

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

Julien Prat*

February 2006.

Abstract

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 evolution through tenure of the cross-sectional distribution of wages. We find that the returns to tenure are slightly higher for workers without tertiary education than for tertiary educated workers.

THEME : Macroeconomics of unemployment and inequality.

KEYWORDS : Job Search, Uncertainty, Structural Estimation.

JEL-CODE : J31, J64.

*I am especially grateful to Giuseppe Bertola and Christopher Flinn for helpful comments. Department of Economics, Vienna University. E-mail : julien.prat@univie.ac.at

1 Introduction

The attachment of workers to their jobs is often supported by active labor market programmes such as wages or employment subsidies. Determining their optimal design crucially depends on the rate at which workers learn as a by-product of work.¹ This is why a large body of empirical research has focused on estimating the rate of learning-by-doing.² Yet, due to the lack of structural approaches, the task of quantifying the aggregate benefits of these active labor market policies remains elusive. This paper attempts to contribute to such an objective by estimating the returns to tenure in an equilibrium set-up.

The proposed framework is an “estimable” version of the Mortensen and Pissarides (1994) framework. To our knowledge, this paper is the first to estimate by maximum likelihood the Mortensen-Pissarides model, despite the fact that it has become the standard framework of analysis for aggregate labor markets. In order to render the model suitable for empirical purposes, we modify the stochastic process that changes jobs’ output: instead of considering Poisson processes, we introduce geometric Brownian motions. Thus the job’s idiosyncratic productivity follows a random walk with constant mean growth rate. Obviously, the deterministic trend captures the average rate at which workers accumulate human capital through learning-by-doing. To preserve the tractable 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.³

Notwithstanding these simplifying assumptions, the empirical analysis raises several challenges. First of all, the rate of learning-by-doing cannot be estimated in a deter-

¹See for example Cossa, Heckman and Lochner (2003) for an analysis of the diverging evaluations of the Earned Income Tax Credit depending on the chosen model for skill-accumulation (on-the-job training versus learning-by-doing).

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

³The Nash-bargaining solution does not take into account 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 assuming that each party receives a fixed share of the expected surplus at any point in time.

ministic set-up. When workers deterministically progress along the learning curve, the likelihood of observing an individual close to the reservation wage converges to zero as tenure accumulates. Given the presence of such observations in the data, the introduction of uncertainty is a prerequisite to the estimation of the learning-by-doing rate. But idiosyncratic shocks lead to endogenous separations. This feature distinguishes our approach from previous structural estimations that were based on the premise that job destruction is induced by an exogenous process.⁴ Endogenous separations greatly complicate the derivations since we have to deduce all the sample paths that breach the reservation threshold. Nevertheless, we show that the convenient analytical properties of Brownian motions allows us to address this potentially daunting problem and consequently to express the joint distribution of wages and job spells in closed-form.

The second main difficulty is due to the non-standard property of the likelihood function. As the reservation wage and consequently the support of the data is a function of the estimated parameters, standard regularity conditions are not satisfied. This peculiarity is well known 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 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 other parameters. This estimation procedure yields consistent estimates for deterministic search-matching 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 the lowest reported wage is not anymore a super-consistent estimator.

Thus 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. Intuitively, firms cannot perfectly counteract random perturbations. Hence, they remain within the neighborhood of the optimal combination of input without ever achieving it

⁴See Eckstein and Wolpin (1995) for an early estimation of an equilibrium search model in a deterministic environment and Flinn (2006) for an analysis of the minimum wage effects.

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 Green's proof to our set-up and establish that, despite the appearance, endogenous separation actually simplifies the analysis since it allows 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. Individuals are placed in two different sub-samples according to their educational attainments. The estimation procedure returns estimates for the rate of learning-by-doing of 2.3% per year for workers without tertiary education and 2% for tertiary educated workers. The two estimated volatility parameters are almost identical with a very low value of around 7%. We assess the ability of the model to fit the cross-sectional joint distributions of wages and jobs spells and find that it fits the data surprisingly well given its parsimonious specification.

The rest of the paper is organized as follows. The proofs of the propositions are collected in the Appendix. 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. Section 5 concludes.

2 The model

We consider a labor market with search frictions where jobs' output are subject to random fluctuations. Our set-up differs in three respects from the one proposed by Mortensen and Pissarides (1994) : firstly we allow initial outputs to differ, secondly we assume that productivity follows a geometric Brownian motion and finally we introduce learning-by-doing. Given that the Current Population Survey (henceforth CPS) does not contain informations on the number of posted vacancies, the data will not allow us to estimate

the parameters of the matching function. Thus we take the rate of contact between searchers and firms as given. Notice that it is quite straightforward to close the model in order to endogenize job creation by assuming calibrated values for the aggregate matching function.

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 ergodic distribution $G(\cdot)$, which is assumed to be continuously differentiable.

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 and instead to keep looking for a better opportunity.

In the case where both parties decide to match, they immediately start to produce and output begins to fluctuate. We abstract from aggregate shocks so that stochastic shocks 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

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

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 increases. It can also be referred to as the rate of job-specific human capital accumulation given that workers become identical after joining the unemployment pool. The parameter σ reflects dispersion: the higher it is, the faster output fluctuates.

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 rate of job turnover 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 purpose of this paper. We 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 .

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. Actually, it is shown in Nagypál (2005) that the wage distribution cannot be expressed in closed-form when workers search on the job and uncertainty is modelled using a diffusion process. These difficulties partly explain why empirical models of employers competition typically assume away idiosyncratic uncertainty.⁵ As we want to investigate the explanatory power of idiosyncratic uncertainty, we make the converse simplification of neglecting on-the-job search 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. 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

⁵A 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.

cannot be the case that one is positive and the other negative. Hence it is clear that 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 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

$$rU = -s + \lambda \int_R^{+\infty} \beta S(P) dG(P) \quad (2)$$

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

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

where it is assumed that firms discount the future at the same rate than workers. The additional term $\delta S(P_t)$ 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 $S(P_t)$ by combining equations (2) and (3). In the stochastic case, we have to solve the partial differential equation satisfied by $S(P_t)$ as explained in the Appendix.

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

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

where α is the negative root of the following quadratic equation

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

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_t , so that R is identical across matches, as one should expect. Its optimal value is given by

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

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 again turn profitable 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, there is no labor-hoarding when productivity is constant and the reservation output is equal to the opportunity cost of employment.

2.3 *The equilibrium*

This section characterizes the equilibrium rate of unemployment and the cross-sectional distributions of jobs spells and wages. It will be shown in section 3 that these statistics have closed-form solutions when the sampling distribution $G(\cdot)$ 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

$$w(P_t) = \beta P_t + (1 - \beta)rU \quad (6)$$

The wage follows from output by a location transformation. So we can restrict our discussion without loss of generality to the cross-sectional distribution of output. 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.⁶ The statistics of interest are derived using a progressive approach: starting from the most informative one, namely the joint distribution of jobs spells and output, we aggregate it step by step in order to obtain the rate of unemployment.

⁶Although conventional for obvious technical reasons, the steady-state assumption is actually quite restrictive. We refer to Jolivet et al. (2004) for empirical evidence in its favor.

Proposition 2 *The measure of ongoing job relationships with current output x and tenure T is given by*

$$v(x, T) = u\lambda \left(\int_R^{+\infty} h(x, T; P) dG(P) \right) \quad (7)$$

where u denotes the rate of unemployment. The analytical expression for $h(x, T; P)$ reads

$$h(x, T; P) = \begin{cases} e^{-\delta T} \left(\frac{e^{-\frac{1}{2} \left(\frac{\ln(x) - \ln(P) - \mu T}{\sigma\sqrt{T}} \right)^2}}{\sigma\sqrt{2\pi T}} - \left(\frac{R}{P} \right)^{\frac{2\mu}{\sigma^2}} \frac{e^{-\frac{1}{2} \left(\frac{\ln(x) + \ln(P) - 2\ln(R) - \mu T}{\sigma\sqrt{T}} \right)^2}}{\sigma\sqrt{2\pi T}} \right); & \text{if } x > R \\ 0; & \text{otherwise.} \end{cases} \quad (8)$$

where $\mu = \zeta - \frac{\sigma^2}{2}$ is the trend of $\ln(P_t)$ and $d = \sqrt{\mu^2 + 2\delta\sigma^2}$.

The first term on the right side of (7) measures the number of contacts between searchers and firms. The function $h(x, T; P)$ 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 .⁷ The unconditional joint density is obtained integrating $h(x, T; P)$ with respect to $G(P)$, since entrants draw their starting output from $G(\cdot)$. The sampling distribution is integrated from R up to infinity because contacts lead to matches solely when P is above R .

The mass 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 measure of ongoing job relationships with current output x is given by*

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

⁷Notice that an econometrician who observes the wages of workers at different points in time of their jobs spells could use equation (8) to compute the likelihood of the sample paths.

where the function $\varphi(x; P)$ reads

$$\varphi(x; P) = \begin{cases} P \left(\frac{x}{P}\right)^{\frac{\mu-d}{\sigma^2}-1} \left(\frac{1-\left(\frac{R}{P}\right)^{\frac{2d}{\sigma^2}}}{d}\right); & \text{if } x \geq P \\ P \left(\frac{x}{P}\right)^{\frac{\mu+d}{\sigma^2}-1} \left(\frac{1-\left(\frac{R}{x}\right)^{\frac{2d}{\sigma^2}}}{d}\right); & \text{if } x \in [R, P] \\ 0; & \text{otherwise} \end{cases} \quad (10)$$

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 $(R/P)^{\frac{\mu+d}{\sigma^2}}$ becomes negligible and the expression of u converges to the standard one.

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

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

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 difficulty has been recognized since a long time and is now gradually overcome by research based on matched employer-employees data (see Cahuc et al., 2005; Dey and Flinn, 2005). Given the one-sided nature of the CPS data, we stick to the usual practice of assuming symmetric bargaining.

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 $Y = \{y_1, y_2, \dots, y_n\}$ 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 report the current hourly wage and job tenure respectively. 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

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

where $\tilde{G}(R) = 1 - G(R)$. The likelihood of observing an employed worker paid wage w is given by $v(x, T)$. The density can be further decomposed reinserting (11) into (9) to

obtain

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

Notice that output is defined as a function of the observed wage. Its implicit value follows from combining (5) with (6)

$$x(w) = \frac{w - (1 - \beta) \left(\frac{\alpha-1}{\alpha} \right) \left(\frac{r+\delta}{r+\delta-\zeta} \right) R}{\beta}$$

Similarly, the joint likelihood of observing a worker paid wage w with a job tenure equal to T is given by

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

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

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

where n_U is the number and U 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. Despite being absent from the analytical expression of the likelihood function, the parameters ζ , σ and Ω are implicitly identified. Whereas ζ and σ determine the values of μ , d , $\varphi(\cdot)$ and $h(\cdot)$, the parametric vector Ω obviously influences $G(\cdot)$. Notice that is treated for the sake of the estimation as a primitive parameter of the model.

The likelihood function is continuously differentiable and its parameters belong to a

compact support. It does not, however, 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. Yet next proposition shows that, under very mild requirements, the estimators satisfy asymptotic normality.

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*

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

converges in probability to Θ_0 so that $\sqrt{n}(\hat{\Theta} - \Theta_0) \xrightarrow{d} N(0, H^{-1}JH^{-1})$ 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 and consequently allows us to characterize the asymptotic property of the estimator by linear approximation. Our problem is slightly less standard than the one considered by Greene because the derivative 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 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. This result has direct relevance for the model studied in this paper but is also of more general interest since it is likely to apply to a wide class of models where exit is endogenized.

3.2 Log-normal sampling distribution

We have characterized the estimation procedure for general sampling distributions. The econometric implementation of the model requires to narrow our analysis to a particular family of distributions. Accordingly, we will hereafter assume that $G(\cdot)$ belongs to the lognormal family. 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” is larger than the lognormal family as it encompasses among others gamma and exponential distributions.⁸ Thus lognormality is eventually justified by its good fit of the data. In our case, the lognormal assumption has a more crucial role. Given the intricate expression of the likelihood function, there is little hope to derive it in closed-form. Yet, when initial outputs are lognormally distributed in the population, the likelihood function has an analytical expression so that approximation errors due to numerical integrations can be avoided.

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

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

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

Given its length, we do not include the expression of $L(\Theta)$ in the main body of the paper. As intricate as it is, it may come as a surprise to the reader that a closed-form expression can be derived. But this result is better understood recalling that geometric Brownian motions are also lognormal processes.

⁸See Flinn (2006) for a careful discussion of the class of functions which satisfy the “recoverability condition”.

4 Empirical results

4.1 *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, the current wages and job spells of the Outgoing Rotation Groups⁹ 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 were 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 split the sample into two sub-samples: one containing workers with a high school graduation diploma or less and another one containing workers with tertiary education. So we consider that the sample is actually composed of two distinct labor markets depending on the education level. While restrictive, this hypothesis is not uncommon in the literature. Its main interest in our case is that it will enable us to estimate two distinct reservation wages. Finally we have trimmed each sub-sample by excluding the top and bottom percentile of the wage distribution. The trimming of the bottom percentile is particularly important for the estimation of the reservation wage since it allows us to avoid implausibly low estimates due to measurement errors.¹⁰

⁹In our case, the Outgoing Rotation Groups are composed of the households that entered the panel in October 2002 and 2003.

¹⁰The trimming of the data is particularly 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 the usual hours of

TABLE 1
DESCRIPTIVE STATISTICS

| | No Tertiary Education | | Tertiary Education | |
|-------------------------|-----------------------|-------------------------------|--------------------|-------------------------------|
| | <i>Mean</i> | <i>Standard Dev. of Mean</i> | <i>Mean</i> | <i>Standard Dev. of Mean</i> |
| Age | 40.4 | .165 | 40.9 | .136 |
| Female | .417 | .007 | .477 | .006 |
| Working week (hours) | 41.4 | .087 | 42.8 | .08 |
| Average spells (months) | | | | |
| Unemployed | 4.5 | .246 | 4.44 | .28 |
| Job | 92.5 | 1.6 | 94.8 | 1.27 |
| Hourly wage | | | | |
| All jobs | 13.7 | .10 | 20.8 | .13 |
| Job spell<1 year | 11.8 | .27 | 17.8 | .36 |
| Entrants | 10.3 | .36 | 16.8 | .73 |
| | | <i>Number of Observations</i> | | <i>Number of Observations</i> |
| Labor Market Position | | | | |
| Unemployed | | 413 | | 316 |
| Employed | | 4336 | | 6534 |
| Reported job spells | | | | |
| Total number | | 3676 | | 5776 |
| Job spell<1 year | | 535 | | 787 |
| Entrants | | 154 | | 195 |
| Sample size | | 4749 | | 6850 |

Descriptive statistics for the two sub-samples are reported in Table 1. Age and Sex are quite similar. Whereas working hours do not differ much, the other work related data exhibit significant discrepancies. As one might expect, employees with a higher education earn on average a higher wage. More interestingly, even though the rate of unemployment is two times higher for workers without tertiary education, the average unemployment durations are almost identical. This points to differences in the matching work per week in order to impute their hourly wage. Obviously this computation interacts potential measurement errors and for some observations lead to extremely low wages.

process between the two groups of workers and provides support to the premise that we are actually considering distinct labor markets.

Table 1 also reports descriptive 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 actually approximates the latter using data with short jobs spells. This points to potential bias as can be seen from the statistics of the wage distribution for entrants.¹¹ The wage distribution of jobs with less than one year of tenure has a higher mean than the entrants distribution. This feature of the data is easily explained by on-the-job search as workers that quit their job select offers which are above their current wage. Since our model excludes on-the-job search, the job-ladder effect is ignored and consequently our estimates of the sampling distribution are biased upwards. This in turn will lead to downward bias for the estimates of the rate of learning-by-doing.

However figures 1 and 2 suggests that these biases are not likely to be severe. Both figures report the non-parametric kernel density estimates of the three wage distributions. Although the entrants distribution differ from the distribution for workers with less than one year of tenure, the discrepancies appear to be relatively mild. To the contrary, the aggregate wage distributions exhibit a much higher dispersion. This feature fits well the model since Brownian motions are diffusion processes, meaning that their distribution becomes more and more dispersed as time elapses.

Turning our attention to the divergences between the sub-samples, we observe that they display the same qualitative features. Yet, we notice two obvious differences. Firstly, the densities are more dispersed for tertiary educated workers. Secondly, their aggregate wage distribution appears to be more closely related to the distribution among job entrants.

4.2 *Estimates*

Table 2 contains the estimated parameters for yearly periods along with their standard deviations. 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

¹¹We are able to identify individuals entering the employment pool by excluding those that were working a year ago.

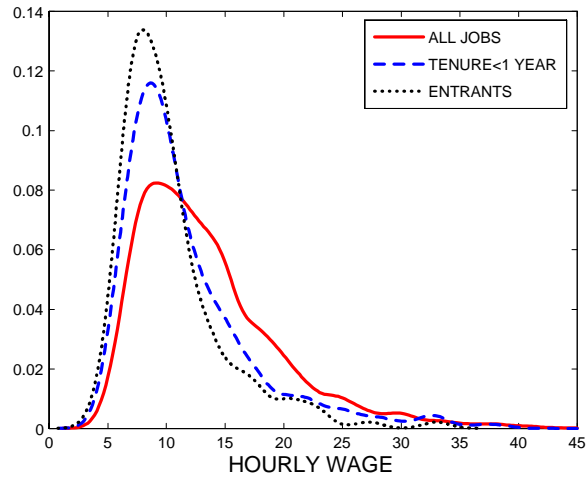


Figure 1: Wage densities for workers without tertiary education.

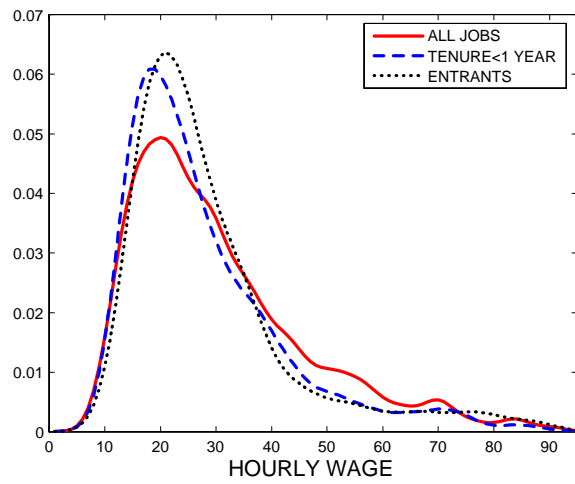


Figure 2: Wage densities for workers with tertiary education.

one proposed in this paper. Notice that in addition to the learning-by-doing rate ζ and variance parameter σ , the model also allows us to estimate the expected value and standard deviation of the reservation wage w_r . On the contrary, the values reported for the deterministic model correspond to the lowest wage in each sub-sample so that their standard deviations are not well 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 expected because 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 for jobs with a tenure close to zero. As shown in Figures 1 and 2, the aggregate distributions are located to the right and exhibit a higher dispersion than the entrants distributions, which is reflected by the different estimates of Σ and ξ .

The exogenous rates of job destruction δ are quite similar in both models. 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. However, for the very small estimates of the variance parameters, endogenous separation is a marginal phenomenon. Accordingly the separation rates are quit similar in both models and imply that the average job spell is close to 92 weeks for workers with tertiary education and to 94 weeks for workers without tertiary education. Thus the estimated average lengths of a job replicate the values of the descriptive statistics reported in Table 1.

The estimated values of λ imply that searchers with and without tertiary education receive job offers every 6 and 4 months respectively. Although the models over-estimate the average unemployment duration for workers without tertiary education, the estimates of their unemployment rate is biased downward to 6.4% while its value in the data is 8.7%. These opposite biases highlight the tension contained in the data between the informations on unemployment duration and aggregate unemployment rate. They point towards violation of the stationarity assumption that is most probably explained by a significant and recent entry into the labor force of workers without tertiary education.

TABLE 2
ESTIMATES

| Parameters | No Tertiary Education | | Tertiary Education | |
|------------------------|-----------------------|----------------|--------------------|-----------------|
| | Stochastic | Deterministic | Stochastic | Deterministic |
| w_r | 3.92 (.182) | 4.11 (___) | 4.84 (.084) | 5.15 (___) |
| ζ | .023 (.0038) | --- | .020 (.0014) | --- |
| σ | .0685 (.0047) | --- | .0705 (.0020) | --- |
| δ | .131 (.0049) | .133 (.071) | .125 (.0292) | .126 (.0811) |
| Σ | 2.84 (.083) | 3.01 (.008) | 3.27 (.044) | 3.41 (.0076) |
| ξ | .464 (.0079) | .526 (.011) | .570 (.004) | .609 (.009) |
| λ | 1.93 (.0062) | 1.94 (.057) | 2.69 (.0315) | 2.66 (.0567) |
| $\ln L$ | -28885 | -29045 | -47192 | -47329 |
| Calibrated Parameters: | $r = 0.05$ | $\beta = 0.5$ | | |

We now turn our attention to the estimates which are specific to the stochastic model. First of all, we notice that the variance parameters σ are quite low. With a standard deviation close to 7%, the stochastic model predicts that sample paths are almost deterministic. Nonetheless, one cannot constrain σ to be equal to zero and at the same time estimate ζ . The reason is that the lower-bound of the joint distribution is an increasing function of tenure in a purely deterministic environment. As a result, the likelihood of observing a job with a tenure equal to T and a wage inferior to $e^{\zeta T} w_r$ is zero. Given that both samples contain such observations, setting σ equal to zero yields a sample likelihood of minus infinity as long as ζ significantly differs from zero. In other words, 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 learning-by-doing rate, the former being necessary to implement the latter.

The rate of learning-by-doing ζ is almost 10% lower for tertiary educated workers. To interpret this finding it is useful to recall that in our model learning is purely match-

specific. Our estimates do not capture the accumulation of general human capital which is known to be higher for skilled workers. Given this limitation, our result concurs with the findings in Dustmann and Meghir (2005) according to which skilled workers have lower returns to tenure. Quantitatively speaking, the model predicts that 10 years of tenure result in a wage increase of 18.8% for low-skilled worker and 16.9% for high-skilled workers. Accordingly our estimates of the cumulative returns to tenure are in the middle of the range separating the low returns obtained by Altonji and Williams (2005) and of the high returns obtained by Topel (1991).

We now turn to the ability of the model to fit the sample information. Of most interest to our analysis are its predictions about the joint distributions of wages and tenure. The data for jobs with a tenure below 1, 5, 10 and 20 years are reported against their simulated counterparts in Figures 3 and 4. Both 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. We also notice that the model matches very well the right tails of the distributions. This is a classical test for models of wage dispersion due to the “heavy tail” property of the data. As explained in Prat (2006), endogenous diffusion of job output implies that 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 Figures 4 and 5 the actual distributions of jobs spells together with their structural estimations. While not as convincing than for wages, the model predictions are still reasonably close to the data. The model systematically under-predicts the rate of job destruction in the first year. As discussed in Ridder and van der Berg (2003), the difficulty to explain the occurrence of many short job spells is also inherent to models of on-the-job-search. Nonetheless, 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 hazard rate of job separation. Unfortunately this would require us to set the idiosyncratic variance σ at a higher value than the one resulting from the estimation.

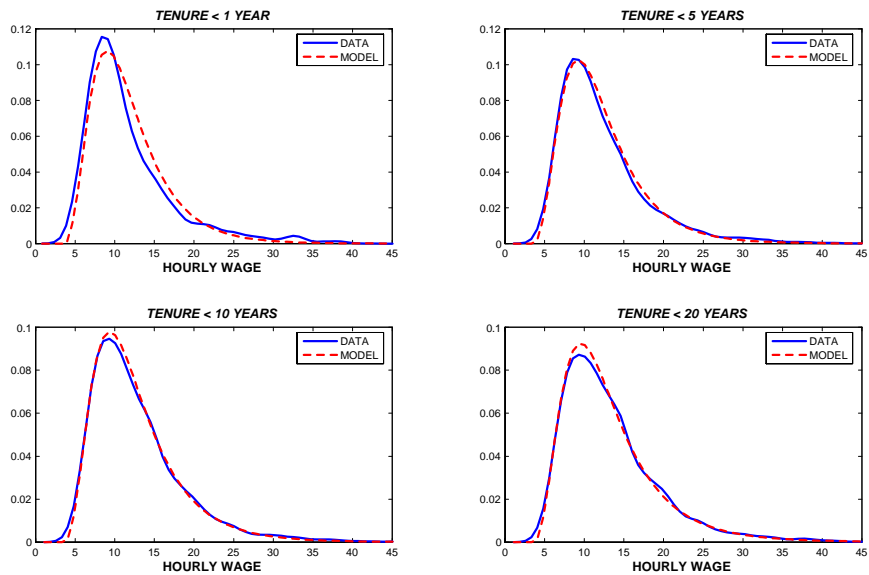


Figure 3: Wage densities at different jobs spells for workers without tertiary education.

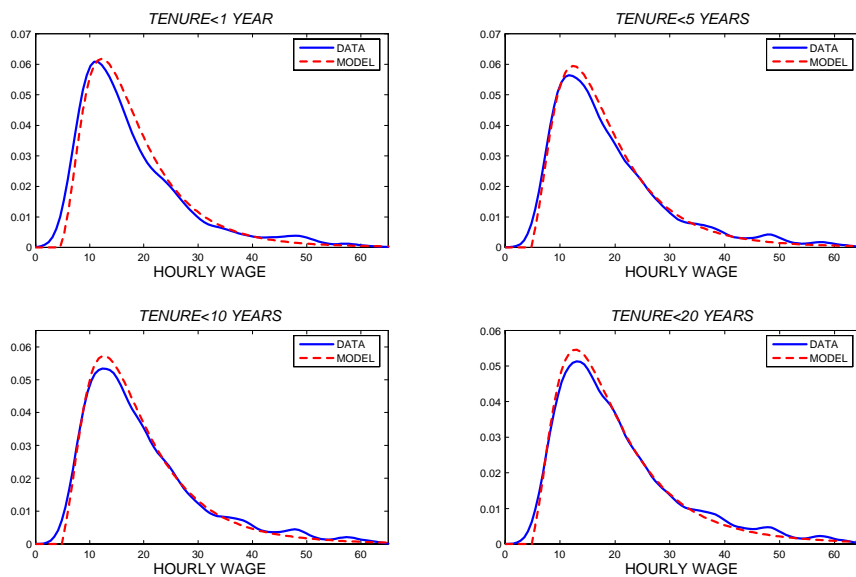


Figure 4: Wage densities at different jobs spells for workers with tertiary education.

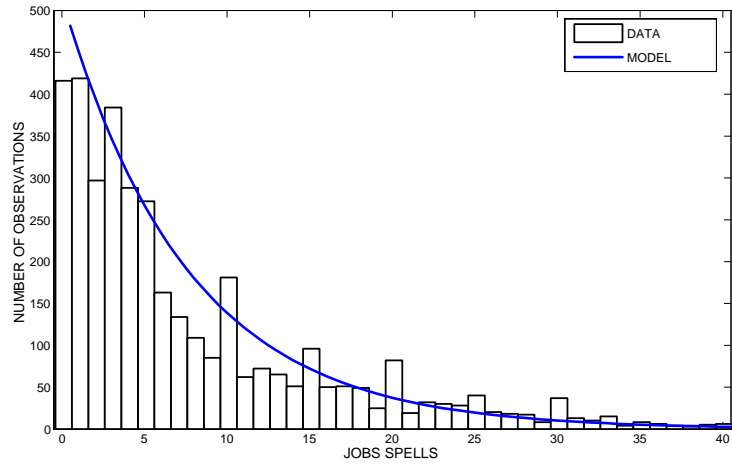


Figure 5: Jobs spells distribution for workers with no tertiary education.

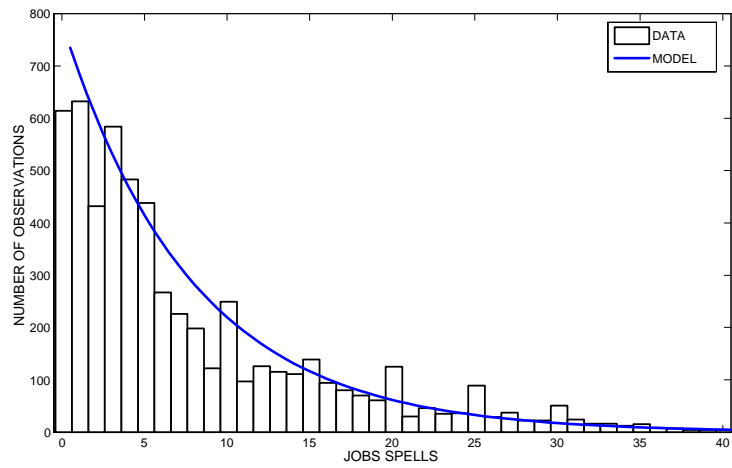


Figure 6: Jobs spells distribution for workers with tertiary education.

5 Conclusion

It has been shown in this paper how the production side of the Mortensen and Pissarides (1994) framework can be estimated by maximum likelihood using cross-sectional data. The analysis establishes that a parsimoniously specified model can convincingly fit the cross-sectional features of the data once learning-by-doing is taken into account. We find that workers without tertiary educations have slightly higher returns to tenure than tertiary educated workers. A concrete contribution of the analysis is to identify the rate of learning-by-doing in an equilibrium set-up, whereas the estimates previously available in the literature are typically based on “reduced-form” estimations. The model could easily be closed by introducing an aggregate matching function and so used to evaluate the impact of learning-by-doing on various labor market policies. Applying the model to cross-sectional data would also be straightforward and certainly helpful to obtain more precise informations on the variance parameter.

As we have deliberately tipped the balance in favor of tractability over realism, the model lends itself naturally to several 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, we suspect that 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 our 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 (Nagypál, 2006; Postel-Vinay and Thuron, 2005; Yamagushi, 2006), the interest of combining both dimensions is now widely recognized. Such a research project raises serious technical challenges and for the moment available estimates are based on indirect inference methods. This paper shows that stochastic calculus might 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 use to researchers interested in other areas than labor economics since our result should apply to a wide class of models with endogenous exit.

APPENDIX

- **Proof of Proposition 1:** We guess that R does not depend on current output and so is the same across jobs. 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

$$S(P_t^i, R) = C(R) \left(\frac{P_t^i}{R} \right)^\alpha + D(R) \left(\frac{P_t^i}{R} \right)^\eta + E_{P_t^i} \left[\int_t^{+\infty} e^{-(r+\delta)(\tau-t)} (P_\tau^i - rU) d\tau \right] \quad (14)$$

where C and D 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 characteristic equation

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

We impose the following boundary conditions on the solution of (14)

$$\begin{cases} \lim_{P_t^i \rightarrow +\infty} S(P_t^i, R) = \frac{P_t^i}{r+\delta-\zeta} - \left(\frac{1}{r+\delta} \right) rU \\ \lim_{P_t^i \rightarrow R_t} S(P_t^i, R) = 0 \end{cases}$$

The first boundary condition implies that we can set $D(R)$ equals to zero in order to eliminate the positive root η . The solution proposed in (4) satisfies both 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 $\frac{\partial S(\cdot, R)}{\partial R} = 0$ when¹²

$$\left. \frac{\partial S(P_t^i, R_t)}{\partial P_t^i} \right|_{P_t^i=R_t} = 0 \quad (15)$$

¹²See Merton(1973), p.171.

It is commonly referred to equation (15) as the smooth-pasting condition. Its solution reads

$$R = \left(\frac{\alpha}{\alpha - 1} \right) \left(\frac{r + \delta - \zeta}{r + \delta} \right) rU \quad (16)$$

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

- **Proof of Proposition 2:** 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

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

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 superior to $t+T$. As death shocks and idiosyncratic fluctuations in output are independent, it follows by complementarity that

$$\begin{aligned} & \Pr \{ P_{t+T}^i \in A \cap \tau_2^i > t \cap \tau_1^i > t \mid P_t^i \} \\ &= (\Pr \{ P_{t+T}^i \in A \mid P_t^i \} - \Pr \{ P_{t+T}^i \in A \cap \tau_2^i \leq t+T \mid P_t^i \}) * \Pr \{ \tau_1^i > t+T \} \end{aligned} \quad (17)$$

where the Borel set $A \subset (R, +\infty)$. These probabilities are more easily computed considering $\ln(P_{t+T}^i/P_t^i)$ since it is a standard Brownian motion starting at 0. Thus the expression of $\Pr \{ P_{t+T}^i \in A \mid P_t^i \}$ is

$$\Pr \{ P_{t+T}^i \in A \mid P_t^i \} = \int_A \frac{e^{-\frac{1}{2} \left(\frac{\ln(x) - \ln(P_t^i) - \mu T}{\sigma \sqrt{T}} \right)^2}}{\sigma \sqrt{2\pi T}} dx \quad (18)$$

where $\mu = \zeta - \frac{\sigma^2}{2}$ is the trend of $\ln(P_t^i)$. The expression of $\Pr \{ P_t^i \in A \cap \tau_2^i \leq t \mid P_{t-T}^i \}$ is easily obtained from the reflection principle when the drift μ is equal to zero. The general expression is derived in Harrison (1985) through a change of measure

$$\Pr \{ P_{t+T}^i \in A \cap \tau_2^i \leq t+T \mid P_t^i \} = \int_A \left(\frac{R}{P_t^i} \right)^{\frac{2\mu}{\sigma^2}} \frac{e^{-\frac{1}{2} \left(\frac{\ln(x) + \ln(P_t^i) - 2\ln(R) - \mu T}{\sigma \sqrt{T}} \right)^2}}{\sigma \sqrt{2\pi T}} dx \quad (19)$$

Equation (8) is obtained substituting (18) and (19) into (17), and multiplying the resulting expression by $\Pr \{\tau_1^i > t + T\} = e^{-\delta T}$. According to Bayes' rule, the unconditional density

$$\begin{aligned} & \Pr \{P_{t+T}^i \in A \cap \tau_2^i > t \cap \tau_1^i > t\} \\ &= \Pr \{P_{t+T}^i \in A \cap \tau_2^i > t \cap \tau_1^i > t \mid P_t^i \in B\} * \Pr \{P_t^i \in B\} \end{aligned}$$

where the Borel set $B \subset (R, +\infty)$. Therefore, under the assumption according to which initial output P_t^i are drawn from $G(\cdot)$, the unconditional density is obtained integrating the conditional density (8) with respect to $G(\cdot)$. Finally the unconditional measure $v(x, T)$ is given by the unconditional density multiplied by the rate of job creation. According to the stationarity assumption, the mass of jobs created at each point in time is constant and equal to $\mu\lambda$.

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

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

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

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

where $d = \sqrt{\mu^2 + 2\delta\sigma^2}$. Using the result in Leland and Toft (1997) according to which for positive values of x

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

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

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

Using this limit to integrate $h(x, T; P)$ with respect to T and ensuring that the integration is always performed over positive values, yields

$$\varphi(x; P) = \int_0^{+\infty} h(x, T; P) dT = \begin{cases} P \left(\frac{x}{P} \right)^{\frac{\mu-d}{\sigma^2}-1} \left(\frac{1 - \left(\frac{R}{P} \right)^{\frac{2d}{\sigma^2}}}{d} \right); & \text{if } x \geq P \\ P \left(\frac{x}{P} \right)^{\frac{\mu+d}{\sigma^2}-1} \left(\frac{1 - \left(\frac{R}{x} \right)^{\frac{2d}{\sigma^2}}}{d} \right); & \text{if } x \in [R, P] \\ 0; & \text{otherwise} \end{cases} \quad (20)$$

The expression of $v(x)$ reported in Proposition 3 follows directly.

- **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 $v(x)$ from R up to infinity. Thus

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

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

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

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

- **Proof of Proposition 5:** When $G(P)$ 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

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

Leibnitz's rule implies that

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

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

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

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

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

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

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

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

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

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 $h(x, T; P)$. Under the parametric assumption that

$G(P)$ is lognormal, it reads

$$\begin{aligned}
& \frac{\int_R^{+\infty} h(x, T, P) dG(P)}{e^{-\delta T}} = \int_R^{+\infty} h(x, T, P) \left(\frac{e^{-\frac{1}{2} \left(\frac{\ln(P) - \Sigma}{\xi} \right)^2}}{\xi \sqrt{2\pi}} \right) d \ln(P) \\
& = \left(\frac{e^{-\frac{1}{2} \left(\frac{B(x, T) (\xi^2 + \sigma^2 T)}{\xi^2 \sigma^2 T} \right)^2}}{\sqrt{2\pi}} \right) \int_R^{+\infty} \left(\frac{e^{-\frac{1}{2} \left(\frac{\ln(P) - (A(x, T) \rho^2(T) + \Sigma(1 - \rho^2(T)))}{\sqrt{(\xi^2 + \sigma^2 T) \rho^2(T) (1 - \rho^2(T))}} \right)^2}}{\xi \sigma \sqrt{2\pi T}} \right) d \ln(P) \\
& - R^{\frac{2\mu}{\sigma^2}} \left(\frac{e^{-\frac{1}{2} \left(\frac{D(x, T) (\xi^2 + \sigma^2 T)}{\xi^2 \sigma^2 T} \right)^2}}{\sqrt{2\pi}} \right) \int_R^{+\infty} \left(\frac{P^{-\frac{2\mu}{\sigma^2}} e^{-\frac{1}{2} \left(\frac{\ln(P) - (C(x, T) \rho^2(T) + \Sigma(1 - \rho^2(T)))}{\sqrt{(\xi^2 + \sigma^2 T) \rho^2(T) (1 - \rho^2(T))}} \right)^2}}{\xi \sigma \sqrt{2\pi T}} \right) d \ln(P) \\
& = \left(\frac{e^{-\frac{1}{2} \left(\frac{B(x, T) (\xi^2 + \sigma^2 T)}{\xi^2 \sigma^2 T} \right)^2}}{\sqrt{2\pi} (\xi^2 + \sigma^2 T)} \right) \tilde{\Phi} \left(\frac{\ln(R) - (A(x, T) \rho^2(T) + \Sigma(1 - \rho^2(T)))}{\sqrt{(\xi^2 + \sigma^2 T) \rho^2(T) (1 - \rho^2(T))}} \right) \\
& - R^{\frac{2\mu}{\sigma^2}} \left(\frac{e^{-\frac{1}{2} \left(\frac{D(x, T) (\xi^2 + \sigma^2 T)}{\xi^2 \sigma^2 T} \right)^2}}{\sqrt{2\pi} (\xi^2 + \sigma^2 T)} \right) e^{\left(\left(-\frac{2\mu}{\sigma^2} \right)^2 \left(\frac{\xi^2 \sigma^2 T}{2(\xi^2 + \sigma^2 T)} \right) + \left(-\frac{2\mu}{\sigma^2} \right) (C(x, T) \rho^2(T) + \Sigma(1 - \rho^2(T))) \right)} \\
& * \tilde{\Phi} \left(\frac{\ln(R) - \left(\left(-\frac{2\mu}{\sigma^2} \right) \left(\frac{\xi^2 \sigma^2 T}{\xi^2 + \sigma^2 T} \right) + C(x, T) \rho^2(T) + \Sigma(1 - \rho^2(T)) \right)}{\sqrt{(\xi^2 + \sigma^2 T) \rho^2(T) (1 - \rho^2(T))}} \right) \quad (21)
\end{aligned}$$

where $\tilde{\Phi}(x) = 1 - \Phi(x)$ and

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

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

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

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

$$\begin{aligned}
\int_R^{+\infty} \varphi(x, P) dG(P) &= \int_x^R \left(\frac{x}{P}\right)^{\frac{\mu+d}{\sigma^2}-1} P \left(\frac{1 - \left(\frac{R}{x}\right)^{\frac{2d}{\sigma^2}}}{d}\right) \left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right) d\ln(P) \\
&\quad + \int_R^{+\infty} \left(\frac{x}{P}\right)^{\frac{\mu-d}{\sigma^2}-1} P \left(\frac{1 - \left(\frac{R}{P}\right)^{\frac{2d}{\sigma^2}}}{d}\right) \left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right) d\ln(P) \\
&= \left(\frac{x^{\frac{\mu-d}{\sigma^2}-1}}{d}\right) e^{\left(\left(\frac{d-\mu}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right) + \left(\frac{d-\mu}{\sigma^2}\right)\Sigma\right)} \left(\begin{array}{c} \Phi\left(\frac{\ln(x) - \left(\left(\frac{d-\mu}{\sigma^2}\right)\xi^2 + \Sigma\right)}{\xi}\right) \\ -\Phi\left(\frac{\ln(R) - \left(\left(\frac{d-\mu}{\sigma^2}\right)\xi^2 + \Sigma\right)}{\xi}\right) \end{array}\right) \\
&\quad + \left(\frac{x^{\frac{\mu+d}{\sigma^2}-1}}{d}\right) e^{\left(\left(\frac{-d-\mu}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right) + \left(\frac{-d-\mu}{\sigma^2}\right)\Sigma\right)} \tilde{\Phi}\left(\frac{\ln(x) - \left(\left(\frac{-d-\mu}{\sigma^2}\right)\xi^2 + \Sigma\right)}{\xi}\right) \\
&\quad + R^{\frac{2d}{\sigma^2}} \left(\frac{x^{\frac{\mu-d}{\sigma^2}-1}}{d}\right) e^{\left(\left(\frac{-d-\mu}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right) + \left(\frac{-d-\mu}{\sigma^2}\right)\Sigma\right)} \tilde{\Phi}\left(\frac{\ln(R) - \left(\left(\frac{-d-\mu}{\sigma^2}\right)\xi^2 + \Sigma\right)}{\xi}\right) \quad (22)
\end{aligned}$$

Finally consider

$$\begin{aligned}
\int_R^{+\infty} \left(1 - \left(\frac{R}{P}\right)^{\frac{\mu+d}{\sigma^2}}\right) dG(P) &= \int_R^{+\infty} \left(1 - \left(\frac{R}{P}\right)^{\frac{\mu+d}{\sigma^2}}\right) \left(\frac{e^{-\frac{1}{2}\left(\frac{\ln(P)-\Sigma}{\xi}\right)^2}}{\xi\sqrt{2\pi}}\right) d\ln(P) \\
&= \tilde{G}(R) - R^{\frac{\mu+d}{\sigma^2}} e^{\left(\left(\frac{-\mu-d}{\sigma^2}\right)^2\left(\frac{\xi^2}{2}\right) + \left(\frac{-\mu-d}{\sigma^2}\right)\Sigma\right)} \tilde{\Phi}\left(\frac{\ln(R) - \left(\left(\frac{-\mu-d}{\sigma^2}\right)\xi^2 + \Sigma\right)}{\xi}\right) \quad (23)
\end{aligned}$$

The closed-form expression of the likelihood function is obtained inserting (21), (22) and (23) into (12). Notice that the unemployment rate also has an analytical solution

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

References

- [1] Altonji, Joseph G. and Robert Shakotko, (1987), “Do Wages Rise with Job Seniority?,” *Review of Economic Studies*, vol. 54, no. 3, pp. 437-59.
- [2] Altonji, Joseph G. and Nicholas Williams, (2005), “Do Wages Rise with Job Seniority? A Reassessment”, *Industrial and Labor Relations Review*, 58, No. 3: 370-397.
- [3] Cahuc, Pierre, Fabien Postel-Vinay and Jean-Marc Robin, (2005), “Wage Bargaining with On-the-job Search: Theory and Evidence”, forthcoming *Econometrica*.
- [4] Cossa, Ricardo, James J. Heckman and Lance Lochner, (2003), “Wage Subsidies and Skill Formation: A Study of the Earned Income Tax Credit” in E. Phelps, ed., *Designing Inclusion*, Cambridge University Press.
- [5] Dey, Matthew and Christopher Flinn, (2005), “An Equilibrium Model of Health Insurance Provision and Wage Determination”, *Econometrica*, 73: 571-627.
- [6] Dustmann Christian and Costas Meghir, (2005), “Wages, Experience and Seniority”, *Review of Economic Studies*, 72, 77-108.
- [7] 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, 263-286.
- [8] Felli Leonardo and Christopher Harris, (1996), “Learning, Wage Dynamics, and Firm-Specific Human Capital”, *Journal of Political Economy*, 104, No. 4: 838-868.
- [9] Flinn, Christopher and James Heckman, (1982), “New Methods for Analyzing Structural Models of Labor Force Dynamics”, *Journal of Econometrics*, 18, 115-168.
- [10] Flinn, Christopher, (2006), “Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates”, Unpublished Manuscript, New-York University.
- [11] Greene, William H., (1980), “Maximum Likelihood Estimation of Econometric Frontier Functions”, *Journal of Econometrics*, 13, 27-56.

- [12] Jolivet, Grégory, Fabien Postel-Vinay and Jean-Marc Robin, (2004), “The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US”, Unpublished Manuscript, CREST-INSEE, Paris.
- [13] 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, No 3: 987-1019.
- [14] Mortensen, Dale T. and Christopher A. Pissarides, (1994), “Job Creation and Job Destruction in the Theory of Unemployment”, *Review of Economic Studies*, 61, 397-415.
- [15] Nagypál, Éva, (2005), “On the Extent of Job-to-Job Transitions”, Unpublished Manuscript, Northwestern University.
- [16] Newey, Whitney K. and Daniel McFadden, (1994), “Large Sample Estimation and Hypothesis Testing”, in *Handbook of Econometrics*, Amsterdam: North-Holland, pp. 2111-245.
- [17] Postel-Vinay, Fabien and Hélène Thuron, (2005), “On-the-job Search, Productivity Shocks, and the Individual Earnings Process”, Unpublished Manuscript, University of Bristol.
- [18] Prat, Julien, (2006), “Job Separation under Uncertainty and the Wage Distribution”, *Contributions to Macroeconomics*, Berkeley Electronic Press, 6(1), Article 2.
- [19] 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.
- [20] Topel, Robert, (1991), “Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority”, *Journal of Political Economy*, 99(1), 145-176.
- [21] Yamaguchi, Shintaro, (2006), “Job Search, Bargaining, and Wage Dynamics”, Unpublished Manuscript, University of Wisconsin, Madison.