡ow and the unemployment stock, bearing in mind that the baseline hazard would somehow re‡ect the heaps in reported durations.

Total unemployment duration is computed as the number of consecutive quarters the individual is observed as unemployed during the survey period, plus the elapsed duration (if any). A spell is considered as completed when either the individual declares to be employed or to have abandoned job search. Durations longer than 16 quarters are treated as censored at 17 quarters.

The hazard model estimated includes personal and family characteristics of individuals such as gender, age, education, marital status and number of dependent children; a set of industry dummies that refer to the last job held; a dummy that indicates whether the worker is receiving unemployment bene…ts; and the cause of employment termination. The inclusion of this last variable should shed light on the e¤ect of past employment histories on future job-…nding prospects. In particular, we distinguish among four possible causes of termination: layo¤, quit, termination of a general …xed-term contract, and termination of a …xed-term contract with a speci…c cause (including seasonal jobs and contracts designed for speci…c projects). According to our data, laid-o¤ workers experience longest unemployment spells, and are least likely to …nd a new job within the survey period, as illustrated in Table 1.

Finally, year dummies capture the e¤ect of business cycle ‡uctuations, and the unemployment rate of the province of residence represents an indirect measure of local labor demand. Unemployment data are obtained from the INE Tempus database, and merged to our sample using the provincial indicator attached to each individual record …le.

Table 2 reports some descriptive statistics of our sample.

5.3 Econometric specification

The structure of our data set requires a discrete time hazard function approach, as outlined in Narendranathan and Stewart (1993). Suppose that the transition out of unemployment is a continuous process with hazard

$$\mu_i(t) = \lambda(t)\exp(x_i'\beta); \tag{33}$$

where $\lambda(t)$ denotes the baseline hazard, x is a vector of time-invariant explanatory variables, and β is a vector of unknown coefficients. The discrete time hazard denotes the probability of a spell of unemployment being completed by time $t+1$, given that it was still continuing at time t. The discrete time hazard is therefore given by

$$h_i(t) = 1 - \exp\left\{-\int_t^{t+1}\mu_i(u)du\right\} = 1 - \exp\{-\exp(x_i'\beta)\Lambda(t)\} \tag{34}$$

where $\Lambda(t) = \int_t^{t+1}\lambda(u)du$ denotes the integrated baseline hazard. We do not specify any functional form for the $\Lambda(t)$, and estimate the model semiparametrically.

In order to assess the likelihood contribution of a single spell, we need to consider the stock nature of our sample.[6] We may observe spells of unemployment that started before the survey period, and we can use self-reported information to find out the quarter in which these spells begun. In this case we need to avoid a stock sample bias (see Lancaster and Chesher, 1983), by conditioning the hazard on the elapsed duration at the first interview date. Suppose that an individual i enters the survey after j_i quarters of unemployment and stays unemployed for another k_i quarters, for a total duration $j_i + k_i$, that can be either censored or uncensored. The individual (log)likelihood contribution is therefore

$$\begin{aligned} L_i &= c_i \ln h_i(j_i + k_i) + \sum_{t=j_i}^{j_i+k_i-1} \ln\{1 - h_i(t)\} \\ &= c_i \ln(1 - \exp[-\exp(x_i'\beta)\Lambda(j_i + k_i)]) - \exp(x_i'\beta)\sum_{t=j_t}^{j_i+k_i-1}\Lambda(t). \end{aligned} \tag{35}$$

where c_i is a censoring indicator that takes the value 1 if the spell is uncensored and zero otherwise.

The model outlined in 35 is further modified in order to take into account the competing-risk nature of our problem. An unemployment spell can terminate with job-finding or alternative states. Given that we are interested in the first type of transition, we need

[6] See also Güell and Petrongolo (1999) for a similar application on the EPA.

to treat durations finishing for other reasons than job finding as censored at time of exit (see Narendranathan and Stewart, 1993). Having said this, the semi-parametric hazard specification (35) used for the single-risk model can be applied for the job finding hazard.

One further step consists in taking into account the effect of unobserved workers' characteristics. This is quite an important issue in our analysis, because more able workers are less likely to be fired and more likely to find a new job if unemployed. But this sort of correlation would tell very little on the signalling effect of the type of employment termination.

We therefore control for the effect of possibly omitted regressors in the exit from unemployment by conditioning the hazard rate on an individual's unobserved characteristics, summarized into the variable v. The conditional (continuous time) hazard rate is then written as $\theta_i(t) = \lambda(t)\exp(x_i'\beta + v_i)$, with v_i independent of x_i and t. In a competing risk framework, allowing a random disturbance term in each of the cause-specific hazards requires an additional assumption, that imposes the independence of these disturbance terms across the cause-specific hazards.[7]

The unconditional hazard (that depends on observable regressors only) is obtained by integrating the conditional one over v, under the assumption that v is distributed as a Gamma variate of unit mean and variance σ^2. Under these assumptions the likelihood is given by[8]

$$L_i = \ln\left[\left(1+\sigma^2\exp(x_i'\beta)\sum_{t=j_i}^{j_i+k_i-1}\gamma(t)\right)^{-1/\sigma^2} - c_i\left(1+\sigma^2\exp(x_i'\beta)\sum_{t=j_i}^{j_i+k_i}\gamma(t)\right)^{-1/\sigma^2}\right]. \quad (36)$$

The baseline hazard can be estimated non-parametrically by maximizing the log-likelihood $L = \sum_{i=1}^{n} L_i$ with respect to the $\gamma(t)$ terms, the vector β and the variance term σ^2. The individual and job-related characteristics included in the vector x_i are treated as time invariant, and are measured at the start of the spell (or at the time of the first interview if the spell had already started).

[7] The alternative approach would be to assume perfect correlation (as opposed to zero correlation) between the cause-specific disturbance terms (see Narendranathan and Stewart, 1993, for a discussion of advantages and disadvantages of the two methods).

[8] See Han and Hausman (1990).

5.4 Results

The results of our estimates are reported in Table 3. Two specifications of our regressions are provided. In the first, we do not allow for unobserved heterogeneity among individuals. In the second, we control for the effect of possibly omitted regressors by including a Gamma-distributed disturbance term.

Looking first at regression I, we find that re-employment probabilities are lower for women than for men, tend to decrease with age, and are enhanced by formal qualifications only when it comes to university education. Being married implies higher job-finding rates, as it does the number of dependent children in the household, due to tighter budget constraints and therefore lower reservation wages for those with numerous families.[9] For a similar argument, the receipt of unemployment benefits reduces the job-finding hazard, through an increase in the reservation wage. Displaced workers from manufacturing, construction and service industries have lower probabilities of re-entering employment than those displaced from agriculture, possibly due to stronger seasonality and turnover in this sector.

Among the controls that indicate the cause of unemployment, we find that those who quit their job or have terminated a fixed-term contract of any kind have significantly higher job-finding rates than those who were laid off. We interpret this result as evidence in favor of the main prediction of the model. Firing workers is costly to employers and having been fired attaches some stigma to the unemployed, who turn out to have worse re-employment prospects than other categories of job losers. Also, interaction terms between age and the reason of employment termination show that such stigma effect is significantly stronger for older workers. Interactions between education and the cause of unemployment have instead a non-significant effect on re-employment rates.

Concerning job quitters, one may argue that they have higher exit rates than laid-off workers simply because they may have quit in order to take up a new job that was already in their plans. However, this interpretation seems hard to justify in a context in which job quitters stay unemployed for more than five quarters on average - and more than one quarter in any case - before starting a new contract. Having said this, it seems more appropriate to interpret quits as a similar case of job loss to contract termination: no firing costs are involved in these cases and, according to our model, no stigma is attached.

Finally, it is worthwhile to mention that cohort dummies imply quite clearly that re-employment probabilities are procyclical, being higher at the 1987 peak (our reference year), declining over the following recession, and rising again with the mid-1990s recovery. Local deviations from the aggregate cycle are captured by the provincial unemployment

[9] We believe that this effect is at work mainly for male re-employment probabilities, see Petrongolo (1999).

rate, that significantly reduces the hazard.

Interestingly enough, regression II - which allows for unobserved heterogeneity among individuals - delivers a vector of estimated coefficients which is qualitatively similar to the one reported in regression I. In other words, lower hazards for laid-off workers do not necessarily reflect lower unobserved ability. We are therefore more inclined to accept an explanation of their lower job-finding prospects based on the signalling value of their employment history.

6 Conclusions

This paper is an attempt to explain some differences in youth and long term unemployment rates across countries. The starting point of the work is the observation that countries with very strict employment protection legislations tend to have higher youth and long term unemployment. This suggested an important link between the problem of learning about workers' productivities and the opportunity cost of hiring when firing costs are high.

The main intuition of the model developed is that in a world of imperfect information firing costs have several effects on unemployment. First, they lower the option value of hiring workers new to the market, hence tend to increase youth unemployment. Second, firing costs affect the expected productivity of the unemployment, by increasing the stigma of being fired and reducing the stigma of not being hired. Finally, they may reduce the profitability of posting new vacancies.

The second part of the work contains evidence that higher firing costs do increase the stigma attached to poor employment histories. Our empirical analysis has focused on Spain, that in 1984 relaxed employment protection legislation through the introduction of fixed-term contracts. Our results show that workers who lost their job through some costly firing procedure have worse re-employment prospects than those who terminated a fixed-term contract or quit their position. This finding is also robust to the introduction of various individual controls that may capture an individual propensity to high turnover, and of a Gamma distributed heterogeneity term.

The evidence presented has policy implications. The introduction of fixed-term contracts may increase the willingness of firms to hire workers even if they do not have a clear perception of their ability, and reduce the stigma that workers would suffer in case of contract termination. However, other effects of fixed-term contracts may represent a strong case against deregulation and the relative relevance of each of them should be carefully assessed. Excessive segmentation of the labor market, unresponsiveness of wages to unemployment and precariousness of employment are some of the damaging effects of fixed-term contracts pointed out in the recent literature (see - among others - Bentolila and Dolado, 1994).

Alternatively, the problem of …ring and stigma could be addressed from a di¤erent angle, by giving workers the opportunity to update their productivity if they feel that their position is at stake. In this perspective, subsidies to on-the-job training would help to reduce the risk of unemployment and stigma in high-…ring-cost countries.

References

[1] Ahn, N. and A. Ugidos (1995), "Duration of Unemployment in Spain: Relative E¤ects of Unemployment Bene…ts and Family Characteristics", Oxford Bulletin of Economics and Statistics 57, 249-264.

[2] Akerlof, G. (1970), "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism", Quarterly Journal of Economics 84, 488-500.

[3] Alba, A. (1998), "Re-employment Probabilities of Young Workers in Spain", Investigaciones Económicas 22, 201-224.

[4] Bentolila, S. and J. Dolado (1994), "Labour Flexibility and Wages: Lessons from Spain", Economy Policy 9, 53-99.

[5] Bentolila, S. and G. Bertola (1990), "Firing Costs and Labour Demand: How Bad is Eurosclerosis?", Review of Economic Studies 57, 381-402.

[6] Bentolila, S. and G. Saint-Paul (1992), "The Macroeconomic Impact of Flexible Labor Contracts with an Application to Spain", European Economic Review 36, 1013-1047.

[7] Bertola, G. (1990), "Job Security, Employment and Wages", European Economic Review 34, 851-886.

[8] Bertola, G. and R. Rogerson (1995), "Institutions and Labor reallocation", European Economic Review 41, 1147-1171.

[9] Bover O., M. Arellano and S. Bentolila (1998), "Unemployment Duration, Bene…t Duration, and the Business Cycle", CEPR Discussion Paper No. 1840.

[10] García Pérez, I. (1997), "Las Tasas de Salida del Empleo y el Desempleo en España, 1978-1993", Investigaciones Económicas 21, 29-53.

[11] Gibbons, R. and L. Katz (1991), "Layo¤s and Lemons", Journal of Labor Economics 9, 351-380.

[12] Greenwald, B. (1986), "Adverse Selection in the Labour Market", Review of Economic Studies 53, 325-347.

[13] Güell-Rotllan, M. and B. Petrongolo (1999), "Workers' Transitions from Temporary to Permanent Employment: The Spanish Case", Centre for Economic Performance Discussion Paper No..

[14] Han, A. and J. Hausman (1990), "Flexible Parametric Estimation of Duration and Competing Risk Models", Journal of Applied Econometrics, Vol. 5, pp. 1-28.

[15] Hopenhayn, H. and R. Rogerson (1993), "Job Turnover and Policy Evaluation: A General Equilibrium Analysis", Journal of Political Economy 101, 915-938.

[16] Jovanovic, B. (1979), "Job Matching and the Theory of Turnover", Journal of Political Economy 87, 972-990.

[17] Jovanovic, B. (1984), "Matching, Turnover, and Unemployment", Journal of Political Economy 92, 108-122.

[18] Lancaster, T. and A. Chesher (1983), "The Estimation of Models of Labour Market Behavior", Review of Economic Studies 50, 609-624.

[19] Lazear, E. (1990), "Job Security Provisions and Employment", Quarterly Journal of Economics 105, 699-726.

[20] Lazear, E. (1995), "Hiring Risky Workers", NBER Working Paper No. 5334.

[21] Machin, S. and A. Manning (1999), "The Causes and Consequences of Long-Term Unemployment in Europe", in O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, Volume 3, Amsterdam: North Holland.

[22] Narendranathan, W. and M. Stewart (1993), "Modelling the Probability of Leaving Unemployment: Competing Risks Models with Flexible Base-line Hazards", Applied Statistics, 42, 63-83.

[23] OECD (1994a), Employment Outlook.

[24] OECD (1994b), The OECD Jobs Study, Part II.

[25] Petrongolo, B (1999), "Re-employment Probabilities and Returns to Matching", CEPR Discussion Paper No 2266.

[26] Pissarides, C. (1990), Equilibrium Unemployment Theory, Oxford: Basil Blackwell.

[27] Riordan, M. and Staiger, R. (1993), "Sectoral Shocks and Structural Unemployment", International Economic Review 34, 611-629.

[28] Spence, M. (1973), "Job Market Signalling", Quarterly Journal of Economics 15, 255-267.

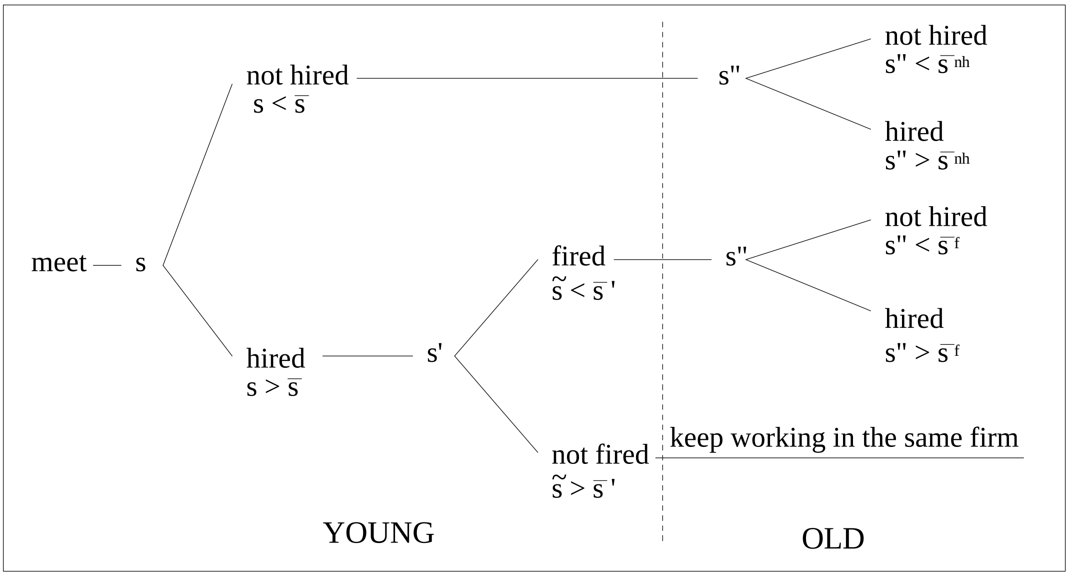

Figure 1: The Signalling Scheme

Table 1: Re-employment prospects according to unemployment cause

Cause of termination	Average duration (quarters)	% that finds a new job
end of seasonal contract	4.4	50.0
end of non-seasonal contract	5.3	42.8
quit	5.5	34.7
laid-off	6.5	33.2

Source: EPA

Table 2: Sample characteristics of the unemployed

	proportion or mean	(Std. dev)
exit in employment	42.0	
exit in non-employment	18.2	
stay unemployed	38.3	
female	42.1	
age 18-29	49.6	
age 30-44	31.4	
age 45+	19.0	
primary education or below	43.8	
secondary education	48.8	
university education	7.4	
married	49.1	
receiving bene…ts	37.8	
agriculture	13.4	
manufacturing	19.3	
construction	17.7	
services	49.5	
terminated seasonal contract	20.6	
terminated non-seasonal contract	52.6	
quit last job held	13.5	
laid o¤	13.3	
Uncensored duration	3.4	(3.2)
Censored duration	6.7	(7.5)
No. of kids	0.8	(1.0)
local unemployment rate	22.3	(0.1)
No. of cases	91,664	

Source: EPA.

Table 3: Maximum likelihood estimates of the transition from unemployment to employment

	I		II	
female	-0.402	(0.012)	-0.410	(0.013)
age 30-44	-0.123	(0.014)	-0.126	(0.014)
age 45+	-0.718	(0.041)	-0.715	(0.041)
secondary education	-0.006	(0.013)	-0.007	(0.013)
college education	0.069	(0.074)	0.053	(0.075)
married	0.028	(0.014)	0.027	(0.014)
number of kids	0.037	(0.006)	0.039	(0.006)
receiving bene…ts	-0.082	(0.011)	-0.086	(0.011)
manufacturing	-0.322	(0.020)	-0.332	(0.021)
construction	-0.161	(0.019)	-0.165	(0.019)
services	-0.347	(0.017)	-0.356	(0.018)
end non-seas contract	0.278	(0.023)	0.286	(0.024)
end non-seas contract £ age 45+	0.473	(0.047)	0.466	(0.048)
end non-seas contract £ college ed.	0.009	(0.090)	0.023	(0.093)
end seas contract	0.139	(0.020)	0.144	(0.021)
end seas contract £ age 45+	0.473	(0.044)	0.466	(0.046)
end seas contract £ college ed.	0.081	(0.081)	0.102	(0.080)
quit	0.083	(0.025)	0.086	(0.026)
quit £ age 45+	0.359	(0.057)	0.351	(0.057)
quit £ college ed.	-0.031	(0.093)	-0.008	(0.095)
year 1988	-0.053	(0.027)	-0.052	(0.028)
year 1989	-0.117	(0.026)	-0.115	(0.027)
year 1990	-0.195	(0.027)	-0.197	(0.028)
year 1991	-0.311	(0.027)	-0.314	(0.028)
year 1992	-0.519	(0.026)	-0.525	(0.027)
year 1993	-0.424	(0.026)	-0.431	(0.027)
year 1994	-0.286	(0.026)	-0.291	(0.027)
year 1995	-0.283	(0.026)	-0.286	(0.027)
year 1996	-0.232	(0.027)	-0.235	(0.028)
ln(unemployment rate)	-0.243	(0.015)	-0.249	(0.016)
¾^2	–		0.238	(0.022)
mean log-likelihood	-0.980		-0.979	
No. of cases	91,664		91,664	

Notes. Reference category: male, aged 18-29, without secondary education, not married, not receiving unemployment bene…ts, previously employed in agriculture, laid o¤, entered survey in 1987. Asymptotic standard errors in brackets.

A Computation of $E[a_i|\bar{s}_i > \bar{s}^0; s_i]$

$E[a_i|\bar{s}_i \geq \bar{s}^0; s_i]$ denotes the expected value of a young worker's ability, conditional on the average of the two signals in the first period of employment being higher than a cutoff $\bar{s}^0$, and on the signal in the current period s. Given the assumptions made in Section (2), a_i, s_i and s_i^0 are distributed as a multivariate normal, i.e. (dropping all subscripts):

$$\begin{pmatrix} a \\ \bar{s} \\ s \end{pmatrix} \sim N_3 \left(\begin{pmatrix} \bar{a} \\ \bar{a} \\ \bar{a} \end{pmatrix} ; \begin{pmatrix} \frac{1}{h_a} & \frac{1}{h_a} & \frac{1}{h_a} \\ \frac{1}{h_a} & \sigma^2_{\bar{s}} & \frac{1}{h_a} + \frac{1}{2h_\varepsilon} \\ \frac{1}{h_a} & \frac{1}{h_a} + \frac{1}{2h_\varepsilon} & \sigma^2_s \end{pmatrix} \right)$$

It is convenient to define four partitions of the covariance matrix, corresponding to the variances and covariances of a and the vector $\mathbf{s} \overset{\text{def}}{=} (\bar{s}; s)$. The four partitions are Σ_{11}, Σ_{12}, Σ_{21}, and Σ_{22}:

$$\Sigma_{11} = \frac{1}{h_a} \qquad \Sigma_{12} = \left(\frac{1}{h_a}; \frac{1}{h_a} \right) \tag{37}$$

$$\Sigma_{21} = \begin{pmatrix} \frac{1}{h_a} \\ \frac{1}{h_a} \end{pmatrix} \qquad \Sigma_{22} = \begin{pmatrix} \sigma^2_{\bar{s}} & \frac{1}{h_a} + \frac{1}{2h_\varepsilon} \\ \frac{1}{h_a} + \frac{1}{2h_\varepsilon} & \sigma^2_s \end{pmatrix} \tag{38}$$

The expected value of a conditional on $\bar{s}$ and s is equal to:

$$\begin{aligned} E[a|\bar{s}; s] &= E(a) + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{s} - E[\mathbf{s}]) && (39) \\ &= \bar{a} + \left(\frac{1}{h_a}; \frac{1}{h_a} \right) \frac{1}{\sigma^2_{\bar{s}}\sigma^2_s - \left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon} \right)^2} \begin{pmatrix} \sigma^2_s & -\frac{1}{h_a} - \frac{1}{2h_\varepsilon} \\ -\frac{1}{h_a} - \frac{1}{2h_\varepsilon} & \sigma^2_{\bar{s}} \end{pmatrix} \begin{pmatrix} \bar{s} - \bar{a} \\ s - \bar{a} \end{pmatrix} \\ &= \bar{a} + \left(\frac{1}{h_a}; \frac{1}{h_a} \right) \begin{pmatrix} (\bar{s} - \bar{a}) \left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon} \right) - (\bar{s} - \bar{a}) \left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon} \right) \\ - \left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon} \right) (s - \bar{a}) + \left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon} \right) (\bar{s} - \bar{a}) \end{pmatrix} \\ &= \bar{a}\frac{h_a}{h_a + 2h_\varepsilon} + \bar{s}\frac{2h_\varepsilon}{h_a + 2h_\varepsilon} = E[a|\bar{s}] && (40) \end{aligned}$$

The expected value of a conditional on $\bar{s} \geq \bar{s}^0$ and s is:

$$E[a|\bar{s} \geq \bar{s}^0; s] = \frac{1}{1 - \Phi(\bar{s}^0; s; \cdot)} \int_{\bar{s}^0}^{\infty} E[a|\bar{s}; s] g(\bar{s}|s) d\bar{s} \tag{41}$$

The distribution of $\bar{s}$ conditional on s is:

$$g(\bar{s}|s) \sim N \left(\frac{1}{2} \bar{a} \left[1 - \frac{\frac{1}{h_a} + \frac{1}{2h_\varepsilon}}{\sigma^2_s} \right] + \frac{1}{2} s \frac{\frac{1}{h_a} + \frac{1}{2h_\varepsilon}}{\sigma^2_s} ; \sigma^2_{\bar{s}} - \frac{\left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon} \right)^2}{\sigma^2_s} \right)$$

or, equivalently:

$$g(\tilde{s}|s) \sim N\left(\frac{1}{2}\frac{h_a}{h_a+h_\epsilon}\bar{a} + \frac{1}{2}\frac{h_a+2h_\epsilon}{h_a+h_\epsilon}s,\ \frac{1}{4}\frac{h_a+2h_\epsilon}{h_\epsilon(h_a+h_\epsilon)}\right) \qquad (42)$$

We can rewrite Equation (41) as follows:

$$\begin{aligned}
E[a|\tilde{s} > \bar{s}'; s] &= \frac{h_a}{h_a+2h_\epsilon}\bar{a} + \frac{2h_\epsilon}{h_a+2h_\epsilon}E[\tilde{s}|\tilde{s} \geq \bar{s}'; s] \\
&= \frac{h_a}{h_a+2h_\epsilon}\bar{a} + \frac{2h_\epsilon}{h_a+2h_\epsilon}\left[\frac{h_a}{h_a+h_\epsilon}\frac{1}{2}\bar{a} + \frac{h_a+2h_\epsilon}{h_a+h_\epsilon}\frac{1}{2}s + \sigma_{\tilde{s}|s}\frac{\phi(\bar{s}'; s; \cdot)}{1-\Phi(\bar{s}'; s; \cdot)}\right] \\
&= \bar{a}\left[\frac{h_a}{h_a+2h_\epsilon}\left(1+\frac{h_a h_\epsilon}{h_a+h_\epsilon}\right)\right] + s\frac{h_\epsilon}{h_a+h_\epsilon} + \frac{2h_\epsilon}{h_a+2h_\epsilon}\sigma_{\tilde{s}|s}\frac{\phi(\bar{s}'; s; \cdot)}{1-\Phi(\bar{s}'; s; \cdot)} \\
&= \bar{a}\frac{h_a}{h_a+h_\epsilon} + s\frac{h_\epsilon}{h_a+h_\epsilon} + \sigma_{\tilde{s}|s}\frac{2h_\epsilon}{h_a+2h_\epsilon}\frac{\phi(\bar{s}'; s; \cdot)}{1-\Phi(\bar{s}'; s; \cdot)}
\end{aligned}$$

where $\sigma_{\tilde{s}|s}$ is the standard error of the distribution of $\tilde{s}$ conditional on s as defined in Equation (42).

B Proof that $\partial\bar{s}/\partial f > 0$

The expected profitability of worker i defined in equation (10) can be rewritten as follows:

$$\begin{aligned}
E[\pi_i|s_i] &= v(\bar{a}; h_a; h_\epsilon; s_i) + \\
&+ \int_{-\infty}^{\infty} \max\left\{-f; \frac{h_a}{h_a+2h_\epsilon}\bar{a} + \frac{2h_\epsilon}{h_a+2h_\epsilon}\frac{s_i+s'}{2}\right\}\phi(s'|s_i)ds'
\end{aligned} \qquad (43)$$

The hiring cutoff for young job applicants $\bar{s}$ is implicitly defined by setting $E[\pi_i|s_i] = 0$, i.e.

$$v(\bar{a}; h_a; h_\epsilon; s_i) + \int_{-\infty}^{\infty} \max\left\{-f; \frac{h_a}{h_a+2h_\epsilon}\bar{a} + \frac{2h_\epsilon}{h_a+2h_\epsilon}\frac{s_i+s'}{2}\right\}\phi(s'|s_i)ds' = 0 \qquad (44)$$

In order to show that $\bar{s}$ increases with f, we are going to prove that, first, $E[\pi_i|s_i]$ is increasing in s_i and second, that it is decreasing in f. Since the expected profitability is increasing in s_i, we can picture it as an upward sloping curve in a graph with s_i on the horizontal axis, and the expected profitability on the vertical axis. Showing that the derivative of the expected profitability decreases with f corresponds to showing that the upward sloping curve of $E[\pi_i|s_i]$ shifts to the right when firing costs go up.

$E[\pi_i|s_i]$ is increasing in s_i since:

$$\frac{\partial E[\pi_i|s_i]}{\partial s_i} = \begin{cases} \frac{h_\varepsilon}{h_a+h_\varepsilon} > 0 & \text{if} \quad -f > \frac{h_a}{h_a+2h_\varepsilon}\hat{a} + \frac{2h_\varepsilon}{h_a+2h_\varepsilon}\frac{s_i+s_i'}{2} \\ \frac{h_\varepsilon}{h_a+h_\varepsilon} + \frac{h_\varepsilon}{h_a+2h_\varepsilon} > 0 & \text{otherwise} \end{cases} \tag{45}$$

$E[\pi_i|s_i]$ is decreasing in f since:

$$\frac{\partial E[\pi_i|s_i]}{\partial f} = \begin{cases} -1 & \text{if} \quad -f > \frac{h_a}{h_a+2h_\varepsilon}\hat{a} + \frac{2h_\varepsilon}{h_a+2h_\varepsilon}\frac{s_i+s_i'}{2} \\ 0 & \text{otherwise} \end{cases} \tag{46}$$

The derivatives in (46) show that $E[\pi_i|s_i]$ is weakly decreasing in f. Since s_i' is unbounded from below, it can always be small enough to make the $\max\{\cdot;\cdot\} = -f$.

C Computation of $E[a_i|s_i < \hat{s}; s_i'']$

$E[a_i|s_i < \hat{s}; s_i'']$ denotes the expected value of a worker's ability conditional on the signal in the previous period being lower than a cutoff $\hat{s}$, and on the signal in the current period s_i''. Given the assumptions made in Section (2), a_i, s_i and s_i'' are distributed as a multivariate normal. Dropping the subscripts, the distribution is as follows:

$$\begin{pmatrix} a \\ s \\ s'' \end{pmatrix} \sim N_3\left(\begin{pmatrix} \hat{a} \\ \hat{a} \\ \hat{a} \end{pmatrix}; \begin{pmatrix} \frac{1}{h_a} & \frac{1}{h_a} & \frac{1}{h_a} \\ \frac{1}{h_a} & \sigma_s^2 & \frac{1}{h_a} \\ \frac{1}{h_a} & \frac{1}{h_a} & \sigma_s^2 \end{pmatrix}\right)$$

It is convenient to define four partitions of the covariance matrix, corresponding to the variances and covariances of a and the vector $\mathbf{s} \stackrel{\text{def}}{=} (s; s'')$. The four partitions are Σ_{11}, Σ_{12}, Σ_{21}, and Σ_{22}:

$$\Sigma_{11} = \frac{1}{h_a} \qquad \Sigma_{12} = \left(\frac{1}{h_a}; \frac{1}{h_a}\right)$$

$$\Sigma_{21} = \begin{pmatrix} \frac{1}{h_a} \\ \frac{1}{h_a} \end{pmatrix} \qquad \Sigma_{22} = \begin{pmatrix} \sigma_s^2 & \frac{1}{h_a} \\ \frac{1}{h_a} & \sigma_s^2 \end{pmatrix}$$

The expected value of a conditional on s and s'' is equal to:

$$\begin{aligned} E[a|s; s''] &= E(a) + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{s} - E[\mathbf{s}]) = \\ &= \hat{a} + \left(\frac{1}{h_a}; \frac{1}{h_a}\right) \frac{1}{\sigma_s^4 - \frac{1}{h_a^2}} \begin{pmatrix} \sigma_s^2 & -\frac{1}{h_a} \\ -\frac{1}{h_a} & \sigma_s^2 \end{pmatrix} \begin{pmatrix} s - \hat{a} \\ s'' - \hat{a} \end{pmatrix} \\ &= \hat{a}\left[1 - \frac{\frac{2}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\right] + (s + s'')\frac{\frac{1}{h_a}}{\sigma_s^2 + \frac{1}{h_a}} \end{aligned} \tag{47}$$

The expected value of a conditional on $s < \hat{s}$ and s'' is:

$$E[a_2|s < \hat{s}; s''] = \frac{1}{\Phi(\hat{s};\cdot)} \int_{-\infty}^{\hat{s}} E[a|s; s'']g(s|s'')ds \tag{48}$$

The distribution of s conditional on s'' is:

$$g(s|s'') \sim N\left(\hat{a}\left[1 - \frac{\frac{1}{h_a}}{\sigma_s^2}\right] + s''\frac{\frac{1}{h_a}}{\sigma_s^2}; \sigma_s^2 - \frac{\frac{1}{h_a^2}}{\sigma_s^2}\right) \tag{49}$$

or

$$g(s|s'') \sim N\left(\hat{a}\frac{h_a}{h_a + h_2} + s''\frac{h_2}{h_a + h_2}; \sigma_s^2 - \frac{\frac{1}{h_a^2}}{\sigma_s^2}\right)$$

We can rewrite equation (48) as follows:

$$\begin{aligned}
E[a|s < \hat{s}; s''] &= \frac{1}{\Phi(\hat{s};\cdot)} \int_{-\infty}^{\hat{s}} E[a|s; s'']f(s|s'')ds \\
&= \left[\hat{a}\left(1 - \frac{\frac{2}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\right) + s''\frac{\frac{1}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\right] + \frac{\frac{1}{h_a}}{\sigma_s^2 + \frac{1}{h_a}} E[s|s < \hat{s}; s''] \\
&= \left[\hat{a}\left(1 - \frac{\frac{2}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\right) + s''\frac{\frac{1}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\right] + \\
&\quad + \frac{\frac{1}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\left[\left(\hat{a}\left(1 - \frac{\frac{1}{h_a}}{\sigma_s^2}\right) + s''\frac{\frac{1}{h_a}}{\sigma_s^2}\right) - \left(\sigma_s^2 - \frac{\frac{1}{h_a^2}}{\sigma_s^2}\right)^{\frac{1}{2}} \frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)}\right] \\
&= \left[\hat{a}\left(1 - \frac{\frac{1}{h_a}}{\sigma_s^2}\right) + s''\frac{\frac{1}{h_a}}{\sigma_s^2}\right] - \sigma_{s|s''}\frac{\frac{1}{h_a}}{\sigma_s^2 + \frac{1}{h_a}}\frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)} \\
&= \hat{a}\frac{h_a}{h_a + h_2} + s''\frac{h_2}{h_a + h_2} - \sigma_{s|s''}\frac{2h_2}{h_a + 2h_2}\frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)}
\end{aligned}$$

where $\sigma_{s|s''}$ is the standard error of the distribution of s conditional on s'' as defined in equation (49)

D Proof that $E[a_i|s_i < \hat{s}; s_i''] < E[a_i|s_i'']$

We want to compare the one-period expected productivity for a worker that met with a firm in the previous period but was not hired, and that sends signal s'' of her ability in the current period, $E[a_i|\text{not hired}; s'']$, with the one-period expected productivity of a worker new to the market, given the same signal s'', $E[a_i|s'']$. From Appendix C:

$$E[a_i|\text{not hired}; s''] = \bar{a}\frac{h_a}{h_a + h_\epsilon} + s''\frac{h_\epsilon}{h_a + h_\epsilon} - \sigma_{s|s''}\frac{2h_\epsilon}{h_a + 2h_\epsilon}\frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)}$$

while

$$E[a_i|s''] = \bar{a}\frac{h_a}{h_a + h_\epsilon} + s''\frac{h_\epsilon}{h_a + h_\epsilon}$$

Therefore, we can rewrite $E[a_i|\text{not hired}; s'']$ as follows:

$$E[a_i|\text{not hired}; s''] = E[a_i|s''] - \sigma_{s|s''}\frac{2h_\epsilon}{h_a + 2h_\epsilon}\frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)} < E[a_i|s''].$$

E Computation of $E[a_i|\bar{s}_i < \bar{s}^0; s_i'']$

$E[a_i|\bar{s}_i < \bar{s}^0; s_i'']$ denotes the expected value of a worker's ability conditional on the average of the two signals in the previous period being lower than a cutoff $\bar{s}^0$, and on the signal in the current period s_i''. Given the assumptions made in Section (2), a_i, s_i and s_i'' are distributed as a multivariate normal, i.e. (dropping the subscripits):

$$\begin{pmatrix} a \\ \bar{s} \\ s'' \end{pmatrix} \sim N_3\left(\begin{pmatrix} \bar{a} \\ \bar{a} \\ \bar{a} \end{pmatrix}; \begin{pmatrix} \frac{1}{h_a} & \frac{1}{h_a} & \frac{1}{h_a} \\ \frac{1}{h_a} & \sigma_{\bar{s}}^2 & \frac{1}{h_a} + \frac{1}{2h_\epsilon} \\ \frac{1}{h_a} & \frac{1}{h_a} + \frac{1}{2h_\epsilon} & \sigma_s^2 \end{pmatrix}\right)$$

The four partitions of the covariance matrix, corresponding to the variances and covariances of a and the vector $\mathbf{s} \stackrel{\text{def}}{=} (\bar{s}; s'')$ are given in equations (37) and (38).

The expected value of a conditional on $\bar{s}$ and s'' is equal to:

$$\begin{aligned}
E[a|\bar{s}; s''] &= E(a) + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{s} - E[\mathbf{s}]) \\
&= \bar{a} + \left(\frac{1}{h_a}; \frac{1}{h_a}\right)\frac{1}{\sigma_{\bar{s}}^2\sigma_s^2 - \left(\frac{1}{h_a} + \frac{1}{2h_\epsilon}\right)^2}\begin{pmatrix} \sigma_s^2 & -\frac{1}{h_a} - \frac{1}{2h_\epsilon} \\ -\frac{1}{h_a} - \frac{1}{2h_\epsilon} & \sigma_{\bar{s}}^2 \end{pmatrix}\begin{pmatrix} \bar{s} - \bar{a} \\ s'' - \bar{a} \end{pmatrix} \\
&= \bar{a} + \left(\frac{1}{h_a}; \frac{1}{h_a}\right)\begin{pmatrix} (s'' - \bar{a})\left(\frac{1}{h_a} + \frac{1}{2h_\epsilon}\right) - (\bar{s} - \bar{a})\left(\frac{1}{h_a} + \frac{1}{2h_\epsilon}\right) \\ -\left(\frac{1}{h_a} + \frac{1}{2h_\epsilon}\right)(s'' - \bar{a}) + \left(\frac{1}{h_a} + \frac{1}{2h_\epsilon}\right)(\bar{s} - \bar{a}) \end{pmatrix} \\
&= \bar{a}\frac{h_a}{h_a + 2h_\epsilon} + \bar{s}\frac{2h_\epsilon}{h_a + 2h_\epsilon} = E[a|\bar{s}]
\end{aligned}$$

The expected value of a conditional on $\tilde{s} < \bar{s}'$ and s'' is:

$$E[a|\tilde{s} < \bar{s}'; s''] = \frac{1}{\Phi(\bar{s}'; s''; \cdot)} \int_{-\infty}^{\bar{s}} E[a|\tilde{s}; s''] f(\tilde{s}|s'') d\tilde{s} \tag{50}$$

The distribution of $\tilde{s}$ conditional on s'' is:

$$f(\tilde{s}|s'') \sim N\left(\frac{1}{2}\bar{a}\left[1 - \frac{\frac{1}{h_a} + \frac{1}{2h_\varepsilon}}{\sigma_s^2}\right] + \frac{1}{2}s''\frac{\frac{1}{h_a} + \frac{1}{2h_\varepsilon}}{\sigma_s^2}; \sigma_{\tilde{s}}^2 - \frac{\left(\frac{1}{h_a} + \frac{1}{2h_\varepsilon}\right)^2}{\sigma_s^2}\right)$$

or, equivalently:

$$f(\tilde{s}|s'') \sim N\left(\frac{1}{2}\frac{h_a}{h_a + h_\varepsilon}\bar{a} + \frac{1}{2}\frac{h_a + 2h_\varepsilon}{h_a + h_\varepsilon}s''; \frac{1}{4}\frac{h_a + 2h_\varepsilon}{h_\varepsilon(h_a + h_\varepsilon)}\right) \tag{51}$$

We can rewrite Equation (50) as follows:

$$\begin{aligned}
E[a|\tilde{s} < \bar{s}'; s''] &= \frac{h_a}{h_a + 2h_\varepsilon}\bar{a} + \frac{2h_\varepsilon}{h_a + 2h_\varepsilon}E[\tilde{s}|\tilde{s} < \bar{s}'; s''] \\
&= \frac{h_a}{h_a + 2h_\varepsilon}\bar{a} + \\
&+ \frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\left[\frac{h_a}{h_a + h_\varepsilon}\frac{1}{2}\bar{a} + \frac{h_a + 2h_\varepsilon}{h_a + h_\varepsilon}\frac{1}{2}s'' - \sigma_{\tilde{s}|s}\frac{\phi(\bar{s}'; s''; \cdot)}{\Phi(\bar{s}'; s''; \cdot)}\right] = \\
&= \bar{a}\left[\frac{h_a}{h_a + 2h_\varepsilon}\left(1 + \frac{h_a h_\varepsilon}{h_a + h_\varepsilon}\right)\right] + s''\frac{h_\varepsilon}{h_a + h_\varepsilon} - \frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\sigma_{\tilde{s}|s}\frac{\phi(\bar{s}'; s''; \cdot)}{\Phi(\bar{s}'; s''; \cdot)} = \\
&= \bar{a}\frac{h_a}{h_a + h_\varepsilon} + s''\frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{\tilde{s}|s}\frac{2h_\varepsilon}{h_a + 2h_\varepsilon}\frac{\phi(\bar{s}'; s''; \cdot)}{\Phi(\bar{s}'; s''; \cdot)}
\end{aligned}$$

where $\sigma_{\tilde{s}|s}$ is the standard error of the distribution of $\tilde{s}$ conditional on s as defined in Equation (51).

F Proof that $E[a_i|\tilde{s}_i < \bar{s}'; s_i''] < E[a_i|s_i'']$

We want to compare the one-period expected productivity for an old worker that worked when young but was fired, and that sends signal s'' in the current period $E[a_i|\text{fired}; s'']$, with the one-period expected productivity for a worker that did not meet with a firm in the previous period, $E[a_i|s'']$.

The proof is similar to that in Appendix D. In Appendices D and E, respectively we computed

$$E[a_i|s''] = \mu \frac{h_a}{h_a + h_\varepsilon} + s'' \frac{h_\varepsilon}{h_a + h_\varepsilon}$$

and

$$E[a|s < \hat{s}^{1\prime}; s''] = \mu \frac{h_a}{h_a + h_\varepsilon} + s'' \frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\phi(\hat{s}^{1\prime}; s''; \cdot)}{\Phi(\hat{s}^{1\prime}; s''; \cdot)}$$

We can rewrite $E[a|s < \hat{s}^{1\prime}; s'']$ as:

$$E[a|s < \hat{s}^{1\prime}; s''] = E[a_i|s''] - \sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\phi(\hat{s}^{1\prime}; s''; \cdot)}{\Phi(\hat{s}^{1\prime}; s''; \cdot)} < E[a_i|s''].$$

G Proof that $\frac{\partial \hat{s}^{nh}}{\partial f} < 0$

The expected profitability of an old worker i who was not hired when young is defined in equation (18). Her hiring cutoff $\hat{s}$ is implicitly defined by equating her expected profitability to zero:

$$\mu \frac{h_a}{h_a + h_\varepsilon} + s'' \frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)} = 0$$

In order to show that $\hat{s}$ increases with f, we are going to prove that, first, $E[\pi_i|s_i'']$ is increasing in s_i'' and second, that it is increasing in f.
$E[\pi_i|s < \hat{s}; s_i'']$ is increasing in s_i'' since:

$$\frac{\partial E[\pi_i|s < \hat{s}; s_i'']}{\partial s_i''} = \frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\partial \left(\frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)} \right)}{\partial s''} > 0 \quad (52)$$

$E[\pi_i|s < \hat{s}; s_i]$ is increasing in f since:

$$\frac{\partial E[\pi_i|s < \hat{s}; s_i'']}{\partial f} = -\sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\partial \left(\frac{\phi(\hat{s}; s''; \cdot)}{\Phi(\hat{s}; s''; \cdot)} \right)}{\partial \hat{s}} \frac{\partial \hat{s}}{\partial f} > 0 \quad (53)$$

H Proof that $\frac{\partial \hat{s}^{f}}{\partial f} > 0$

The expected profitability of an old worker i who was fired when young is defined in equation (20). Her hiring cutoff $\hat{s}^f$ is implicitly defined by equating her expected profitability to zero:

$$\mu \frac{h_a}{h_a + h_\varepsilon} + s_i'' \frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\phi(\hat{s}^{1\prime}; s_i''; \cdot)}{\Phi(\hat{s}^{1\prime}; s_i''; \cdot)} = 0$$

In order to show that $\bar{s}^f$ increases with f, we are going to prove that, first, $E[a_i|s_i < \bar{s}'; s_i'']$ is increasing in s_i'' and second, that it is decreasing in f.
$E[a_i|s_i < \bar{s}'; s_i'']$ is increasing in s_i since:

$$\frac{\partial E[a_i|s_i < \bar{s}'; s_i'']}{\partial s_i''} = \frac{h_\varepsilon}{h_a + h_\varepsilon} - \sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\partial \left(\frac{\phi(\bar{s}'; s''; \cdot)}{\Phi(\bar{s}'; s''; \cdot)} \right)}{\partial s''} > 0 \qquad (54)$$

$E[a_i|s_i < \bar{s}'; s_i'']$ is decreasing in f since:

$$\frac{\partial E[a_i|s_i < \bar{s}'; s_i'']}{\partial f} = - \sigma_{s|s''} \frac{2h_\varepsilon}{h_a + 2h_\varepsilon} \frac{\partial \left(\frac{\phi(\bar{s}'; s''; \cdot)}{\Phi(\bar{s}'; s''; \cdot)} \right)}{\partial \bar{s}'} \frac{\partial \bar{s}}{\partial f} < 0 \qquad (55)$$