EFRW 7/25/02
MOSCARINI 3:15PM

SKILL AND LUCK IN THE THEORY OF TURNOVER*

GIUSEPPE MOSCARINI†
Yale University, Department of Economics

First draft: April 2001. This draft: May 2002.

Abstract

This paper investigates the joint implications of search and matching frictions in labor markets for wage inequality, and quantifies the average amount and the distribution of specific job-matching capital, vulnerable to exogenous job destruction. Workers differ both *ex-ante* in their average individual productivity (skill) and *ex-post* in their luck when matching with employers, learn over time the quality of the match, bargain on a wage, and search on and off the job for new employers. Conditional on skills, learning and selection map gaussian output noise into an equilibrium stationary and ergodic wage distribution which is unimodal and right-skewed, with a Paretian right tail. In contrast to the predictions of the standard search environment with observable productivity, high idiosyncratic technological risk hinders learning and sorting, and tends to reduce wage inequality. When parameterized to match observed aggregate worker flows, the model accurately predicts the observed wage loss following job destruction and hazard rates of separation as a function of tenure. The average amount of matching capital, vulnerable to job destruction, is then quantified at over a year worth of wages. Across skills, more able workers are more willing to tolerate mismatch to avoid unemployment; hence on average they experience a longer tenure, a more pronounced within-skill wage dispersion, a lower relative wage and welfare loss from displacement, a lower entry rate into unemployment and unemployment rate, higher job-to-job quitting rates with associated larger wage raises. Less skilled workers are dismissed earlier and then need to try more jobs or to get luckier to stay employed.

Keywords: wage distribution, job matching, job stability, specific human capital, worker flows, unemployment, Bayesian learning, ergodic analysis.

JEL Classification: C73, D31, D83, E24, J63, J64.

*I thank Yale University and the Alfred P. Sloan Foundation for financial support; Boyan Jovanovic, Alan Manning, Julia Lane and seminar participants at UW Madison, Rochester, Yale, LSE, Northwestern, NYU, MIT and SOLE for comments that are incorporated in this version; David Neumark for kindly sharing data on job tenure. All errors are mine.

†PO Box 208268, New Haven, CT 06520-8268, USA. Phone (203) 432-3596. Fax (203) 432-5779. E-mail: giuseppe.moscarini@.yale.edu, URL www.econ.yale.edu/~mosca/mosca.htm

1. Introduction

Job destruction is a major concern for economists, politicians, and the general public alike, who all view it as a traumatic event of serious welfare consequences – hence the name "destruction". Several sources of welfare loss are indirect and relate to the unemployment spell of the worker: uninsured income loss, skill loss, negative social externalities. Other sources refer to job-specific capital that evaporates upon a permanent separation.

This paper provides the first thorough characterization of the equilibrium of a frictional labor market in the presence of incomplete information about match quality. A parsimonious analytic framework captures the continuous *ex-post* sorting and labor reallocation to more productive matches, reconciles the intricate patterns of observed worker flows and wage dynamics (both within and across jobs), accounts for the main stylized facts on wage inequality, and sheds new light on unobserved heterogeneity in structural wage equations. Building on this new framework, we evaluate quantitatively the average size and the cross-sectional distribution of matching capital in the sense of Jovanovic (1979), namely the knowledge that firm and worker accumulate over time about their match and that is lost upon separation. This magnitude provides a lower bound to the implied productivity decline that causes the separation, if this is efficient, and to the potential deadweight welfare loss for society from an inefficient separation.

There are two reasons to focus on job-matching and learning as a form of specific capital. First, ample empirical evidence speaks to its importance as an engine of worker turnover and wage dynamics.[1] However, contrary to the claims made in this literature, this evidence is also consistent with other sources of *ex-post* heterogeneity and specific human capital. Our equilibrium analysis of job-matching provides a comparative statics prediction that is unique to an incomplete information environment: more pronounced idiosyncratic technological uncertainty throws sand in the works of learning and sorting, and reduces wage inequality. Second, the presence of learning guides the modeler in the formalization of idiosyncratic productivity uncertainty. This source of risk has attracted increasing attention in the incomplete-market macroeconomic literature, but is hard to measure empirically. In this respect, learning imposes some general restrictions that are independent of the productivity

[1] Lane and Parkin (1998) find strong support for the Jovanovic matching model in the turnover patterns of a major accounting firm. Altonji and Pierret (2001) find in the NLSY79 evidence of "statistical discrimination": firms hire workers based on easily observable characteristics – such as education – and then base their wage and promotion policies increasingly on what they learn from each employee's performance. Nagypal (2000) finds in French matched employer-employee data that learning about match quality vastly dominates learning-by-doing as a source of specific human capital accumulation. The job turnover literature offers ample evidence that a displacement begets subsequent job instability, which indirectly suggests the importance of specific learning. Stevens (1997) finds in the PSID that the long-run wage loss following a displacement increases in the number of intervening addtional job switches.

process – most notably, posterior beliefs about match quality are martingales – and thus have robust implications for the correlation between tenure, wages, on-the-job search, and other observables.

The first contribution of this paper is the introduction and the analysis of a tractable equilibrium model of the labor market, which incorporates both search frictions and learning about match quality. The model is qualitatively consistent with a wide rage of empirical regularities concerning worker flows and stocks as well as wages, both within and across worker skill categories. In particular, the model allows for closed-form aggregation and yields a realistic shape of the wage distribution: as a consequence of Bayesian learning and selection, the wage distribution is skewed to the right and has a "fat" Paretian right tail, even conditional on worker skills. The analytic solution also allows to uncover the new effect of output risk on wage inequality mentioned earlier, and sheds new light on the equilibrium interaction between general and specific human capital. The deep parameters of the model are also amenable to direct structural estimation, by simple maximization of the closed-form equilibrium wage likelihood function; this is the focus of a companion paper.

The second contribution is a parameterization of the model that successfully matches many empirical observations about labor market flows and wage inequality, with a considerable degree of "over-identification power": the calibration of a set of scalar parameters allows to correctly predict several moments of the data as well as entire distributions, of wages and separation rates. This empirical performance makes the model a potentially useful tool for the quantitative analysis of labor market policies and the role played by churning and turnover in promoting productivity growth.

The third contribution is an estimate of welfare effects: the average level of matching human capital, vulnerable to exogenous job destruction, exceeds one year of wages, and is larger for less skilled individuals. We abstract from other (orthogonal) potential sources of welfare loss, such as unemployment stigma, aversion to uninsurable income risk, and specific training: hence, here we consider just one side of the coin – better, of the die.

A key, but natural and rather uncontroversial assumption drives the distributional results of the model: more skilled workers have a larger wedge between productivity and opportunity cost of work. Hence they compress wages, because they have more to lose from rejecting a wage offer to stay unemployed; if dismal output suggests unlucky matching, they wait longer before separating from the firm and instead search on the job. In other words, they partially offset bad matching luck with their strong general human capital and better "tolerate mismatch", so they experience a lower entry rate into unemployment and unemployment rate, and a longer average tenure. Less skilled workers give the (false) impression of impatience when experimenting with jobs, due to their relative comparative disadvantage in market ac-

tivities. A bad match pays too meager a salary to justify the time and effort subtracted from leisure and job search. Fewer of the unskilled work, but those who do are matched better on average with their employers; they need to get luckier to stay employed. Displaced workers suffer a spell of unemployment and a decline in wage on average, but not always. More skilled individuals – prone to mismatch – experience a higher within-skill wage dispersion, a smaller wage loss from displacement and a larger wage gain when quitting to another job. Less skilled workers – who work continuously only when matched very well – have less to gain from quitting and more to lose from a layoff.

Section 2 reviews some related literature, Section 3 illustrates the model, Section 4 the bargaining equilibrium, Section 5 the effects of tenure on wages and turnover, Section 6 the stationary and ergodic wage distribution, Section 7 the general equilibrium of the model, Sections 8 its quantitative implications, Section 9 concludes with some directions for future research, an Appendix collects proofs of the propositions and illustrates the data used for the quantitative exercise.

2. Related Literature

Macroeconomic studies of worker turnover and labor market flows typically fail to provide realistic (or any) predictions on the wage distribution and dynamics, while microeconomic investigations of tenure, wages and turnover neglect aggregate implications for unemployment and job flows. Both strands of the literature downplay the equilibrium interaction between general and specific human capital. This paper aims to start filling this gap; therefore, it is closely related to a large macro and microeconomic literature, and it important to place it in the proper context.

First, we review the stylized facts that the present analysis captures in a unified framework. More productive (skilled, educated, experienced) individuals enjoy higher employment rates and hours: firms consider more able workers also more profitable, and employ them more.[2] Pronounced differences are observed in worker mobility. On-the-job search, which accounts for a significant fraction of all separations and hires (Blanchard and Diamond (1990), Fallick and Fleischman (2001)), produces fewer contacts but more job offers than unemployed job search, and is more effective for high-wage, mature, long-tenured workers (Pissarides and Wadsworth (1994) for UK, Blau and Robins (1990) for the US, Belzil (1996) for Canada), al-

[2]The orders of magnitude for workers of different ages and education levels are similar in all industrialized countries. At any point in time, "primary" workers experience from 5% to 45% higher labor participation rates, from 3% to 20% lower unemployment rates and from two to ten times lower entry rates into unemployment, as well as midly higher exit rates from unemployment and from 1.2 to two times as many weekly hours worked. This is true even within industry or sector. Moscarini (1996) surveys the relevant empirical evidence in this respect and on cyclical patterns.

though it declines in intensity with tenure (Pissarides 1994). More educated workers change occupation less frequently, even controlling for unobservable worker heterogeneity correlated with education (Moscarini and Vella (2000)).

On the microeconomic side, wage dynamics and tenure provide information on specific human capital and on the welfare effects of job mobility that this paper addresses. Wages correlate positively with education and tenure (Topel 1991), and separation rates negatively with tenure (Jovanovic (1979), Farber (1994)). Quits from job to job are associated to wage raises. In the empirical search literature, the wage attached to a job is traditionally decomposed into a worker-specific and a firm-specific component (Flinn (1986)). Recently, matched employer-employee panel data have allowed to decompose within-skill and within-firm unexplained wage dispersion (what we call "job matching") into worker and firm unobserved attributes. Postel-Vinay and Robin (2000) find that total log-wage dispersion as well as its "worker component" rise with the mean log-wage across occupational categories.

Motivated by the widespread concern about job loss mentioned at the outset, a substantial empirical literature aims to assess the wage and earnings loss suffered by a worker after a displacement (Ruhm 1991, Jacobson *et alii* 1993, Stevens 1997). Special attention is paid to the persistence of wage losses, as a clue of the implied reductions in permanent income ad thus welfare – what we should ultimately care about. In particular, Stevens uniquely emphasizes the temporal dependence of layoffs. However, this empirical work can assess (if any) the earnings loss to the worker only, conditional on few observable characteristics. The evaluation of the total size of this welfare loss, to firm and worker, and especially of its distribution across worker skill groups requires a structural, equilibrium analysis. As hours and wages vary considerably among workers, for reasons that are in part *ex-ante* predictable, before the job is filled, and in part revealed only *ex-post*, we should presume the same about job-matching capital. Extant dynamic equilibrium models of frictional labor markets often contradict the main stylized facts on the wage distribution, such as skewness and fat right tail. Hence, their predictions for the size and distribution of specific rents among active jobs should be taken with a pinch of salt.

Jovanovic (1984) introduces job search in his seminal 1979 matching model, and characterizes the individual employment relationship in market equilibrium, but does not draw distributional or just quantitative implications, as done here. Indeed, we provide one of the very first quantitative evaluations of learning-based matching models, in particular of their distributional implications within and across skill classes.

The equilibrium search literature has focused on *ex-post* job heterogeneity to explain aggregate job flows. In Mortensen and Pissarides (1994) [MP94], the benchmark framework for the study of labor market flows, firm-specific productivity is perfectly observed, starts

from the highest possible level and is subject to infrequent shocks; thus, wages do rise not with tenure but do rise after a displacement. Of course, a different productivity process might get around these counterfactual predictions. But, even assuming that the model be equally tractable, the distributional implications would always differ from those of this learning model. In particular, higher idiosyncratic output noise obviously amplifies wage inequality in a MP94 environment. By contrast, the same uncertainty clouds the intrinsic inequality in match outcomes, and compresses the wage structure near the starting wage. More generally, this model aims to bring the MP94 program a step further, anchoring it to empirical evidence on wages.

In Pissarides (1994), learning-by-doing on the job promotes wage growth and consequently temporary on-the-job search; the alternative source explored here, learning about match quality, encompasses and extends the relevant implications of that model. For instance, equilibrium wage dispersion here obtains even conditional on tenure. Pries (2001) shows that the job instability following a layoff in a job matching framework can explain the hysteresis of unemployment. His framework is greatly simplified to achieve tractable aggregate dynamics, hence remote from any realistic wage dispersion. Outside the wage bargaining tradition, Mortensen (1998) shows that a wage-posting search equilibrium with firm-specific upfront investments may generate a typical wage distribution, unimodal and with a long right tail. However, by construction, wage-posting models cannot account for within-job wage dynamics, in particular for the robust wage-tenure correlation.

In the early 1990s, economists detected a decade-long increase in the inequality of US wages and European unemployment rates across distinct classes of workers. This fact attracted new attention to *ex-ante* skill inequality in frictional labor markets, a line pursued here with regards to matching capital. Most authors focus on low-frequency shifts in wage inequality rather than on short-term labor market dynamics (Acemoglu (1998)). Mortensen and Pissarides (1999) introduce perfect labor market segmentation by skill levels in their 1994 framework, but focus mainly on the cross-sectional pattern of (un)employment, not of tenure and wages. In Neal (1998), a technological complementarity between general skills and specific training leads more educated workers to accumulate more training and to move less across jobs; here, a positive correlation between the two forms of human capital is not wired into technology, and yet emerges in equilibrium. In Krause (1999) workers are *ex-ante* homogeneous, while firms differ *ex-ante* in their entry costs, which affect an upfront investment in worker specific training, arguably producing inter-industry wage differentials.

3. The Economy

A consumption good is produced in continuous time by pairwise firm-worker matches (*jobs*). The average productivity of each match $f(x,\theta)$ is an increasing function of two orthogonal time-invariant factors: a worker-specific component x (*skill*), transferable across jobs and observable *ex-ante* by both parties, and a match-specific component θ (*match quality*), which is *ex-ante* uncertain and captures the experience good nature of a job. Upon matching, worker and firm share a common prior belief on θ which is independent of their past histories and is concentrated on two points: $p_0 = \Pr(\theta = \theta_H) = 1 - \Pr(\theta = \theta_L) \in (0,1)$, where θ_L denotes a "bad" match and $\theta_H (> \theta_L)$ a "good" match.

The performance of the match is also subject to two additional and orthogonal sources of idiosyncratic noise. First, cumulate output in $[0,t]$ is a normal random variable, a Brownian Motion with uncertain drift $f(x,\theta)$ and known infinitesimal variance σ^2:

$$y_t = f(x,\theta)\,t + \sigma Z_t \sim N\left(f(x,\theta)\,t, \sigma^2 t\right).$$

Here Z_t is a Wiener process, a continuous additive noise that keeps θ hidden and creates an inference problem. Over time, parties observe output realizations $\langle y_t \rangle$, generating a filtration $\{\mathcal{F}_t^y\}$, and update in a Bayesian fashion their belief from the prior p_0 to the posterior $p_t \equiv \Pr(\theta = \theta_H \mid \mathcal{F}_t^y)$. The second, more drastic source of idiosyncratic shocks is a Poisson jump process forcing jobs out of business at rate $\delta > 0$.

The economy is populated by a continuum of *ex-ante* identical firms, of mass large enough to ensure free entry, and by a continuum of *ex-ante* heterogeneous workers, with ordered skill types $x \in \mathbb{X} \subset \Re$ distributed by a given and known density μ. A jobless worker x enjoys a flow value of leisure $b(x)$, while idle firms get zero flow returns. Workers and firms are risk-neutral optimizers and discount future payoffs at rate $r > 0$.

Assumption 1. *For every skill $x \in \mathbb{X}$: $f(x,\theta_L) < b(x) < f(x,\theta_H)$.*

Assumption 2. *For every match outcome $\theta \in \{\theta_L, \theta_H\}$: $f(x,\theta) - b(x)$ is increasing in x.*

By Assumption 1, the matching choice is non trivial: a match should be dissolved if and only if $\theta = \theta_L$, when it produces less than the joint value of inactivity $b(x)$. *De facto*, parties perform a sequential probability ratio test of simple hypotheses on the viability of the match. By Assumption 2, more skilled workers have a comparative advantage in market activities. A skill-inelastic value of leisure $b(x) = b$ would trivially satisfy this assumption.

A worker x is hired at finite Poisson rate $\lambda(x)$ when unemployed, and at rate $\psi\lambda(x)$ when searching on the job. In both cases job search is costless, except for discounting. Here

ψ is the chance at every point in time that an employed worker who wants a new job has the opportunity to actively search for one. Every new match, whether the worker joins from unemployment or from another job, restarts from a common prior chance p_0 of success $(\theta = \theta_H)$. Search frictions create rents that the parties split according to a generalized Nash bargaining rule.

In this frictional environment with free entry in job creation, the influence of the general equilibrium on each employment relationship is summarized by *one number*, the worker matching rate $\lambda(x)$. This fact greatly simplifies the analysis, and allows to postpone til later the details of the matching process determining the function $\lambda(x)$ in equilibrium. The next few sections analyze the model in steady state abstracting from *ex-ante* worker heterogeneity, which is reintroduced in Section 7. For notational simplicity, until then we omit skill x and set $f(x, \theta) = \theta$, $\lambda(x) = \lambda$.

4. Wage Bargaining and Job Separation

The natural state variable of the bargaining game is the belief p_t that a match is productive $(\theta = \theta_H)$. By Theorem 9.1 in Liptser and Shyryaev (1977), conditional on the output process y, beliefs evolve from any initial $p_0 \in (0, 1)$ as a martingale diffusion solving:

$$dp_t = p_t(1 - p_t)sd\bar{Z}_t \tag{4.1}$$

where

$$s \equiv \frac{\theta_H - \theta_L}{\sigma}$$

is the *signal/noise ratio* of output, and

$$d\overline{Z}_t \equiv \frac{1}{\sigma}\left[dy_t - p_t\theta_H dt - (1 - p_t)\theta_L dt\right]$$

is the *innovation* process, the normalized difference between realized and unconditionally expected flow output. This is a standard Wiener process w.r. to the filtration $\{\mathcal{F}_t^y\}$. Intuitively, beliefs move faster the more uncertain match quality (the term $p(1-p)$ peaks at $p = 1/2$), and the more informative production, as measured by s.[3]

We may then write Bellman equations for worker and firm given an arbitrary wage function $w : [0, 1] \to \Re$ of the belief p, to be pinned down by the bargaining equilibrium. We assume that an employer can commit not to match outside offers received by its employees

[3]In this binary structure, unlike in the gaussian model of Jovanovic (1979, 1984), belief precision does not necessarily increase over time as evidence accumulates. However, the qualitative implications of Jovanovic's model depend on the martingale property of beliefs and on optimal selection, *not* on the specific match distribution assumed. In fact, these properties survive intact in this binary framework (see Proposition 2). By contrast, aggregation is tractable in the binary structure, not in the gaussian model.

who search on the job. We will see that a firm has in fact a strict *ex-ante* incentive to make such a commitment, and often no *ex-post* temptation to renege on it.[4]

When unemployed a worker enjoys a value U, and when working at belief p of a good match a value $W(p)$. These values solve the Hamilton-Jacobi-Bellman (HJB) equations:[5]

$$\begin{aligned} rU &= b + \lambda[W(p_0) - U] \\ rW(p) &= w(p) + \Sigma(p)W''(p) - \delta[W(p) - U] + \psi\lambda \max\langle W(p_0) - W(p), 0\rangle \end{aligned} \tag{4.2}$$

where

$$\Sigma(p) \equiv \frac{1}{2}s^2p^2(1-p)^2$$

is half the *ex-ante* variance of the change in posterior beliefs, roughly speaking "the speed of learning" about match quality. The opportunity cost of unemployment, rU, equals its flow benefit b plus the capital gain $W(p_0) - U$ accruing from a new match at rate λ, with prior belief p_0 that it will be successful. Similarly, the opportunity cost of working in a job that is productive with updated chance p, namely $rW(p)$, equals the flow wage $w(p)$, plus a diffusion-learning term $\Sigma(p)W''(p)$, minus the capital loss following exogenous separation at rate δ, plus the capital gain following a profitable quit to another job, which resets the prior to p_0. The learning speed $\Sigma(p)$ is converted into payoff units by the convexity of the Bellman value $W''(p)$, because information (here in the form of output) spreads posterior beliefs and empowers more informed decisions by the worker.

The worker optimally quits to unemployment at every belief $\underline{p}_w \in [0,1]$ such that $W(\underline{p}_w) = U$ (*value matching*) and $W'(\underline{p}_w) = 0$ (*smooth pasting*), and resumes search on the job whenever $W(p)$ falls short of the value $W(p_0)$ that the worker can obtain from a fresh start at a new firm. In this case the worker gains exactly $W(p_0) - W(p)$, given that her current employer is committed not to match outside offers. We assume that W is increasing and show that there exists a bargaining equilibrium that verifies this conjecture.

The problem of the firm is similar. The value of a vacancy is zero by free entry, a restriction that will be used later to close the model. The value of an active job to the employer producing at belief p solves the HJB equation

$$rJ(p) = \bar{y}(p) - w(p) + \Sigma(p)J''(p) - \delta J(p) - \psi\lambda J(p)\mathbb{I}\{W(p) \leq W(p_0)\}] \tag{4.3}$$

with $\mathbb{I}\{\cdot\}$ an indicator function, so $\mathbb{I}\{W(p) \leq W(p_0)\} = 1$ if and only if the worker seeks

[4]Absent outside interference, such as firing taxes and subsidies, due to the private Pareto efficiency of Nash Bargaining, no side-payments occur in equilibrium when both parties are idle and meet, or when they are matched and abandon production.

[5]Unless otherwise noted, Karlin and Taylor (1981) is the main reference for the standard technical results in diffusion theory exploited in this paper.

outside offers. The opportunity cost of production $rJ(p)$ equals expected flow output

$$\bar{y}(p) \equiv p\theta_H + (1-p)\theta_L$$

net of wage payments $w(p)$, plus returns from learning $\Sigma(p)J''(p)$, minus expected capital losses due to exogenous separation ($\delta J(p)$) and to a quit by the worker to another job ($\psi\lambda J(p)$ when $W(p) \leq W(p_0)$ and the worker keeps searching). The firm optimally fires the worker at every belief $\underline{p}_f \in [0,1]$ such that $J(\underline{p}_f) = 0$ and $J'(\underline{p}_f) = 0$.

The Nash bargaining solution gives the worker a fraction $\beta \in (0,1)$ of total match surplus:

$$J(p) = \frac{1-\beta}{\beta} \cdot [W(p) - U]. \tag{4.4}$$

Hence parties agree to separate at the same belief(s) $\underline{p} = \underline{p}_w = \underline{p}_f$ and become idle. This standard result fails when the worker quits to another job: while the worker forfeits positive rents $W(p) - U > 0$ for even larger ones in the new match, her employer suffers a loss $J(p) \propto W(p) - U > 0$. As discussed in Moscarini (2001) in a conceptually similar framework, an employer has a strict *ex-ante* incentive to commit not to match outside offers to its employees, and nearly no incentive to deviate from it *ex-post*, despite the loss $J(p)$.

To better appreciate the incentives to commit, consider first the case $p < p_0$: a worker employed at belief p seeks another *better* job, a fresh start at p_0. The new job can be easily shown to generate a bigger "pie", by $p_0 > p$: hence a counteroffer by the current employer could never outbid (*e.g.*, in an auction) the new employer, but would just strengthen her employee's bargaining position vis-à-vis the new firm. In turn, this effect would raise the worker returns to on-the-job search, her propensity to pursue outside offers when employed, the likelihood of a quit, and finally the associated expected loss for the current employer. If instead $p \geq p_0$ and the worker seeks an alternative match that is *worse* than her current job, her employer's commitment not to respond to outside offers makes the quitting threat empty: left alone to bargain with a new firm re-starting from belief p_0, the worker earns just the usual bargaining rents $W(p_0)$, smaller than the $W(p)$ she is leaving behind. If the worker carries out her non-credible threat, then her employer operating at $p > p_0$ will be tempted to respond and to outbid the poacher. Again, succumbing to this temptation would reduce profits ex ante. Reputation as a commitment device, albeit often abused, appears natural in this case.[6]

[6]Pissarides (1994) posits a different solution: employed search takes place as long as it creates a positive surplus for the current match, to be abandoned, while the costs and returns of the new employer do not play a role. Felli and Harris (1996) characterize the unique, perfect sequential equilibrium of the repeated "poaching" game without any commitment but with perfect recall, or equivalently without search frictions. In Postel-Vinay and Robin (2000) the new employer has bargaining power, but each job move entails a wage gain because the worker's outside option is the current job, not unemployment. They find that the wage distribution implied by the model fits well evidence from a matched employer-employee French panel.

Observe that (4.4) implies $\mathbb{I}\{W(p) \leq W(p_0)\} = \mathbb{I}\{J(p) \leq J(p_0)\}$ and $J''(p) = W''(p) \cdot (1-\beta)/\beta$. Using these facts and (4.4) into the HJB equations (4.2) and (4.3) and a fair amount of (omitted) algebra yield a simple and intuitive expression for the equilibrium wage:

$$w(p) = \beta \bar{y}(p) + (1-\beta)b + \beta\lambda J(p_0)(1 - \psi \mathbb{I}\{J(p) \leq J(p_0)\}). \tag{4.5}$$

The wage includes (a share of) flow expected output $\bar{y}(p)$ and outside options, both exogenous b and endogenous from unemployed job search $\lambda J(p_0)$, and is reduced by $\beta\psi\lambda J(p_0)$ when the match looks unpromising ($W(p) \leq W(p_0)$ or $J(p) \leq J(p_0)$) in order to compensate the firm for the resulting potential loss of a still valuable employee. The wage is linearly increasing in beliefs and jumps up at p_0 as the worker ceases search on-the-job and the firm no longer faces the loss. Employed search improves the worker's outside option at the expense of joint match surplus.

Finally, using (4.4), (4.5) and boundaries turns the firm's HJB equation (4.3) into:

$$J(p) = \frac{(1-\beta)[\bar{y}(p) - b] + \Sigma(p) J''(p) - \beta\lambda J(p_0)(1 - \psi \mathbb{I}\{J(p) \leq J(p_0)\}),}{r + \delta + \psi\lambda \mathbb{I}\{J(p) \leq J(p_0)\}} \tag{4.6}$$

subject to value matching and smooth pasting at $\underline{p}$. An additional boundary condition is

$$J(1) = (1-\beta)\frac{\theta_H - b}{r+\delta} - \frac{\beta\lambda}{r+\delta} J(p_0),$$

because the worker would never quit a most productive match ($W(1) > W(p_0)$). This allows to solve for the value function, which equals the sum of the present discounted value of flow returns and of the option value of separating should things go wrong, including a direct quit for $p \leq p_0$. So let:

$$\alpha_0 \equiv \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta+\psi\lambda)}{s^2}}; \quad \alpha_1 \equiv \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta)}{s^2}}$$

with $\alpha_0 > \alpha_1 > 1$.

Proposition 1. (Bargaining Equilibrium and Optimal Stopping) *A firm and a worker match sharing a common prior belief p_0 of a good outcome, continuously observe output in $[0,t]$, update the posterior belief p_t according to (4.1), renegotiate the wage $w(p)$ according to (4.5), and separate when the posterior declines to a low cutoff $\underline{p} \in (0, p_0)$ (when the wage falls to a reservation value $w(\underline{p})$). The worker searches on the job for a new match at prior p_0 (again) if and only if $p_t \leq p_0$. The value function of the firm is the increasing and convex function of beliefs:*

$$\begin{aligned} J(p) = {} & \frac{(1-\beta)[\bar{y}(p) - b] - \beta\lambda J(p_0)(1 - \psi \mathbb{I}\{\underline{p} \leq p \leq p_0\})}{r + \delta + \psi\lambda \mathbb{I}\{\underline{p} \leq p \leq p_0\}} \\ & + \left[c_{0J} p^{1-\alpha_0}(1-p)^{\alpha_0} + k_{0J} p^{\alpha_0}(1-p)^{1-\alpha_0}\right] \mathbb{I}\{\underline{p} \leq p \leq p_0\} \\ & + c_{1J} p^{1-\alpha_1}(1-p)^{\alpha_1} \mathbb{I}\{p_0 < p \leq 1\}, \end{aligned}$$

where the coefficients c_{0J}, k_{0J}, c_{1J}, the initial value $J(p_0) = J(p_0-)$ and the optimal stopping point $\underline{p} \in (0, p_0)$ uniquely solve the system of five algebraic equations:

$$J(\underline{p}) = 0 \qquad J'(\underline{p}+) = 0 \qquad J(p_0-) = J(p_0+) \qquad J'(p_0-) = J'(p_0+)$$
$$J(p_0+) = \frac{(1-\beta)[\bar{y}(p_0) - b] + (r+\delta)\, c_{1J} p_0^{1-\alpha_1}(1-p)^{\alpha_1}}{r + \delta + \beta\lambda}.$$

5. Tenure, Wages and Turnover

The bargaining equilibrium described in Proposition 1 implies a stochastic process for the worker's employment status and, conditional on employment, for the belief of a good match p_t. The belief starts from p_0, evolves as the diffusion (4.1) following output realizations, is "killed" at rate δ by exogenous destruction and is absorbed into unemployment for a random duration of mean $1/\lambda$. The same happens if dismal output drives the belief down to $\underline{p}$ and leads parties to separate and to restart search. If $p_t \leq p_0$ the worker also seeks outside job offers, and finds one at rate $\psi\lambda$, resetting the belief to p_0.

Before presenting the novel contributions of this paper, we verify that the model predicts some qualitative correlations between tenure, wages and employed search that are observed in the data and that are central to previous turnover models, most notably Jovanovic (1984) and Pissarides (1994). In addition, we find a closed-form expression for expected residual tenure $\mathcal{T}(p)$ as a function of the current belief p (or wage). In the absence of endogenous separation at $\underline{p}$, $\mathcal{T}(p)$ should be equal to $1/\delta$ for $p > p_0$ when outside offers are rejected, and to $1/(\delta + \psi\lambda)$ for $p \leq p_0$ when they are accepted. But the match also terminates endogenously, when the posterior falls to $\underline{p}$. Overall, $\mathcal{T}(p)$ solves:

$$\Sigma(p)\mathcal{T}''(p) - (\delta + \psi\lambda\mathbb{I}\{\underline{p} \leq p \leq p_0\})\mathcal{T}(p) = -1.$$

Lemma 1. (Expected Tenure) *The expected future duration of a match is the increasing and convex function of beliefs that the match is productive:*

$$\begin{aligned}\mathcal{T}(p) \;&=\; \frac{1}{\delta + \psi\lambda}\left\{1 + c_{0\mathcal{T}} p^{1-\alpha_0}(1-p)^{\alpha_0} + k_{0\mathcal{T}} p^{\alpha_0}(1-p)^{1-\alpha_0}\right\}\mathbb{I}\left\{\underline{p} \leq p \leq p_0\right\} \\ &\quad + \frac{1}{\delta}\left\{1 + c_{1\mathcal{T}} p^{1-\alpha_1}(1-p)^{\alpha_1}\right\}\mathbb{I}\left\{p_0 < p \leq 1\right\},\end{aligned}$$

where $\{c_{0\mathcal{T}}, k_{0\mathcal{T}}, c_{1\mathcal{T}}\}$ solve $\mathcal{T}(\underline{p}) = 0$, $\mathcal{T}(p_0-) = \mathcal{T}(p_0+)$, $\mathcal{T}'(p_0-) = \mathcal{T}'(p_0+)$.

Standard in Bayesian learning, in expectation with respect to current beliefs p_t, posterior beliefs $p_{t+\Delta t}$ are a martingale: $\mathbb{E}[p_{t+\Delta t} | 0 \leq p_{t+\Delta t} \leq 1, p_t] = p_t$ for all $\Delta t \geq 0$. But, if we

condition on match continuation from t to $t + \Delta t > t$, the belief is a strict *sub*martingale, because it is bounded below by $\underline{p} > 0$ and reflects only good output outcomes. In fact:

$$\begin{aligned} \mathbb{E}[p_{t+\Delta t}|\text{match continues in } [t, t+\Delta t], p_t] &> \mathbb{E}[p_{t+\Delta t}|\underline{p} \leq p_{t+\Delta t} \leq 1, p_t] \\ &> \mathbb{E}[p_{t+\Delta t}|0 \leq p_{t+\Delta t} \leq 1, p_t] \\ &= p_t \end{aligned}$$

where the first inequality holds because quits occur only for low beliefs ($p \leq p_0$), and the chance of exogenous match dissolution is independent of p_t, so match continuation is more likely for high beliefs in $[t, t + \Delta t]$; the second follows from $\underline{p} > 0$; and the equality is the martingale property. Hence the confidence in a good match rises on average with tenure, although there is always a positive chance of a decline. Also standard in Bayesian learning, the value function is convex in beliefs p, hence a submartingale too. The flow wage (4.5) is linear in beliefs, due to the combined assumptions of expected utility and linear sharing rule, hence it is also a submartingale for continuing matches. We summarize these findings:

Proposition 2. (Tenure, Wages, and Search Behavior) *Unconditionally on true match quality, but conditional on match continuation, the human wealth of the employed worker $W(\cdot)$, her flow wage $w(\cdot)$, and the rents of her employer $J(\cdot)$ rise on average with tenure. On-the-job search is more common among low-tenured workers.*

6. The Ergodic Wage Distribution

The first contribution of this paper is an explicit solution for the stationary distribution of wages, conditional on skills, which exhibits the right qualitative properties to fit the data. This fit is hardly a major strength of extant models of frictional labor market. The analytic solution for the wage distribution also uncovers some new sources of wage inequality, and makes quantitative evaluation simple and transparent.

The stochastic process describing the evolution of the belief of a good match in equilibrium is clearly Markovian and strongly recurrent. Therefore, the stationary density is also ergodic: from any non-degenerate prior, the posterior belief converges to a random variable with density denoted by ϕ, if this exists, support $[\underline{p}, 1]$, and total mass equal to total employment, plus an atom of unemployment. In a large population of workers, ϕ can be interpreted also as the ergodic and stationary cross-sectional distribution of workers (beliefs, wages) among jobs. For the following results we refer to Feller (1954)'s classic treatment of diffusions on an interval and to his physical interpretation of the dynamic equations. Imposing stationarity in the Fokker-Planck (Kolmogorov forward) equation of the process, which

describes the dynamics of the transition density, we obtain the following equation for the stationary and ergodic density ϕ of the belief process:

$$\frac{d^2}{dp^2}[\Sigma(p)\phi(p)] - (\delta + \psi\lambda\mathbb{I}\{\underline{p} \leq p \leq p_0\})\phi(p) = 0, \tag{6.1}$$

subject to the following three boundary conditions:

1. no time spent at the separation boundary $\underline{p} > 0$: $\Sigma(\underline{p})\phi(\underline{p}+) = 0$, thus by $\Sigma(\underline{p}) > 0$,

$$\phi(\underline{p}+) = 0;$$

2. balance of total flows (respectively) in and out of employment:

$$\Sigma(p_0)[\phi'(p_0-) - \phi'(p_0+)] = \psi\lambda \int_{\underline{p}}^{p_0} \phi(p)dp + \delta \int_{\underline{p}}^{1} \phi(p)dp + \Sigma(\underline{p})\phi'(\underline{p}+);$$

3. balance of total flows (respectively) in and out of unemployment:

$$\delta \int_{\underline{p}}^{1} \phi(p)dp + \Sigma(\underline{p})\phi'(\underline{p}+) = \lambda(\mu - \int_{\underline{p}}^{1} \phi(p)dp).$$

The first boundary condition is standard for "attainable" boundaries, which can be hit in finite time with positive probability and are either absorbing or reflecting. The second boundary condition equates the total inflow into employment on the LHS to the total outflow, due to (in order) quits to other jobs, exogenous job destructions and quits to unemployment at $\underline{p}$. The third boundary condition equates the flow into unemployment, both involuntary due to job dissolution at rate δ and voluntary through the separation boundary $\Sigma(\underline{p})\phi'(\underline{p}+)$, to the outflow, equal to the exit rate times unemployment. This is a standard restriction in search models, which gives rise to a Beveridge curve. The total flow in or out of employment exceeds the flow in or out of unemployment by an amount equal to job-to-job quits, because these are the only separations that do not entail an unemployment spell.[7]

To solve (6.1) for ϕ, let:

$$\gamma_0 \equiv \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(\delta + \psi\lambda)}{s^2}}; \qquad \gamma_1 \equiv \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2\delta}{s^2}}$$

where $\gamma_0 > \gamma_1 > 1$.

[7]Although the second condition may appear equivalent to the third, if the second boundary condition is violated then the distribution ϕ will generate different flows of quits $\psi\lambda \int_{\underline{p}}^{p_0} \phi(p)dp$ and new hires from employment, namely total hires $\Sigma(p_0)[\phi'(p_0-) - \phi'(p_0+)]$ minus exits from unemployment $\lambda(\mu - \int_{\underline{p}}^{1} \phi(p)dp)$.

Proposition 3. (Ergodic Distribution of Beliefs) *The ergodic and stationary density of beliefs on match quality equals, for $p \in [\underline{p}, 1]$:*

$$\begin{aligned}\phi(p) &= c_{0\phi} p^{-1-\gamma_0}(1-p)^{\gamma_0 - 2}\left[\left(\frac{1-\underline{p}}{\underline{p}}\frac{p}{1-p}\right)^{2\gamma_0 - 1} - 1\right]\mathbb{I}\left\{\underline{p} \le p \le p_0\right\} \\ &\quad + c_{1\phi} p^{-1-\gamma_1}(1-p)^{\gamma_1 - 2}\mathbb{I}\left\{p_0 < p \le 1\right\}\end{aligned}$$

where the coefficients $c_{0\phi}$ and $c_{1\phi}$ are the unique and positive solution of the linear algebraic system $\Xi(c_{0\phi}, c_{1\phi})' = (\lambda\mu, 0)'$ (the matrix Ξ is appendicized in Equation (A.3)). This density is increasing in $[\underline{p}, p_0]$ and continuous with a kink at p_0. For $(p_0, 1]$ it is decreasing if $\gamma_1 \ge 2$, namely if the rate of attrition exceeds the squared signal/noise ratio of output $\delta \ge s^2$, it is U-shaped if $\min\{3p_0 - 1, 1\} < \gamma_1 < 2$, and it is increasing if $1 < \gamma_1 \le \min\{3p_0 - 1, 1\}$.

Bayesian learning and Nash bargaining produce a (piece-wise) Lévy-stable distribution ϕ of beliefs, of the Lévy-Pareto *type*. This expression for ϕ is easy to interpret and, as we will see in Section 8, considerably simplifies the aggregation of the model. The interpretation of ϕ is empirically more meaningful in wage space. Invert the wage function (4.5) to obtain beliefs as an affine increasing function of wages (recall that wages are increasing in beliefs and jump up at p_0): $w(p) = \mathfrak{w}_{\mathbb{I}\{\underline{p} \le p \le p_0\}} + \beta\sigma s p$ where:

$$\mathfrak{w}_{\mathbb{I}\{\underline{p} \le p \le p_0\}} \equiv \beta\theta_L + (1-\beta)\, b + \beta\lambda J(p_0)\left(1 - \psi\mathbb{I}\left\{\underline{p} \le p \le p_0\right\}\right..$$

Then for $w \ge \underline{w} \equiv \mathfrak{w}_1 + \beta\sigma s\underline{p}$ and $w_0 \equiv \mathfrak{w}_1 + \beta\sigma s p_0$ the wage density is:

$$\chi(w) = \frac{1}{\beta\sigma s} \cdot \phi\left(\frac{w - \mathfrak{w}_{\mathbb{I}\{w \le w_0\}}}{\beta\sigma s}\right). \tag{6.2}$$

This scale-location transformation preserves the properties of ϕ for the wage density χ after a normalization $\beta\sigma s = 1$. In particular, both ϕ and χ belong to the Pareto type and have a *fat right tail*, as the numerical examples of Section 8 will confirm graphically. Quits to other jobs and to unemployment weed out disproportionately bad matches, censor the left tail, and skew the distribution to the right.[8]

The *rate of decay of the right tail* of the distribution, namely $\gamma_1 - 2$, increases with the ratio δ/s^2. When jobs are at high risk of exogenous destruction (δ is large), or when the

[8]Roy (1951) explains skewness by the self-selection of workers who know *ex-ante* their comparative advantages to work in different sectors. Under the assumption of log-normal bi-variate skills, Roy produced a wage distribution that lacks the Pareto-like tail and fails to fit the US wage distribution. Heckman and Sedlacek (1985) amend this shortcoming of the Roy model by introducing unobserved worker heterogeneity. Moscarini (2001) analyzes the effects of search frictions on self-selection in the Roy model. Here selection occurs *ex-post*, after matching with a firm; skewness and Pareto tail arise for any initial distribution of beliefs, due to the ergodic nature of ϕ.

output process is very noisy and uninformative and beliefs move slowly (σ is large and then the signal/noise ratio s is low), the selection process has no time to produce its effects. As we will see shortly, the parameter values that make the model's predictions consistent with observed workers flows imply $\delta \simeq s^2$, or $\gamma_1 \simeq 2$, hence a unique mode and declining but "fat" right tail in the wage distribution.

Proposition 3 also yields a new prediction, unique to this incomplete information model. Paradoxically, a "noisy" economy is "sclerotic": high idiosyncratic output uncertainty unrelated to firm and worker characteristics (high δ and σ) clouds the intrinsic inequality in productivities (θ) and prevents it from showing through equilibrium prices. The distribution of wages remains concentrated around the starting value. So income inequality tends to be dampened, rather than enhanced, by high idiosyncratic risk.[9] This reduction in wage inequality has a large efficiency cost due to large noise and poor sorting. This prediction is in striking contrast with that of a search model with perfectly observable match-specific productivity, such as Mortensen-Pissarides (1994). In that environment, a larger variance of idiosyncratic output shocks raises the incentives to maintain the job active in order to save on new search costs; in turn, this lowers the destruction cutoff, widens the range of wages and thus their inequality.

7. The Matching Function and General Equilibrium

The description of the economy is completed by a frictional matching process, and the equilibrium is closed by a free entry condition that determines job-finding rates $\lambda(x)$, so far treated parametrically. There are two main specifications of the matching process, depending on the interpretation of skills.

A simple solution, which ignores the complications stemming from the endogenous skill distribution of unemployment, is to assume perfectly segmented labor markets. For instance, suppose each occupation is characterized by a minimum required skills $\bar{x}$: the production function is $f(x, \theta|\bar{x}) = f(\min\langle x - \bar{x}, 0\rangle, \theta)$ with $f(0, \theta|\bar{x}) = 0$ and $f_1 > 0$, as in Mortensen and Pissarides (1999). They show that high$-x$ workers self-select in high-$\bar{x}$ jobs, so we may speak of an $\bar{x}-$market. Each exit rate $\lambda(x)$ is then determined by a separate random matching process in the corresponding segmented market, and there is no direct interaction between different skills. Given a functional form for the matching function, we can use a free entry condition to back out the vacancy-posting costs that are necessary in each market

[9] In the limit as output noise (σ) and drastic shocks (δ) vanish, every worker finds a good match fairly quickly and keeps it. So eventually wages tend to concentrate on the high-end, and inequality declines. This "equality at the top", however, is quite far from empirically plausible parameterizations of the model, as suggested by the numerical examples of the Section 8.

to produce in equilibrium empirically plausible exit rates.

Here we analyze the case of *competition for jobs among different skills.* If average output depends smoothly on worker skills, for example $f(x,\theta) = x + \theta$, then markets do not segment, but all workers apply to and compete for the same jobs, and the skill composition of unemployment matters for firms. The following description of non-sequential search by firms draws in part from Moscarini (2001). Let

$$a(x) \equiv \left[\mu(x) - \int_{\underline{p}(x)}^{1} \phi(p|x)\,dx\right] + \psi \int_{\underline{p}(x)}^{p_0} \phi(p|x)\,dx$$

denote the measure (density) of job applicants among $x-$workers, both unemployed – in square brackets – and employed but mismatched and having an opportunity for search with probability ψ. Each of the v open vacancies "fishes" a sample of applications, of random size and skill composition, from the distribution $a(x)$. The firm then chooses and optimally selects the applicant x who guarantees the largest rents (PDV of profits) $J(p_0|x)$. Hence, the matching rate $\lambda(x)$ for a worker x depends negatively on measure of more attractive workers. More precisely, let

$$m(x) \equiv \lambda(x)\,a(x)$$

denote the flow of workers of type x hired at every point in time, and

$$\pi(x) \equiv \int_{\{y:J(p_0|y)\geq J(p_0|x)\}} m(y)dy$$

denote the total measure of hired workers who are more profitable than x. It follows that only a total of $v - \pi(x)$ vacancies remain available for x workers. Among those vacancies, we assume that x workers match randomly. We posit a standard isoelastic CRS matching function: for $\eta \in (0,1)$, $m_0 > 0$,

$$\lambda(x) = \frac{m(x)}{a(x)} = \frac{m_0\,[v-\pi(x)]^{\eta}\,[a(x)]^{1-\eta}}{a(x)} = m_0 \left[\frac{v-\pi(x)}{a(x)}\right]^{\eta}.$$

In a symmetric equilibrium where all firms use the same strategy, the flow chance for a firm of hiring *some* worker equals $\int_{\mathbf{X}} \lambda(y)\,a(y)\,dy/v$, total hires divided by vacancies. Conditional on hiring a worker, the probability density of hired skills is $\lambda(x)\,a(x) \,/\, \int_{\mathbf{X}} \lambda(y)\,a(y)\,dy$. Free entry guarantees that the flow cost κ of keeping a vacancy open (advertising, interviewing etc.) equals the expected value of a new hire times the hiring rate:

$$\kappa = \frac{\int_{\mathbf{X}} \lambda(y)\,a(y)\,dy}{v} \int_{\mathbf{X}} J(p_0|x) \frac{\lambda(x)\,a(x)}{\int_{\mathbf{X}} \lambda(y)\,a(y)\,dy} dx = \int_{\mathbf{X}} J(p_0|x) \frac{\lambda(x)\,a(x)}{v} dx.$$

Rather than digging into existence and uniqueness in general, in the next section we implement this matching process in numerical examples to obtain quantitative predictions.

8. Quantitative Implications

The second goal of this paper is to turn the model into a quantitative tool, useful for a variety of applications. To this purpose, we parameterize the model to match quantitatively a wide range of stylized facts with considerable "over-identifying power": namely, the model should replicate many more empirical moments and distributions than the free parameters it contains. The third and final goal of the paper is to apply this tool to evaluate the welfare loss caused by an exogenous job destruction, through the evaporation of matching capital.

8.1. Parameterization and (Over-)Identifying Restrictions.

We study a version of the model with two types of skills, Low and High (x_L and x_H). The production function is additive, $f(x,\theta) = x+\theta$, with little loss in generality – we can always define a worker skill to be her ex ante expected productivity, and treat the rest as additive orthogonal zero-mean noise. We look for a Ranking Equilibrium, in which more productive workers are hired first: we conjecture $J(p_0|x_H) > J(p_0|x_L)$ and

$$\lambda(x_H) = \left[\frac{v}{a(x_H)}\right]^\eta; \qquad \lambda(x_L) = \left[\frac{v - \lambda(x_H)\, a(x_H)}{a(x_L)}\right]^\eta.$$

We verify ex post that initial values are ordered as conjectured.

This version of the model is parameterized at a monthly frequency, so that its aggregate predictions match some broad empirical facts about worker stocks and flows and wage inequality in the US in the last three decades. The data sources and the exact definitions of the various statistics in the model and in the data are illustrated in Appendix A.2.

The parameters to be chosen are described in Table 1. Some are simple normalizations: the labor force is 100, skills $x_L = 0.4$ and $x_H = 0.6$, match outcomes $\theta_H = -\theta_L = 0.5$. Normalizing the inter-skill productivity difference to 0.2 is equivalent to pin down the productivity scale, while its location depends on the value of leisure b. Since the prior belief will be set to $p_0 = 0.5$, match outcomes must sum up to zero by definition, so that their prior expectation is zero; their size can be chosen freely because only the ratio of $\theta_H - \theta_L$ to the standard deviation of output σ matters for decisions and outcomes. Three parameters are directly constrained by direct or indirect evidence. The discount rate is $r = 0.004$, corresponding to a 5% annual interest rate; the vacancy-elasticity $\eta = 0.3$ is chosen below more standard estimates in the [0.4,0.6] range (Petrongolo and Pissarides 2001) because the model's definition of unemployment is more extensive than in the data (see Appendix A.2); the proportion of low skills in the population equals 70%, the average proportion of the US labor force holding a High School degree or less in the last three decades.

The remaining eight parameters can be chosen freely, and are pinned down to match a variety of facts, as detailed in the last two columns of Table 2. Eight empirical observations serve as "identifying restrictions"; the qualitative properties of the wage distribution, the ninth datapoint, and the entire hazard rate of separation as a function of tenure (Figure A.1) serve as "over-identifying" restrictions. Unlike the statistics in Table 2, the hazard rate cannot be computed in closed-form and immediately evaluated. Rather, we derive the hazard rate from a 10^7-step simulation of a discrete-time version of the same model, where the step Δt is one day. The belief is replaced by a $\Delta t-$discrete time Markov process constructed so as to converge in distribution to the true belief diffusion as $\Delta t \to 0$.

Among the parameters that are selected by the identification procedure, in Table 1 notice in particular the exogenous attrition rate $\delta = 0.0129$ (a job is expected to last 77.5 months or 6.45 years unless the worker quits earlier); $\psi = 0.24$ (an employed worker who would like to switch jobs has an opportunity to search on average a quarter of the time, compared to a jobless individual); and $b = 0.32$, implying that wedge between general human capital x and value of leisure is 3.5 times larger for highly skilled workers ($x_H - b = 0.28$) than for low skills ($x_L - b = 0.08$).

The model correctly predicts more empirical observations than free parameters. To assess wage inequality and skewness, the statistics (Average wage − median wage)/Stdev(wages) is chosen because invariant to affine transformation in wages. Figure A.3 reports the ergodic belief distribution for low-skilled workers (the wage distribution is just an affine rescaling). The total mass of the distribution is the employment rate. For comparison, Figure A.4 reports the belief distribution from the same simulation of the model used to predict the hazard rate of separations in Figure A.1. In both cases it emerges that the distribution is right-skewed and has a fat and leveling right tail: in fact the chosen parameters imply $\gamma_1 = 1.96$, close to the critical value of 2 (see Proposition 3). From Table 2, the median wage is always below the mean wage. All these properties are characteristic of the US empirical wage distribution, both conditional and unconditional on observable worker skills.

The predicted hazard rate of separations in Figure A.1 tends to modestly overestimate the observed one on average, and the more so the higher is tenure. The reason is that agents in the model have an infinite time horizon to experiment with new jobs; in real life, as tenure proceeds and retirement approaches, the incentive to switch job decline, and the hazard rate of separation with it.

8.2. Cross-Skill Inequality and the Welfare Consequences of Job Destruction

The described parameterization of the model, in the light of its empirical performance, can be considered informative for additional predictions, such as cross-sectional patterns and welfare consequences of job mobility, also reported in Table 2.

In the absence of job upgrading and selection through quits to other jobs and to unemployment, the employed population would be split evenly above and below $p_0 = .5$. Notice *the strength of selection*: of the 90.32 workers employed at any point in time, only about 21% (=18.88/90.32) stick to their job although prior expectations have been disappointed, while the remaining 79% appear more productive than when they started, hence earn a wage above the initial value $w(p_0)$. Of the 18.88 "disappointed" workers, about one fourth can search for a job upgrading at every point in time, hence "Effective on-the-job searchers" are correctly predicted to about 5% of employment. As a consequence, the average paid wage (weighted across skills by employments) is roughly 16% larger than the average initial wage $w(p_0)$ (weighted across skills by new hires). Though the worker receives only a 40% share in bargaining, the starting wage is almost equal to initial expected flow output, just like in an economy without search frictions (Jovanovic 1979), because of the forward-looking returns to remain matched. The average relative wage gain from a direct quit is $1 - \mathbb{E}[w(p_0)/w(p)|\underline{p} \leq p \leq p_0]$, because only workers employed at $p \leq p_0$ search on the job and, when successful, restart from $w(p_0)$. This is estimated at a much more modest number than the wage loss from displacement; the model allows only for limited gains from a quit, while job destruction can hit even the most established of matches.

Figure A.2 illustrates the implied value function of a firm employing a low-skill worker, which is strikingly steep: the maximum present profits $J(1|x_L)$ are about 10 times bigger than the initial value $J(p_0|x_L)$. In contrast, flow output may only roughly double from a prior expectation of $x_L = 0.4$ to a theoretical maximum of $x_L + \theta_H = 0.9$. Average match surplus among active jobs is about four times bigger than the initial one. This value represents the expected *welfare loss from exogenous job destruction* for firm and worker, and far exceeds on average one year worth of wages. This is a large number, if we consider that it includes only the permanent income loss, and abstracts from additional sources of welfare loss, such as risk aversion and unemployment stigma.

Next, we may turn to distributional predictions. The key Assumption 2 implies that *more skilled workers are more willing to mismatch*, as they have more to lose from not working. Thus, the stopping belief $\underline{p}$ is decreasing in skills. A larger proportion of skilled workers (21.40% vs. 17.80% of the labor force) prefer to remain on a job that pays below going starting wages, and to search on the job rather than become unemployed. The lower unemployment rate of skilled workers reflects this attitude as well as the Ranking in hiring

performed by firms. In fact, skilled workers quit 50% less frequently to unemployment, and quit to other jobs almost 80% more often. Their job-finding rate is 50% higher, and Ranking hypothesis is confirmed by the ranking of the initial values.

Cross-skill wage inequality is compressed, relative to skill inequality, by the different attitudes to mismatch. The cross-skill differential in either the starting or the average wage is smaller than the $x_H - x_L = 0.2$ skill difference. Within-skill wage dispersion is larger for skilled workers, and this is due entirely to their higher propensity to mismatch, not to technological reasons, as skill x and luck θ are additive in technology. Unskilled workers are matched better on average if employed; hence they lose more of their wage from a displacement and gain slightly less from a direct quit to another job. Finally, the welfare loss from a displacement is a much bigger multiple of average monthly wages for unskilled workers, as they accumulate more learning conditional on being employed.

To recap, the most pronounced inequality across skills is observed in unemployment rates, quits, both to unemployment and to other jobs (in opposite directions), and the relative welfare loss following job destruction. Job-finding rates and most wage measures are instead compressed across skills. More skilled workers choose to protect themselves from unemployment by paying a price in terms of matching quality. Unskilled workers need to get lucky to stay employed, search more from unemployment, hence when employed they must be matched better.

9. Conclusions and Directions for Future Research

This paper introduces a rich yet tractable equilibrium model of the labor market which is consistent both qualitatively and quantitatively with a wide range of stylized facts on worker turnover and wage inequality. The model is used to evaluate the deadweight welfare loss from job destruction, which is found to exceed on average one year worth of wages, and to decline in a worker's general human capital, relative to her average wages. The results warn against the standard practice of using wage changes associated to job changes as measures of specific human capital, as the correspondence between wages and welfare is loose, due to the complicated dynamics of work careers.

The main methodological contribution is a closed-form solution of the stationary wage distribution, which uncovers new comparative statics effects and makes the model a convenient tool for the investigation of labor market dynamics and wage inequality.

The model suggests many different applications and extensions. An important one that is worth mentioning here is the possibility of structural estimation. The decomposition of worker productivity in skill and luck is directly inspired by the standard econometric spec-

ification of wage equations. This model provides a novel structural specification. Consider the wage equation inclusive of skills:

$$w(p|x) = \beta p f(x, \theta_H) + \beta(1-p) f(x, \theta_L) + (1-\beta) b(x) + \beta\lambda(x)\, J(p_0|x)(1 - \psi \mathbb{I}\{\underline{p} \leq p \leq p_0\}).$$

When bringing this equation to the data, x summarizes worker observables (easily extended to a vector) and p is the residual. We detect fixed effects not only through the mean (value of leisure $b(x)$ and value of search $\beta\lambda J(p_0|x)$), but more generally via the whole error distribution $\phi(p|x)$. At the very least, errors are heteroskedastic. The deep parameters of the model may be estimated directly by maximizing the closed-form likelihood function ϕ.

From a theoretical viewpoint, the natural next step is the introduction of firm heterogeneity. The possibility that more productive firms employ more skilled workers appears quite consistent with microeconometric evidence on factor substitution, and gives rise to assortative matching as a mechanism for wage compression alternative to Assumption 2. The interaction between this type of *ex-ante* sorting and *ex-post* matching is central to the growing literature on empirical wage equations with matched employer-employee data, but we know nothing about its equilibrium properties.

A. Appendix

A.1. Proofs of the Propositions

Proposition 1. J is the value function of an optimal Bayesian experimentation problem with flow returns $\bar{y}(p)$ that are linear in beliefs p by the expected utility hypothesis: ergo J is convex in beliefs by a standard improvement argument. It follows that $\underline{p}$ is unique and J is everywhere continuous and almost everywhere *twice* differentiable. So standard HJB Verification Theorems for optimal stopping problems apply, including value matching and smooth pasting at $\underline{p}$ (Shyryaev (1978), 3.8). Since J is convex, non-negative (a firm could always separate to obtain zero), and flat at the lower bound $\underline{p}$ where it is also zero, it must be globally increasing where strictly positive, so $\mathbb{I}\{J(p) \leq J(p_0)\} = \mathbb{I}\{p \leq p_0\}$ and it is optimal to stop on-the-job search at and only at p_0. Continuity of J and J' at p_0 are value matching and smooth pasting conditions for this stopping choice.

By direct verification and using $\mathbb{I}\left\{\underline{p} \leq p \leq p_0\right\} = \mathbb{I}\left\{J(p) \leq J(p_0)\right\}$, the general solution to the HJB Equation (4.6) is:

$$J(p) = \frac{(1-\beta)[\bar{y}(p) - b] - \beta\lambda J(p_0)(1 - \psi\mathbb{I}\left\{\underline{p} \leq p \leq p_0\right\})}{r + \delta + \psi\lambda\mathbb{I}\left\{\underline{p} \leq p \leq p_0\right\}} + c_{iJ}p^{1-\alpha_i}(1-p)^{\alpha_i} + k_{iJ}p^{\alpha_i}(1-p)^{1-\alpha_i}.$$

for $i = \mathbb{I}\left\{p_0 < p \leq 1\right\}$. A useful fact is:

$$J'(p) = \frac{(1-\beta)\left(\theta_H - \theta_L\right)}{r + \delta + \psi\lambda\mathbb{I}\left\{\underline{p} \leq p \leq p_0\right\}} + c_{iJ}p^{-\alpha_i}(1-p)^{\alpha_i - 1}(1 - \alpha_i - p) + k_{iJ}p^{\alpha_i - 1}(1-p)^{-\alpha_i}(\alpha_i - p).$$

Imposing

$$J\left(1\right) = \frac{(1-\beta)(\theta_H - b) - \beta\lambda J(p_0)}{r+\delta} < \infty$$

and $p_0 < 1$ yields $k_{1J} = 0$, else the term $k_{1J}p^{\alpha_1}(1-p)^{1-\alpha_1}$ would explode to $\pm\infty$ as $p \uparrow 1$ by $\alpha_1 > 1$, violating $J(p) \in [0, J\left(1\right)]$.

Next, fix an arbitrary positive value $J(p_0)$ and consider the linear system of three equations in $\{c_{0J}, k_{0J}, c_{1J}\}$, in order:

1. $J(p_0-) = J(p_0)$:

$$\begin{aligned} 0 = {} & \left(r + \delta + \beta\lambda + \psi\lambda(1-\beta)\right) J(p_0) - (r + \delta + \psi\lambda)p_0^{1-\alpha_0}(1-p_0)^{\alpha_0}c_{0J} \\ & -(r + \delta + \psi\lambda)p_0^{\alpha_0}(1-p_0)^{1-\alpha_0}k_{0J} - (1-\beta)[m(p_0) - b]. \end{aligned}$$

2. Value matching for stopping on-the-job search at p_0, $J(p_0) = J(p_0+)$:

$$(r + \delta + \beta\lambda)J(p_0) - (r+\delta)\,p_0^{1-\alpha_1}(1-p_0)^{\alpha_1}c_{1J} = (1-\beta)[m(p_0) - b].$$

3. Smooth pasting for stopping on-the-job search at p_0, $J'(p_0+) = J'(p_0-)$:

$$p_0^{\alpha_0-1}(1-p_0)^{-\alpha_0}(\alpha_0-p_0)k_{0J} - p_0^{-\alpha_1}(1-p_0)^{\alpha_1-1}(1-\alpha_1-p_0)c_{1J}+$$

$$+p_0^{-\alpha_0}(1-p_0)^{\alpha_0-1}(1-\alpha_0-p_0)c_{0J} = \frac{(1-\beta)(\theta_H-\theta_L)}{r+\delta} - \frac{(1-\beta)(\theta_H-\theta_L)}{r+\delta+\psi\lambda}$$

Then plug back the guess $J(p_0)$ and the solution $\{c_{0J}, k_{0J}, c_{1J}\}$ into value matching at separation, $J(\underline{p}) = 0$:

$$-\beta\lambda(1-\psi)J(p_0)+c_{0J}\underline{p}^{1-\alpha_0}(1-\underline{p})^{\alpha_0}(r+\delta+\psi\lambda)+k_{0J}\underline{p}^{\alpha_0}(1-\underline{p})^{1-\alpha_0}(r+\delta+\psi\lambda) = -(1-\beta)[m(\underline{p})-b]$$

and smooth pasting $J'(\underline{p}) = 0$:

$$\frac{(1-\beta)(\theta_H-\theta_L)}{r+\delta+\psi\lambda} + c_{0J}\underline{p}^{-\alpha_0}(1-\underline{p})^{\alpha_0-1}(1-\alpha_0-\underline{p}) + k_{0J}\underline{p}^{\alpha_0-1}(1-\underline{p})^{-\alpha_0}(\alpha_0-\underline{p}) = 0.$$

Finally iterate over values of $J(p_0)$ until the last two equations yield the same value of $\underline{p}$.

Lemma 1. For the ODE solved by $\mathcal{T}(p)$ see Karlin and Taylor (1986, Chapt. 15). The general solution in the claim can be verified directly. The boundary conditions are those stated in the claim plus $\mathcal{T}(1) = 1/\delta$, because a job that is good for sure can be destroyed only exogenously ($p = 1$ is an absorbing belief as $\Sigma(1) = 0$). By the same reasoning as in the previous proof this implies $k_{1\mathcal{T}} = 0$. To see why $\mathcal{T}$ is increasing, notice that $\mathcal{T}(\cdot) \leq 1/\delta$ because job destruction is always a risk, with strict inequality somewhere, so from the general solution we get $c_{1\mathcal{T}} < 0$. Next:

$$\begin{aligned}\mathcal{T}'(p) &= \frac{1}{\delta+\psi\lambda}\left[c_{0\mathcal{T}}p^{-\alpha_0}(1-p)^{\alpha_0-1}(1-\alpha_0-p) + k_{0\mathcal{T}}p^{\alpha_0-1}(1-p)^{-\alpha_0}(\alpha_0-p)\right]\mathbb{I}\left\{\underline{p} \leq p \leq p_0\right\} \\ &\quad +\frac{1}{\delta}c_{1\mathcal{T}}p^{-\alpha_1}(1-p)^{\alpha_1-1}(1-\alpha_1-p)\mathbb{I}\{p_0 < p \leq 1\}\end{aligned}$$

so $\mathcal{T}'(p) > 0$ for $p > p_0$ by $1-\alpha_1-p < 1-\alpha_1 < 0$. By contradiction, suppose $0 \geq \mathcal{T}'(p)$ for some $p \in (\underline{p}, p_0)$. Since $\mathcal{T}'(\underline{p}) \geq 0$ and $\mathcal{T}'(p_0-) = \mathcal{T}'(p_0+) > 0$, by continuity of $\mathcal{T}'$ in $[\underline{p}, p_0]$ and the Mean Value Theorem either there is only one such p, with $\mathcal{T}'(p) = 0$, equivalent to:

$$c_{0\mathcal{T}} = -k_{0\mathcal{T}}\left(\frac{p}{1-p}\right)^{2\alpha_0-1}\frac{\alpha_0-p}{\alpha_0+p-1}, \tag{A.1}$$

in which case the claim obtains, or there are two roots of (A.1). But tedious algebra shows that the RHS of (A.1) is either globally increasing or decreasing in p, according to the sign of $k_{0\mathcal{T}}$, hence (A.1) may have at most one root.

Proposition 3. Let $\nu(p) = p^2(1-p)^2\phi(p)$; we obtain from (6.1) a familiar-looking ODE:

$$p^2(1-p)^2\nu''(p) = \frac{2(\delta + \psi\lambda\mathbb{I}\{p \leq p_0\})}{s^2}\nu(p).$$

The general solution is:

$$\nu(p) = \nu_i(p) = \tilde{c}_{\phi i}p^{1-\gamma_i}(1-p)^{\gamma_i} + \tilde{k}_{\phi i}p^{\gamma_i}(1-p)^{1-\gamma_i}$$

for $i = \mathbb{I}\{p > p_0\}$. Therefore the ergodic density is:

$$\phi_i(p) = \nu(p)p^{-2}(1-p)^{-2} = \tilde{c}_{i\phi}p^{-1-\gamma_i}(1-p)^{\gamma_i-2} + \tilde{k}_{i\phi}p^{\gamma_i-2}(1-p)^{-1-\gamma_i}$$

the sum of Inverted-Beta-1 functions. Integrating $\phi_1(p)$ from between p_0 and 1 yields an exploding second term in ϕ, because $\int_{p_0}^{p}(1-x)^{-1-\gamma_1}dx = (\gamma_1)^{-1}(1-x)^{-\gamma_1}|_{p_0}^{p} \to \infty$ as $p \uparrow 1$ by $\gamma_1 > 0$. Hence we must have $\tilde{k}_{1\phi} = 0$ for $0 < \int_{\underline{p}}^{1}\phi(z)\,dz < \mu < \infty$. By contrast, $\int_{p_0}^{1}(1-x)^{\gamma_1-2}dx < \infty$ by $2-\gamma_1 < 1$, or $\gamma_1 > 1$, so $\tilde{c}_{1\phi}$ can be non-zero. Next let:

$$\xi_1 = \left(\frac{1-\underline{p}}{\underline{p}}\right)^{2\gamma_0-1}$$

By the change of variable $p' = p/(1-p)$:

$$\begin{aligned}
\xi_2 &\equiv \int_{\underline{p}}^{p_0} p^{-1-\gamma_0}(1-p)^{\gamma_0-2}dp = \int_{\frac{\underline{p}}{1-\underline{p}}}^{\frac{p_0}{1-p_0}}\left(\frac{p'}{1+p'}\right)^{-1-\gamma_0}\left(\frac{1}{1+p'}\right)^{\gamma_0-2}\frac{dp'}{(1+p')^2} \\
&= \int_{\frac{\underline{p}}{1-\underline{p}}}^{\frac{p_0}{1-p_0}}(p')^{-1-\gamma_0}(1+p')dp' = \frac{\left(\frac{\underline{p}}{1-\underline{p}}\right)^{-\gamma_0} - \left(\frac{p_0}{1-p_0}\right)^{-\gamma_0}}{\gamma_0} + \frac{\left(\frac{\underline{p}}{1-\underline{p}}\right)^{1-\gamma_0} - \left(\frac{p_0}{1-p_0}\right)^{1-\gamma_0}}{\gamma_0-1}.
\end{aligned}$$

Notice that the mean of ϕ can be obtained analogously. Similarly:

$$\xi_3 \equiv \int_{\underline{p}}^{p_0} p^{\gamma_0-2}(1-p)^{-1-\gamma_0}dp = \frac{\left(\frac{p_0}{1-p_0}\right)^{\gamma_0-1} - \left(\frac{\underline{p}}{1-\underline{p}}\right)^{\gamma_0-1}}{\gamma_0-1} + \frac{\left(\frac{p_0}{1-p_0}\right)^{\gamma_0} - \left(\frac{\underline{p}}{1-\underline{p}}\right)^{\gamma_0}}{\gamma_0}.$$

$$\begin{aligned}
\xi_4 &\equiv \int_{p_0}^{1} p^{-1-\gamma_1}(1-p)^{\gamma_1-2}dp = \frac{1}{\gamma_1}\left(\frac{p_0}{1-p_0}\right)^{-\gamma_1} + \frac{1}{\gamma_1-1}\left(\frac{p_0}{1-p_0}\right)^{1-\gamma_1} \\
\xi_5 &\equiv \frac{s^2}{2}\underline{p}^{-\gamma_0}(1-\underline{p})^{\gamma_0-1}(1-2\gamma_0) \\
\xi_6 &\equiv \frac{s^2}{2}\left[p_0^{-\gamma_0}(1-p)^{\gamma_0-1}(3p_0-1-\gamma_0) - \xi_1 p_0^{\gamma_0-1}(1-p)^{-\gamma_0}(3p_0-2+\gamma_0)\right].
\end{aligned}$$

The boundary conditions then read:

1) no time spent at the separation boundary $\phi(\underline{p}+) = 0$:

$$\tilde{k}_{0\phi} = -\xi_1 \tilde{c}_{0\phi}. \tag{A.2}$$

which implies $\tilde{c}_{0\phi}\tilde{k}_{0\phi} < 0$. Using (A.2) replace $\tilde{k}_{0\phi}$ out. Thus:

$$\int_{\underline{p}}^{1} \phi(p)dp = \tilde{c}_{0\phi}\left(\xi_2 - \xi_1\xi_3\right) + \tilde{c}_{1\phi}\xi_4$$

$$\begin{aligned}
\Sigma(\underline{p})\phi'(\underline{p}+) &= \tilde{c}_{0\phi}\frac{s^2\underline{p}^2(1-\underline{p})^2}{2}\left[\underline{p}^{-2-\gamma_0}(1-\underline{p})^{\gamma_0-3}(3\underline{p}-1-\gamma_0) + \tilde{k}_{0\phi}\underline{p}^{\gamma_0-3}(1-\underline{p})^{-2-\gamma_0}(3\underline{p}+\gamma_0-2)\right] \\
&= \tilde{c}_{0\phi}\frac{s^2}{2}\underline{p}^{-\gamma_0}(1-\underline{p})^{\gamma_0-1}\left[3\underline{p}-1-\gamma_0-(3\underline{p}+\gamma_0-2)\right] = \tilde{c}_{0\phi}\xi_5.
\end{aligned}$$

2) Balance of flows in and out of employment:

$$\tilde{c}_{0\phi}\xi_6 = \psi\lambda\tilde{c}_{0\phi}\left(\xi_2 - \xi_1\xi_3\right) + \delta[\tilde{c}_{0\phi}\left(\xi_2 - \xi_1\xi_3\right) + \tilde{c}_{1\phi}\xi_4] + \tilde{c}_{0\phi}\xi_5.$$

3) Balance of flows in and out of unemployment.

$$\lambda[\mu - \tilde{c}_{0\phi}\left(\xi_2 - \xi_1\xi_3\right) - \tilde{c}_{1\phi}\xi_4] = \delta[\tilde{c}_{0\phi}\left(\xi_2 - \xi_1\xi_3\right) + \tilde{c}_{1\phi}\xi_4] + \tilde{c}_{0\phi}\xi_5.$$

To obtain the expression in claim, let $c_{0\phi} = -\tilde{c}_{0\phi}$ and $c_{1\phi} = \tilde{c}_{1\phi}$. So $\Xi\binom{c_{0\phi}}{c_{1\phi}} = \binom{\lambda\mu}{0}$, where:

$$\Xi \equiv \begin{pmatrix} -\left(\xi_2 - \xi_1\xi_3\right)(\lambda+\delta) - \xi_5, & (\lambda+\delta)\,\xi_4 \\ \xi_6 - (\psi\lambda+\delta)(\xi_2 - \xi_1\xi_3) - \xi_5, & \delta\xi_4 \end{pmatrix}. \tag{A.3}$$

The solution is:

$$\begin{aligned}
c_{0\phi} &= \mu\frac{\lambda\delta}{\left(\xi_2 - \xi_1\xi_3\right)(\lambda+\delta)\,\psi\lambda + \lambda\xi_5 - (\lambda+\delta)\,\xi_6} \\
c_{1\phi} &= c_{0\phi}\frac{\xi_5 + (\psi\lambda+\delta)(\xi_2 - \xi_1\xi_3) - \xi_6}{\delta\xi_4}
\end{aligned}$$

Finally, a substantial amount of algebra (omitted) shows that the boundary conditions also imply that the density is continuous at p_0: $\phi(p_0-) = \phi(p_0+)$, therefore the inflow at p_0 creates a kink but not a jump in the density. This also implies $c_{0\phi}, c_{1\phi} > 0$ by a simple contradiction argument using $c_{0\phi}k_{0\phi} > 0$ found earlier.

A.2. Data Sources and Quantitative Evaluation of the Model

We draw monthly worker flows between Employment (E), Unemployment (U) and Not in the labor force (N) from Blanchard and Diamond (1990) [BD90], based on the 1968-1986 Annual Demographic March files of the Current Population Survey. Two corrections are necessary to make the model comparable with BD90's data. First, the distinction between U and N is not made here while it is crucial to BD90. They document that the UE flow (from U to E) is approximately equal to the NE flow, and that a subset of N not much smaller than total unemployment U, say $\hat{\mathrm{N}}$, declare to "want a job", Hence $\hat{\mathrm{N}}$ is added to formal unemployment U to form a "Jobless" category, calibrated at 9.5% of the labor force, also extended to include $\hat{\mathrm{N}}$. This group of people $\hat{\mathrm{N}}$ is also assumed to produce the entire NE flow reported by BD90, which is added up to the EU flow to obtain a total flow in and out of joblessness of 2.1% of the labor force. A caveat is that the evidence of BD90, based on the CPS and the LRD (manufacturing only), does not describe a steady state, because employment rose over the period. The LRD allows BD90 to distinguish also between layoffs, about 1.3% of employment and therefore (by extrapolation to the economy as a whole) about 1.2% of the labor force, while total quits are roughly 1.8%, giving an overall 3% average separation rate. Of quits, direct and indirect evidence in BD90 points to an educated guess of a 50-50 split between quits to unemployment and direct quits to other jobs.

These facts are supplemented by the more recent, reliable and detailed evidence presented by Fallick and Fleischman (2001) [FF01]. FF01 exploit the "dependent interview" techniques adopted for the CPS since 1994, and document worker flows in 1994-2000. They find a monthly flow of workers from employer to employer (EE flow) of the order of 2.7% of employment. This number overstates direct job-to-job quits because some movements entail an intervening spell of unemployment shorter than a month (hence unobserved). Since about 40% of new unemployed usually find a job in the first month, this cuts the proportion of job-to-job quitters to 1.6%. As FF01 point out, some EE movements occur when a multiple job holder separates from his/her main employer and keeps the other job. These should not be counted as quits. Finally, 1994-2000 were years of expansion, when direct quits are know to exceed the overall time average. Since the measurement of EE flows in FF01 is direct, while those in BD90 are just an indirect and noisy imputation, we raise the target quit rate to 1.1% of the labor force. The figures for quits to unemployment and layoffs in FF01 are comparable to BD90's, after aggregating U and $\hat{\mathrm{N}}$ into the relevant jobless pool.

FF01 also find that 5% of employees report searching on the job at every point in time, in line with evidence for the UK from Pissarides and Wadsworth (1994). In the model we interpret this proportion as representing the "effective employed job searchers", namely the fraction ψ of potential employed job searchers (of mismatched workers) who have an

actual opportunity to search at each point in time. Once the opportunity is there, these job searchers face the same job-finding rate λ as the unemployed of equal skills.

Evidence on the shape of the wage distribution and the (mean-median)/stdev statistics is obtained from the March CPS files; for the latter statistics, we took simple averages for the period 1985-1995.

The average short-run wage loss from job displacement is drawn from Stevens (1997). She finds that the wage loss at impact, on which we focus here, is roughly the same whether or not the job loss involved a plant closure, an event that necessarily implies a job destruction.

The average recruiting cost in terms of months of wages in the model is computed as the vacancy posting cost multiplied by the expected duration of a vacancy, divided by the average wage. The corresponding empirical observation is a typical number found repeatedly in the literature (*e.g.*. Abowd and Kramarz 1997), with regards to recruiting and training costs.

The empirical hazard rate of separation reported in Figure A.1, dashed line, is computed as the rate of decrement of the empirical tenure frequency distribution. The tenure data are from Diebold, Neumark, and Polsky (1997) [DNP97], drawn from the 1987 CPS Tenure Supplement and adjusted for "heaping". This methodology to compute the hazard rate is valid under the assumption that the tenure distribution is stable; see DNP97 for the issues involved in this computation and for corroborating evidence of such stability. The model hazard rate is computed in the theoretically correct manner from the distribution of completed tenures generated by the simulation of the two-skill model: distribution divided by survival function. Rather than fitting the same parametric duration-dependence model to the data and to the model and comparing the two estimates, we just report the raw numbers and let the reader eyeball the congruence between the two series.

References

ABOWD, JOHN AND FRANCIS KRAMARZ, 1997, "The Cost of Hiring and Separations", NBER WP 6110.

ACEMOGLU, DARON, 1998, "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality", *Quarterly Journal of Economics*, 113(4), 1055-89.

ALTONJI, JOSEPH AND CHARLES PIERRET, 2001, "Employer Learning and Statistical Discrimination", *Quarterly Journal of Economics*, 116(1), 313-350.

BELZIL, CHRISTIAN, 1996, "Relative Efficiencies and Comparative Advantages in Job Search", *Journal of Labor Economics*, 14(1), 154-173.

BLANCHARD, OLIVIER AND PETER DIAMOND, 1990, "The Cyclical Behavior of the Gross Flows of U.S. Workers", *Brookings Papers on Economic Activity*, 2, 85-155. [BD90]

BLAU, DAVID AND PHILIP ROBINS, 1990, "Job Search Outcomes for the Employed and the Unemployed", *Journal of Political Economy* 98(3), 637-655.

DIEBOLD, FRANCIS, DAVID NEUMARK, AND DANIEL POLSKY 1997, "Job Stability in the United States", *Journal of Labor Economics*, 15(2), 206-233.

FALLICK, BRUCE AND CHARLES FLEISCHMAN, 2001, "The Importance of Employer-to-Employer Flows in the US Labor Market", mimeo, FED Board of Governors.

FARBER, HENRY, 1994, "The Analysis of Interfirm Worker Mobility", *Journal of Labor Economics*, 12(4), 554–593.

FELLER, WILLIAM, 1954, "Diffusion Processes in One Dimension", *Transactions of the American Mathematical Society*, 1-31.

FELLI, LEONARDO AND CHRISTOPHER HARRIS, 1996, "Learning, Wage Dynamics, and Firm-Specific Human Capital", *Journal of Political Economy*, 104(4), 838-868.

FLINN, CHRISTOPHER, 1986, "Wages and Job Mobility of Young Workers", *Journal of Political Economy*, 94(3) Part I, S88-S110.

HECKMAN, JAMES AND GUILHERME SEDLACEK, 1985, "Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market", *Journal of Political Economy*, 93(6), 1077-1125.

JACOBSON, LOUIS, ROBERT LALONDE AND DANIEL SULLIVAN, 1993, "Earnings Losses of Displaced Workers", *American Economic Review*, 83(4), 685-709.

JOVANOVIC, BOYAN, 1979, "Job Matching and the Theory of Turnover", *Journal of Political Economy*, 87(5) Part I, 972-990.

JOVANOVIC, BOYAN, 1984, "Matching, Turnover and Unemployment", *Journal of Political Economy*, 92(1), 108-122.

LANE, JULIA AND MICHAEL PARKIN, 1998, "Turnover in an Accounting Firm", *Journal of Labor Economics*, 16(4), 702-717.

KARLIN, SAMUEL AND HOWARD TAYLOR, 1981, *A Second Course in Stochastic Processes*, Academic Press.

KRAUSE, MICHAEL, 1999, "Inter-Industry Wage Differentials and the Creation and Destruction of Jobs", Ph.D. dissertation, Yale University, Department of Economics.

LIPTSER, ROBERT AND ALBERT SHYRYAEV, 1977, *Statistics of Random Processes*, 2nd volume, Springer-Verlag.

MORTENSEN, DALE AND CHRISTOPHER PISSARIDES, 1994, "Job Creation and Job Destruction in the Theory of Unemployment", *Review of Economic Studies*; 61(3), 397-415.

MORTENSEN, DALE AND CHRISTOPHER PISSARIDES, 1999, "Unemployment Responses to 'Skill-Biased' Technology Shocks: the Role of Labour Market Policy", *Economic Journal*; 109(455), 242-265.

MORTENSEN, DALE, 1998, "Equilibrium Unemployment with Wage Posting: Burdett-Mortensen Meet Pissarides", mimeo Northwestern University.

MOSCARINI, GIUSEPPE, 1996, "Worker Heterogeneity and Job Search in the Flow Approach to Labor Markets: a Theoretical Analysis", unpublished Ph.D. dissertation, MIT.

MOSCARINI, GIUSEPPE, 2001, "Excess Worker Reallocation". *Review of Economic Studies,* July, 68(3), 593-612.

MOSCARINI, GIUSEPPE AND FRANCIS VELLA, 2000, "Worker Mobility and Aggregate Labor Reallocation: Evidence from the NLSY79". Mimeo Yale University and NYU.

NAGYPAL, EVA, 2000, "Learning-By-Doing Versus Learning About Match Quality: Can We Tell Them Apart?". Mimeo Stanford University.

NEAL, DEREK, 1998, "The Link Between Ability and Specialization", *The Journal of Human Resources,* Vol. XXXIII, Winter, 174–200.

OSBERG, LARS, 1993, "Fishing in Different Pools: Job-Search Strategies and Job-Finding Success in Canada in the Early 1980s", *Journal of Labor Economics*, 11(2), 348-386.

PETRONGOLO, BARBARA AND CHRISTOPHER PISSARIDES, 2001, "Looking into the Black Box: A Survey of the Matching Function", *J. of Economic Literature,* 39(2), 390-431.

PISSARIDES, CHRISTOPHER, 1994, "Search Unemployment with On-the-Job Search", *Review of Economic Studies,* 61(3), 457-475.

PISSARIDES, CHRISTOPHER AND JONATHAN WADSWORTH, 1994, "On-the-job Search: Some Empirical Evidence from Britain", *European Economic Review* 38(2), 385-401.

POSTEL-VINAY, FABIEN AND JEAN-MARC ROBIN, 2000, "Wage Dispersion with Heterogeneous Firm Technologies and Worker Abilities: An Equilibrium Search Model for Matched Employer-Employee Data". Forthcoming in *Econometrica.*

PRIES, MICHAEL, 2001, "Persistence of Employment Fluctuations: A Model of Recurring Job Loss", mimeo University of Maryland, Department of Economics

RUHM, CHRISTOPHER, 1991, "Are Workers Permanently Scarred by Job Displacement?", *American Economic Review* 81(1), 319-324.

SHYRYAEV, ALBERT, 1978, *Optimal Stopping Rules*, Springer-Verlag.

STEVENS, ANN, 1997, "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses", *Journal of Labor Economics*, 15(1), 165-188.

TOPEL, ROBERT, 1991, "Specific Capital, Mobility and Wages: Wages Rise with Job Seniority", *Journal of Political Economy*, 99(1), 145-176.

Table 1. Parameterization of the Model

Technology	$f(x,\theta) = x + \theta$
NORMALIZATIONS	
Low skill x_L	0.4
High skill x_H	0.6
Good match outcome θ_H	0.5
Bad match outcome θ_L	-0.5
Labor force	100
CONSTRAINED (ESTIMATED) PARAMETERS	
Rate of time preference (monthly)	0.004
Proportion of low skills	0.7
Matching function v – elasticity η	0.3
FREE PARAMETERS	
Exogenous job destruction rate δ (monthly)	0.0129
Relative efficiency of employed search ψ	0.24
Output noise σ	8.5
Value of leisure b	0.32
Prior belief p_0	0.5
Worker bargaining share β	0.4
Matching function scale m_0	0.23
Flow recruiting cost κ	0.85

Table 2: Predictions of the Model and "(Over-)Identifying Restrictions"

	MODEL			DATA
	Low skill $x_L = 0.4$	High skill $x_H = 0.6$	Economy	
Stopping belief $\underline{p}$	0.39	0.36		
WORKER STOCKS				
Jobless	7.91	1.77	9.68	
Employed below starting wage $w(p_0)$	12.46	6.42	18.88	
Employed above starting wage $w(p_0)$	49.63	21.81	71.44	
Total	70	30	100	
WORKER STOCKS (% OF LF)				
Jobless	11.3	5.90	9.68	9.5
Employed below starting wage $w(p_0)$	17.80	21.40	18.88	
Employed above starting wage $w(p_0)$	70.90	72.70	71.44	
"Effective" on-the-job searchers (% of empl.)	3.36	1.65	5.01	5
MONTHLY WORKER FLOWS (% OF LF)				
Quits to joblessness	1.09	0.53	0.91	0.9
Exogenous displacements	1.15	1.22	1.17	1.2
(Total inflow into joblessness)	(2.24	1.75	2.08	2.1)
Job-to-job quits	0.86	1.54	1.07	1.1
(Total separations)	(3.10	3.29	3.15	3.2)
Hires from joblessness	2.24	1.75	2.08	2.1
Monthly job-finding rate $\lambda(x)$	0.2	0.3	0.22	
MONTHLY WAGES				
Skewed distribution, Paretian right tail	yes	yes	yes	yes
Starting wage $w(p_0)$	0.41	0.56	0.46	
Average wage $\mathbb{E}[w(p)\|p \geq \underline{p}]$	0.478	0.660	0.534	
Median wage	0.463	0.646	0.519	
Standard deviation of wages	0.06	0.07	0.068	
(Aver. wage − Median wage)/Stdev(wages)			0.22	0.19
% wage loss from displacement	14.6	13.9	14.3	13.8
% wage gain from job-to-job quit	3.5	3.6	3.5	
VALUES AND WELFARE (AS MULTIPLES OF AVERAGE WAGES)				
Average recruiting cost			2.47	2.5
Total surplus from a new job	3.36	3.76	3.47	
Average match surplus	15.33	12.75	14.50	
Average returns from learning	12	9		

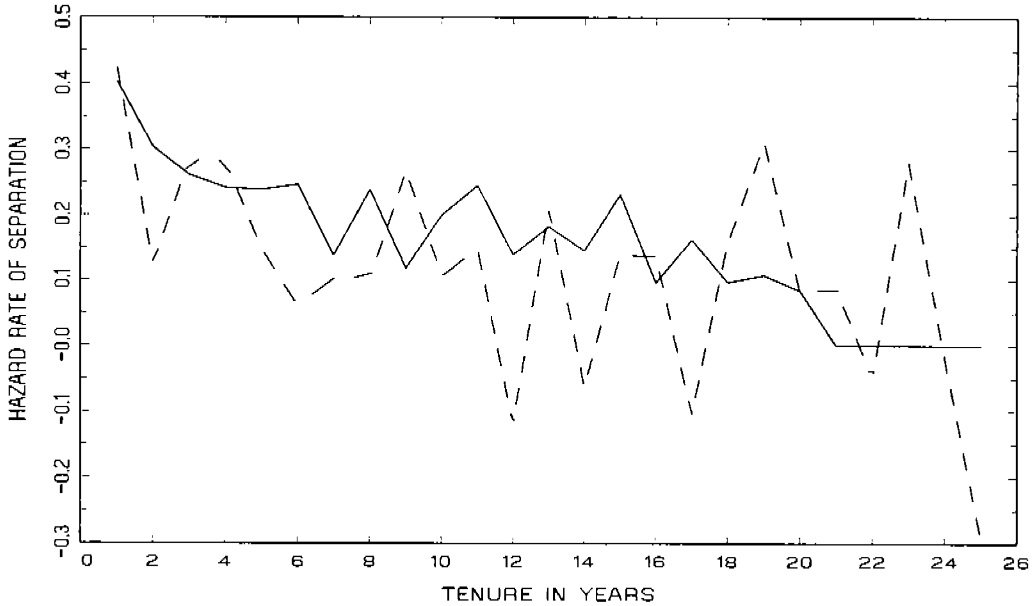

Figure A.1: Hazard rate of separation: model (solid line) and data (dashed line).

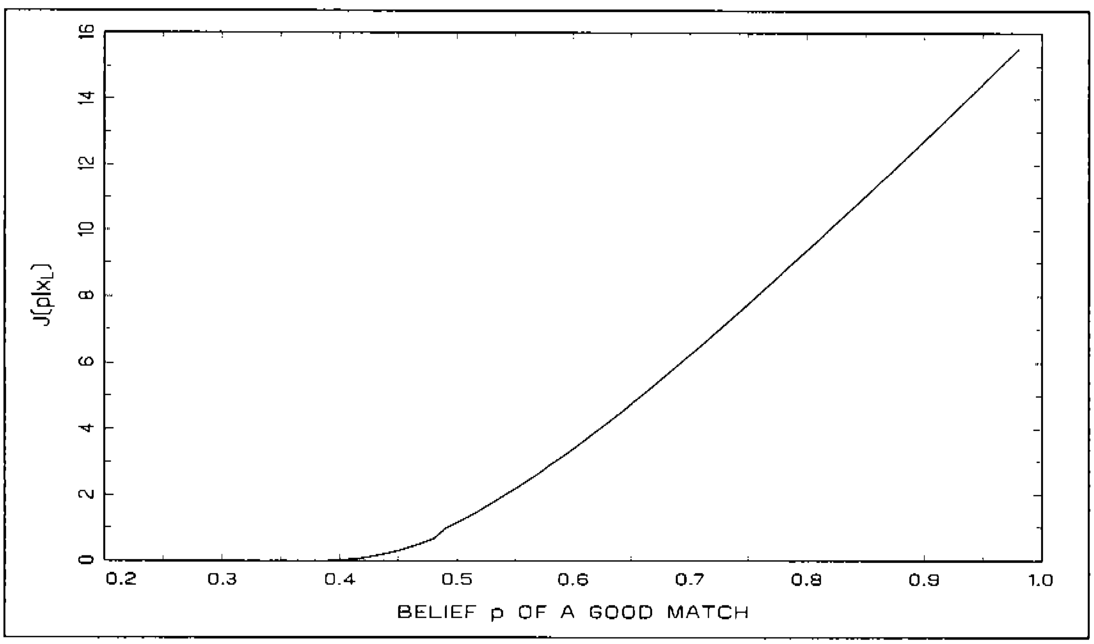

Figure A.2: The value function of a firm employing a low-skill worker.

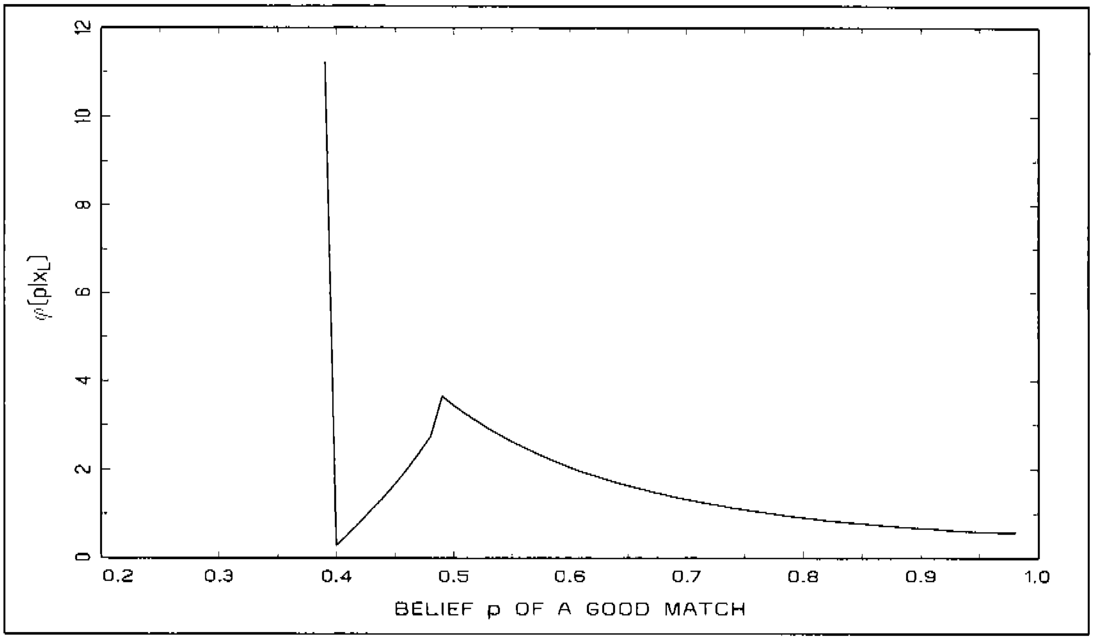

Figure A.3: The ergodic and stationary density of beliefs on match quality for low-skill workers. The atom at the lower bound is the stationary measure of low-skill jobless workers.

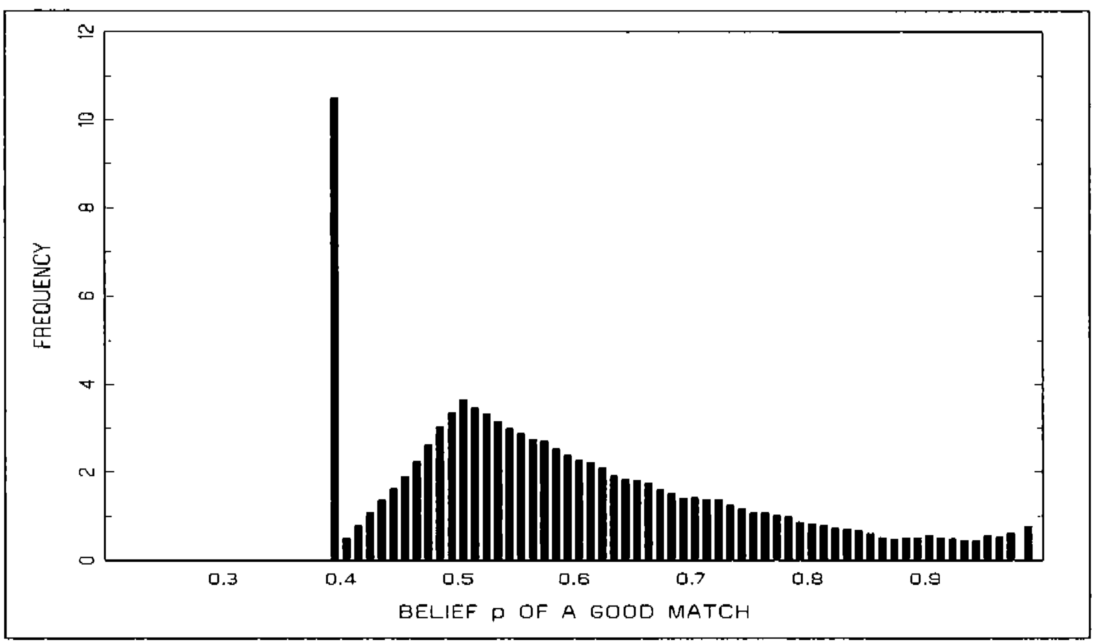

Figure A.4: Frequency distribution of simulated belief process for low-skill workers.