IZA DP No. 2780

The Rate of Learning-by-Doing: Estimates from a Search-Matching Model

Julien Prat

May 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Rate of Learning-by-Doing:
Estimates from a Search-Matching Model

Julien Prat
*University of Vienna
and IZA*

ABSTRACT

The Rate of Learning-by-Doing: Estimates from a Search-Matching Model[*]

We construct and estimate by maximum likelihood an equilibrium search model where wages are set by Nash bargaining and idiosyncratic productivity follows a geometric Brownian motion. The proposed framework enables us to endogenize job destruction and to estimate the rate of learning-by-doing. Although the range of the observations is not independent of the parameters, we establish that the estimators satisfy asymptotic normality. The structural model is estimated using Current Population Survey data on accepted wages and employment durations. We show that it captures almost perfectly the joint distribution of wages and job spells. We find that the rate of learning-by-doing has an important positive effect on aggregate output and a small impact on employment.

JEL Classification: J31, J64

Keywords: job search, human capital, uncertainty, structural estimation

Corresponding author:

Julien Prat
Department of Economics
University of Vienna
Hohenstaufengasse 9
A-1010 Wien
Austria
E-mail: julien.prat@univie.ac.at

1   Introduction

The Mortensen and Pissarides (1994) model (thereafter MP94) with endogenous job destruction is the canonical macro-framework for the analysis of labor markets. Given its considerable influence, numerous papers have assessed its ability to explain aggregate outcomes and business cycle fluctuations. Comparatively little is known about its capacity to fit the cross-sectional features of the data. This gap in the literature is due to the fact that jobs' outputs follow stochastic paths in MP94, whereas structural estimations of equilibrium search models typically assume that they remain constant through time.[1] This paper bring together the two approaches by using maximum likelihood to estimate the MP94 model with endogenous job destruction. In order to preserve the aggregation properties that are required for equilibrium analysis, we follow Mortensen and Pissarides in assuming that productivity and thus human capital are purely match-specific. We also consider that firms and workers cannot commit so that wages are set by Nash bargaining. Accordingly, our first contribution is to show that the standard MP94 model can indeed capture the joint distribution of wages and job spells, and to identify the parameter values required to achieve a good fit.

Most interestingly, the structural model identifies the rate at which matches' outputs increase with tenure or, for brevity, the rate of learning-by-doing (thereafter LBD). A large body of empirical research has focused on estimating the rate of LBD.[2] This paper attempts to contribute to this line of research by proposing a structural approach that makes it possible to quantify the aggregate impact of LBD. We find that it has a significant positive effect on aggregate output but a small effect on employment.

The rate of LBD cannot be estimated in a purely deterministic set-up because the

---

[1] See Eckstein and Wolpin (1995) for an early estimation of an equilibrium search model in a deterministic environment with symmetric Nash-bargaining and Flinn (2006) for an analysis of the effects of the minimum wage. Only recently has the literature begun to address the observed pattern of wage dynamics. One way to reconcile theory and evidence can be found in the new line of research (Cahuc et al., 2006; Dey and Flinn, 2005) where workers bring alternative employers into Bertrand price competition with their current employer. Although the game of alternating offers effectively generates upward sloping wage profiles, this prediction is not due to increases in output but to the gradual appropriation of the job's rent by the worker. This paper focuses instead on the mechanism prevailing in the equilibrium theory of unemployment whereby changes in jobs' productivities determine wage dynamics.

[2] The related literature is too extensive to be comprehensively reported. An arbitrary sample includes the seminal paper by Jovanovic and Mincer (1981) and more recent contributions by Altonji and Shakotko (1987), Topel (1991), Altonji and Williams (2005), Dustmann and Meghir (2005).

lower bound of the wage distribution increases with tenure when workers regularly progress along the learning curve. Given that cross-sectional data always contain individuals with significant seniority and earnings close to the reservation wage, the estimated rate of LBD necessarily collapses to zero in a deterministic set-up. Hence, one needs to introduce some noise around the learning curve. A possibility is to allow for measurement errors. This approach, however, has the major drawback of ignoring the interaction between job destruction and the rate of LBD. On the other hand, a higher rate of productivity growth naturally lowers the incidence of endogenous job separations. Thus our structural model takes into account the positive effect of LBD on job retention.[3]

Endogenous separations greatly complicate the derivation of the likelihood function because we have to deduce all the sample paths that breach the reservation productivity. We show that this problem can be solved through the introduction of geometric Brownian motions. We therefore assume that jobs' idiosyncratic productivities follow random walks with constant growth rate. Such a specification of the stochastic process is justified by several empirical studies substantiating that a random walk with transitory shocks accurately captures the dynamics of log-wages.[4] This paper illustrates that this specification also has great advantages for structural estimation since it allows us to solve for the equilibrium in closed-form and so to estimate the model by full-information methods.[5] Thus, although we are sympathetic to the idea that more sophisticated ARMA processes have a higher accuracy, we propose geometric Brownian motions as a useful first-order approximation.

The other main technical difficulty is due to the non-standard properties of the likelihood function: the reservation wage and consequently the support of the data is a function of the estimated parameters. This peculiarity of equilibrium search models is well know since Flinn and Heckman (1982). In order to circumvent it, they proposed to evaluate the likelihood function in two steps. First of all the reservation wage is set equal to the lowest

---

[3]The positive effect of skill accumulation on job retention is commonly invoked to justify training programs. An important example is provided by the official goals of the Workforce Investment Act of 1998. For an econometric assessment of the relationship between job stability and training programs, see Winter-Ebmer and Zweimüller (1996).

[4]This finding has been uncovered by Topel and Ward (1992) using the LEED (longitudinal employee-employer data) over the period going from 1957-1972. It has been recently confirmed by Buhai and Teulings (2005) using a dataset based on a PSID extract of 18 waves, covering the years 1975 through 1992.

[5]In a recent paper, Lopes de Melo (2006) also uses stochastic calculus to estimate an equilibrium search model. Although his model shares similarities to the one considered here, he focuses on the learning process about the quality of the firm-worker match.

wage observed in the sample. Since the lowest wage is a super-consistent estimator, one can treat the estimated reservation wage as being equal to its true value when evaluating the remaining parameters. This estimation procedure yields consistent estimates for deterministic search-marching models. When job destruction is endogenous, however, workers and firms separate precisely at the reservation wage. As a result, the likelihood of observing the reservation wage is equal to zero and so the lowest reported wage is not anymore a super-consistent estimator.

Hence, we have to rely on a different estimation method than the two-step approach proposed by Flinn and Heckman (1982). Our problem bears similarities to the estimation of optimal production frontiers. Optimal frontiers models also imply that the range of the observations changes with the parameters being estimated. Moreover, they share with our model the additional implication that agents are never exactly on the optimal frontier. As firms cannot perfectly counteract random perturbations, they remain within the neighborhood of the optimal combination of inputs without ever achieving it perfectly. Given that the estimation of optimal frontiers is one of the most popular area of applied econometrics, great attention has been devoted to the econometric solutions for this kind of problem. In an influential paper, Greene (1980) showed that when the likelihood of observing the actual boundary of the distribution is equal to zero, standard regularity conditions need not be satisfied in order to produce standard asymptotic distribution results. We adapt Greene's proof to our set-up and establish that, despite the appearance, endogenous separation actually simplifies the analysis since it enables us to estimate the likelihood function as if it were completely standard.

After having analyzed the equilibrium of the economy and derived the properties of the likelihood function, we estimate the model using data from the January 2004 supplement of the Current Population Survey. We restrict our attention to workers without tertiary education because the estimates do not capture the accumulation of general human capital which is known to be much more significant for skilled workers.[6] The estimation procedure returns estimates for the rate of LBD of around 2% per year. We assess the ability of the model to fit the joint distribution of wages and job spells and find that it reproduces the

---

[6]See for example Dustmann and Meghir (2005) for evidence according to which the acquisition of general skills is important for skilled workers whereas unskilled workers benefit primarily from being attached to a particular firm.

data surprisingly well given its parsimonious specification. Then we use the estimates to characterize the impact of the rate of LBD. We show that it shifts to the right the wage distribution and significantly increases its dispersion.

The rest of the paper is organized as follows. Section 2 lays out the set-up and characterizes the equilibrium. The econometric procedure and the asymptotic properties of the estimates are detailed in Section 3. Section 4 describes the data and discusses the estimation results. In section 5 we introduce an aggregate matching function to close the model and evaluate the impact of LBD on the equilibrium. Section 6 concludes and the Appendix contains the proofs of the propositions.

2   The model

We consider a labor market with search frictions where jobs' output are subject to random fluctuations. The set-up differs in three respects from the one proposed by Mortensen and Pissarides (1994): firstly we allow initial productivities to differ, secondly we assume that output follows a geometric Brownian motion and finally we introduce LBD. Given that the Current Population Survey (CPS hereafter) does not contain information on the number of posted vacancies, the data will not allow us to estimate the parameters of the matching function. Thus we take as given the rate of contact between firms and job seekers, and postpone the introduction of the aggregate matching function to section 5.

2.1   *The production process*

Consider a market in which homogenous workers, who live forever, are either employed or looking for a job. Each competitive firm has one job which can be either filled or vacant. Firms use only labor to produce a unique multi-purpose good. When an unemployed worker meets a firm with a vacant job, they sample a positive output for their match. The initial productivity is a random draw from the exogenous distribution $G(\cdot)$, which is assumed to be continuously differentiable. In the remainder of the paper, we will refer to $G(\cdot)$ as the sampling distribution.

Both parties instantaneously observe the initial productivity. Then the firm can decide whether or not to make a job offer. If the firm "passes" on the applicant, it does not incur

any specific cost for doing so and it continues to keep its vacancy open to other workers. Similarly, the worker can choose to refuse the job offer if he prefers to search for a better opportunity.

In the case where both parties decide to match, they immediately start to produce and output begins to fluctuate. We do not consider aggregate shocks so that stochastic fluctuations are uncorrelated across jobs. The stochastic process that changes the idiosyncratic output is a geometric Brownian motion. Thus its law of motion is given by

$$\frac{dP_t^i}{P_t^i} = \zeta dt + \sigma dB_t^i \tag{1}$$

where dB_t^i is the increment of a standard Brownian motion. The subscript i indexes jobs. In the remainder of the text we will neglect it when not necessary. According to (1), the expected output at time $t + T$ of a job with current output P_t is equal to $P_t e^{\zeta T}$. Hence ζ is the rate at which productivity on-the-job increases. The acquired skills are purely job-specific since workers become identical when they return to the unemployment pool. The parameter σ reflects dispersion: the higher it is, the faster output fluctuates. We also introduce an exogenous source of uncertainty such that jobs are forced out of business when hit by random shocks that arrive at the Poisson rate δ.

The introduction of Brownian motions contrasts with the standard practice of considering Poisson processes. Whereas Brownian motions have continuous sample paths, Poisson processes are by definition discontinuous. It is explained in Prat (2006) why Brownian motions deliver more accurate predictions about the hazard rate of job separation and the shape of the wage distribution. It is also shown in Prat (2006) how most of the statistics of interest can be derived in closed-form using stochastic calculus. As we will see in section 3, the convenient analytical properties of Brownian motions are crucial for the empirical implementation of the model.

We also assume that workers do not receive alternative job offers while employed. Thus we do not consider on-the-job search so that trade in the labor market is completely separated from production. This restriction is imposed due to technical reasons as it is notoriously involved to combine idiosyncratic uncertainty with on-the job-search.[7] These

---

[7]For example, Nagypál (2005) establishes that the wage distribution cannot be expressed in closed-form when workers search on-the-job and uncertainty is modelled using a diffusion process.

6

difficulties partly explain why empirical models of employers competition typically assume away idiosyncratic uncertainty.[8] We make the converse assumption and leave to further research the task of devising a comprehensive model.

2.2 *Optimal job separation*

Because trading in the labor market is a costly process, matched pairs have to share a quasi-rent. We assume a Nash-bargaining rule whereby each party obtains a constant share of the job's surplus $S(P_t)$ at each point in time.[9] The rent of each party is defined as the difference between the asset value obtained by participating in the match and the disagreement outcome of continued search. Since the two rents remain proportional, it cannot be the case that one is positive and the other negative. In other words workers and firms always separate by common agreement.

Let U denote the steady-state expected value of search by an unemployed worker. The search effort costs the worker s and he meets at the flow rate λ a firm with an open vacancy. The contact leads to a match if the initial output drawn from the sampling distribution $G(\cdot)$ is at least as great as the reservation output R. Under the assumption that workers are risk-neutral and that they discount the future at rate r, U satisfies the following equation

$$rU = -s + \lambda \int_{R}^{+\infty} \beta S(P) dG(P) \tag{2}$$

where β denotes the worker's bargaining power. As opposed to the labor force whose size is fixed and normalized to one, new firms enter the market until arbitrage opportunities are exhausted. Thus free-entry ensures that the firm's outside option is equal to zero and

---

[8]A notable exception is the recent paper by Postel-Vinay and Thuron (2005). They estimate a model with i.i.d. productivity shocks, on-the-job search and wage renegotiation by mutual consent. Given the complexity of their set-up, they do not incorporate human capital accumulation and take job separation as exogenous. Nevertheless, the likelihood function of the model cannot be analytically characterized so they have to rely on Optimal Minimum Distance estimation.

[9]The Nash-bargaining solution assumes away the difficulty of relating wages to job-specific human capital. As explained in Felli and Harris (1996), wages increase with human capital to the extent that workers are able to appropriate some of the return. As specific human capital enhances the worker's productivity only in its current working place, it is not clear why the worker should receive any of the return on it. We do not address this issue and instead follow the typical practice of considering that each party receives a fixed share of the expected surplus at any point in time.

the total surplus of the match can be decomposed in the following way

$$rS\left(P_t\right) = P_t - rU - \delta S\left(P_t\right) + \frac{E}{dt}\left[dS\left(P_t\right)\right] \tag{3}$$

where it is assumed that firms discount the future at the same rate than workers. The term $\delta S\left(P_t\right)$ corresponds to the loss incurred by both parties when the job is hit by an exogenous destruction shock. Notice also that the surplus evolves through time due to output fluctuations. In the deterministic case one can immediately solve for $S\left(P_t\right)$ by combining equations (2) and (3). In the stochastic case we have to solve the partial differential equation satisfied by $S\left(P_t\right)$, as explained in the Appendix.

Proposition 1 *The expected surplus of a match with current output P and reservation output R is given by*

$$S\left(P;R\right) = \frac{P}{r+\delta-\zeta} - \left(\frac{1}{r+\delta}\right)rU - \left[\frac{R}{r+\delta-\zeta} - \left(\frac{1}{r+\delta}\right)rU\right]\left(\frac{P}{R}\right)^{\alpha} \tag{4}$$

where α is the negative root of the following quadratic equation

$$\alpha^2 \frac{\sigma^2}{2} + \alpha\left(\zeta - \frac{\sigma^2}{2}\right) - r - \delta = 0$$

One can solve for the optimal reservation output using a standard first-order condition with respect to R. The resulting solution is homogenous of degree zero in P, so that R is identical across matches, as one should expect. Its optimal value is given by

$$R = \left(\frac{\alpha}{\alpha-1}\right)\left(\frac{r+\delta-\zeta}{r+\delta}\right)rU \tag{5}$$

From the definition of α, it is easily seen that R is upper-bounded by the opportunity cost of employment rU. Within the neighborhood of R, output is too low to cover costs but the job might turn profitable again thanks to future shocks. Therefore the worker and the firm procrastinate up to the point where the value of waiting equals the operational losses. On the contrary when productivity is constant, there is no labor-hoarding and the reservation output is equal to the opportunity cost of employment.

8

2.3 The equilibrium

This section characterizes the equilibrium rate of unemployment and the joint distribution of job spells and wages. It will be shown in section 3 that these statistics have closed-form solutions when the sampling distribution is lognormal. But for the moment we keep the analysis as general as possible by not imposing any parametric assumption. First of all, we notice that the Nash-bargaining problem is satisfied if and only if wages are such that

$$w(P_t) = \beta P_t + (1 - \beta)rU \tag{6}$$

The wage follows from output by a location transformation. So the discussion can be restricted to the output distribution, without loss of generality. The derivations are based on the premise that the labor market is in steady state so that job flows are constant and balance at all time.[10] The statistics of interest are derived using a progressive approach: starting from the most informative one, namely the joint distribution of job spells and output, we aggregate it step by step in order to obtain the rate of unemployment.

Proposition 2 *The joint density of output x and tenure T is given by*

$$\upsilon(x, T) = u\lambda \left(\int_R^{+\infty} \psi(x, T; P) \, dG(P) \right) \tag{7}$$

where u denotes the rate of unemployment. For $x \in [R, +\infty)$, the function $\psi(x, T; P)$ is

$$\psi(x, T; P) = \left(\frac{e^{-\delta T}}{x} \right) \left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(x) - \ln(P) - \mu T}{\sigma\sqrt{T}} \right)^2}}{\sigma\sqrt{2\pi T}} - \left(\frac{R}{P} \right)^{\frac{2\mu}{\sigma^2}} \frac{e^{-\frac{1}{2}\left(\frac{\ln(x) + \ln(P) - 2\ln(R) - \mu T}{\sigma\sqrt{T}} \right)^2}}{\sigma\sqrt{2\pi T}} \right) \tag{8}$$

where $\mu = \zeta - \frac{\sigma^2}{2}$.

The first term on the right hand side of (7) measures the number of contacts between job seekers and firms. The function $\psi(x, T; P)$ is the conditional joint density of current output and tenure given initial output. From the set of sample paths starting from P and reaching

---

x after tenure T, it deduces all those that breach the separation threshold R.[11] Given that initial productivities are drawn from the sampling distribution, the unconditional joint density is obtained integrating $\psi(x, T; P)$ with respect to $G(P)$. The integration is performed from R up to infinity because contacts lead to matches solely when P is above R. The density of jobs with a given current output is readily obtained from (7) after having integrated tenure from 0 up to infinity. The following proposition shows that the resulting integral can be expressed analytically.

Proposition 3 *The density of output x is given by*

$$v(x) = u\lambda \left(\int_R^{+\infty} \varphi(x; P) dG(P) \right) \tag{9}$$

where the function $\varphi(x; P)$ is defined for $x \in [R, +\infty)$ and is equal to

$$\varphi(x; P) = \begin{cases} P^{-1} \left(\frac{x}{P}\right)^{\frac{\mu - \gamma}{\sigma^2} - 1} \left(\frac{1 - \left(\frac{R}{P}\right)^{\frac{2\gamma}{\sigma^2}}}{\gamma} \right) ; \text{if } x > P \\ \\ P^{-1} \left(\frac{x}{P}\right)^{\frac{\mu + \gamma}{\sigma^2} - 1} \left(\frac{1 - \left(\frac{R}{x}\right)^{\frac{2\gamma}{\sigma^2}}}{\gamma} \right) ; \text{if } x \in [R, P] \end{cases} \tag{10}$$

and $\gamma = \sqrt{\mu^2 + 2\delta\sigma^2}$.

In a similar way, the aggregate rate of employment follows integrating equation (9) from R up to infinity. Again the calculation leads to a closed-form expression that is given in Proposition 4. The expression is reminiscent of the equilibrium rate of unemployment under certainty. Actually, when uncertainty vanishes so that σ goes to zero, the term $(R/P)^{\frac{\mu+\gamma}{\sigma^2}}$ also converges to zero. Thus the expression of u converges to the standard solution under certainty.

Proposition 4 *The equilibrium rate of unemployment is equal to*

$$u = \frac{\delta}{\delta + \lambda \int_R^{+\infty} \left(1 - \left(\frac{R}{P}\right)^{\frac{\mu+\gamma}{\sigma^2}} \right) dG(P)} \tag{11}$$

---

[11]Notice that an econometrician who observes the workers' wages at different points in time of their jobs spells could use equation (8) to compute the likelihood of their sample paths.

In this section we have presented the statistics that will be useful for the econometric estimation. Next section details the econometric procedure and analyzes the property of the likelihood function.

3   Estimation procedure

We now discuss how to estimate the model's parameters. The searching costs parameter s is not identified because it enters the likelihood function only through its impact on R. As explained below, we will treat the reservation output as if it were an endogenous parameter to be estimated. After all the parameter estimates have been obtained, the equilibrium conditions can be used to retrieve the implied searching costs. Conversely, the values of r and β have to be fixed prior to the estimation. While not so problematic for r since it has been estimated with precision in other research, the calibration of β is more unsettling. Although the bargaining power is theoretically identified due to the highly non-linear likelihood function, trials show that in practice the model fails to pin it down. In the absence of informations on firm profits, it is not surprising that the dataset does not allow us to recover both sizes and allocations of the jobs' surpluses. This difficultly is well-known and is now gradually overcome by research based on matched employer–employees data (see Cahuc et al., 2006). Given the one-sided nature of the CPS data, we stick to the usual practice of assuming symmetric bargaining and then perform some robustness test with respect to β.

3.1   *The likelihood function*

Following these preliminary steps, the likelihood of the sample can be expressed as a function of the remaining set of parameters. We slightly restrict the generality of the problem by assuming that the sampling distribution $G(\cdot)$ can be completely parametrized in terms of a finite-dimensional vector Ω so that the set of estimated parameters $\Theta = \{\zeta, \delta, \sigma, \Omega, R, \lambda\}$.

The likelihood of the sample is computed as follows. Let Y denote the set of observations, so that $Y \equiv \{y_1, y_2, ..., y_n\}$ where n is the total number of workers in the sample. The individual observations are defined using three variables: w^i, T^i, τ^i. The variables w^i and

T^i are the current hourly wage and job tenure. In the case where worker i fails to report the length of his job spell, T^i is obviously ignored. If worker i is currently searching for a job, y^i is set equal to the unemployment duration τ^i. The likelihood function is therefore made of three distinct components. The individual contribution of a job searcher is equal to the density associated with an on-going unemployment spell of length τ conditional on unemployment times the probability of observing an unemployed worker

$$f(\tau, u) = f(\tau \mid u)u = \left(\lambda \overline{G}(R)e^{-\lambda \overline{G}(R)\tau} \right) u = \frac{\delta \lambda \overline{G}(R)e^{-\lambda \overline{G}(R)\tau}}{\delta + \lambda \int_R^{+\infty} \left(1 - \left(\frac{R}{P} \right)^{\frac{\mu+\gamma}{\sigma^2}} \right) dG(P)}$$

where $\overline{G}(P) \equiv 1 - G(P)$. The likelihood of observing an employed worker paid wage w is given by $\upsilon(x(w))$. The expression can be further decomposed reinserting (11) into (9) to obtain

$$f(w, e) = \upsilon(x(w)) = \frac{\delta \lambda \left(\int_R^{+\infty} \varphi(x(w); P) dG(P) \right)}{\delta + \lambda \int_R^{+\infty} \left(1 - \left(\frac{R}{P} \right)^{\frac{\mu+\gamma}{\sigma^2}} \right) dG(P)}$$

Notice that output is defined as a function of the observed wage. Its implicit value follows from combining (5) with (6). Similarly, the joint likelihood of observing a worker paid wage w with a job tenure equal to T is given by

$$f(w, T, e) = \upsilon(x(w), T) = \frac{\delta \lambda \left(\int_R^{+\infty} \psi(x(w), T; P) dG(P) \right)}{\delta + \lambda \int_R^{+\infty} \left(1 - \left(\frac{R}{P} \right)^{\frac{\mu+\gamma}{\sigma^2}} \right) dG(P)}$$

Putting together these three components, the log likelihood for the observed sample reads

$$
\begin{aligned}
\ln L(\Theta, Y) = {} & n \left(\ln(\lambda) + \ln(\delta) - \ln \left(\delta + \lambda \int_R^{+\infty} \left(1 - \left(\frac{R}{P} \right)^{\frac{\mu+\gamma}{\sigma^2}} \right) dG(P) \right) \right) \\
& + n_U \ln \left(\overline{G}(R) \right) - \lambda \overline{G}(R) \sum_{i \in U} \tau^i + \sum_{i \in W} \ln \left(\int_R^{+\infty} \varphi(x(w^i); P) dG(P) \right) \\
& + \sum_{i \in H} \ln \left(\int_R^{+\infty} \psi(x(w^i), T^i; P) dG(P) \right) \quad (12)
\end{aligned}
$$

where n_U is the number and U is the set of indices of job searchers in the sample, W is the

set of indices of employees who only report their current wage and H is the set of indices of employees who report both wage and job spell. Although the parameters ζ and Ω do not appear in the analytical expression of the likelihood function, they are implicitly identified: ζ determines the values of μ, γ, $\varphi(\cdot)$ and $\psi(\cdot)$, while the parametric vector Ω obviously characterizes $G(\cdot)$. Notice that, for the sake of the estimation, the reservation output is treated as a primitive parameter of the model.

The likelihood function is continuously differentiable and its parameters belong to a compact support. Nevertheless, it does not satisfy all the standard requirements for a well-behaved likelihood function since the support of the distribution of the data is a function of the parameters. Furthermore the likelihood of observing the reservation wage is equal to zero. As explained in the introduction, this feature implies that we cannot use the smallest observed wage as a super-consistent estimator. Nevertheless, as stated in Proposition 5, the estimators satisfy asymptotic normality under standard requirements.

Proposition 5 *Suppose that (i) The parameter space Γ is compact and contains an open neighborhood of the true value Θ_0 of the population parameter; (ii) The sampling distribution $G(P)$ is continuously differentiable. Then the maximum likelihood estimator*

$$\widehat{\Theta} = \arg\max_{\Theta \in \Gamma} \ln L\left(\Theta, Y\right)$$

converges in probability to Θ_0 so that $\sqrt{n}\left(\widehat{\Theta} - \Theta_0\right) \xrightarrow{d} N\left(0, H^{-1}JH^{-1}\right)$ where H is the Hessian of the likelihood function and J is the information matrix.

The proof of Proposition 5 relies on the fact that $f(w(R), e)$ and $f(w(R), T, e)$ are both equal to zero. As shown in Greene (1980), this property justifies the interchange of the order of integration and differentiation so that the asymptotic property of the estimator can be characterized by linear approximation. Our problem is slightly less standard than the one considered by Greene because the derivatives of the density functions with respect to Θ are not equal to zero when evaluated at the reservation output. Thus the interchange of the order of integration and differentiation is justified solely for the first derivative. This is why the hessian matrix H is not equal to $-J$ so that the asymptotic covariance matrix cannot be simplified and set equal to J^{-1}. But, as explained in Newey and McFadden

(1994), the information matrix equality is not essential to asymptotic normality. The only complication is technical and due to the more intricate form of the asymptotic variance.

3.2 *Lognormal sampling distribution*

We have characterized the estimation procedure for general sampling distributions. The econometric implementation of the model requires to narrow the analysis to a particular family of distributions. Accordingly we will hereafter assume that $G(\cdot)$ is lognormal. Lognormal distributions are commonly assumed because they satisfy the "recoverability condition" defined by Flinn and Heckman (1982), meaning that their location and scale parameters can be recovered from truncated observations. The class of functions which satisfy the "recoverability condition" also encompasses, among others, gamma distributions.[12] Thus lognormality is eventually justified by its good fit of the data. In our case this assumptions has a more crucial role. Given the intricate expression of the likelihood function, there is little hope to derive it in closed-form. Solely when initial productivities are lognormally distributed in the population, does the likelihood function admits an analytical expression so that approximation errors due to numerical integrations can be avoided. Given its length, we do not include the expression of $L(\Theta)$ in the body of the paper.[13]

Proposition 6 *Under the assumption that the initial productivities are drawn from a lognormal distribution, so that*

$$dG(P) = \frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{P\xi\sqrt{2\pi}}dP \tag{13}$$

the likelihood functions $L(\Theta)$ has a closed-form solution. The resulting expression is reported in Appendix.

---

[12]See Flinn (2006) for a careful discussion of the class of functions which satisfy the "recoverability condition".

[13]The derivation of an analytical expression for the likelihood function is made possible by the fact that geometric Brownian motions are also lognormally distributed.

4 Empirical results

4.1 *Data*

We now estimate the model using cross-sectional data. Whereas most surveys systematically ask unemployed workers to report the time they have been searching for a job, employees are rarely asked the length of their job spells. As a result, data on job durations are scarcer than data on unemployment durations. A notable exception is the January/February supplement of the Current Population Survey. The CPS is structured as a rotating panel with 4 months of participation, 8 months without interviews, and 4 more months of participation after which the household is taken out of the panel. In January or February, the current wages and job spells of the Outgoing Rotation Groups[14] are collected. More precisely, employees are asked the following question: How long have you been working continuously for your present employer?

Hence the job tenure supplement provides data on both job spells and wages for a supposedly random sample of one fourth from the January 2004 CPS. We use data on males and females with an age between 20 and 65 years. Since our model does not include a state of non participation to the labor market, we have restricted our sample to individuals who indicated that they were currently employed or actively searching for a job. For the same reason, we have excluded individuals observed as self-employed, working part time or employed in the non-civilian labor force. After excluding observations with missing wage data, we restricted the sample to workers with a high school graduation diploma or less. Finally we have trimmed the sub-sample by excluding observations below the bottom percentile of the wage distribution. The trimming is particularly important for the estimation of the reservation wage since it allows us to avoid implausibly low estimates due to measurement errors.[15]

---

[14]In our case, the Outgoing Rotation Groups are composed of the households that entered the panel in October 2002 and 2003.

[15]The trimming of the data is especially useful given the nature of the observations. For the workers which are not paid on an hourly basis, we have divided their gross weekly wage by their usual hours of work per week in order to impute their hourly wage. This computation obviously interacts potential measurement errors and for some observations leads to implausibly low hourly wages.

TABLE I

DESCRIPTIVE STATISTICS[*]

	All	Males	Females
Age	40.475 (11.403)	39.841 (11.316)	41.358 (11.467)
Working week (hours)	41.411 (5.750)	42.324 (6.720)	40.163 (3.723)
Average spells (months)			
Unemployed	4.537 (5.011)	4.438 (4.956)	4.704 (5.117)
Job	92.586 (97.160)	96.773 (101.926)	86.601 (89.604)
Hourly wage			
All jobs	13.777 (6.982)	14.932 (7.458)	12.199 (5.922)
Job spell<1 year	11.829 (6.308)	12.801 (6.843)	10.394 (5.109)
Entrants	10.301 (4.486)	11.059 (5.140)	9.439 (3.436)
	Number of Observations		
Labor Market Position			
Unemployed	413	260	153
Employed	4336	2504	1832
Reported job spells			
Total number	3676	2163	1513
Job spell<1 year	535	319	216
Entrants	154	82	72
Sample size	4749	2764	1985

* Standard Deviations in Parenthesis.

Descriptive statistics are reported in Table 1. It contains statistics for jobs with a reported tenure below one year. Their wage distribution will be very close to the estimated sampling distribution because the estimation procedure approximates the latter using observations with short job spells. This points to potential bias since the distribution of wages among job entrants has a lower mean than the distribution of wages among workers with less than one year of tenure.[16] This feature of the data is easily explained by on-the-job

---

[16]We are able to identify individuals entering the employment pool by excluding from the set of workers
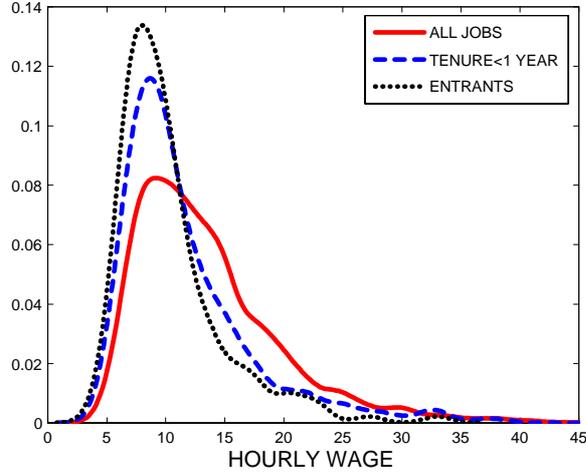
Figure 1: Wage densities.

search as employees select offers which are above their current wage. Given that our model excludes on-the-job search, the job-ladder effect is ignored and consequently our estimates of the location and scale parameters of the sampling distribution will be biased upwards. As shown in section 4.3 where we analyze the robustness of the estimation procedure, this leads to downward biases for the estimate of the rate of LBD.

The non-parametric kernel density estimates of the three wage distributions are reported in Figure 1. As expected, the distribution of wages among job entrants is located to the left and exhibits slightly less dispersion than the distribution of wages among workers with less than one year of tenure. On the contrary, the dispersion of the aggregate wage distribution is much higher. This feature fits well the model since Brownian motions are diffusion processes, meaning that their distributions become more and more dispersed as time elapses.

4.2   *Estimates*

The estimated parameters and their standard deviations are reported in Table 2. We also estimate the deterministic model using the procedure devised by Flinn and Heckman (1982). Table 2 makes clear that their approach is nested into the one proposed in this

with less than a year of tenure those that were working a year ago.

17

paper. Notice that, in addition to the rate of LBD ζ and the variance parameter σ, the model also allows us to estimate the value and standard deviation of the reservation wage w_r. Conversely, the value reported for the deterministic model corresponds to the lowest wage in the sample so that its standard deviation is not defined.

The estimates of Σ and ξ imply that the mean and dispersion of the sampling distributions are higher in the deterministic model than in the stochastic model. This result is intuitive since the deterministic model is based on the premise that the sampling distribution and the aggregate distribution are one and the same. To the opposite, the estimation of the stochastic model sets the parameters Σ and ξ so as to fit the distribution of wages among jobs with a short tenure and so yields smaller values for both parameters.

Conversely, the exogenous rate of job destruction δ is similar. This is somewhat surprising since the stochastic model generates endogenous separations. Thus one might expect that the rate of exogenous job destruction would be significantly lower in a stochastic environment. For the very small estimate of the variance parameter, however, endogenous separation is a marginal phenomenon. Accordingly the separation rates are quite similar in both models with an estimated average length of a job close to 90 months.

The estimated values of λ imply that job searchers receive a job offer every 6 months. Although the model over-estimates the average unemployment duration, the predicted unemployment rate is equal to 6.4% whereas its value in the data is 8.7%. These opposite biases suggest that the sample does not completely satisfy the stationarity assumption, most probably because of a significant and recent entry of workers into the labor force.

We now turn our attention to the estimates that are specific to the stochastic model. First of all, we notice that the variance parameter σ is quite low. With a standard deviation close to 3.5%, the model predicts that the sample paths are nearly deterministic. Nonetheless, one cannot set σ to zero and at the same time estimate ζ. In a deterministic environment the lower-bound of the joint distribution is an increasing function of tenure and consequently the likelihood of observing a worker with a seniority equal to T and a wage inferior to $e^{\zeta T} w_r$ is zero. Given that the sample contains such observations as long as ζ significantly differs from zero, the deterministic model necessarily collapses to the case where the sampling and aggregate wage distributions are indistinguishable. Thus there is a fundamental link between the introduction of uncertainty and the estimation of the LBD

rate, the former being necessary to implement the latter.[17]

<div align="center">

TABLE II

MODEL ESTIMATES[*]

</div>

Parameters	$\beta = 0.5$		$\beta = 0.4$	
	Stochastic	Deterministic	Stochastic	Deterministic
w_r	3.76 (.180)	4.11 (___)	3.60 (.213)	4.11 (___)
ζ	.0204 (.001)	___	.0207 (.001)	___
σ	.0350 (.006)	___	.0267 (.012)	___
δ	.133 (.002)	.133 (.002)	.133 (.002)	.133 (.002)
Σ	2.86 (.018)	3.01 (.008)	3.05 (.023)	3.19 (.008)
ξ	.488 (.008)	.526 (.006)	.506 (.008)	.550 (.006)
λ	1.95 (.073)	1.95 (.081)	1.94 (.073)	1.95 (.085)
$\ln L$	-28874	-29045	-29837	-30000

Discount rate: $r = 5\%$

* Standard Errors in Parenthesis.

The estimated rate of LBD is close to 2%. This translates into a smaller wage growth of around 1.75% per year because the constant outside option accounts for a substantial share of wages. The model predicts that ten years of tenure raises the average wage by about 18.8% so that our estimate of the cumulative returns to tenure lies in the range separating the low returns obtained by Altonji and Williams (2005) from the high returns obtained by Topel (1991). Notice that given the low value of the variance, non-random selection in who acquires seniority is not a source of important bias.

Table 2 also contains the estimates when β is equal to 0.4 instead of 0.5. A bargaining power of 0.4 is in the range of the estimates obtained by Flinn (2006) using wage-share informations for CPS sample members between 16 and 24 years of age. Of all the estimates,

---

[17]An alternative strategy is to introduce measurement errors. As discussed in the introduction, it has the drawback of neglecting the interactions between job separation and LBD. Furthermore, the likelihood function with normally distributed measurement errors cannot be expressed in closed-form. Thus, although this alternative strategy is arguably more stylized, it has to be implemented through numerical integration.

the most sensitive to the change in β are the parameters determining the shape of the sampling distribution of output. A lower value of β implies that the worker receives a smaller share of the job's output, so the observed dispersion of wages must result from a higher degree of productivity dispersion. This is why both Σ and ξ substantially increase. Conversely, the estimate of ζ is quite robust to variations in the bargaining power. It slightly increases because workers must learn at a higher rate to benefit from a given pay raise.

We now consider the ability of the model to fit the sample information. Of most interest to our analysis are its predictions about the joint distribution of wages and job spells. The data for jobs with a tenure below 1, 5 and 10 years as well as the aggregate distribution are reported against their simulated counterparts in Figure 2. The panels illustrate the ability of the model to fit almost perfectly the gradual increase in dispersion of the cross-sectional distributions. Yet, careful inspection shows that the simulation tends to be a little bit less responsive to changes in tenure. More precisely, the mean of the wage distribution among workers with less than one year of tenure is slightly higher than in the data. Next section proposes an alternative estimation procedure which reduces this discrepancy and evaluates its impact on ζ. We also notice that the model matches very well the right tails of the distributions. This a classical test for models of wage dispersion due to the "heavy tail" property of the data. As explained in Prat (2006), the cross-sectional distribution aggregates underlying distributions with right-tails of Pareto functional form. Therefore it is not surprising that the model easily fits the wage distribution at high quantiles.

Since we have excluded on-the-job-search, one may wonder whether the fit of the wage distribution is achieved at the expense of the turnover process. To address this potential concern, we report in Figure 3 the actual distribution of job spells together with the structural estimation. While not as convincing than for wages, the simulation is still reasonably close to the data. Given that endogenous separations are quite rare, the stochastic and deterministic models have very similar predictions about the distribution of job spells. Thus the hazard rate of job separation is almost flat. This result is somewhat disappointing since it is shown in Prat (2006) that the framework estimated in this paper has the potential to fit the hump-shaped hazard rate of job separation that is observed in the data. Unfortunately this would require to set the idiosyncratic variance σ to a higher value than the one resulting from the estimation.
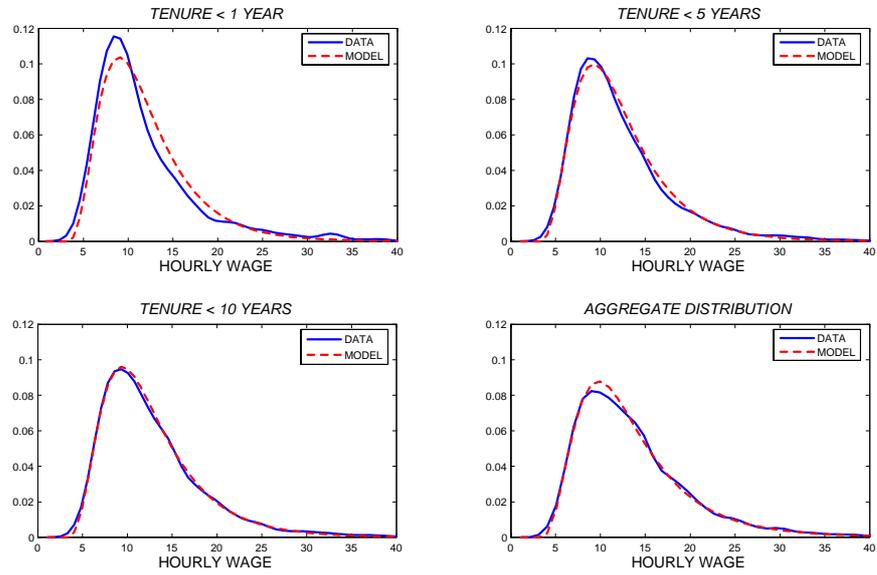
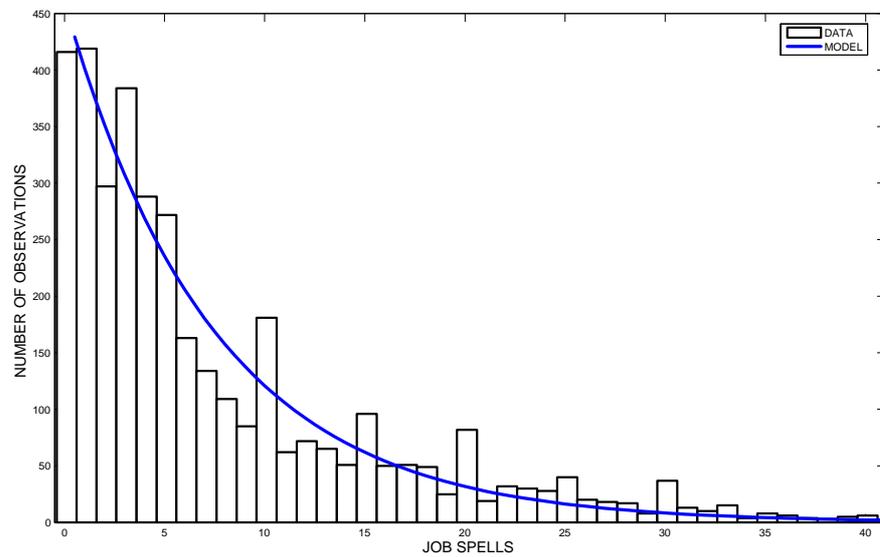Figure 2: Wage densities at different job spells.

Figure 3: Job spells distribution.

21

4.3 Robustness

In this section we assess the robustness of the estimation procedure. We focus on the biases induced by the exclusion of on-the-job search. Obviously a complete evaluation would require to devise a comprehensive model but such a project is beyond the scope of this paper. Nevertheless, we are able to partly control for the impact of the job ladder effect on the estimation of the sampling distribution. By focusing on employees who were unemployed a year ago, we can infer the actual wage distribution among job entrants. Thus we can set $\Omega = \{\Sigma, \xi\}$ so as to fit the entrants distribution for every given value of the vector of remaining parameters $\Theta_1 \equiv \{\zeta, \delta, \sigma, R, \lambda\}$ and then maximize the sample likelihood with respect to Θ_1.

The resulting estimators are not "two-step" estimators as the optimal parametric vector $\Omega = \{\Sigma, \xi\}$ changes with Θ_1. In other words, we cannot estimate Ω independently because the mapping between wages and productivity depends on Θ_1. Instead we define $\Omega(\Theta_1, Y_1)$ as a function of Θ_1 and of the sub-set of observations $Y_1 \subseteq Y$ used to infer the shape of the sampling distribution. The procedure shares some formal similarities with concentrating the likelihood function but it is substantially different since we do not set $\Omega(\Theta_1, Y_1)$ to maximize the likelihood of the sample but instead to approximate available information about the sampling distribution.[18] This approach is particularly justified if one suspects that the model is somehow misspecified. Then it is well known that full-information estimation can be a source of significant bias. By constraining Ω to fit the entrants distribution and imposing its value afterwards, we are able to reduce the size of the bias.[19]

We call the values resulting from this procedure restricted estimates. They are reported in Table 3 for the two distinct restrictions where Y_1 contains either the wages of job entrants or the wages of workers with less than one year of tenure. In the first case, the experiment gives us a sense of how estimates are biased by the job-ladder effect. As expected, both

---

[18]Notice that when Y_1 is equal to Y, the restricted estimation is equivalent to concentrating the likelihood function so that it coincides with the full estimation procedure.

[19]Since the properties used in the proof of Proposition 5 are satisfied, the estimators are still consistent and normally distributed although not efficient. The asymptotic variance of $\widehat{\Omega}$ is derived using the delta method so that

$$Var\left(\widehat{\Omega}\right) = \left.\frac{\partial \Omega(\Theta_1, Y_1)}{\partial \Theta_1'}\right|_{\Theta_1 = \widehat{\Theta}_1'}' \left(H^{-1}JH^{-1}\right) \left.\frac{\partial \Omega(\Theta_1, Y_1)}{\partial \Theta_1'}\right|_{\Theta_1 = \widehat{\Theta}_1'}$$

Σ and ξ significantly decrease. This leads to noticeable increases in both ζ and σ: given that the wage distribution among job entrants has a lower mean and dispersion than the one resulting from the full-estimation, the fit of the sample information is achieved through a higher rate of LBD and more idiosyncratic uncertainty. The estimated value of 3.4% for ζ should be interpreted as an upper-bound since the job-ladder effect is excluded from the sampling distribution but not from the aggregate distribution. Accordingly a complete evaluation would require to explicitly model on the job search and the restricted estimation suggests that this would lead to substantial adjustments of the estimates.

In the second case $\Omega(\Theta_1, Y_1)$ fits the distribution of wages among workers with less than one year of tenure. As can be seen from Figure 2, this restriction is motivated by the fact that the full-estimation uses Ω to improve the fit of the aggregate distribution at the cost of slightly over-estimating the mean of the wage distribution among jobs with a tenure below one year. Controlling for this bias leads to small but noticeable increases in ζ.

TABLE III

RESTRICTED ESTIMATES[*]

Parameters	$\beta = 0.5$		$\beta = 0.4$	
	Entrants	Tenure < 1Year	Entrants	Tenure < 1Year
w_r	3.79 (.398)	3.58 (.409)	3.58 (.129)	3.48 (.115)
ζ	.034 (.005)	.0223 (.001)	.034 (.001)	.0228 (.0006)
σ	.059 (.002)	.0271 (.002)	.055 (.001)	.0276 (.0003)
δ	.132 (.005)	.133 (.002)	.132 (.003)	.133 (.002)
Σ	2.67 (.037)	2.84 (.032)	2.86 (.014)	3.02 (.010)
ξ	.477 (.019)	.514 (.016)	.491 (.007)	.531 (.005)
λ	1.94 (.082)	1.96 (.076)	1.93 (.073)	1.96 (.075)
ln L	-29009	-28892	-29974	-29856
Discount rate: $r = 5\%$				

* Standard Errors in Parenthesis.

5 The effect of learning-by-doing

In this section we introduce an aggregate matching function to close the model and evaluate
the impact of the rate of LBD on labor market outcomes. We assume that the matching
process is similar to the one described in Pissarides (2000). Since the matching function
has become the workhorse for the study of equilibrium unemployment, the exposition can
be brief. Firms post vacancies that are randomly matched and incur a flow cost equals to
c. The number of job matches per unit of time is a function of the number of vacancies
and job seekers. When the aggregate matching function is homogenous of degree one, the
rate at which a vacancy meets a worker only depends on the unemployment rate u and on
the ratio v of vacant jobs divided by the size of the labor force. The transition rate for
vacancies is given by a function $q(\theta)$ where the labor market tightness parameter θ denotes
the vacancy-unemployment ratio. Similarly, jobs seekers meet firms at the rate $\theta q(\theta)$.

As opposed to the labor force whose size is fixed and normalized to one, new firms enter
the market until arbitrage opportunities are exhausted. Therefore the *Free-Entry condition*
is given by

$$c = q(\theta) \int_R^{+\infty} (1-\beta) S(P) dG(P) \tag{14}$$

Similarly we can replace in (2) the exogenous contact rate λ by $\theta q(\theta)$ to obtain

$$rU = -s + \theta q(\theta) \int_R^{+\infty} \beta S(P) dG(P)$$

Reinserting (14) into the previous equation allows us to solve for the asset value of being
unemployed as a function of the labor market tightness

$$rU = -s + c\theta \left(\frac{\beta}{1-\beta} \right)$$

The job's surplus follows from replacing the previous equation into (3). Accordingly the
Optimal Separation rule is such that

$$R = \left(\frac{\alpha}{\alpha - 1} \right) \left(\frac{r + \delta - \zeta}{r + \delta} \right) \left(-s + c\theta \left(\frac{\beta}{1-\beta} \right) \right) \tag{15}$$

The equilibrium values of the two endogenous variables θ and R are determined by the

equilibrium conditions (14) and (15).[20] Given that the aggregate matching function defines a one-to-one mapping between λ and θ, a parametric assumption allows us to retrieve the values of the labor market tightness and searching costs using the estimates reported in the previous section. As it is common in the literature, we assume that the matching function is Cobb-Douglas. We further restrict our attention to the case where the allocation is efficient and consequently use the "Hosios condition" to set the elasticity of the matching function equal to β, so that $q(\theta) = \theta^{-1/2}$.

Table 4 contains the implied costs of search and equilibrium labor market tightness for the deterministic and stochastic models as well as for the restricted estimates. The high values of the equilibrium tightness is unreasonable if interpreted as the ratio of vacancies to job seekers. Thus one should interpret θ as measuring the ratio of recruitment effort to search effort. The relatively high searching costs are required to offset the important surpluses of the matches in the right tail of the sampling distribution.

TABLE IV

POINT ESTIMATES OF REMAINING VARIABLES[*]

	Deterministic	Stochastic	Entrants	Tenure < 1Year
Tightness				
θ	3.86 (.320)	3.80 (.286)	3.86 (.321)	3.95 (.302)
Flow Costs of Search				
s	98.2 (4.36)	92.7 (4.45)	80.2 (7.29)	96.0 (6.37)
c	26.8 (1.26)	25.4 (1.25)	21.8 (2.23)	25.2 (1.85)

* Standard Errors in Parenthesis.

By keeping the values of the environmental parameters constant and varying the rate of LBD, we can simulate its impact on labor market outcomes. The results are reported in Figure 4. The upper-left panel contains the effect on the aggregate wage distribution. Not surprisingly, a higher rate of LBD increases the mass in the right-tail. But this does not necessarily leads to higher inequality because the left tail of the wage distribution is truncated by the increase in the reservation wage. The ratio of standard deviation to

---

[20]We refer the reader to Prat (2006) for a proof of the existence and uniqueness of the equilibrium when the sampling distribution $G(\cdot)$ is degenerate.
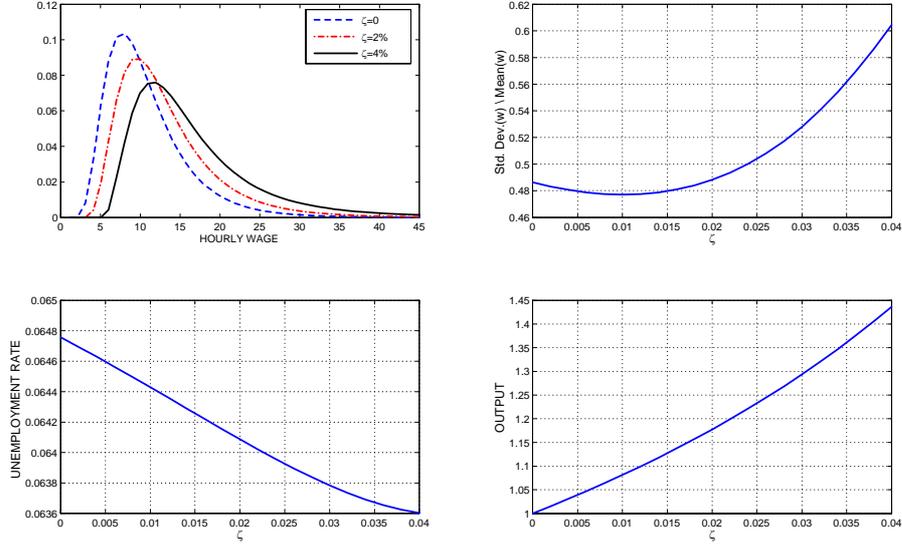
Figure 4: Effect of LBD on the equilibrium.

average wage reported in the upper-right panel shows that the latter effect dominates when the rate of LBD is close to zero.

The unemployment rate as a function of ζ is reported in the lower-left panel. As expected, the function is decreasing but its elasticity is probably below what one's intuition might suggest. To understand why this is the case, it is useful to recall that the opportunity cost of employment rU is equal to $-s + c\theta\beta/(1-\beta)$. This implies that, for the estimated value of the recruitment costs, the impact of θ on the worker's outside option is amplified by more than one order of magnitude. Hence, small adjustments of the labor market tightness have drastic effects on jobs' surpluses and so unemployment remains remarkably stable. Obviously the rigidity of the unemployment rate is even more pronounced in the deterministic model since its estimated recruitment costs are higher.[21]

The lower-right panel contains a plot of the aggregate output as a function of the LBD rate. We normalize aggregate output to one when $\zeta = 0$ for ease of interpretation. The

[21]See Hornstein et al. (2006) for a detailed analysis of the discrepancy between the value of non-market activity derived from structural estimations and the one required to fit business-cycle fluctuations of the unemployment rate. Our analysis shows that controlling for the effect of tenure on wage dispersion reduces the gap, but only modestly.

model predicts that an increase of the LBD rate from 0 to 4% raises aggregate output by around 45%. These gains arise due to three reinforcing effects: (i) the direct impact of LBD obviously leads to a higher average output for a given job spell, (ii) the increase of the reservation wage implies that, *ceteris paribus*, ongoing job relationships have a higher average productivity, (iii) the higher rate of employment mechanically raises aggregate output. Decomposing the relative importance of these effects show that the first one accounts for nearly 97% of the total gains whereas the increase in the reservation wage explains most of the remaining 3%.

6 Conclusion

It has been shown in this paper how the production side of the Mortensen and Pissarides model with endogenous job separation can be estimated by maximum likelihood using cross-sectional data. The analysis establishes that the parsimoniously specified model convincingly fits the joint distribution of wages and job spells once LBD is taken into account. A concrete contribution of the analysis is to identify the rate of LBD in an equilibrium set-up, whereas the estimates available in the literature are typically based on "reduced-form" estimations. Introducing an aggregate matching function allows us to close the model. We find that the rate of LBD has a significantly positive effect on aggregate output and a small impact on employment.

The structural model could easily be estimated using panel data. Actually, Proposition 2 already contains the expression of the likelihood function for the jobs' sample paths. This explains why we have decided to devise an estimation procedure for cross-sectional data, since it is a more comprehensive task. From an empirical perspective, however, panel data would certainly be desirable. They would make it possible to take into account the patterns of wage dynamics. We conjecture that estimating the model using panel data would raise the idiosyncratic variance and thus the rate of endogenous separation.

As we have deliberately tipped the balance in favor of tractability over realism, the model also lends itself naturally to several theoretical extensions. We conclude by briefly discussing some of these. The most obvious refinement would be to introduce general human capital. Although not so demanding at the conceptual level, this extension will come at the cost of closed-form solutions. More interesting is the introduction of on-the-

job search since it would connect the model with the burgeoning econometric literature based on employers competition. Until recently, uncertainty and on-the-job search have been considered in isolation. But, as attested by a series of recent papers (Postel-Vinay and Thuron, 2005; Yamagushi, 2006), the importance of combining both dimensions is now widely recognized. Such a research project raises serious technical challenges. For the moment, these structural models treat job separations as exogenous and available estimates are based on indirect inference methods. This paper suggests that stochastic calculus can help to alleviate some of the difficulties. Finally, we also hope that the derivations of the asymptotic properties of the estimators would be of some interest to researchers working in other areas than labor economics since our result can be applied to a wide class of models with endogenous exit.

Appendix

- **Proof of Proposition 1:** We guess that R does not depend on current output. Then the Bellman equation satisfied by the surplus within the continuation region follows by Ito's Lemma

$$(r + \delta)S(P_t^i, R) = P_t^i - rU + \zeta P_t^i S_1(P_t^i, R) + \frac{(P_t^i \sigma)^2}{2} S_{11}(P_t^i, R)$$

where number subscripts denote the partial derivatives of the function. It is well known that the general solution of this partial differential equation is of the form

$$S(P_t^i, R) = C(R)\left(\frac{P_t^i}{R}\right)^\alpha + D(R)\left(\frac{P_t^i}{R}\right)^\eta + E_{P_t^i}\left[\int_t^{+\infty} e^{-(r+\delta)(\tau - t)}\left(P_\tau^i - rU\right)d\tau\right] \tag{16}$$

where $C(R)$ and $D(R)$ are some constants of integration which do not depend on the current state P_t^i, while α and η are respectively the negative and positive roots of the quadratic equation

$$\frac{\sigma^2}{2}\chi(\chi - 1) + \chi\zeta - r - \delta = 0$$

The values of $C(R)$ and $D(R)$ are pinned down by the two boundary conditions. Firstly, when P_t^i diverges to infinity, the values of the option to separate converges to zero. This implies that we can set $D(R)$ equals to zero in order to eliminate the positive root η. The second boundary, $\lim_{P_t^i \to R} S(P_t^i, R) = 0$, follows from the definition of the reservation threshold R. One can easily verify that the solution proposed in (4) satisfies the differential equation and boundary conditions. The optimal reservation productivity is set so as to maximize the surplus. Since current revenues are independent of the reservation productivity, it can be shown that $\partial S(P_t^i, R)/\partial R = 0$ when[22]

$$\left.\frac{\partial S(P_t^i, R)}{\partial P_t^i}\right|_{P_t^i = R} = 0 \tag{17}$$

---

[22] See Merton(1973), p.171.

It is commonly referred to equation (17) as the smooth-pasting condition. Its solution reads

$$R = \left(\frac{\alpha}{\alpha - 1}\right)\left(\frac{r + \delta - \zeta}{r + \delta}\right)rU \tag{18}$$

which is equivalent to (5) and verifies our guess that R is independent of P_t^i.

- **<u>Proof of Proposition 2:</u>** Consider a match i that is operational at date t. We define the stopping time τ_1^i as the time of arrival of the first exogenous destruction shock and

$$\tau_2^i = \min\left\{\tau > t : P_\tau^i = R\right\}$$

So τ_2^i is the first time at which the job would have been endogenously destroyed. Hence, job i is operational at time $t+T$ if and only if τ_1^i and τ_2^i are both superior to $t+T$. As destruction shocks and idiosyncratic fluctuations are independent, it follows by complementarity that

$$\Pr\left\{P_{t+T}^i \in A \cap \tau_2^i > t + T \cap \tau_1^i > t + T \middle| P_t^i\right\}$$
$$= \left(\Pr\left\{P_{t+T}^i \in A \middle| P_t^i\right\} - \Pr\left\{P_{t+T}^i \in A \cap \tau_2^i \le t + T \middle| P_t^i\right\}\right) * \Pr\left\{\tau_1^i > t + T\right\} \tag{19}$$

where the Borel set $A \subset (R, +\infty)$. These probabilities are more easily computed considering $\ln(P_{t+T}^i)$ since it is a standard Brownian motion. Thus

$$\Pr\left\{\ln(P_{t+T}^i) \in A \middle| P_t^i\right\} = \int_A \frac{e^{-\frac{1}{2}\left(\frac{\ln(x) - \ln\left(P_t^i\right) - \mu T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}} d\ln(x) \tag{20}$$

where $\mu = \zeta - \frac{\sigma^2}{2}$ is the trend of $\ln\left(P_\tau^i\right)$. When μ is equal to zero, the expression of $\Pr\left\{\ln P_{t+T}^i \in A \cap \tau_2^i \le t \middle| P_t^i\right\}$ is easily obtained from the reflection principle. The general expression is derived in Harrison (1985) through a change of measure

$$\Pr\left\{\ln(P_{t+T}^i) \in A \cap \tau_2^i \le t + T \middle| P_t^i\right\} = \int_A \left(\frac{R}{P_t^i}\right)^{\frac{2\mu}{\sigma^2}} \frac{e^{-\frac{1}{2}\left(\frac{\ln(x) + \ln\left(P_t^i\right) - 2\ln(R) - \mu T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}} d\ln(x) \tag{21}$$

Expressing the densities (20) and (21) in terms of x instead of $\ln(x)$, substituting the resulting expressions into (19) and multiplying by $\Pr\left\{\tau_1^i > t + T\right\} = e^{-\delta T}$ implies that

30

for all $x > R$

$$\int_A \psi(x, T; P)\, dx \equiv \Pr\left\{ P_{t+T}^i \in A \cap \tau_2^i > t + T \cap \tau_1^i > t + T \mid P_t^i = P \right\}$$

$$= \int_A e^{-\delta T} \left(\frac{1}{x}\right) \left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(x) - \ln(P) - \mu T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}} - \left(\frac{R}{P}\right)^{\frac{2\mu}{\sigma^2}} \frac{e^{-\frac{1}{2}\left(\frac{\ln(x) + \ln(P) - 2\ln(R) - \mu T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}} \right) dx$$

According to Bayes' rule, the unconditional density is given by

$$\Pr\left\{ P_{t+T}^i \in A \cap \tau_2^i > t + T \cap \tau_1^i > t + T \right\}$$
$$= \Pr\left\{ P_{t+T}^i \in A \cap \tau_2^i > t + T \cap \tau_1^i > t + T \mid P_t^i \in B \right\} * \Pr\left\{ P_t^i \in B \right\}$$

where the Borel set $B \subset (R, +\infty)$. Therefore, under the assumption according to which the initial output P_t^i is drawn from $G(\cdot)$, the unconditional density is equal to

$$\Pr\left\{ P_{t+T}^i \in A \cap \tau_2^i > t + T \cap \tau_1^i > t + T \right\} = \int_A \left(\frac{\int_R^{+\infty} \psi(x, T; P)\, dG(p)}{1 - G(R)} \right) dx$$

Finally the measure $\upsilon(x, T)$ is given by the unconditional density multiplied by the rate of job creation. According to the stationarity assumption, the job creation flow is constant and equal to $u\lambda(1 - G(R))$ which yields the expression in Proposition 2.

- **Proof of Proposition 3:** By definition, the mass of jobs with current output equal to x is given by

$$\upsilon(x) \equiv \int_0^{+\infty} \upsilon(x, T) dT = u\lambda \int_0^{+\infty} \left(\int_R^{+\infty} \psi(x, T; P)\, dG(P) \right) dT$$

Reversing the order of the integrals allows us to find an analytical solution for $\upsilon(x)$. A few algebra yields

$$\psi(x, T; P) = P^{-1} \left(\frac{x}{P}\right)^{\frac{\mu + \gamma}{\sigma^2} - 1} \left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(x) - \ln(P) + \gamma T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}} - \left(\frac{R}{P}\right)^{-\frac{2\gamma}{\sigma^2}} \frac{e^{-\frac{1}{2}\left(\frac{\ln(x) + \ln(P) - 2\ln(R) + \gamma T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}} \right)$$

where $\gamma = \sqrt{\mu^2 + 2\delta\sigma^2}$. Using the result in Leland and Toft (1997) according to which

for positive values of x

$$\int_0^\tau \frac{e^{-\frac{1}{2}\left(\frac{\ln(x)+\gamma T}{\sigma\sqrt{\tau}}\right)^2}}{\sigma\sqrt{2\pi T}} dT = \left(\frac{1}{\gamma}\right)\left(-\Phi\left(\frac{-\ln(x)-\gamma\tau}{\sigma\sqrt{\tau}}\right) + x^{-\frac{2\gamma}{\sigma^2}}\Phi\left(\frac{-\ln(x)+\gamma\tau}{\sigma\sqrt{\tau}}\right)\right)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, we obtain

$$\lim_{\tau\to+\infty}\int_0^\tau \frac{e^{-\frac{1}{2}\left(\frac{\ln(x)+\gamma\tau}{\sigma\sqrt{\tau}}\right)^2}}{\sigma\sqrt{2\pi\tau}} dT = \frac{x^{-\frac{2\gamma}{\sigma^2}}}{\gamma}$$

Using this limit to integrate $\varphi(x;P) \equiv \int_0^{+\infty}\psi(x,T;P)dT$ and ensuring that the integration is always performed over positive values, yields the expression of $\upsilon(x)$ reported in Proposition 3.

- **Proof of Proposition 4:** Given that the size of the labor force has been normalized to one, the rate of employment is equal to the integral of $\upsilon(x)$ from R up to infinity. Thus

$$\begin{aligned}
1-u &= \int_R^{+\infty}\upsilon(x)dx \\
&= u\lambda\int_R^{+\infty}\left(\int_R^{+\infty}\varphi(x;P)dG(P)\right)dx = u\lambda\int_R^{+\infty}\left(\int_R^{+\infty}\varphi(x;P)dx\right)dG(P)
\end{aligned}$$

Integrating $\varphi(x;P)$ with respect to x is straightforward though tedious. It yields

$$\int_R^{+\infty}\varphi(x;P)dx = \left(\frac{1}{\delta}\right)\left(1-\left(\frac{R}{P}\right)^{\frac{\mu+\gamma}{\sigma^2}}\right)$$

The expression of the unemployment rate u is immediately obtained reinserting this solution into the previous equation and simplifying.

- **Proof of Proposition 5:** When $G(\cdot)$ is continuously differentiable, the density functions $f(w,e)$, $f(w,T,e)$ and consequently the likelihood function $L(\Theta)$ are also continuously differentiable. Since

$$\int_{w_r(\Theta)}^{+\infty}f(w,e;\Theta)dw = 1$$

32

where $w_r(\Theta)$ is the reservation wage. Leibnitz's rule implies that

$$
\begin{aligned}
\int_{w_r(\Theta)}^{+\infty} \frac{\partial f(w,e;\Theta)}{\partial \Theta} dw &= \frac{\partial}{\partial \Theta} \int_{w_r(\Theta)}^{+\infty} f(w,e;\Theta) dw + f(w_r(\Theta),e;\Theta) \frac{dw_r(\Theta)}{d\Theta} \\
&= \frac{\partial}{\partial \Theta} \int_{w_r(\Theta)}^{+\infty} f(w,e;\Theta) dw = 0
\end{aligned}
$$

The second equality holds because $\varphi(R;P) = 0$ for all P, so that $f(w_r(\Theta),e;\Theta) = 0$. Similarly, since $\psi(R(\Theta),T;P) = 0$ for all P and T, we have $f(w_r(\Theta),T,e;\Theta) = 0$. Thus

$$
\int_{w_r(\Theta)}^{+\infty} \frac{\partial f(w,e;\Theta)}{\partial \Theta} dw = \int_{w_r(\Theta)}^{+\infty} \frac{\partial f(w,T,e;\Theta)}{\partial \Theta} dw = 0
$$

Therefore the order of integration can be reversed and the central limit theorem yields

$$
\frac{1}{\sqrt{n}} \left(\sum_{i=1}^{n} \frac{\partial \ln f(\Theta, y_i)}{\partial \Theta} \right) \xrightarrow{d} N(0, J)
$$

where J is the information matrix. Since the estimator $\widehat{\Theta}$ is consistent, by the law of large number

$$
-\frac{1}{n} \left(\sum_{i=1}^{n} \frac{\partial^2 \ln f(\Theta, y_i)}{\partial \Theta \partial \Theta'} \right) \xrightarrow{p} H
$$

where H is the Hessian matrix. Notice, that the Hessian matrix is not equivalent to the information matrix as

$$
\begin{aligned}
\int_{w_r(\Theta)}^{+\infty} \frac{\partial^2 f(w,e;\Theta)}{\partial \Theta \partial \Theta'} dw &= \frac{\partial}{\partial \Theta'} \int_{w_r(\Theta)}^{+\infty} \frac{\partial f(w,e;\Theta)}{\partial \Theta} dw + f_\Theta(w_r(\Theta),e;\Theta) \frac{dw_r(\Theta)}{d\Theta} \\
&= f_\Theta(w_r(\Theta),e;\Theta) \frac{dw_r(\Theta)}{d\Theta} \neq 0
\end{aligned}
$$

Given that the likelihood function satisfies all the other regularity conditions, asymptotic efficiency and asymptotic normality of the maximum likelihood estimator are established.

- **Proof of Proposition 6:** The proof of proposition 6 follows by direct calculation. Given that the algebra is tedious, we decompose the solution in several steps. First consider the integral with respect to $\psi(x,T;P)$. Under the parametric assumption that $G(P)$ is

lognormal, it reads

$$\frac{\int_R^{+\infty} x\psi\left(x,T;P\right)dG(P)}{e^{-\delta T}} = \int_R^{+\infty} \left(\frac{x\psi\left(x,T;P\right)}{e^{-\delta T}}\right)\left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right)d\ln(P)$$

$$= \left(\frac{e^{-\frac{1}{2}\left(\frac{B(x,T)\left(\xi^2+\sigma^2T\right)}{\xi^2\sigma^2T}\right)^2}}{\sqrt{2\pi}}\right)\int_R^{+\infty}\left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\left(A(x,T)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)\right)}{\sqrt{(\xi^2+\sigma^2T)\rho^2(T)(1-\rho^2(T))}}\right)^2}}{\xi\sigma\sqrt{2\pi T}}\right)d\ln(P)$$

$$-R^{\frac{2\mu}{\sigma^2}}\left(\frac{e^{-\frac{1}{2}\left(\frac{D(x,T)\left(\xi^2+\sigma^2T\right)}{\xi^2\sigma^2T}\right)^2}}{\sqrt{2\pi}}\right)\int_R^{+\infty}\left(\frac{P^{-\frac{2\mu}{\sigma^2}}e^{-\frac{1}{2}\left(\frac{\ln(P)-\left(C(x,T)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)\right)}{\sqrt{(\xi^2+\sigma^2T)\rho^2(T)(1-\rho^2(T))}}\right)^2}}{\xi\sigma\sqrt{2\pi T}}\right)d\ln(P)$$

$$= \left(\frac{e^{-\frac{1}{2}\left(\frac{B(x,T)\left(\xi^2+\sigma^2T\right)}{\xi^2\sigma^2T}\right)^2}}{\sqrt{2\pi\left(\xi^2+\sigma^2T\right)}}\right)\Phi\left(\frac{-\ln\left(R\right)+A\left(x,T\right)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)}{\sqrt{\left(\xi^2+\sigma^2T\right)\rho^2(T)\left(1-\rho^2(T)\right)}}\right)$$

$$-R^{\frac{2\mu}{\sigma^2}}\left(\frac{e^{-\frac{1}{2}\left(\frac{D(x,T)\left(\xi^2+\sigma^2T\right)}{\xi^2\sigma^2T}\right)^2}}{\sqrt{2\pi\left(\xi^2+\sigma^2T\right)}}\right)e^{\left(\left(-\frac{2\mu}{\sigma^2}\right)^2\left(\frac{\xi^2\sigma^2T}{2\left(\xi^2+\sigma^2T\right)}\right)+\left(-\frac{2\mu}{\sigma^2}\right)\left(C(x,T)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)\right)\right)}$$

$$*\Phi\left(\frac{-\ln\left(R\right)+\left(\frac{-2\mu}{\sigma^2}\right)\left(\frac{\xi^2\sigma^2T}{\xi^2+\sigma^2T}\right)+C\left(x,T\right)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)}{\sqrt{\left(\xi^2+\sigma^2T\right)\rho^2(T)\left(1-\rho^2(T)\right)}}\right) \tag{22}$$

where

$$\rho^2(T) = \frac{\xi^2}{\xi^2+\sigma^2T}$$

$$A\left(x,T\right) = -\mu T + \ln(x) \qquad B\left(x,T\right) = \frac{-\left(A(x,T)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)\right)^2}{+A(x.T)^2\rho^2(T)(1-\rho^2(T))}$$

$$C(x,T) = 2\ln(R) + \mu T - \ln(x) \quad D(x,T) = \frac{-\left(C(x,T)\rho^2(T)+\Sigma\left(1-\rho^2(T)\right)\right)^2}{+C(x,T)^2\rho^2(T)(1-\rho^2(T))}$$

Now consider the integral with respect to $\varphi(x; P)$

$$\int_R^{+\infty} \varphi(x; P)dG(P) = \int_R^x P^{-1}\left(\frac{x}{P}\right)^{\frac{\mu-\gamma}{\sigma^2}-1}\left(\frac{1-\left(\frac{R}{P}\right)^{\frac{2\gamma}{\sigma^2}}}{\gamma}\right)\left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right)d\ln(P)$$

$$+\int_x^{+\infty} P^{-1}\left(\frac{x}{P}\right)^{\frac{\mu+\gamma}{\sigma^2}-1}\left(\frac{1-\left(\frac{R}{x}\right)^{\frac{2\gamma}{\sigma^2}}}{\gamma}\right)\left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right)d\ln(P)$$

$$=\left(\frac{x^{\frac{\mu-\gamma}{\sigma^2}-1}}{\gamma}\right)e^{\left(\left(\frac{\gamma-\mu}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right)+\left(\frac{\gamma-\mu}{\sigma^2}\right)\Sigma\right)}\left(\begin{array}{c}\Phi\left(\frac{\ln(x)-\left(\left(\frac{\gamma-\mu}{\sigma^2}\right)\xi^2+\Sigma\right)}{\xi}\right)\\-\Phi\left(\frac{\ln(R)-\left(\left(\frac{\gamma-\mu}{\sigma^2}\right)\xi^2+\Sigma\right)}{\xi}\right)\end{array}\right)$$

$$+\left(\frac{x^{\frac{\mu+\gamma}{\sigma^2}-1}}{\gamma}\right)e^{\left(\left(\frac{-\gamma-\mu}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right)+\left(\frac{-\gamma-\mu}{\sigma^2}\right)\Sigma\right)}\Phi\left(\frac{-\ln(x)+\left(\frac{-\gamma-\mu}{\sigma^2}\right)\xi^2+\Sigma}{\xi}\right)$$

$$-R^{\frac{2\gamma}{\sigma^2}}\left(\frac{x^{\frac{\mu-\gamma}{\sigma^2}-1}}{\gamma}\right)e^{\left(\left(\frac{-\gamma-\mu}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right)+\left(\frac{-\gamma-\mu}{\sigma^2}\right)\Sigma\right)}\Phi\left(\frac{-\ln(R)+\left(\left(\frac{-\gamma-\mu}{\sigma^2}\right)\xi^2+\Sigma\right)}{\xi}\right) \quad (23)$$

Finally consider

$$\int_R^{+\infty}\left(1-\left(\frac{R}{P}\right)^{\frac{\mu+\gamma}{\sigma^2}}\right)dG(P) = \int_R^{+\infty}\left(1-\left(\frac{R}{P}\right)^{\frac{\mu+\gamma}{\sigma^2}}\right)\left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right)d\ln(P)$$

$$=\overline{G}(R) - R^{\frac{\mu+\gamma}{\sigma^2}}e^{\left(\left(\frac{-\mu-\gamma}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right)+\left(\frac{-\mu-\gamma}{\sigma^2}\right)\Sigma\right)}\Phi\left(\frac{-\ln(R)+\left(\frac{-\mu-\gamma}{\sigma^2}\right)\xi^2+\Sigma}{\xi}\right) \quad (24)$$

The closed-from expression of the likelihood function is obtained inserting (22), (23) and (24) into (12). Notice that the unemployment rate also has an analytical solution

$$u = \frac{\delta}{\delta + \lambda\left(\overline{G}(R) - R^{\frac{\mu+\gamma}{\sigma^2}}e^{\left(\left(\frac{-\mu-\gamma}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right)+\left(\frac{-\mu-\gamma}{\sigma^2}\right)\Sigma\right)}\Phi\left(\frac{-\ln(R)+\left(\frac{-\mu-\gamma}{\sigma^2}\right)\xi^2+\Sigma}{\xi}\right)\right)}$$

References

[1] Altonji, Joseph G., and Robert Shakotko. 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies* 54 (3): 437-59.

[2] Altonji, Joseph G., and Nicholas Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *Industrial and Labor Relations Review* 58 (3): 370-397.

[3] Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage Bargaining with On-the-job Search: Theory and Evidence." *Econometrica* 74 (2): 323-365.

[4] Dey, Matthew, and Christopher Flinn. 2005. "An Equilibrium Model of Health Insurance Provision and Wage Determination." *Econometrica* 73 (March): 571-627.

[5] Dustmann, Christian, and Costas Meghir. 2005. "Wages, Experience and Seniority." *Review of Economic Studies* 72 (1): 77-108.

[6] Eckstein, Zvi, and Kenneth I. Wolpin. 1995. "Duration to first job and return to schooling: estimates from a search-matching model." *Review of Economic Studies* 62 (2): 263-286.

[7] Felli, Leonardo, and Christopher Harris. 1996. "Learning, Wage Dynamics, and Firm-Specific Human Capital." *Journal of Political Economy* 104 (4): 838-868.

[8] Flinn, Christopher, and James Heckman. 1982. "New Methods for Analyzing Structural Models of Labor Force Dynamics." *Journal of Econometrics* 18: 115-168.

[9] Flinn, Christopher. 2006. "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates." forthcoming *Econometrica*.

[10] Greene, William H. 1980. "Maximum Likelihood Estimation of Econometric Frontier Functions." *Journal of Econometrics* 13: 27-56.

[11] Harrison, Michael J. 1985. *Brownian Motion and Stochastic Flow Systems.* Malabar, Florida: E. Krieger publishing company.

[12] Hornstein, Andreas, Krusell, Pers, and Gianluca Violante. 2006. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." Unpublished Manuscript, New-York University.

[13] Jovanovic, Boyan, and Jacob Mincer. 1981. "Labor Mobility and Wages." In *Studies in Labor Markets*, edited by Sherwin Rosen, New York: National Bureau of Economic Research.

[14] Jolivet, Grégory, Fabien Postel-Vinay, and Jean-Marc Robin. 2005. "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US." Forthcoming *European Economic Review*.

[15] Leland, Hayne E., and Klaus B. Toft. 1997. "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads." *The Journal of Finance* 51 (3): 987-1019.

[16] Lopes de Melo, Rafael. 2006. "Wage Dynamics and Distribution in an Equilibrium Search Model." Unpublished Manuscript, Yale University.

[17] Merton, Robert C. 1973. "The Theory of Rational Option Pricing." *The Bell Journal of Economics and Management Science* 4 (1): 141-184.

[18] Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3): 397-415.

[19] Nagypál, Éva. 2005. "On the Extent of Job-to-Job Transitions." Unpublished Manuscript, Northwestern University.

[20] Newey, Whitney K., and Daniel McFadden. 1994. "Large Sample Estimation and Hypothesis Testing." In *Handbook of Econometrics*, Amsterdam: North-Holland, pp. 2111-245.

[21] Pissarides, Christopher. 2000. *Equilibrium Unemployment Theory.* 2nd. Ed., Cambridge Mass.: MIT Press.

[22] Postel-Vinay, Fabien, and Hélène Thuron. 2005. "On-the-job Search, Productivity Shocks,and the Individual Earnings Process." Unpublished Manuscript, University of Bristol.

[23] Prat, Julien. 2006. "Job Separation under Uncertainty and the Wage Distribution." *Contributions to Macroeconomics*, Berkeley Electronic Press, 6 (1), Article 2.

[24] Ridder, Geert, and Gerard J. van den Berg. 2003. "A Cross-Country Comparison of Labor Market Frictions." *Journal of the European Economic Association* 1 (1): 224-244.

[25] Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy* 99 (1): 145-176.

[26] Winter-Ebmer, Rudolf and Josef Zweimüller. 1996. "Manpower Training and Employment Stability." *Economica* 63: 113-130.

[27] Yamaguchi, Shintaro. 2006. "Job Search, Bargaining, and Wage Dynamics." Unpublished Manuscript, University of Wisconsin, Madison.