A Theory of Dynamic Selection in the Labor Market*

Pooya Molavi†

First draft: March 2016
This draft: December 2018

Abstract

I propose an equilibrium search and matching model with permanent worker heterogeneity, asymmetric information, and endogenous separations and study the dynamics of adverse selection in the labor market. The interaction between asymmetric information and endogenous separations leads to a cyclical adverse selection problem that has testable predictions both for the aggregate variables and for individual workers' outcomes. First, a deterioration in the distribution of ability in the pool of the unemployed leads firms to raise their hiring standards, thus resulting in shifting out of the Beveridge curve. Second, if the separation rate is log-supermodular (log-submodular) in productivity and ability, the pool of the unemployed becomes more (less) adversely selected in downturns. Third, firms rationally discriminate against the long-term unemployed by demanding more unequivocally positive signals of their ability before hiring them. Fourth, this scarring effect is more (less) severe for lower-ability workers and after deeper recessions if the separation rate is log-supermodular (log-submodular). I conclude by providing conditions on the fundamentals of the economy that lead to log-supermodular and log-submodular separation rates.

*An earlier version of this paper was circulated under the title "Adverse Selection in the Labor Market and Slow Recoveries." I am grateful to Daron Acemoglu, Marios Angeletos, David Autor, David Berger, Peter Diamond, Chishio Furukawa, Arda Gitmez, Guido Menzio, Emi Nakamura, Alan Olivi, Christina Patterson, Will Rafey, Alp Simsek, Juuso Toikka, Olivier Wang, Iván Werning, Nathan Zorzi, and participants in MIT Macro Lunch and MIT Third-Year Lunch for helpful comments that significantly improved the paper.

†Department of Economics, Massachusetts Institute of Technology.

1 Introduction

Although the Great Recession officially ended in the June of 2009, the U.S. unemployment rate continued to increase until hitting a high of 10 percent in the October of 2009. It took five more years for employment to recover to the levels economists consider normal. The sluggish recovery in the labor market combined with a fast recovery in other indicators of economic activity such as productivity and corporate profits has lead economists to refer to the U.S. experience as a “jobless recovery.” What is more, aggregate statistics obscure the different experiences of individual workers. Low-skilled workers bore the brunt of the slack labor market. While the unemployment rate for workers with a Bachelor’s degree remained below 5 percent throughout the recession and the recovery that ensued, the unemployment rate among those with less than a high-school degree remained above 8 percent well into the recovery. Education is not the only dimension along which workers experienced heterogeneous outcomes. Kroft, Lange, and Notowidigdo (2013) provide experimental evidence for “unemployment scarring”: the long-term unemployed experience diminished employment prospects regardless of their educational attainment.

This paper develops a theory of dynamic selection in the labor market that can provide an explanation for jobless recoveries and unemployment scarring. I propose a search and matching model featuring persistent differences among workers in their ability, asymmetric information about worker ability, and layoff decisions by firms. Workers are heterogeneous in their ability. A worker’s ability is perfectly observable to the worker and any firm that has previously hired the worker, but other firms only observe noisy signals of the worker’s ability. Firms are affected both by an aggregate productivity shock, which follows a diffusion process, and by idiosyncratic cost shocks, which arrive at a Poisson rate independently of other variables. Sufficiently adverse cost shocks can lead firms to cease production and lay off their workforce. Although the shocks are symmetrical across firms, they affect firms differentially depending on the quality of their labor force. Firms that employ higher-ability workers are more likely to withstand the adverse cost shocks without going under. Furthermore this difference in the resilience of firms can vary with variations in productivity. The model delivers rich dynamics in the composition of the unemployed pool and observable statistics of the labor market, which are laid out in a number of formal propositions.

In Section 4, I provide a characterization of the way the composition of the unemployed pool affects the firms’ hiring decisions through its affect on both the value of a vacancy and the firms’ belief about the ability of a typical job applicant. To quantify the changes in the composition of the pool, I use a strengthening of the notion of first-order stochastic dominance called monotone likelihood ratio property (MLRP). I show that a deterioration in the quality of the pool in the sense of MLRP leads firms to post fewer vacancies and to demand a more positive signal of a worker’s ability before hiring the worker. When the pool becomes more adversely selected, these two effects conspire to depress the job-finding rates of all workers irrespective of their ability.

In Section 5, I characterize the response of the composition of the pool to changes in the aggregate productivity. A decline in productivity can lead to a deterioration or an improvement of the pool in the sense of MLRP depending on auxiliary functional-form assumptions on the fundamentals of the economy. Sufficient conditions for a compositional deterioration of the pool with a decline in a productivity are, first, that the signals of ability observed by the firms are not too precise, and second, that a worker's separation rate is log-supermodular in productivity and the worker's ability. I argue in Section 7 of the paper that the latter assumption is indeed satisfied under the most natural specifications of the wage rate: both a fully rigid real wage rate and a wage that is affine in the worker's output result in a log-supermodular separation rate. Therefore in the most plausible specifications of the model the composition of the unemployed pool is procyclical when measured in the sense of MLRP.

These results rationalize a number of well-documented empirical facts. The model provides an explanation for upskilling—the increase in recessions in job requirements—as arising from cyclical variations in the composition of the pool. The compositional deterioration in recessions leads firms to demand more unequivocally positive signals from job applicants. The model also generates a shifting out of the Beveridge curve in recessions. The increase in hiring requirements in recessions leads to lower job-filling rates relative to what would prevail in an economy with an acyclical composition of the pool and the same unemployment rate as the model economy, thus causing an outward shift of the Beveridge curve. The shift in the Beveridge curve in the aftermath of the Great Recession is evident in Figure 1.

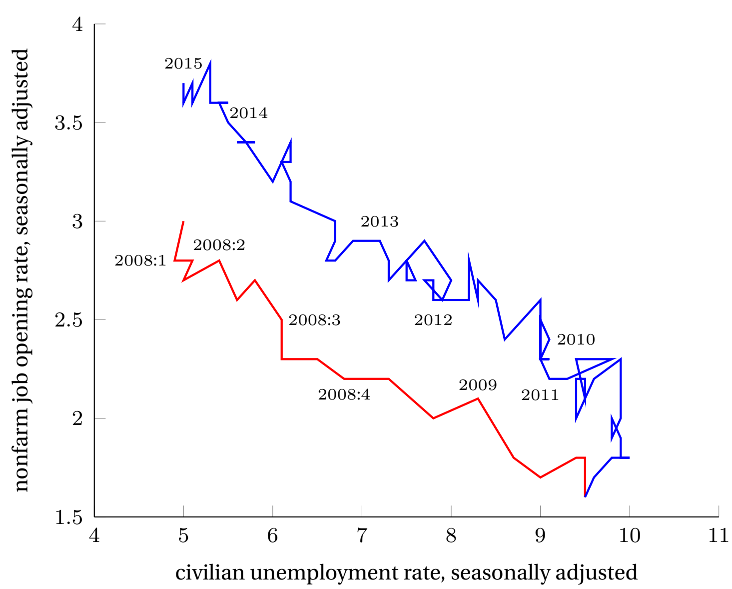


FIGURE 1. The US Beveridge curve: December 2007–November 2015
Source: Bureau of Labor Statistics.

The model can also explain the consequences of job loss for individual workers. In Section 6, I extend the model by assuming that firms can observe their job applicants' time of entry into unemployment. This extension of the model generates unemployment scarring through the following mechanism. Higher-ability workers have higher hazard rates of exit from unemployment. Therefore the distribution of ability in a cohort of workers who have entered unemployment at the same time declines over time in the sense of MLRP. Knowing about this compositional deterioration over time, firms impose more stringent hiring requirements on the workers who have experienced longer unemployment spells, thus depressing their job-finding prospects. This mechanism is reminiscent of the one used in the statistical discrimination literature à la Coate and Loury (1993) and Rosén (1997); it provides an explanation for unemployment scarring that is complementary to the hypothesis of human capital loss. The model can also be seen as providing a microfoundation for the assumption of ranking of workers by firms based on unemployment duration that was proposed by Blanchard and Diamond (1994).

In Section 6, I use the model to conduct comparative statics analyses of scarring. I show that scarring is more severe for lower-ability workers and for workers who are separated in recessions. Lower-ability workers are more likely to generate signals that barely meet the firms' hiring requirements, so the increase in requirements with unemployment duration afflicts them more. Scarring is also more severe in downturns, given the assumption that a worker's separation rate is log-supermodular in productivity and the worker's ability. Under this condition, the distribution of ability in cohorts that become unemployed in recessions is lower in the sense of MLRP. Therefore, firms impose more requirements on job candidates that have entered unemployment in downturns.

In Section 8, I discuss how the predictions of the model stack up against empirical evidence and discuss two additional extensions of the model. Mueller (2015) finds that recessions lead to improvements in the composition of ability in both the unemployed and employed populations, pointing to an increase in sorting in recessions. I show that, once the firing of workers is taken into account, my model can generate an *increase* in recessions in the average unemployed worker's ability—consistent with Mueller's finding—and a *deterioration* in the ability in the sense that is relevant for the firms' hiring decisions.

I also show that the predictions of the model are broadly consistent with stylized facts on the cyclicality of quits, layoffs, and separations. The layoff rate and layoff-separation ratio are both countercyclical in the model, in agreement with the findings of Davis and Haltiwanger (1990, 1992) and Blanchard and Diamond (1990). Quits are acyclical in the model by construction, whereas they are procyclical in data. However, as I argue in subsection 8.3, the model can easily be extended to accommodate procyclical quits without affecting any of its other predictions. The modified model can also generate acyclical average separation rates in concordance with the new view of the labor market championed by Hall (2005b) and Shimer (2005). This extension illustrates that the cyclicality of the separation rate plays no role in the mechanism introduced in this paper. What is important is rather the cyclicality of the slope of the separation rate with respect to ability.

Related literature. In addition to the studies mentioned above, the paper connects and contributes to several strands of research in labor economics and macroeconomics. The paper belongs to the general class of models that view recessions as times of reorganization. Important contributions to this literature include the works of Hall (1991), Caballero and Hammour (1994, 1996), Koenders and Rogerson (2005), Philippon (2006), Pries (2008), Jaimovich and Siu (2012), Berger (2015), Jackson and Tebaldi (2015), and Restrepo (2015). The great majority of these contributions are in the context of models with ex ante identical workers in which the composition of the unemployed pool is invariant throughout the course of a cycle; notable exceptions are Jaimovich and Siu (2012), Restrepo (2015), and Pries (2008). In Jaimovich and Siu (2012) and Restrepo (2015) workers with different skills enter unemployment differentially due to an interaction of structural and cyclical factors—secular technological progress renders low-skilled jobs obsolete and this process is accelerated in recessions. In my model, in contrast, compositional changes are purely due to cyclical variations in labor productivity.

Perhaps most closely related to this paper is the work by Pries (2008). He considers an extension of the standard search and matching model with low-productivity and high-productivity workers. His paper however is different from the current study along two important dimensions. First, in Pries (2008) firms simply post a vacancy and hire any and all workers who apply, whereas in my model firms additionally observe a signal of the worker's ability and decide whether to hire the applicant. This feature of the model is crucial for several of its predictions. Cyclical variations in the threshold used by the firms in their hiring decisions enables the model to explain phenomena such as upskilling, a procyclical recruiting intensity, movements of the Beveridge curve, and unemployment scarring. Second, while Pries's exercise is quantitative, the conclusions of the current paper are all analytical. The use of analytical tools enables me to discuss the countervailing economic forces that are in play and to establish general principles that underpin the mechanism of this paper. Using analytical techniques also allows me to uncover insights that are absent in a stylized model with two types. For instance, as mentioned before, in my model you can have a simultaneous increase in the average ability in the pool and a deterioration of the ability in the sense that is relevant to the firms' hiring decisions. This is not possible in a model with two types.

This paper also contributes to the literature that studies the consequences of job loss and unemployment. In a seminal contribution, Jacobson, LaLonde, and Sullivan (1993) show that job loss leads to a persistent decline in future earnings. This finding has been replicated more recently in a number of papers, among them contributions by Sullivan and von Wachter (2009) and Davis and von Wachter (2011). The consequences of job loss are not limited to earning losses. Kroft, Lange, and Notowidigdo (2013) and Jarosch (2014) find that unemployment also leads to a decline in future employment. The loss in earnings and employment resulting from unemployment is termed unemployment scarring. My model provides an explanation of scarring as arising from statistical discrimination by the firms. Jarosch and Pilossoph (2015) also propose a model that generates unemployment scarring due to statistical discrimination. But in contrast to this paper, they do not explicitly model the cyclical variations in the composition of the unemployed pool. A closely related literature studies the effects of entry into unemployment in recessions. Consistent with the prediction of this paper, Davis, Faberman, and Haltiwanger (2012) find that in the U.S. labor market job loss in recessions is associated with more severe scarring than job less in expansions. Kahn (2010) and Oreopoulos, von Wachter, and Heisz (2012) find a similar negative effect on workers of entering the labor force in a recession.

This paper also belongs to a body of work on asymmetric information in the labor market that goes back to Greenwald (1986), Lockwood (1991), and Gibbons and Katz (1991) and includes important contributions by Acemoglu and Pischke (1998), Guerrieri (2007, 2008), Nakamura (2008), Moen and Rosén (2011), Kahn (2013), and Kahn and Lange (2014). But to the best of my knowledge, this is the first paper in this literature that uses an equilibrium search model to study the dynamics of adverse selection in the labor market. Finally, the model studied in this paper belongs to the class of search and matching models with heterogeneous workers such as the models proposed in a number of important contributions by Acemoglu (1996, 1999, 2001). But the current paper focuses on a very different set of questions than those studied by Acemoglu. He studies the effects of directed technological change and changes in the supply of skills on the long-run composition of jobs in the economy, whereas this paper is concerned with business cycle fluctuations in the compositions of the unemployment pool.

2 Model

I consider a labor search model in continuous time. The economy is populated by a unit measure of workers and a large measure of firms. Workers and firms are both risk-neutral and discount the future at rate ρ.

Workers are heterogeneous in their effective labor endowment, or ability, y. Ability is distributed in the population according to p.d.f. g supported on an interval of positive reals $Y = [\underline{y}, \overline{y}]$. Workers of ability y supply y units of labor inelastically.

There is a single final good, which can be used for consumption and investment. Production takes place inside single-worker firms using labor and capital. A worker needs to work on exactly one machine in order to be productive. Machines cost k units of the final good each. A worker of ability y working on a machine produces output at rate yA. The labor productivity, A, is common to all firms and follows a geometric Brownian motion with drift μA, volatility $\sqrt{2\mu}A$, and reflecting barriers at some arbitrary $\underline{A}$ and $\overline{A}$.[1] I assume that $\rho > \mu$ to ensure that the expected present discounted value of A would remain finite absent the reflecting barriers.

The economy is subject to matching frictions. Let $u(y)$ denote the unemployment rate among workers of ability y. When fraction $\mathbb{E}u$ of workers are unemployed and there are v vacancies in the economy, matches are created at rate $M(\mathbb{E}u, v)$.[2] The matching function, M, is continuously differentiable, strictly increasing in both arguments, satisfies the Inada conditions, and exhibits constant returns to scale. Given the assumptions on the matching function, each vacancy is contacted by unemployed workers at rate $q(\theta) = M\left(\theta^{-1}, 1\right)$ and each worker encounters vacancies at rate $p(\theta) = M(1, \theta) = m(1/q(\theta))$, where $\theta = v/\mathbb{E}u$ denotes the labor market tightness and m is a strictly increasing function.

Firms have imperfect information about the ability of individual workers when they make hiring decisions. When a worker is matched to a firm the firm observes a noisy signal, $\omega \in \mathbb{R}$, of the worker's ability, y, and decides whether to hire the worker. Workers who are not hired return to the unemployed pool. I assume that ω is i.i.d. conditional on the ability and is drawn from a differentiable p.d.f. $\ell(\omega|y)$, which satisfies the strict monotone likelihood ratio property (MLRP)—so higher signals are better news about the worker's ability.[3] Firms learn the abilities of their employees immediately after hiring them. The asymmetric information between past and future employers of a worker about his ability is the key ingredient of the model that enables it to explain the phenomena discussed in the introduction.

Matches are destroyed due to quits and layoffs.[4] Workers have to quit their jobs if they are hit by a shock that arrives at Poisson rate ς.[5] They may also be laid off when the firms that hire them experience an adverse idiosyncratic shock. I model this firm-specific shock as a machine breakdown that occurs at Poisson rate δ independently of other random variables. Once a firm experiences a

[1]The relationship between the drift and the volatility of the geometric Brownian motion guarantees that A has no time trend.

[2]Note that the unemployment rate in this economy is equal to $\mathbb{E}u$, the expected value of $u : Y \to [0, 1]$ with respect to g, where $\mathbb{E}$ denotes the expectation operator corresponding to p.d.f. g.

[3]Likelihood function $\ell(\omega|y)$ satisfies the strict monotone likelihood ratio property if $\frac{\ell(\omega|y)}{\ell(\omega|y')} > \frac{\ell(\omega'|y)}{\ell(\omega'|y')}$ for all $\omega > \omega'$ and $y > y'$.

[4]I make the following distinction between layoffs and firings. A layoff represents an attempt by a firm to reduce employment due to reasons that are not directly attributable to the laid off worker's performance; layoffs are accompanied by job destruction in my model. Firing of a worker, on the other hand, is a direct response to the worker's poor performance and so is not accompanied by job destruction: the job previously occupied by a fired worker turns into a vacancy following the firing of the worker. Throughout the majority of the paper, I assume that firms face a firing cost that is sufficiently high to deter firing. In subsection 8.1, I consider an extension of the model wherein firing arises in equilibrium due to a moral hazard problem on the workers' side.

[5]The quits in my model are better thought of as due to reasons, such as irresolvable workplace conflict or the need to accompany a spouse to a new location, which are largely independent of the wage and the state of the economy. What I exclude is quitting a low-paying job to search for a job higher on the job ladder—there is no job ladder in this economy since a worker receives the same wage regardless of where he is employed. The main qualitative predictions of the model would however be little changed by incorporating job heterogeneity and on-the-job search.

machine breakdown, it can either repair the machine or cease production and lay off its employee. The scrap value of a broken machine is zero; the cost of repairing the machine is a random variable c, which is distributed according to a differentiable p.d.f. h independent of other variables.[6]

Wages are determined by a wage rule as in Hall (2005a). The wage rate for a worker of ability y when the labor productivity is A is given by a twice continuously differentiable function $w(y, A)$. The wage could depend on ability and the labor productivity, but it cannot be a function of the composition of ability in the pool of the unemployed.[7] This assumption significantly simplifies the analysis by making the job creation and layoff decisions of firms independent of how they expect the composition of ability in the unemployed pool to evolve over time. I assume that both the wage rate $w(y, A)$ and a firm's flow profit $\pi(y, A) = yA - w(y, A)$ are positive and strictly increasing in y and A.

3 Equilibrium

In this section I derive the equations that describe the equilibrium of the model. I first characterize the separation rate and the job-finding rate—the ins and outs of unemployment—as functions of ability, labor productivity, and the composition of the pool of unemployed. The dynamics of the economy are then described by a dynamical system in which the state space is the unemployment rates for workers of every ability and the rate of change in the unemployment rate among workers of ability y is given by their net flow into unemployment.

I start by characterizing the value of vacancies. Let $V(\lambda, A)$ denote the expected net present value of a vacancy when the distribution of ability in the pool of the unemployed is given by λ and labor productivity is given by A. Free entry and exit of firms pins down the value of vacancies: any inactive firm can create a vacancy by purchasing a machine that costs k units of the final good and any firm with no worker and a machine can sell the machine for k units of the final good. Therefore, whenever some hiring is taking place in the economy, $V(\lambda, A)$ must be identically equal to k. I assume that the parameters of the model are such that this is always the case.[8]

[6]One special case of the model is one in which the cost of repair is concentrated on a large number $\overline{c}$. Given such a specification, machines would almost never be repaired. The model would then be isomorphic to a model with stochastic depreciation of capital at rate δ. Another special case is one in which the cost of repair is concentrated on a small number $\underline{c}$. Then firms would almost always repair a broken machine and the model would be equivalent to a model with a flow maintenance cost equal to $\delta\underline{c}$.

[7]A particular case of my wage rule is the fully rigid real wage proposed by Hall (2005a) and Shimer (2005, 2012b). Hall and Milgrom (2008), Gertler and Trigari (2009), and Kennan (2010) provide microfoundations for rigid real wages. However, see Pissarides (2009) for an argument against the use of rigid wages in search models.

[8]A sufficient condition for there to always be some hiring is that $(\pi(\overline{y}, \underline{A}) + \varsigma k)/(\rho + \varsigma + \delta)$ is larger than k. Under this condition a job employing a worker of the highest ability, $\overline{y}$, is at all times more valuable to a firm than, k, the cost of a machine. This observation together with the Inada conditions and the presence of quits imply that there is always some hiring. To see why, suppose to the contrary that no firm is hiring at some point in time. This would imply that the market tightness is equal to zero, so given that the matching function satisfies the Inada conditions, if a firm were to post a vacancy, it would be contacted by workers infinitely fast. But due to quits there are some workers of ability $\overline{y}$ in the pool of the unemployed. Therefore by being extremely

I next characterize the present value of profits of a firm that has already hired a worker of ability $y > 0$. The flow profit of the firm is given by $\pi(y, A) = yA - w(y, A)$. Denote the expected present discounted value of the firm's profits by $\Pi(y, A)$. The value of profits is affected by machine breakdowns and separations due to quits and layoffs. The firm lays off its employee and ceases production in the event of a machine breakdown if and only if the cost of repair, c, is larger than the present value of the firm's profits, $\Pi(y, A)$. Taking layoffs and quits into consideration, the value function needs to satisfy the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\rho\Pi(y, A) = \pi(y, A) + \mu A \frac{\partial \Pi(y, A)}{\partial A} + \mu A^2 \frac{\partial^2 \Pi(y, A)}{\partial A^2} + \varsigma(k - \Pi(y, A)) - \delta \mathbb{E}_{\tilde{c}} \min\{\tilde{c}, \Pi(y, A)\}, \quad (1)$$

where I have used the fact that $V(\lambda, A) \equiv k$ to substitute k for the value of a vacancy. The first three terms on the right-hand side of equation (1) are standard terms that arise from an application of Itô's lemma. The fourth term represents the capital gain (or less) the firm experiences when its employee quits his job. The last term represents the capital loss from a machine breakdown: the firm incurs a cost of c if it decides to repair the machine and loses the entire value of the firm, $\Pi(y, A)$, if it decides against a repair. The value of the firm needs to additionally satisfy the following boundary conditions:[9]

$$\frac{\partial \Pi(y, A)}{\partial A}\Big|_{A=\underline{A}} = \frac{\partial \Pi(y, A)}{\partial A}\Big|_{A=\overline{A}} = 0. \quad (2)$$

The assumption that $\pi(y, A)$ is continuous and strictly increasing in y and A ensures that $\Pi(y, A)$ is also continuous and strictly increasing in y and A over $Y \times (\underline{A}, \overline{A})$.

The separation rate is given by the sum of quits and layoffs. The quit rate is fixed at ς. The layoff rate for a worker of ability y when the labor productivity is A is given by $\delta(1 - H(\Pi(y, A)))$, where H denotes the c.d.f. corresponding to h. Thus the total separation rate for workers of ability y is given by

$$s(y, A) = \varsigma + \delta\left(1 - H(\Pi(y, A))\right). \quad (3)$$

The separation rate, $s(y, A)$, is strictly decreasing in y and A since $\Pi(y, A)$ is strictly increasing in its arguments.

I next characterize the hiring policy of a firm with a functioning machine that is matched to a worker of unknown ability. The firm observes an imperfect signal, ω, of the worker's ability, y, that is distributed according to p.d.f. $\ell(\omega|y)$. By Bayes' rule, the present value of the profits the firm is expected to obtain by hiring the worker is given by

$$J(\omega, u, A) = \frac{\displaystyle\int \Pi(\tilde{y}, A)\ell(\omega|\tilde{y})u(\tilde{y})d\tilde{y}}{\displaystyle\int \ell(\omega|\tilde{y})u(\tilde{y})d\tilde{y}}, \quad (4)$$

selective an entrant firm can meet and hire a worker of ability $\overline{y}$ immediately after investing k in the purchase of a machine. Therefore vacancy posting is a strictly profitable deviation.

[9] See for instance Dixit (1993).

where $u = \{u(y)\}_{y \in Y}$ denotes the profile of unemployment rates across different abilities. The distribution of ability in the pool of the unemployed affects the firm's prior and posterior beliefs about the ability of the worker: when unemployed workers are expected to have lower abilities on average, the firm interprets good signals more pessimistically.

The expected value of hiring a worker only depends on the composition of the unemployed pool and not on the unemployment rate. This can be seen by examining equation (4): a doubling of the unemployment rate across all abilities results in a doubling of both the numerator and the denominator of the expression on the right-hand side but leaves the expression on the left-hand side unchanged. The reason for this invariance is that proportional changes in the number of unemployed workers across all levels of ability have no effect on the likelihood that a randomly selected worker is of a given ability. I use this observation to normalize u to a probability distribution λ (that integrates to one) and express J as a function of λ. More formally, define

$$\lambda(y) = \frac{u(y)g(y)}{\mathbb{E}u} = \frac{u(y)g(y)}{\displaystyle\int u(\tilde{y})g(\tilde{y})d\tilde{y}}. \tag{5}$$

Then $J(\omega, \lambda, A) = J(\omega, u, A)$.

The firm hires the matched worker whenever doing so is expected to generate profits valued in excess of the resale value of the machine owned by the firm. An operational machine can always be sold for k units of the final good. So the firm hires the worker if and only if it observes a signal ω about the ability of the worker that satisfies $J(\omega, \lambda, A) \geq k$. Given that $\ell(\omega|y)$ satisfies the strict MLRP and Π is a strictly increasing function of ability, $J(\omega, \lambda, A)$ is strictly increasing in ω for any nondegenerate distribution of ability.[10] So for all λ and A there exits a unique threshold $\underline{\omega}(\lambda, A)$ satisfying

$$J(\underline{\omega}(\lambda, A), \lambda, A) = k \tag{6}$$

such that firms hire the matched workers if they generate a signal larger than $\underline{\omega}(\lambda, A)$.[11] I denote the probability that a worker of ability y generates a signal larger than $\underline{\omega}(\lambda, A)$ by $P(y, \lambda, A) = 1 - L(\underline{\omega}(\lambda, A)|y)$, where $L(\cdot|y)$ is the c.d.f. of ω conditional on y.

I next study the vacancy-creation decisions of firms. Since $V(\lambda, A)$ is independent of A and λ, the value of a vacancy satisfies the following simple HJB equation:

$$\rho V(\lambda, A) = q(\theta(\lambda, A)) \int (\Pi(\tilde{y}, A) - V(\lambda, A)) P(\tilde{y}, \lambda, A)\lambda(\tilde{y})d\tilde{y},$$

where the equation is used to define the market tightness as a function only of λ and A. With some abuse of notation let $q(\lambda, A) = q(\theta(\lambda, A))$. The above equation can be rewritten in a more convenient form in terms of $q(\lambda, A)$:

$$\frac{1}{q(\lambda, A)} = \frac{1}{\rho k} \int (\Pi(\tilde{y}, A) - k) P(\tilde{y}, \lambda, A)\lambda(\tilde{y})d\tilde{y}. \tag{7}$$

[10] For a proof, see, for example, Milgrom (1981).

[11] I set $\underline{\omega}(\lambda, A) = -\infty$ if $\lim_{\omega \to -\infty} J(\omega, \lambda, A) > k$ and set $\underline{\omega}(\lambda, A) = \infty$ if $\lim_{\omega \to \infty} J(\omega, \lambda, A) < k$.

The rate at which unemployed workers encounter vacancies is given by

$$p(\lambda, A) = m\left(\frac{1}{\rho k}\int \left(\Pi(\tilde{y}, A) - k\right) P(\tilde{y}, \lambda, A)\lambda(\tilde{y})d\tilde{y}\right). \tag{8}$$

The job-finding rate for a worker of ability y is the product of the rate at which the worker encounters a firm and the probability that he clears the firm's hiring threshold by generating a signal larger than $\underline{\omega}$. The former is given by $p(\lambda, A)$ while the latter is given by $P(y, \lambda, A)$. Thus the job-finding rate for a worker of ability y when the distribution of ability in the unemployed pool is λ and the labor productivity is A is given by

$$f(y, \lambda, A) = p(\lambda, A)P(y, \lambda, A). \tag{9}$$

The distribution of ability in the unemployed pool affects both terms that appear in the job-finding rate. The number of vacancies firms create and so the rate of contact between firms and workers is a function of the quality of the pool they face. This is how λ affects p. Additionally the firms' hiring threshold is a function of their prior belief about the quality of a typical matched worker: when firms expect lower-ability workers, more unequivocally good signals are needed to convince them otherwise. This is the channel through which λ affects P and so the job-finding rate.

Once the separation and job-finding rates are determined, the dynamics of the model are easy to describe. The unemployment rate among workers of ability y evolves according to the following differential equation:

$$\frac{d}{dt}u_t(y) = s(y, A_t)(1 - u_t(y)) - f(y, \lambda_t, A_t)u_t(y), \tag{10}$$

where A_t, u_t, and λ_t denote the values of A, u, and λ at time t. The definition of equilibrium is straightforward.

Definition 1. Given a path $\{A_t\}_{t\geq 0}$ for the labor productivity and an initial condition $\{u_0(y)\}_{y\in Y}$, an equilibrium is a path for the unemployment rate $\{u_t(y)\}_{y\in Y, t>0}$ that satisfies equation (10) at all times and functions Π, s, J, $\underline{\omega}$, q, p, and f that solve the functional equations (1)–(9).

I have defined equilibrium as a set of deterministic objects that do not directly refer to the underlying stochastic process for labor productivity that drives the model economy. The stochasticity is however incorporated in my definition of equilibrium in two ways. First, functions Π, s, J, $\underline{\omega}$, q, p, and f are the solutions to a set of functional equations that are derived by making use of the properties of the Brownian motion that describes the evolution of A. Second, the equilibrium path of the unemployment rate, $\{u_t(y)\}_{y\in Y, t>0}$, depends on the realized path of labor productivity, $\{A_t\}_{t\geq 0}$, by way of equation (10). This dependence defines a mapping from the space of paths of A to the space of paths of $\{u(y)\}_{y\in Y}$. I could have alternatively defined equilibrium as this mapping from paths to paths. However such a choice would have required introducing additional notation without yielding any additional mathematical or conceptual fruit.

4 Adverse Selection and Hiring

In this section I characterize the effect of changes in the composition of the pool of the unemployed on the firms' vacancy-posting and hiring decisions. I show that a worsening of the pool leads to fewer vacancies and more exacting hiring requirements, forces which both contribute to a smaller job-finding rate. But before formally presenting the results I need to define the sense in which one pool is worse than another pool. In other words I need a (partial) order on the set of distributions of ability, ΔY. Throughout the rest of the paper I use the following partial order on ΔY.

Definition 2. Let λ, λ' be two probability distributions over Y. λ is *larger in the sense of MLRP* than λ' if $\lambda(y)/\lambda'(y)$ is an increasing function of y.

It is useful to make a note of the relationship between the MLRP and FOSD partial orders: if λ is larger than λ' in the sense of MLRP, it is also larger than λ' in the sense of first-order stochastic dominance. The converse is however not true. The following proposition is the main result of this section; it summarizes how the observable statistics of the labor market vary as λ and A change.

Proposition 1. *Let $\underline{\omega}$, θ, and f denote the equilibrium functions that determine the hiring threshold, the market tightness, and the job-finding rate, respectively.*

(a) $\underline{\omega}(\lambda, A)$ *is decreasing in* λ *and* A.

(b) $\theta(\lambda, A)$ *is increasing in* λ *and* A.

(c) $f(y, \lambda, A)$ *is increasing in* y, λ, *and* A.

Firms relax their hiring standards both with an increase in λ and an increase in A. The reason for the decrease in the threshold with an increase in A is straightforward. When the labor productivity improves, the value of a job employing a worker of any ability increases while the costs of hiring remain unchanged. So it becomes profitable for firms to hire some of the workers whom they were previously unwilling to hire. The fact that the signal is informative of ability then immediately implies that firms must be willing to be more forgiving when it comes to lower signals when the economy improves.

It is only slightly harder to see why the hiring threshold is decreasing in the distribution of ability among the unemployed. Intermediately high signals are either generated by high-ability workers who got unlucky or low-ability workers who got lucky. The firms' prior belief about the ability of a typical worker determines how these possibilities are weighted. In an adversely selected labor market in which λ is relatively low in the sense of MLRP, firms interpret intermediate signals as being the result of luck and not indicative of high ability. The hiring requirements thus become increasingly insurmountable as the distribution of ability in the unemployed population deteriorates. This lowers the probability of being hired conditional on being matched for workers of all abilities. An increase in λ has the opposite effect of increasing the probability of being hired.

Firms create fewer vacancies in adversely selected labor markets or when A is low since they anticipate imposing stringent hiring standards and being unable to fill the vacancy with a suitable employee. This diminished incentive to post vacancies results in a slack labor market and a slow rate of exit from unemployment for all workers. This is the key mechanism through which the model generates slow recoveries.

5 Dynamics of Adverse Selection

Section 4 discussed the implications of shifts in the composition of the unemployed pool for the firms' hiring decisions and workers' job-finding prospects, but it was silent on the economic conditions that contribute to such aggregate compositional changes. In this section I show that changes in productivity can lead to variations in the composition of the unemployed pool over the course of a business cycle. I start by characterizing the short-term response of the distribution of ability among unemployed workers to an increase in the labor productivity and use this characterization to highlight the conflicting economic forces that are in play. The composition of the pool is altered by changes along the hiring margin and the layoff margin that affect workers differentially depending on their ability. I then introduce additional assumptions that allow me to obtain unambiguous results about the effect of increases in labor productivity on the composition of the pool.

I build up toward the main characterization results by manipulating the law of motion of the unemployment rate. Recall that the unemployment rate among workers of ability y evolves according to the following differential equation:

$$\frac{d}{dt}u_t(y) = s(y, A_t)(1 - u_t(y)) - f(y, \lambda_t, A_t)u_t(y).$$

Equation (5) defines λ as a function of u, so with some abuse of notation, I can write f as a function $f(y, u, A)$ of u. The above equation then can be rewritten as

$$\frac{d}{dt}u_t(y) = \Gamma(y, u_t, A_t)u_t(y), \tag{11}$$

where $\Gamma(y, u, A)$ is defined as

$$\Gamma(y, u, A) = s(y, A)\left(\frac{1}{u(y)} - 1\right) - f(y, u, A). \tag{12}$$

$\Gamma(y, u, A)$ represent the net flow rate into unemployment for workers of ability y when the state of the economy is described by u and A. Integrating equation (11) with respect to $g(y)$ I get

$$\frac{d}{dt}\mathbb{E}u_t = \left[\int \Gamma(\tilde{y}, u_t, A_t)\lambda_t(\tilde{y})d\tilde{y}\right]\mathbb{E}u_t. \tag{13}$$

Combining with equation (11),

$$\frac{d}{dt}\lambda_t(y) = \left[\Gamma(y, u_t, A_t) - \int \Gamma(\tilde{y}, u_t, A_t)\lambda_t(\tilde{y})d\tilde{y}\right]\lambda_t(y). \tag{14}$$

Let $\Lambda_t(y)$ denote the complementary c.d.f. corresponding to probability measure $\lambda_t(y)$.[12] Integrating equation (14) with respect to y I get

$$\frac{d}{dt}\Lambda_t(y) = \int_y \lambda_t(\tilde{y})\Gamma(\tilde{y}, u_t, A_t)d\tilde{y} - \left(\int_y \lambda_t(\tilde{y})d\tilde{y}\right)\left(\int \Gamma(\tilde{y}, u_t, A_t)\lambda_t(\tilde{y})d\tilde{y}\right).$$

Let $\mathbb{E}_{\lambda_t}$ denote the expectation operator corresponding to probability measure λ_t. The above equation can be rewritten in terms of $\mathbb{E}_{\lambda_t}$ as

$$\frac{d}{dt}\Lambda_t(y) = \left[\mathbb{E}_{\lambda_t}\left[\Gamma(\tilde{y}, u_t, A_t)|\tilde{y} \geq y\right] - \mathbb{E}_{\lambda_t}\left[\Gamma(\tilde{y}, u_t, A_t)\right]\right]\Lambda_t(y). \tag{15}$$

An increase in $\Lambda(y)$ for all y corresponds to an increase in the ability of unemployed workers in the sense of first-order stochastic dominance. Therefore if the right hand side of equation (15) is positive for all y at time t, then $\lambda_t(y)$ is increasing at time t in the sense of first-order stochastic dominance. A sufficient condition for this to be the case is that $\Gamma(y, u_t, A_t)$ is increasing in y. One can in fact prove the following stronger result.

Proposition 2. *Consider an equilibrium $\{u_t\}_{t\geq 0}$ corresponding to some path of labor productivity $\{A_t\}_{t\geq 0}$. If $\Gamma(y, u_t, A_t)$ is increasing in y at time t, then the distribution of ability in the unemployed pool improves in the sense of MLRP between t and $t + \delta t$ for sufficiently small δt.*

The proposition is true almost by definition. $\Gamma(y, u, A)$ is defined as the rate of change in the unemployment rate of workers of ability y when the state of the economy is given by u and A. If $\Gamma(y, u_t, A_t)$ is increasing in y, higher-ability workers experience a larger proportional increase in their unemployment rate at time t relative to lower-ability workers. Therefore the unemployed pool improves in the sense of MLRP.

Although mathematically trivial, Proposition 2 is useful in characterizing the response of the economy to shocks to the labor productivity. Consider an economy in which the initial profile of unemployment rate is $\{u_0(y)\}_{y\in Y}$ and the path of labor productivity is given by $\{A_t\}_{t\geq 0}$. Now suppose that the path of labor productivity is perturbed to $\{A'_t\}_{t\geq 0}$, where $A'_t = A_t + \delta A_t$ for some small positive $\{\delta A_t\}_{t\geq 0}$. This increase in labor productivity can induce an improvement or a deterioration in the quality of the unemployed pool depending on the shape of the Γ function. The next proposition delineates this dependence.

Proposition 3. *Suppose that $\partial^2_{yA}\Gamma(y, u_0, A_0)$ is positive (negative) for all y. A small increase in labor productivity at time zero from A_0 to A'_0 results in an improvement (decline) in the quality of the pool of the unemployed in the sense of MLRP at time δt for δt sufficiently small.*

[12]That is, $\Lambda_t(y) = \int_y \lambda_t(\tilde{y})d\tilde{y}$.

The result is intuitive in light of Proposition 2. If $\partial^2_{yA}\Gamma(y, u_0, A_0)$ is positive for all y, an increase in A results in an increase in the slope of $\Gamma(y, u_0, A_0)$ with respect y. But Proposition 2 implies that a more positively sloped $\Gamma(y, u_0, A_0)$ is associated with a better unemployed pool at time δt. Thus a higher A leads to an improved pool when y and A appear as complements in the net rate of flow into unemployment, $\Gamma(y, u, A)$. A formal proof is provided in the appendix.

To get a sense of various forces that contribute to changes in the composition of the unemployed pool when A is changed, it is informative to look at the terms that comprise $\partial^2_{yA}\Gamma(y, u, A)$:

$$\begin{aligned}\frac{\partial^2\Gamma(y,u,A)}{\partial y\partial A} = &-\frac{\partial s(y,A)}{\partial A}\frac{u'(y)}{u^2(y)} + \frac{\partial^2 s(y,A)}{\partial A\partial y}\left(\frac{1}{u(y)} - 1\right) \\ &+ \frac{\partial p(\lambda,A)}{\partial A}\frac{L(\underline{\omega}(\lambda,A)|y)}{\partial y} + p(\lambda,A)\frac{\partial \ell(\underline{\omega}(\lambda,A)|y)}{\partial y}\frac{\partial \underline{\omega}(\lambda,A)}{\partial A}\end{aligned} \tag{16}$$

The first two terms represent the differential impact of an increase in the labor productivity on the flow of workers of different abilities into unemployment. The first term captures the differential exposure of workers of different abilities to a uniform change in the separation rate that arises from differences in their respective employment rates: the decrease in the separation rate lowers the flow into unemployment relatively more for workers of ability y for whom the employment rate, $1 - u(y)$, is relatively larger. This term is positive (negative) when the unemployment rate is increasing (decreasing) in ability.

The second term is the differential change in the separation rate across abilities from an increase in A. It is positive (negative) if the separation rate is supermodular (submodular).[13] If s is supermodular, when A declines firms with lower-ability workers lay off their employees at a higher rate compared to firms with higher-ability workers.

The third and fourth terms represent the differential change in the flow of workers of different abilities out of unemployment when the labor productivity increases. The former is the effect of the improvement in the state of the economy on the firms' incentives to post vacancies and the resulting tightening of the labor market and the differential effect of a tighter labor market on workers of different abilities. The increase in vacancies is more beneficial for higher-ability workers who are more likely to generate sufficiently good signals that allow them to be hired. This term is always negative since $p(\lambda, A)$ is increasing in A by Proposition 1 and $\ell(\omega|y)$ satisfies the strict MLRP.

The last term captures the differential effect of a loosening of the hiring standards that follows an increase in A across workers with different abilities. This term is more positive (negative) for workers who were more likely to generate signals just below (above) the hiring threshold prior to the increase in A. It is always positive for some workers and negative for others.

[13] A function $s(y, A)$ is supermodular if $s(\max\{y, y'\}, \max\{A, A'\}) + s(\min\{y, y'\}, \min\{A, A'\}) \geq s(y, A) + s(y', A')$ for all y, y', A, and A'. The function is strictly supermodular if the above inequality is strict whenever $y > y'$ and $A < A'$. For a twice continuously differentiable function $s(y, A)$, supermodularity is equivalent to $\partial^2 s(y, A)/\partial y\partial A \geq 0$ and strict supermodularity is equivalent to $\partial^2 s(y, A)/\partial y\partial A > 0$. A function $s(y, A)$ is submodular (strictly submodular) if $-s(y, A)$ is supermodular (strictly supermodular).

The effect of an increase in the labor productivity on the distribution of ability in the unemployed pool is therefore in general ambiguous. However there are particular cases in which unambiguous results can be obtained. In the rest of this section I focus on one such case in which the signals that firms observe are noisy and the economy is close to a steady state as defined next.

Definition 3. A *steady state equilibrium* corresponding to labor productivity A^* is an equilibrium in which $A_t = A^*$ and $u_t = u^*$ at all times.

Steady state equilibrium is a useful definition both conceptually and mathematically. It describes a zero-probability event in which the realized path of labor productivity is forever constant at some A^*.[14] Despite being a zero probability event, a steady state equilibrium is a good approximation to the behavior of the economy over short horizons if the drift and volatility of the labor productivity are sufficiently small.[15] More importantly for the purpose of this section, it proves to be a useful benchmark for the initial state of the economy starting from which one can obtain unambiguous results about the effect of changes in A on λ.

Equation (16) can be simplified when the economy is at a steady state. When u^* and A^* constitute a steady state equilibrium, they satisfy the following relation:[16]

$$\frac{1}{u^*(y)} - 1 = \frac{f(y, u^*, A^*)}{s(y, A^*)}. \tag{17}$$

Substituting for u^* from the above equality in equation (16) and some simple algebra imply that in a steady state,

$$\begin{aligned} \frac{\partial^2 \Gamma(y, u^*, A^*)}{\partial y \partial A} &= f(y, u^*, A^*) \frac{\partial \log s(y, A^*)}{\partial y \partial A} + \left(\frac{\partial p(\lambda^*, A^*)}{\partial A} - \frac{p(y, \lambda^*)}{s(y, A^*)} \frac{\partial s(y, A^*)}{\partial A} \right) \frac{\partial L(\underline{\omega}(\lambda^*, A^*)|y)}{\partial y} \\ &+ p(\lambda^*, A^*) \frac{\partial \ell(\underline{\omega}(\lambda^*, A^*)|y)}{\partial y} \frac{\partial \underline{\omega}(\lambda^*, A^*)}{\partial A}. \end{aligned} \tag{18}$$

The first term is positive if s is strictly log-supermodular and negative if it is strictly log-submodular, the second term is always negative, and the last term takes on both positive and negative values.[17] Therefore the net effect of an increase in A on λ is still ambiguous. But as I show in the next proposition, the first term is dominant if the signals observed by the firms are sufficiently noisy in the sense that is made precise next.

[14] Keep in mind however that even in a steady state equilibrium firms base their hiring and layoff decisions on the correct assumption that the labor productivity follows a geometric Brownian motion with positive drift and volatility.

[15] One can in fact show if $A_0 = A^*$ and $u_0 = u^*$, then μ can be chosen sufficiently small to guarantee that u_t is arbitrarily close to u^* for an arbitrarily long period of time and with a probability arbitrarily close to 1.

[16] Note that the steady state unemployment rate is decreasing in ability since the separation rate is decreasing in ability and the job-finding rate is increasing in ability.

[17] A function $s(y, A)$ is log-supermodular (strictly log-supermodular) if $\log s(y, A)$ is supermodular (strictly supermodular). A function $s(y, A)$ is log-submodular (strictly log-submodular) if $\log s(y, A)$ is submodular (strictly submodular). Also see footnote 5.

Definition 4. The signal structure ℓ_σ is *additive* with precision σ^{-1} if $\omega = y + \sigma\varepsilon$, where ε is a random variable supported on the entire reals and distributed according to c.d.f. Φ with a differentiable and bounded density that is symmetric around zero.

The conditions imposed on Φ ensure that endogenous variables change continuously with changes in σ. They are satisfied by most symmetric distributions including the normal distribution. The next proposition shows that the first term in equation (18) is dominant whenever σ is sufficiently large.

Proposition 4. *Fix the labor productivity A^* and all the fundamentals of the economy except for the signal structure and assume that $\int \Pi(\tilde{y}, A^*)g(\tilde{y})d\tilde{y} > k$. Consider a class of additive signal structures $\{\ell_\sigma\}$ that are parameterized by σ. If s is strictly log-supermodular (log-submodular), then there exists some $\sigma^* < \infty$ such that, in any economy with signal structure ℓ_σ with $\sigma > \sigma^*$, starting from the steady state equilibrium corresponding to A^*, a small increase in A at time zero results in an increase (decrease) in the sense of MLRP in the quality of the pool of the unemployed at time δt.*

The intuition for the proposition is as follows. The assumption that $\int \Pi(\tilde{y}, A^*)g(\tilde{y})d\tilde{y} - k$ is positive guarantees that firms are willing to hire a worker who is drawn from a pool that is representative of the population at large. Therefore, given that the signals are not very informative, firms use a low hiring threshold and do not vary the threshold by much when the labor productivity changes. So changes in A mainly affect the composition of the pool through one margin—the separation rates of low and high-ability workers respond asymmetrically to changes in the labor productivity. If s is log-supermodular, an increase in A lowers the separation rates of lower-ability workers more compared to those of higher-ability workers thus causing an improvement in the quality of the pool of the unemployed. In Section 7 I present conditions on the fundamentals of the economy that give rise to log-supermodular and log-submodular separation rates.

6 Scarring

In this section, I extend the model introduced in Section 2 by assuming that firms can observe the time of each worker's latest entry into unemployment and show that the model can generate unemployment scarring—diminished employment prospects of workers who experience a long unemployment spell. In the extended model firms rationally expect the population of workers who have experienced longer unemployment spells to be more adversely selected. They therefore subject such workers to more stringent hiring standards. This statistical discrimination by firms gives rise to lower job-finding rates for the long-term unemployed. In subsection 6.2 I show that scarring can also result from job loss in a recession. I use the model to argue that, if the separation rate is log-supermodular, workers who become unemployed during recessions have worse job-finding prospects than otherwise identical workers who become unemployed during booms, even controlling for labor productivity and market tightness after separation.

6.1 Unemployment scarring

Consider a cohort of unemployed workers who have all entered unemployment at time τ. I let λ_t^τ denote the distribution of ability in the cohort and let θ_t denote the market tightness at time t.[18] The hiring threshold used by the firms when they are matched with a worker from this cohort is given by $\underline{\omega}(\lambda_t^\tau, A_t)$.[19] A worker of ability y who has become unemployed at time τ generates a signal that exceeds the hiring threshold at time t with probability $P(y, \lambda_t^\tau, A_t) = 1 - L(\underline{\omega}(\lambda_t^\tau, A_t)|y)$ and exits unemployment at rate $f(y, \lambda_t^\tau, A_t, \theta_t) = p(\theta_t)P(y, \lambda_t^\tau, A_t)$. The following result is a corollary of the fact that $p(\theta)$ and $P(y, \lambda, A)$ are positive and increasing functions.

Proposition 5. *Let $f(y, \lambda_t^\tau, A_t, \theta_t)$ denote the time t job-finding rate for an unemployed worker of ability y who has entered unemployment at time τ. $f(y, \lambda_t^\tau, A_t, \theta_t)$ is increasing in ability, y, the distribution of ability in the worker's cohort, λ_t^τ, labor productivity, A_t, and labor-market tightness, θ_t.*

Proposition 5 is the extended-model counterpart of part (c) of Proposition 1, but there is a noteworthy difference between the two. In the baseline model the distribution of ability in the pool of the unemployed and labor productivity uniquely determine the market tightness through the free-entry condition, whereas in the current model λ_t^τ and θ_t need not satisfy any joint restriction whatsoever. This is due to the fact that λ_t^τ is the distribution of ability in only one of a continuum of cohorts that make up the pool while θ_t is determined by the distribution of ability in the entire pool. Therefore λ_t^τ and θ_t can independently affect the job-finding rate, $f(y, \lambda_t^\tau, A_t, \theta_t)$. The proposition shows that an increase in either variable leads to an increase in the job-finding rate.

I next characterize the evolution of the job-finding rate for a worker who has become unemployed at time τ. The worker's job-finding rate can vary over time both due to changes in A_t and θ_t and due to time variations in λ_t^τ that result from differential exit rates of different workers in the cohort. As illustrated by Proposition 5, changes in A_t and θ_t affect the job-finding rate in a straightforward manner. Here I abstract away from such changes by assuming that A_t and θ_t are constant after time τ. This simplification enables me to focus on the decline in the job-finding rate that results from changes in the cohort composition. The rate of change in the worker's job-finding rate along a path with constant productivity and market tightness is given by the following expression:

$$\frac{d}{dt} \log f(y, \lambda_t^\tau, A_t, \theta_t)\bigg|_{A_t = A_\tau, \theta_t = \theta_\tau \ \forall t \geq \tau} = \frac{-\ell(\underline{\omega}(\lambda_t^\tau, A_\tau)|y)}{1 - L(\underline{\omega}(\lambda_t^\tau, A_\tau)|y)} \left\langle \frac{\partial}{\partial \lambda} \underline{\omega}(\lambda_t^\tau, A_\tau), \frac{d}{dt} \lambda_t^\tau \right\rangle, \tag{19}$$

where $d\underline{\omega}(\lambda, A)/d\lambda$ denotes the derivative of $\underline{\omega}$ with respect to λ and $\langle \, , \, \rangle$ denotes the inner product.[20]

[18] θ_t is an endogenous process as in the baseline model, but this fact is immaterial for the analyses of this section for the following reason. I study a single cohort of workers who have all entered unemployment at some time τ, and any such cohort has zero measure given that the model is in continuous time.

[19] $\underline{\omega}$ is the function defined in equation (6).

[20] More formally, $\partial\underline{\omega}(\lambda, A)/\partial\lambda$ is the Fréchet derivative of $\underline{\omega}$ with respect to λ, and $\langle \varphi, \nu \rangle = \int \varphi(y) d\nu(y)$ for any function $\varphi : Y \to \mathbb{R}$ and signed measure ν over Y.

Equation (19) has an intuitive interpretation. The rate of change of the job-finding rate is the product of two terms. The first is the sensitivity of the job-finding rate to changes in the hiring threshold. This term depends on the worker's ability through its effect on the hazard rate of signal distribution. The second term is the time derivative of the hiring threshold. This term is in turn the product of the sensitivity of the hiring threshold to the distribution of ability, and the time derivative of the distribution of ability. By Proposition 1, $\underline{\omega}(\lambda, A)$ is decreasing in λ. Therefore equation (19) implies that the job-finding rate declines over time if the cohort becomes more adversely selected over time. This is indeed the case as I argue below.

I proceed by characterizing the dynamics of λ_t^τ. Let $n_t^\tau(y)$ denote the time t measure of unemployed workers of ability y who have entered unemployment at time τ. The evolution of $n_t^\tau(y)$ is described by the following differential equation

$$\frac{d}{dt} n_t^\tau(y) = -f(y, \lambda_t^\tau, A_t, \theta_t) n_t^\tau(y). \tag{20}$$

Let $n_t^\tau = \int n_t^\tau(\tilde{y}) d\tilde{y}$ denote the measure of unemployed workers who have entered unemployment at time τ. Integrating equation (20) with respect to y and using the fact that $\lambda_t^\tau(y) = n_t^\tau(y)/n_t^\tau$ imply that

$$\frac{d}{dt} n_t^\tau = -\left[\int f(y, \lambda_t^\tau, A_t, \theta_t) \lambda_t^\tau(\tilde{y}) d\tilde{y}\right] n_t^\tau. \tag{21}$$

Equations (20) and (21) imply that

$$\frac{d}{dt} \lambda_t^\tau(y) = -\left[f(y, \lambda_t^\tau, A_t, \theta_t) - \int f(y, \lambda_t^\tau, A_t, \theta_t) \lambda_t^\tau(\tilde{y}) d\tilde{y}\right] \lambda_t^\tau(y). \tag{22}$$

Equation (22) describes the evolution of the distribution of ability in the cohort that has entered unemployment at time τ. The share of workers in this cohort who have ability y decreases over time if workers of ability y have a higher-than-average rate of exit from unemployment. The following proposition is a simple corollary of equation (22) and the fact that lower-ability workers are less likely to exit unemployment at any given time.

Proposition 6. *The ability distribution of the cohort of unemployed workers who have entered unemployment at time τ worsens over time in the sense of MLRP.*

The intuition for this result is as follows. Higher-ability workers are more likely to generate sufficiently favorable signals to get hired when they are matched with a firm. Therefore they always have higher job-finding rates than lower-ability workers who have become unemployed at the same time. As the higher-ability workers in a cohort leave the cohort at higher rates compared to the lower-ability workers, the cohort becomes more adversely selected over time.

Equation (19) and Proposition 6 illustrate the mechanism through which the model generates unemployment scarring. Consider again the cohort of workers who have all entered unemployment at time τ. With the passage of time the distribution of ability in the cohort worsens in the sense of MLRP.

Firms respond to the deterioration of ability in the cohort by interpreting the signals generated by its members more pessimistically and thus increasing their hiring threshold. This in turn depresses the job-finding prospects of any remaining workers of the cohort irrespective of their ability. The result is a scarring of the workers who have been long-term unemployed.

Equation (19) also illustrates that the rate of decline of the job-finding rate is larger for workers of ability y for whom the hazard rate $\ell(\underline{\omega}(\lambda_t^\tau, A_t)|y)/(1 - L(\underline{\omega}(\lambda_t^\tau, A_t)|y))$ is larger. Since ℓ satisfies the MLRP, $\ell(\omega|y)/(1 - L(\omega|y))$ is decreasing in y. This observation proves the following proposition.

Proposition 7. *The rate of decline in the job-finding rate with unemployment duration is larger for workers who have lower abilities.*

The proposition shows that unemployment scarring afflicts lower-ability workers more severely. Lower-ability workers are more likely to be hired by generating signals that barely pass the hiring threshold, so the decline in the hiring threshold resulting from the deterioration in the composition of the cohort affects them to a greater extent.

6.2 Recessions and scarring

The state of the economy at the time of separation can influence the separated worker's job-finding prospects as well. In this subsection I show that entering unemployment during a recession can be a further source of scarring. If the separation rate is log-supermodular in ability and productivity, workers who get unemployed in recessions end being part of a more adversely selected cohort. Firms therefore will impose higher hiring thresholds on cohorts who enter unemployment in recessions thereby diminishing the job-finding prospects of the workers in the cohort. The following proposition states the general result.

Proposition 8. *Suppose that the separation rate is log-supermodular (log-submodular). All else equal, workers who become unemployed at a time of low labor productivity experience lower (higher) job-finding rates immediately after the start of their unemployment spell compared to workers who become unemployed at a time of high labor productivity.*

The "all else equal" quantifier is important for the interpretation of the result. To better see this point, assume that the separation rate is log-supermodular and consider two economies: one in which labor productivity at time τ is given by A_τ and another in which it is given by $A'_\tau > A_\tau$. The proposition states that workers in the unprimed economy who lose their jobs at time τ have lower job-finding rates than otherwise identical workers in the primed economy who lose their jobs at time τ—*even if the labor productivity and market tightness follow identical paths after time τ in the unprimed and primed economies.* In other words workers' job-finding rates in this model exhibit path dependence.

Lower-ability workers in a cohort impose a negative externality on the higher-ability workers in the cohort. Since firms cannot perfectly observe the ability of individual workers, they use their prior information about the distribution of ability in a worker's cohort to make inferences about the quality of the worker. So being part of an adversely selected cohort is associated with a lower rate of exit from unemployment. If separation rate is log-supermodular, cohorts that enter unemployment in downturns are more adversely selected—so their members experience lower job-finding rates. This is how the model can generate scarring of workers who become unemployed in recessions.

7 Separation Rate

In the last two sections I argued that whether the separation rate is log-supermodular or log-submodular determines whether the pool of the unemployed becomes more adversely selected in downturns and whether job loss in recessions leads to scarring. In this section I find two sets of conditions on the fundamentals of the economy that give rise to log-supermodular and log-submodular separation rates and discuss their plausibility.

To be able to make inroads into characterizing the way labor productivity and ability determine the separation rate, I need to make functional-form assumptions on the wage rate and the distribution of repair costs. I assume that the cost of repair is distributed according to a Pareto distribution with shape parameter $\alpha > 0$ and scale parameter $\underline{c}$, where $\underline{c}$ is chosen sufficiently small to ensure that firms always repair their machine if the cost of repair is $\underline{c}$.[21] Pareto distribution is chosen for the sake of computational tractability. The result that follows can be generalized to any distribution with a monotonically decreasing density.

I additionally assume that the wage received by a worker of ability y is given by the sum of a base wage $w_0(y)$, which is independent of labor productivity, and a bonus $\beta(y)A$, which varies proportionally with A. Another way of thinking about the wage rule is as the first-order approximation to a more general wage rule. The particular shape of the function $w(y, A)$ is nonetheless unimportant for the result that follows. What turns out to be important is whether the flow profit of the firm, $\pi(y, A) = yA - w(y, A)$, is supermodular or submodular in y and A. The following proposition delineates this point.

Proposition 9. *Assume that the cost of repair is distributed according to a Pareto distribution and the wage rate is given by* $w(y, A) = w_0(y) + \beta(y)A$ *and that* $\Pi(y, A)$ *and its first two derivatives vary smoothly with* δ*.*

(a) Suppose that $\pi(y, A)$ *is strictly supermodular in* y *and* A*. Given any open interval* $\mathcal{A} \subset [\underline{A}, \overline{A}]$*, if the rate of machine breakdown,* δ*, and the slope of the base wage,* $|w_0'|$*, are sufficiently small and*

[21] For this to be the case it is sufficient that $\underline{c}$ is smaller than $(\pi(\underline{y}, \underline{A}) + \varsigma k)/(\rho + \varsigma + \delta)$, a lower bound on the value of a firm that hires a worker of the lowest ability.

the distribution of the cost of repair has a sufficiently thin tail, then the separation rate, $s(y, A)$, is strictly log-supermodular in (y, A) over $Y \times \mathcal{A}$.

(b) *Suppose that $\pi(y, A)$ is strictly submodular in y and A. Given any open interval $\mathcal{A} \subset [\underline{A}, \overline{A}]$, if the rate of machine breakdown, δ, and the slope of the base wage, $|w_0'|$, are sufficiently small and the distribution of the cost of repair has a sufficiently thin tail, then the separation rate, $s(y, A)$, is strictly log-submodular in (y, A) over $Y \times \mathcal{A}$.*

To gain some intuition for this result, it is useful to consider a special case of the model in which $w_0(y) = \varsigma k$ and $\beta(y) = \beta_0 y$ for some constant $\beta_0 \in (0, 1)$. Then for small values of δ the expected present value of a firm's profits is approximately equal to $(1-\beta_0)yA/(\rho+\varsigma)$. So a one percent increase in labor productivity leads to a one percent increase in the value of all firms and an α percent decrease in the probability of layoff for all workers independently of their ability. But this α percent decrease in layoffs represents a larger proportional decrease in total separations for lower-ability workers for whom layoffs constitute a larger share of separations vis-à-vis quits. Therefore the decrease in total separations resulting from the increase in A is larger in percentage terms for lower-ability workers than their higher-ability counterparts. That is, the separation rate is log-supermodular in labor productivity and ability. The proof amounts to formalizing this intuition for the general model and in both the supermodular and submodular cases.

When the separation rate is log-supermodular, the model presented in this paper generates jobless recoveries and scarring due to job loss in recessions. Suppose that the labor productivity declines sharply. The decline results in an increase in layoffs and a decline in new hires. But Proposition 4 implies that layoffs increase relatively more for firms employing lower-ability workers, thus leading to a deterioration in the quality of the pool of the unemployed. Proposition 1 implies that the decline in the quality of the pool induces firms to post fewer vacancies and to impose more stringent hiring requirements on unemployed workers, further depressing the job-finding rates of all unemployment workers. This slowdown in hiring persists past the recovery in the productivity and until the pool returns to its pre-recession composition. The result is a slow recovery in the employment rate. Proposition 8 implies that the population of workers who become unemployed in a downturn is more adversely selected than that of workers who become unemployed in booms, so the former group experiences a more severe unemployment scarring.

But is the separation rate more likely to be log-supermodular or log-submodular based on a priori reasoning? Proposition 9 shows that the answer to this question depends on whether labor productivity and ability appear as complements or substitutes in the firms' flow profits. This in turn depends on the exact specification of the wage rule. Two of the most natural ways of specifying the wage rule are the constant wage rate, $w(y, A) = w_0$, and a wage rate that is affine in the worker's flow output, $w(y, A) = w_0 + \beta_0 yA$. Both of these specifications lead to supermodular flow profits and log-supermodular separation rates. In fact, any wage rule of the form $w(y, A) = w_0(y) + \beta(y)A$ in which the bonus rises more slowly with ability than the total output generates a log-supermodular separation rate. This observation gives some credence to the hypothesis that the separation rate is log-supermodular. The question of whether the separation rate is log-supermodular or log-submodular is however ultimately an empirical question that requires more data to be brought to bear to be answered satisfactorily.

8 Discussion

In this section I discuss a number of empirical studies that speak to different aspects of the model presented in the paper. I exclusively focus on the specification of the model in which the separation rate is log-supermodular in labor productivity and ability since only such a specification can explain the phenomena that the paper aims to explain.

8.1 Cyclical variations in the composition of the pool

Cyclical variations in the composition of the unemployed and employed populations have been documented in several papers. The most relevant studies are by Solon, Barsky, and Parker (1994) and Mueller (2015). Solon, Barsky, and Parker (1994) find using data from the Panel Study of Income Dynamics (PSID) that the typical employed worker is more skilled in recessions. This is consistent with the prediction of my model. Mueller (2015) uses micro data from the Current Population Survey (CPS) to show that in recessions both the pool of the unemployed and the population of the employed shift toward higher-ability workers—when ability is proxied for by the logarithm of the wage or its Mincer residual. He concludes that it must be the case that sorting increases in recessions: the typical worker who moves from employment to unemployment in a recession is less skilled than the typical worker who remains employed throughout the recession but more skilled than the typical worker who was unemployed before the recession.

Although the evidence may seem to contradict the view presented in this paper that a worsening of adverse selection in recessions is a contributor to slow recoveries, a more careful inquiry proves otherwise. The discrepancy between the prediction of the model and the stylized facts presented above stems from the assumption maintained throughout the paper that separations only consist of quits, and layoffs that are accompanied by firm shutdowns. In practice firing of ill-suited workers is an important component of total separations for workers in the bottom of the ability distribution. This loss in realism was tolerated in the previous sections of the paper for the benefit of obtaining sharp analytical results, which highlighted important economic forces. Here I introduce firings into the model and show that, once firings are taken into account, the model has no difficulty explaining the evidence documented by Mueller (2015).

I extend the baseline model by introducing a moral hazard problem à la Shapiro and Stiglitz (1984) and allowing the firms to use firings to mitigate the problem. I assume that production requires workers to exert effort. The flow output produced by a worker of ability y is given by eyA, where $e \in \{0,1\}$ indicates the worker's choice of effort. Workers are heterogeneous in their cost of effort, z, with $z = \underline{z} > 0$ for some workers and $z = \overline{z} > \underline{z}$ for the remaining ones. A worker's choice of effort is not observable to the firm, but workers who shirk run the risk of being caught and getting fired. Firms have access to an imperfect monitoring technology that enables them to detect shirking: There is a signal that indicates shirking and arrives at Poisson rate ζ whenever the worker is shirking and at zero rate whenever he is exerting effort. The signal is verifiable and is grounds for the worker to be fired at no cost to the firm. I continue to assume that no-fault firings are sufficiently costly that firms abstain from using them in equilibrium.

The rest of the model is as in the baseline model with the following minor modifications. Ability and cost of effort are now distributed in the population according to p.d.f. $\hat{g}$ supported on $Y \times \{\underline{z}, \overline{z}\}$. A firm that is matched to a worker with ability y and cost of effort z observes a signal, which is drawn from p.d.f. $\hat{\ell}(\omega|y,z)$. I assume that the signal contains information on both components of the worker's type. Firms learn both the ability of their employees and their cost of effort immediately after hiring them. At the beginning of the employment relationship the worker and the firm sign a wage contract that specifies a wage rate $\hat{w}(y,A,z)$ to be paid by the firm to the worker for as long as the employment relationship lasts. The wage can be a function of the worker's cost of effort, but it cannot depend on the worker's (unverifiable) realized choice of effort. This leaves firing as the firms' only recourse in the event of a worker's shirking.

Despite its complexity, the model delivers sharp predictions regarding cyclical variations in the composition of the pool when $\underline{z}$ is small and $\overline{z}$ and ζ are both large. Having a small $\underline{z}$ implies that workers who have a low cost of effort exert effort in equilibrium. Having a large $\overline{z}$, on the other hand, guarantees that efficiency wages cannot be used to induce the workers with high cost of effort to exert effort. Therefore, in equilibrium such workers shirk and get fired at rate ζ. The role of assuming a large ζ is twofold. It implies that the firms' hiring decisions are little affected by the presence of workers with high cost of effort—since such workers are fired quickly and with no cost to the firm. It also implies that the unemployment rate among workers with high cost of effort is close to one irrespective of labor productivity.

Given these assumptions, the average ability among unemployed workers has a simple expression. Let $\hat{\lambda}$ denote the joint distribution of ability and the cost of effort among unemployed workers, let $\overline{n} = \int \hat{g}(\tilde{y}, \overline{z}) d\tilde{y}$ denote the fraction of workers in the population who have high cost of effort, and let u denote the total unemployment rate. The average ability among unemployed workers approximately equals

$$\frac{\overline{n}}{\mathsf{u}} \int \tilde{y}\hat{g}(\tilde{y}, \overline{z}) d\tilde{y} + \left(1 - \frac{\overline{n}}{\mathsf{u}}\right) \int \tilde{y}\hat{\lambda}(\tilde{y}, \underline{z}) d\tilde{y}.$$

The average ability is the weighted average of two terms: (i) the average ability among the unemployment workers who have high cost of effort, $\int \tilde{y}\hat{g}(\tilde{y}, \overline{z})d\tilde{y}$, and (ii) the average ability among the unemployed workers with low cost of effort, $\int \tilde{y}\hat{\lambda}(\tilde{y}, \underline{z})d\tilde{y}$. The first term is constant over the course of a business cycle. The second term decreases in recessions as the pool of the unemployed workers who have low cost of effort becomes more adversely selected. This was the only effect that was present in the previous sections of the paper in which no worker had high cost of effort and so $\overline{n}$ was zero. When $\overline{n}$ is positive, however, there is a second effect that can go in the opposite direction: workers with high cost of effort are always unemployed while the number of workers with low cost of effort who are unemployed increases in recessions. Therefore in recessions the weight given to the second term increases. If $\overline{n}$ is sufficiently large and workers with high cost of effort have a sufficiently low average ability compared to workers with low cost of effort, the net result is an increase in recessions in the average ability of unemployed workers and an increase in sorting.

Yet the presence of workers with high cost of effort does not interfere with the mechanism through which the model generates slow employment recoveries. Such workers are less likely than workers with low cost of effort to get hired when they are matched with a firm and are quick to get fired if they are ever hired. So they do not influence the firms' hiring decisions by much. It is only the distribution of ability among unemployed workers with low cost of effort that enters the firms' decision problems on whether to post a vacancy and what hiring threshold to use when they do post a vacancy. In recessions the distribution of ability among unemployed workers with $z = \underline{z}$ deteriorates in the sense of MLRP—just as in the baseline model—thus leading firms to post fewer vacancies and raise their hiring threshold and generating a slow recovery.

I end this subsection with a remark on its takeaway. The model analyzed above is clearly special along several dimensions. But it serves to illustrate an important point that goes beyond this stylized model: measures of composition that are relevant for the firms' hiring decisions can move negatively with measures of composition that are observed by the econometrician. Firms only respond to changes in the distribution of ability among the subset of unemployed workers who are hard to screen and costly to fire once employed—such workers tend to lie in the middle of the ability distribution. Measures such as the average log-wage, on the other hand, are coarse statistics of the distribution of ability in the entire pool.

8.2 Upskilling and downskilling

The model predicts that the hiring threshold used by firms varies countercyclically. In recent contributions, Modestino, Shoag, and Ballance (2015a,b) and Hershbein and Kahn (2016) use the data on the education and experience requirements of online job postings collected by Burning Glass Technologies (BGT) to study the relationship between the unemployment rate and job requirements. The papers establish a causal positive effect of an increase in a county's unemployment rate on the education and experience requirements of jobs located in the county. The effect is estimated off the variation over time in job requirements within a firm-job-county triple. The authors term the increase in requirements with an increase in the unemployment rate *upskilling*; a reversal of upskilling following a recovery is termed *downskilling*.

Hershbein and Kahn (2016) and Modestino, Shoag, and Ballance (2015a) discuss two models that can rationalize upskilling. Hershbein and Kahn (2016) argue that upskilling represents restructuring of production by firms in downturns when the opportunity cost of doing so is relatively lower. Firms postpone adoption of routine-replacing technologies until a downturn forces them to cut costs and improve efficiency. Modestino, Shoag, and Ballance (2015a) provide evidence of a reversal in upskilling following the tightening of the labor market between 2012 and 2014. They also show that their main findings are robust to restricting the sample to tradable industries. This evidence casts some doubt on the explanation of upskilling put forward by Hershbein and Kahn (2016). Modestino, Shoag, and Ballance (2015b) use a partial equilibrium model to rationalize upskilling and downskilling. In their model the cost of maintaining a vacancy is constant while the option value of waiting for a better worker increases with slack in the labor market. When the unemployment rate increases, firms take advantage of the slack by demanding more from their potential hires.

The mechanism proposed by Modestino, Shoag, and Ballance (2015b) runs into difficulties, however, in general equilibrium where the number of vacancies is determined by free entry of firms. In a slack labor market entry by firms erodes any potential gain in the option value of waiting for a better worker. This intuition is easily formalized in a version of my model with no heterogeneity in worker ability. Recall that the threshold used by firms, $\underline{\omega}(\lambda, A)$, is only a function of labor productivity, A, and the composition of the pool, λ—and not a function of the unemployment rate. With homogeneous workers, λ reduces to a constant degenerate probability distribution, so the hiring threshold only changes with variations in labor productivity.

The model in the current paper provides an explanation of upskilling and downskilling that works in general equilibrium. A temporary decline in labor productivity leads to a persistent increase in the unemployment rate and a persistent worsening of the distribution of ability in the unemployed pool. While the increase in the unemployment rate has no effect on the firms' hiring threshold, the deterioration of the distribution of ability leads to an increase in the threshold. The result is a positive co-movement of the unemployment rate and the hiring requirements, which is mediated through changes in the composition of the pool.

8.3 Job destruction, separations, quits, and layoffs

Much has been made of the distinction between the separation rate and the job destruction rate and their cyclicalities. In a number of influential contributions, Davis and Haltiwanger (1990, 1992) and Blanchard and Diamond (1990) demonstrate using firm-level data that the job destruction rate is highly countercyclical. They conclude that the increase in unemployment in recessions results more from an increase in job destruction than a decrease in job creation. In a more recent development, Hall (2005b) and Shimer (2012a) articulate a "new view" of the labor market according to which separations are largely acyclical and play a minor role in explaining the volatility of the unemployment rate. Davis however takes issues with Hall's conclusions in his discussion of Hall (2005b). Among other comments, he draws attention to the distinction between layoffs and quits and the cyclical variations in the composition of separations. Quits are procyclical while layoffs are countercyclical. The layoff-separation ratio is highly countercyclical ranging in value from less than 0.2 in booms to more than 0.7 in recessions. Quits and layoffs moreover have very different effects on the workers' employment prospects. Layoffs are associated with larger earning losses and are followed by significantly longer unemployment spells than are quits.

A divergence between the cyclicalities of quits and layoffs is also an important feature of the model presented in this paper. The quit rate in the model is independent of ability and acyclical while the layoff rate is decreasing in ability and countercyclical. The layoff-separation ratio is therefore countercyclical in accordance with the stylized fact discussed in the previous paragraph. It is exactly this combination of ability-independent and acyclical quits and ability-dependent and countercyclical layoffs that enables the model to generate a countercyclical and log-supermodular separation rate, a worsening of adverse selection in recessions, and slow employment recoveries.

But while the log-supermodularity of the separation rate is crucial for the mechanism highlighted in the paper, its countercyclicality plays no role. It is rather an artifact of the counterfactual assumption of a constant quit rate, which was made for the sake of analytical tractability. Incorporating quits, which vary endogenously with the state of economy, would have considerably complicated the model beyond the point where an analytical characterization is possible. One can however get a sense of what to expect in a setting with procyclical quits by studying a simple extension of the model in which the quit rate is increasing in labor productivity. The quit rate can be chosen in a way that the resulting average separation rate is mostly acyclical—consistent with the new view of the labor market—but such that the separation rate continues to be log-supermodular—as in the baseline model. Since the results of the paper only make use of the log-supermodularity assumption, none of them would be affected by this modification. In Appendix C, I formally explore this extension of the model and show that the conclusions of the paper are robust to the introduction of a procyclical quit rate.

8.4 Recruiting intensity and movements of the Beveridge curve

Economists have paid much attention to the shifting out in the Great Recession of the Beveridge curve—the negative empirical relationship between the unemployment rate and the vacancy rate over the course of a business cycle. This shift represents a deterioration of the aggregate matching process in the economy. The outward shift of the Beveridge curve is however not unique to this recession. Diamond and Şahin (2015) find using unemployment data from the BLS and historical data on vacancy rates constructed by Barnichon (2010) that the Beveridge curve shifted out in seven out of eight completed business cycles since 1950. In three of the recessions the Beverige curve shifted back inward at the end of the recovery.

Cyclical variations in the matching process are also evident in other labor market statistics. Using establishment-level data in the Job Openings and Labor Turnover Survey (JOLTS), Davis, Faberman, and Haltiwanger (2012, 2013) find that the job-filling rate increases by less in the aftermath of the Great Recession than is implied by standard search theoretic models. They attribute the discrepancy to variations in recruiting intensity—instruments other than vacancies (such as screening methods and hiring standards) that are used by firms to influence their job-filling rates. Their constructed index of recruiting intensity is highly cyclical, falling by more than 20 percent from early 2007 to late 2009. The authors show that recruiting intensity partly explains the recent breakdown in the matching function, delivers a better-fitting empirical Beveridge curve, and accounts for a large share of fluctuations in aggregate hires.

The model presented in this paper generates cyclical variations in recruiting intensity and movements of the Beveridge curve that resemble the patterns documented by Davis, Faberman, and Haltiwanger (2012, 2013) and Diamond and Şahin (2015). The hiring threshold used by the firms in the model co-moves negatively with changes in the composition of the unemployed pool. In downturns the pool becomes more adversely selected in the sense of MLRP, leading firms to respond by raising their hiring standards. The model thus generates a procyclical recruiting intensity—when the recruiting intensity is measured as the inverse of the hiring threshold used by the firms. The model also generates a shifting out of the Beveridge curve in recessions. The rise in the hiring standards in recessions results in a decline in the probability that a match leads to a hire, thus leading to an increase in the number of vacancies that are needed in a recession to bring about a certain decline in the unemployment rate. This deterioration in the matching process constitutes an outward shift in the Beveridge curve.

9 Concluding Remarks

In this paper I developed an equilibrium search model with worker heterogeneity and adverse selection in the labor market. The model can provide a unified explanation of a number of hitherto unrelated empirical findings. Selective layoff of less-productive workers in recessions leads to a worsening of the distribution of ability in the unemployed pool. This compositional deterioration leads to scarring of workers who lose their jobs in recessions, upskilling by firms, a lowering of recruiting intensity, a shifting out of the Beveridge curve, and a jobless recovery. All these theoretical predictions are consistent with the findings of large empirical literatures. Simple extensions of the model can additionally match the stylized facts on cyclical changes in the average ability of unemployed workers and the cyclicality of separations, quits, and layoffs.

The framework can help uncover policy implications that are elusive in models without heterogeneity. The model economy is inefficient due both to undirected search and asymmetric information. This is in sharp contrast to the efficiency result of Hosios (1990) for directed search models. In addition to the usual inefficiencies of undirected search models, inefficiencies arise in this model since firms do not internalize the informational externalities working through layoffs: a layoff not only affects the firm and the worker involved, but also affects other job seekers by changing the composition of the pool of the unemployed and generating information. This informational externality provides a role for the policy to improve efficiency by influencing the extent to which employers can condition their hiring and layoff decisions on information about individual workers. In future work I intend to investigate the question of optimal policy in this framework.

A Proofs

A.1 Proof of Proposition 1

I prove the proposition through a sequence of simple lemmas.

Lemma 1. $J(\omega, \lambda, A)$ *is strictly increasing in* ω *and increasing in* λ *and* A*.*

Proof. $J(\omega, \lambda, A)$ is strictly increasing in ω since $\ell(\omega|y)$ satisfies the strict MLRP and $J(\omega, \lambda, A)$ is the expectation of $\Pi(y, A)$, which is a strictly increasing function of y, with respect to the nondegenerate distribution λ and conditional on ω.

Let λ_1 and λ_2 be two probability distributions over Y such that λ_1 is larger than λ_2 in the sense of MLRP. Let $\lambda_1(\cdot|\omega)$ and $\lambda_2(\cdot|\omega)$ denote the posterior distributions conditional on ω corresponding to λ_1 and λ_2. By Bayes' rule,

$$\frac{\lambda_1(y|\omega)}{\lambda_2(y|\omega)} = \frac{\frac{\ell(\omega|y)\lambda_1(y)}{\int \ell(\omega|\tilde{y})\lambda_1(\tilde{y})d\tilde{y}}}{\frac{\ell(\omega|y)\lambda_2(y)}{\int \ell(\omega|\tilde{y})\lambda_2(\tilde{y})d\tilde{y}}} = \frac{\int \ell(\omega|\tilde{y})\lambda_2(\tilde{y})d\tilde{y}}{\int \ell(\omega|\tilde{y})\lambda_1(\tilde{y})d\tilde{y}} \frac{\lambda_1(y)}{\lambda_2(y)},$$

which is an increasing function of y by assumption. Therefore, $\lambda_1(\cdot|\omega)$ is larger than $\lambda_2(\cdot|\omega)$ in the sense of MLRP and so in the sense of first-order stochastic dominance. The lemma then follows the fact that $J(\omega, \lambda, A)$ is the expected value of $\Pi(y, A)$ with respect to $\lambda(y|\omega)$ and that Π is increasing in both of its arguments. □

The following is an immediate corollary of the lemma.

Corollary 1. $\underline{\omega}(\lambda, A)$ *is decreasing in its arguments.*

I can use this corollary to prove the following lemma.

Lemma 2. $P(y, \lambda, A)$ *is increasing in its arguments.*

Proof. Note that $P(y, \lambda, A) = 1 - L(\underline{\omega}(\lambda, A)|y)$. Since ℓ satisfies the MLRP, L is decreasing in y and P is increasing in y. Since $\underline{\omega}(\lambda, A)$ is decreasing in λ and A, P is increasing in λ and A. □

Lemma 3. $q(\lambda, A)$ *is decreasing in its arguments.*

Proof. The Fréchet derivative of $q(\lambda, A)$ with respect to the probability measure λ is a function whose value at y' is given by

$$-\frac{1}{\rho k} q(\lambda, A)^2 P(y', \lambda, A)\Psi(y', A),$$

where $\Psi(y,A) = \Pi(y,A) - k$.[22] Consider a small change in λ from λ_0 to $\lambda_1 = \lambda_0 + \delta\lambda$ such that λ_1 is larger than λ_0 in the sense of MLRP. The resulting change in q is given by

$$\begin{aligned}\delta q &= -\frac{1}{\rho k} q(\lambda_0, A)^2 \int P(\tilde{y}, \lambda_0, A)\Psi(\tilde{y}, A)\delta\lambda(\tilde{y})d\tilde{y} \\ &= -\frac{1}{\rho k} q(\lambda_0, A)^2 \left[\int P(\tilde{y}, \lambda_0, A)\Psi(\tilde{y}, A)\lambda_1(\tilde{y})d\tilde{y} - \int P(\tilde{y}, \lambda_0, A)\Psi(\tilde{y}, A)\lambda_0(\tilde{y})d\tilde{y}\right]. \qquad (23)\end{aligned}$$

I need to show that the term in brackets is positive. In order to do this, I prove a few statements about the following auxiliary function:

$$\Omega(y) = P(y, \lambda_0, A)\Psi(y, A) - \frac{\rho k}{q(\lambda_0, A)}.$$

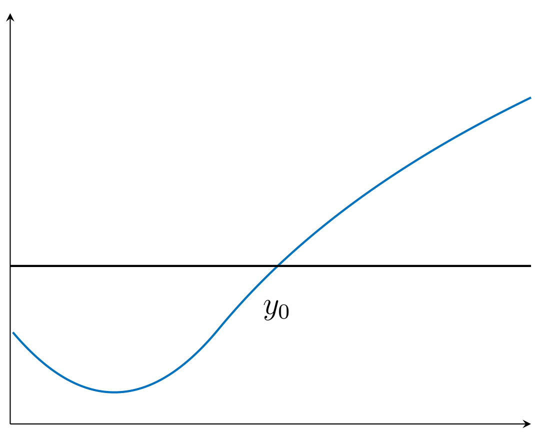
To simplify the exposition I assume that $\Omega(y)$ is differentiable in y, but the properties I prove below are true even without differentiability. First, note that

$$\int \Omega(\tilde{y})\lambda_0(\tilde{y})d\tilde{y} = \int P(\tilde{y}, \lambda_0, A)(\Pi(\tilde{y}, A) - k)\lambda_0(\tilde{y})d\tilde{y} - \frac{\rho k}{q(\lambda_0, A)} = 0, \qquad (24)$$

where I am using the fact that $q(\lambda_0, A)$, by definition, satisfies the free entry condition (7). Second, the derivative of $\Omega(y)$ is given by

$$\Omega'(y) = \frac{\partial P(y, \lambda_0, A)}{\partial y}\Psi(y, A) + P(y, \lambda_0, A)\frac{\partial \Psi(y, A)}{\partial y}.$$

Note that Π and so Ψ are strictly increasing functions of y. So the second term in the above expression is always positive. Furthermore, $\max_{y \in Y} \Psi(y, A)$ is positive, so the first term is also positive for sufficiently large y. Therefore, there exists some y_0 such that $\Omega(y) < 0$ for all $y < y_0$ and $\Omega(y) > 0$ for all $y > y_0$. Finally,



[22] For details of this derivation see Appendix B.1.

$$
\begin{aligned}
\int \Omega(\tilde{y})\lambda_1(\tilde{y})d\tilde{y} &= \int \Omega(\tilde{y})\frac{\lambda_1(\tilde{y})}{\lambda_0(\tilde{y})}\lambda_0(\tilde{y})d\tilde{y} \\
&= \int^{y_0} \Omega(\tilde{y})\frac{\lambda_1(\tilde{y})}{\lambda_0(\tilde{y})}\lambda_0(\tilde{y})d\tilde{y} + \int_{y_0} \Omega(\tilde{y})\frac{\lambda_1(\tilde{y})}{\lambda_0(\tilde{y})}\lambda_0(\tilde{y})d\tilde{y} \\
&\geq \int^{y_0} \Omega(\tilde{y})\frac{\lambda_1(y_0)}{\lambda_0(y_0)}\lambda_0(\tilde{y})d\tilde{y} + \int_{y_0} \Omega(\tilde{y})\frac{\lambda_1(y_0)}{\lambda_0(y_0)}\lambda_0(\tilde{y})d\tilde{y} \\
&= \frac{\lambda_1(y_0)}{\lambda_0(y_0)} \int \Omega(\tilde{y})\lambda_0(\tilde{y})d\tilde{y} = 0,
\end{aligned} \tag{25}
$$

where in the inequality I am using the assumption that $\lambda_1(y)/\lambda_0(y)$ is increasing in y. Equation (23) can be rewritten in terms of Ω as

$$
\delta q = -\frac{1}{\rho k}q(\lambda_0, A)^2 \left[\int \Omega(\tilde{y})\lambda_1(\tilde{y})d\tilde{y} - \int \Omega(\tilde{y})\lambda_0(\tilde{y})d\tilde{y}\right] \leq 0,
$$

where I am using equations (24) and (25). This proves that $q(\lambda, A)$ decreases when λ is increased in the sense of MLRP.

The derivative of $q(\lambda, A)$ with respect to A is given by[23]

$$
\frac{\partial q(\lambda, A)}{\partial A} = -\frac{q(\lambda, A)^2}{\rho k} \int P(\tilde{y}, \lambda, A)\frac{\partial \Psi(\tilde{y}, A)}{\partial A}\lambda(\tilde{y})d\tilde{y},
$$

which is negative since Π and so Ψ are increasing functions of A. This completes the proof of the lemma. □

Proof of proposition 1. The first part of the proposition is proved in corollary 1. The second part of the proposition is a consequence of the fact that θ is a decreasing function of q. This also implies that $p(\lambda, A)$ is increasing in its arguments. The last part of the proposition is due to $P(y, \lambda, A)$ and $p(\lambda, A)$ both being positive and increasing functions of their arguments. □

A.2 Proof of Proposition 2

By equation (14), the logarithm of the share of unemployed workers who have ability y at time $t + \delta t$ is given by

$$
\log \lambda_{t+\delta t}(y) = \log \lambda_t(y) + \int_t^{t+\delta t} \left[\Gamma(y, u_{\tilde{t}}, A_{\tilde{t}}) - \int \Gamma(\tilde{y}, u_{\tilde{t}}, A_{\tilde{t}})\lambda_{\tilde{t}}(\tilde{y})d\tilde{y}\right] d\tilde{t}.
$$

Note that the integral is not an Itô integral; it is a Riemann integral of the sample path of a diffusion process. It is easy to verify that $\Gamma(y, u_t, A_t)$ has a continuous path whenever A_t has a continuous path. But A_t is a diffusion process with paths that are continuous almost surely. Therefore for small δt the above equation can be approximated by

$$
\log \lambda_{t+\delta t}(y) \approx \log \lambda_t(y) + \left[\Gamma(y, u_t, A_t) - \int \Gamma(\tilde{y}, u_t, A_t)\lambda_t(\tilde{y})d\tilde{y}\right] \delta t.
$$

[23] For details of this derivation see Appendix B.2.

Let y and y' be two arbitrary abilities. The above equation implies that

$$\log \lambda_{t+\delta t}(y) - \log \lambda_{t+\delta t}(y') \approx \log \lambda_t(y) - \log \lambda_t(y') + \left[\Gamma(y, u_t, A_t) - \Gamma(y', u_t, A_t)\right] \delta t.$$

Therefore,

$$\log \frac{\lambda_{t+\delta}(y)}{\lambda_t(y)} - \log \frac{\lambda_{t+\delta}(y')}{\lambda_t(y')} \approx \left[\Gamma(y, u_t, A_t) - \Gamma(y', u_t, A_t)\right] \delta t.$$

If $\Gamma(y, u_t, A_t)$ is an increasing function of y, the right-hand side of the above equation will be positive for all $y > y'$. So the left-hand side will also be positive for all $y > y'$. This implies that $\lambda_{t+\delta t}(y)/\lambda_t(y)$ is an increasing function of y. Thus, $\lambda_{t+\delta t}$ is larger than λ_t in the sense of MLRP.

□

A.3 Proof of Proposition 3

By equation (14), the logarithm of the share of unemployed workers at time δt who have ability y can be approximated for small δt by

$$\log \lambda_{\delta t}(y) \approx \log \lambda_0(y) + \left[\Gamma(y, u_0, A_0) - \int \Gamma(\tilde{y}, u_0, A_0)\lambda_0(\tilde{y}) d\tilde{y}\right] \delta t.$$

Therefore, for all y and y',

$$\begin{aligned} \log \frac{\lambda_{\delta t}(y)}{\lambda_0(y)} - \log \frac{\lambda_{\delta t}(y')}{\lambda_0(y')} &\approx \left[\Gamma(y, u_0, A_0) - \Gamma(y', u_0, A_0)\right] \delta t \\ &= \int_{y'}^{y} \left.\frac{\partial \Gamma(y, u_0, A_0)}{\partial y}\right|_{y=\tilde{y}} d\tilde{y}\delta t. \end{aligned}$$

Now consider an alternative scenario in which the state of the economy at time zero is increased to $A_0 + \delta A_0$ for some small positive δA_0 while u_0 (and hence λ_0) is kept fixed. Let $\lambda'_{\delta t}$ denote the logarithm of the distribution of unemployed workers at time δt under this alternative scenario. By a similar argument,

$$\begin{aligned} \log \frac{\lambda'_{\delta t}(y)}{\lambda_0(y)} - \log \frac{\lambda'_{\delta t}(y')}{\lambda_0(y')} &\approx \int_{y'}^{y} \left.\frac{\partial \Gamma(y, u_0, A_0 + \delta A_0)}{\partial y}\right|_{y=\tilde{y}} d\tilde{y}\delta t \\ &\approx \int_{y'}^{y} \left.\frac{\partial \Gamma(y, u_0, A_0)}{\partial y}\right|_{y=\tilde{y}} d\tilde{y}\delta t + \int_{y'}^{y} \left.\frac{\partial^2 \Gamma(y, u_0, A)}{\partial A \partial y}\right|_{y=\tilde{y}, A=A_0} d\tilde{y}\delta t \delta A. \end{aligned}$$

Therefore,

$$\begin{aligned} \log \frac{\lambda'_{\delta t}(y)}{\lambda_{\delta t}(y)} - \log \frac{\lambda'_{\delta t}(y')}{\lambda_{\delta t}(y')} &= \left(\log \frac{\lambda'_{\delta t}(y)}{\lambda_0(y)} - \log \frac{\lambda'_{\delta t}(y')}{\lambda_0(y')}\right) - \left(\log \frac{\lambda_{\delta t}(y)}{\lambda_0(y)} - \log \frac{\lambda_{\delta t}(y')}{\lambda_0(y')}\right) \\ &= \int_{y'}^{y} \left.\frac{\partial^2 \Gamma(y, u_t, A)}{\partial A \partial y}\right|_{y=\tilde{y}, A=A_t} d\tilde{y}\delta t \delta A. \end{aligned}$$

If $\partial^2 \Gamma(y, u, A)/\partial A \partial y$ is positive, then the right-hand side is positive for all $y > y'$, which implies that $\lambda'_{\delta t}(y)/\lambda_{\delta t}(y)$ is increasing in y. Therefore, if $\partial^2 \Gamma(y, u, A)/\partial A \partial y$ is positive at $u = u_t$ and $A = A_0$, then $\lambda_{\delta t}$ increases in the sense of MLRP when A_0 is increased to $A_0 + \delta A_0$. □

A.4 Proof of Proposition 4

I prove the proposition for the case that $s(y, A)$ is log-supermodular. The case of a log-submodular separation rate is proved identically. Let $\Psi(y, A) = \Pi(y, A) - k$. Given that the monitoring structure is additive,

$$L_\sigma(\omega|y) = \Phi\left(\frac{\omega - y}{\sigma}\right),$$

and so

$$\ell_\sigma(\omega|y) = \frac{1}{\sigma}\Phi'\left(\frac{\omega - y}{\sigma}\right),$$

I use a σ subscript to make the dependence of endogenous variables on σ clear. The indifference equation that defines $\underline{\omega}_\sigma$ can be written as

$$\int \Psi(\tilde{y}, A^*)\Phi'\left(\frac{\underline{\omega}_\sigma(\lambda^*_\sigma, A^*) - \tilde{y}}{\sigma}\right)\lambda^*_\sigma(\tilde{y})d\tilde{y} = 0, \tag{26}$$

where λ^*_σ denotes the steady state distribution of ability among the unemployed corresponding to A^* when the monitoring structure is given by ℓ_σ. The free entry condition is given by

$$q_\sigma(\lambda^*_\sigma, A^*)\int \Psi(\tilde{y}, A^*)\left(1 - \Phi\left(\frac{\underline{\omega}_\sigma(\lambda^*_\sigma, A^*) - \tilde{y}}{\sigma}\right)\right)\lambda^*_\sigma(\tilde{y})d\tilde{y} = \rho k. \tag{27}$$

I first show that, as σ goes to infinity, $\underline{\omega}_\sigma(\lambda^*_\sigma, A^*) \to -\infty$ and $(\underline{\omega}_\sigma(\lambda^*_\sigma, A^*) - y)/\sigma \to -\infty$, and $q_\sigma(\lambda^*_\sigma, A^*)$ and $p_\sigma(\lambda^*_\sigma, A^*)$ remain bounded. Consider a sequence $\{\sigma_k\}$ that goes to infinity. I consider various possible cases for $\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*)$.

Suppose that there is a subsequence $\{j_k\}$ over which $\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*)$ is convergent. Then,

$$\Phi'\left(\frac{\underline{\omega}_{\sigma_{j_k}}(\lambda^*_{\sigma_{j_k}}, A^*) - y}{\sigma_{j_k}}\right) \to \Phi'(0).$$

Since Y is compact, by Prokhorov's theorem, $\{\lambda^*_{\sigma_{j_k}}\}$ has a weak* limit, λ^*_∞, possibly by going to another subsequence. Therefore, since Ψ is continuous in y and bounded, in the limit, equation (26) is given by

$$\int \Psi(\tilde{y}, A^*)\Phi'(0)\lambda^*_\infty(\tilde{y})d\tilde{y} = 0,$$

and since $\Phi'(0) > 0$,

$$\int \Psi(\tilde{y}, A^*)\lambda^*_\infty(\tilde{y})d\tilde{y} = 0.$$

Furthermore,

$$\left(1 - \Phi\left(\frac{\underline{\omega}_{\sigma_{j_k}}(\lambda^*_{\sigma_{j_k}}, A^*) - \tilde{y}}{\sigma_{j_k}}\right)\right) \to (1 - \Phi(0)) > 0.$$

The last two equations together with the free entry equation (27) and Prokhorov's theorem imply that, over a further subsequence, $q_{\sigma_{j_k}}(\lambda^*_{\sigma_{j_k}}, A^*) \to \infty$, and so $p_{\sigma_{j_k}}(\lambda^*_{\sigma_{j_k}}, A^*) \to 0$ and $f_{\sigma_{j_k}}(y, \lambda^*_{\sigma_{j_k}}, A^*) \to 0$ for all y. This together with equation (17) implies that $\lambda^*_\infty(y) = g(y)$. Therefore,

$$\int \Psi(\tilde{y}, A^*)g(\tilde{y})d\tilde{y} = \int \Psi(\tilde{y}, A^*)\lambda^*_\infty(\tilde{y})d\tilde{y} = 0,$$

which is a contradiction with the assumption that $\int \Psi(\tilde{y}, A^*)g(\tilde{y})d\tilde{y} > 0$.

Therefore, the absolute value of $\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*)$ must grow without bound. Since y is bounded and Φ' has a compact range, the following sequence must have a convergent subsequence whose limit is independent of y:

$$\Phi'\left(\frac{\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*) - y}{\sigma_k}\right)$$

Let Φ'_∞ denote this limit. Using equation (26), and again using Prokhorov's theorem and possibly going to a subsequence, I get that

$$\Phi'_\infty \int \Psi(\tilde{y}, A^*)\lambda^*_\infty(\tilde{y})d\tilde{y} = 0.$$

If $\Phi'_\infty > 0$, then $(\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*) - y)/\sigma_k$ must have a finite limit over the subsequence, because Φ is assumed to have support on the entire reals. Then by an argument similar to the one used above, I reach a contradiction with the free entry condition.

Therefore, it must be the case that $\Phi'_\infty = 0$, and so that the absolute value of $(\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*) - y)/\sigma_k$ grows unbounded on any subsequence. Suppose that $(\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*) - y)/\sigma_k$ goes to $+\infty$ over some subsequence. Then $\Phi((\underline{\omega}_{\sigma_k}(\lambda^*_{\sigma_k}, A^*) - y)/\sigma_k)$ goes to one over that subsequence. This would again imply using the free entry condition that $q_{\sigma_k}(\lambda^*_{\sigma_k}, A^*) \to \infty$ and so $\lambda^*_{\sigma_k} \to g$ in the weak* topology over some subsequence. Therefore,

$$\int \Psi(\tilde{y}, A^*)\lambda^*_{\sigma_k}(\tilde{y})d\tilde{y} \to \int \Psi(\tilde{y}, A^*)\lambda^*_\infty(\tilde{y})d\tilde{y} = \int \Psi(\tilde{y}, A^*)g(\tilde{y})d\tilde{y} > 0.$$

This contradicts the hypothesis that firms are optimally using a threshold that grows to $+\infty$, because they can strictly benefit from using a threshold that is equal to $-\infty$ instead. So, it must be the case that

$$\lim_{\sigma\to\infty} \frac{\underline{\omega}_\sigma(\lambda^*_\sigma, A) - y}{\sigma} = -\infty.$$

It is also easy to see why $q_\sigma(\lambda^*_\sigma, A^*)$ and $p_\sigma(\lambda^*_\sigma, A^*)$ must remain bounded as $\sigma \to \infty$. Suppose that $q_\sigma(\lambda^*_\sigma, A^*)$ goes to infinity. Then $p_\sigma(\lambda^*_\sigma, A^*)$ and $f_{\sigma_{j_k}}(y, \lambda^*_{\sigma_{j_k}}, A^*)$ go to zero and λ^*_σ converges to g in the weak* topology. This implies that the returns to posting additional vacancies is strictly positive in the limit $\sigma \to \infty$, a contradiction to the assumption that there are a vanishing number of vacancies in the limit. Suppose that $p_\sigma(\lambda^*_\sigma, A^*)$ goes to infinity. Then the net present value of a vacancy is negative. This contradicts the existence of a positive number of vacancies in the limit $\sigma \to \infty$.

I next argue that the last term in equation (18) goes to zero as σ goes to infinity. Implicitly differentiating equation (26) with respect to A and evaluating at A^* results in

$$\int \frac{\partial \Psi(\tilde{y}, A^*)}{\partial A} \Phi' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y} + \frac{1}{\sigma} \frac{\partial \underline{\omega}_\sigma(\lambda_\sigma^*, A^*)}{\partial A} \int \Psi(\tilde{y}, A^*) \Phi'' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y} = 0.$$

Therefore,

$$\frac{\partial \underline{\omega}_\sigma(\lambda_\sigma^*, A^*)}{\partial A} = \frac{-\sigma \displaystyle\int \frac{\partial \Psi(\tilde{y}, A^*)}{\partial A} \Phi' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y}}{\displaystyle\int \Psi(\tilde{y}, A^*) \Phi'' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y}}.$$

So the last term in equation (18) is given by

$$\frac{\partial \ell(\underline{\omega}_\sigma(\lambda_\sigma^*, A^*)|y)}{\partial y} \frac{\partial \underline{\omega}_\sigma(\lambda_\sigma^*, A^*)}{\partial A} = \frac{1}{\sigma} \Phi'' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - y}{\sigma} \right) \frac{\displaystyle\int \frac{\partial \Psi(\tilde{y}, A^*)}{\partial A} \Phi' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y}}{\displaystyle\int \Psi(\tilde{y}, A^*) \Phi'' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y}}.$$

Given that y belongs to a compact set, as σ goes to infinity, λ_σ^* converges to some λ_∞^* in the weak* topology, possibly by going to a subsequence. Therefore, since $(\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - y)/\sigma \to -\infty$ and so $\Phi''\left((\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - y)/\sigma\right) \to 0$, and using the dominated convergence theorem, I have that

$$\frac{\displaystyle\int \Psi(\tilde{y}, A^*) \Phi'' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y}}{\Phi'' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - y}{\sigma} \right)} \to \int \Psi(\tilde{y}, A^*) \lambda_\infty^*(\tilde{y}) d\tilde{y} > 0,$$

where the inequality follows from the free entry condition. This together with the assumptions that y belongs to a compact set and Φ' is bounded proves that the last term in equation (18) goes to zero as σ goes to infinity.

I next focus on the second term. First note that[24]

$$\frac{\partial p_\sigma(\lambda_\sigma^*, A^*)}{\partial A} = \left[\frac{1}{\rho k} m' \left(\frac{1}{\rho k} \int \Psi(\tilde{y}, A^*) \left(1 - \Phi \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \right) \lambda_\sigma^*(\tilde{y}) d\tilde{y} \right) \right] \times \left[\int \left(1 - \Phi \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - \tilde{y}}{\sigma} \right) \right) \frac{\partial \Psi(\tilde{y}, A^*)}{\partial A} \lambda_\sigma^*(\tilde{y}) d\tilde{y} \right],$$

which remains bounded as σ goes to infinity given that y belongs to a compact set and $q_\sigma(\lambda_\sigma^*, A^*)$ remains bounded as σ goes to infinity. I have also shown that $p_\sigma(\lambda_\sigma^*, A^*)$ remains bounded as σ goes to infinity. Therefore, since $s(y, A^*)$ is positive and independent of σ, the following expression remains bounded as σ goes to infinity:

$$\frac{\partial p_\sigma(\lambda_\sigma^*, A^*)}{\partial A} - \frac{p_\sigma(y, \lambda_\sigma^*)}{s(y, A^*)} \frac{\partial s(y, A^*)}{\partial A}$$

Finally,

$$\frac{\partial L_\sigma(\underline{\omega}_\sigma(\lambda_\sigma^*, A^*)|y)}{\partial y} = -\frac{1}{\sigma} \Phi' \left(\frac{\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - y}{\sigma} \right),$$

[24]The details of this derivation are presented in Appendix B.2.

which goes to zero as σ goes to infinity since Φ' is bounded. Therefore, the second term in equation (18) also goes to zero as σ goes to infinity.

The first term in equation (18) is itself the product of two terms: The first one is $f_\sigma(y, \lambda_\sigma^*, A^*) = p_\sigma(\lambda_\sigma^*, A^*)(1 - \Phi((\underline{\omega}_\sigma(\lambda_\sigma^*, A^*) - y)/\sigma)$, which gets close to $p_\sigma(\lambda_\sigma^*, A^*)$ and so remains positive as σ goes to infinity. The second one is independent of σ and positive by strict log-supermodularity of $s(y, A)$. So the first term in equation (18) dominates the expression for sufficiently large σ. This completes the proof of the proposition. □

A.5 Proof of Proposition 6

Let $l_t^\tau(y; y') = \log(\lambda_t^\tau(y)/\lambda_t^\tau(y'))$. To prove the proposition, it is sufficient to show that $l_t^\tau(y; y')$ is decreasing over time at all t and for all $y > y'$. By equation (22),

$$\frac{d}{dt} l_t^\tau(y; y') = -p(\theta_t) \left[P(y, \lambda_t^\tau, A_t) - P(y', \lambda_t^\tau, A_t) \right],$$

which is negative for all $y > y'$, all A_t, and all $t \geq \tau$, since $p(\theta_t)$ is positive and P is an increasing function of y by Lemma 2. □

A.6 Proof of Proposition 8

I prove the proposition for the case that $s(y, A)$ is log-supermodular. The case of a log-submodular separation rate is proved identically. The initial distribution of ability in the cohort separated at time τ is given by

$$\lambda_\tau^\tau(y) = \frac{s(y, A_\tau)(1 - u_\tau(y))}{\displaystyle\int s(\tilde{y}, A_\tau)(1 - u_\tau(\tilde{y}))d\tilde{y}}. \tag{28}$$

Let $y > y'$. By equation (28),

$$\log \frac{\lambda_\tau^\tau(y)}{\lambda_\tau^\tau(y')} = \left[\log s(y, A_\tau) - \log s(y', A_\tau)\right] + \left[\log(1 - u_\tau(y)) - \log(1 - u_\tau(y'))\right].$$

The first term is strictly increasing in A_τ since s is strictly log-supermodular. The second term is independent of A_τ since u_τ is predetermined at time τ. Therefore, an increase in A_τ results in an increase in $\lambda_\tau^\tau(y)/\lambda_\tau^\tau(y')$ for all $y > y'$, and so an increase in λ_τ^τ in the sense of MLRP.

Now consider a continuous path for labor productivity $\{A_t\}_{t\geq 0}$ and another continuous path $\{A_t'\}_{t\geq 0}$ that is identical to $\{A_t\}_{t\geq 0}$ except in the interval $(\tau - \epsilon, \tau + \epsilon)$ over which $A_t' > A_t$. I keep $A_\tau' - A_\tau$ constant as I make ϵ small. Let λ_t^τ and $(\lambda')_t^\tau$ denote the time t distribution of ability in cohort τ given $\{A_t\}_{t\geq 0}$ and $\{A_t'\}_{t\geq 0}$ respectively. By assumption $A_{\tau+\epsilon} = A_{\tau+\epsilon}'$. Furthermore, if ϵ is sufficiently small, then $\theta_{\tau+\epsilon}$ is arbitrarily close to $\theta_{\tau+\epsilon}'$. Therefore, by Proposition 5, ϵ can be chosen sufficiently small such that

$$f\left(y, (\lambda')_\tau^\tau, A_\tau', \theta_\tau'\right) > f(y, \lambda_\tau^\tau, A_{\tau+\epsilon}, \theta_{\tau+\epsilon}),$$

Given that the distribution of ability in a cohort changes continuously over time and f is continuous in λ, if ϵ is sufficiently small, then

$$f\left(y, (\lambda')^{\tau}_{\tau+\epsilon}, A'_{\tau+\epsilon}, \theta'_{\tau+\epsilon}\right) > f(y, \lambda^{\tau}_{\tau+\epsilon}, A_{\tau+\epsilon}, \theta_{\tau+\epsilon}).$$

□

A.7 Proof of Proposition 9

I first prove a lemma about the shape of the value function in a model with no machine breakdowns (that is, with $\delta = 0$). The proposition is then proved using a perturbation method.

Lemma 4. *Let $\Pi_0(y, A)$ be the solution to the HJB equation (1) with $\delta = 0$, $w(y, A) = w_0(y) + \beta(y)A$, and $h(c) = \alpha \underline{c}^{\alpha} c^{-\alpha-1}$, and let $\gamma_0(y, A) = 1 - H(\Pi_0(y, A))$.*

1. *Suppose that $\partial^2_{yA}\pi(y, A) > 0$. There exists $\alpha^*, K > 0$ such that if $\alpha > \alpha^*$ and $w_0'(y) < K$ for all y, then $\partial^2_{yA}\gamma_0(y, A) > 0$ for all $(y, A) \in Y \times \left(\underline{A}, \overline{A}\right)$.*

2. *Suppose that $\partial^2_{yA}\pi(y, A) < 0$. There exists $\alpha^*, K > 0$ such that if $\alpha > \alpha^*$ and $w_0'(y) > -K$ for all y, then $\partial^2_{yA}\gamma_0(y, A) < 0$ for all $(y, A) \in Y \times \left(\underline{A}, \overline{A}\right)$.*

Proof. To simplify the notation let $r = \rho + \varsigma$ denote the effective interest rate. When $\delta = 0$ the HJB equation can be written as

$$r\Pi(y, A) = (y - \beta(y))A + \varsigma k - w_0(y) + \mu A \frac{\partial \Pi(y, A)}{\partial A} + \mu A^2 \frac{\partial^2 \Pi(y, A)}{\partial A^2}. \tag{29}$$

Any solution to equation (29) is given by

$$\Pi_0(y, A) = \frac{y - \beta(y)}{r - \mu} A + \frac{\varsigma k - w_0(y)}{r} + B_1(y) A^{-\theta} + B_2(y) A^{\theta},$$

where $\theta = \sqrt{r/\mu} > 1$ and $B_1(y)$ and $B_2(y)$ are determined by the boundary conditions:

$$B_1(y) = \frac{y - \beta(y)}{\theta(r - \mu)} \left(\frac{\overline{A}^{2\theta} \underline{A}^{\theta+1} - \overline{A}^{\theta+1} \underline{A}^{2\theta}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}} \right),$$
$$B_2(y) = \frac{-(y - \beta(y))}{\theta(r - \mu)} \left(\frac{\overline{A}^{\theta+1} - \underline{A}^{\theta+1}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}} \right).$$

Define

$$B_1 = \frac{\overline{A}^{2\theta} \underline{A}^{\theta+1} - \overline{A}^{\theta+1} \underline{A}^{2\theta}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}},$$
$$B_2 = \frac{\overline{A}^{\theta+1} - \underline{A}^{\theta+1}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}},$$

and let

$$\kappa(y) = \varsigma k - w_0(y),$$
$$\Xi(A) = \frac{r}{r-\mu}\left(A + \frac{1}{\theta}B_1 A^{-\theta} - \frac{1}{\theta}B_2 A^{\theta}\right).$$

Then $\Pi_0(y, A)$ can be written compactly as

$$\Pi_0(y, A) = \frac{1}{r}\left(\frac{\partial \pi(y, A)}{\partial A}\Xi(A) + \kappa(y)\right),$$

where I am using the fact that $y - \beta(y) = \partial\pi(y, A)/\partial A$. The layoff rate is given by

$$\gamma_0(y, A) = \underline{c}^{\alpha}\Pi_0(y, A)^{-\alpha}.$$

So

$$\begin{aligned}\frac{\partial^2\gamma_0(y, A)}{\partial y \partial A} &= \frac{\alpha\underline{c}^{\alpha}}{\Pi_0(y, A)^{\alpha+2}}\left[(1+\alpha)\frac{\partial \Pi_0(y, A)}{\partial y}\frac{\partial \Pi_0(y, A)}{\partial A} - \Pi_0(y, A)\frac{\partial^2\Pi_0(y, A)}{\partial y \partial A}\right] \\ &= \frac{\alpha\underline{c}^{\alpha}\Xi'(A)}{r^2\Pi_0(y, A)^{\alpha+2}}\left[\alpha\frac{\partial\pi(y, A)}{\partial A}\frac{\partial^2\pi(y, A)}{\partial y\partial A}\Xi(A) + (1+\alpha)\kappa'(y)\frac{\partial \pi(y, A)}{\partial A} - \kappa(y)\frac{\partial^2\pi(y, A)}{\partial y \partial A}\right].\end{aligned} \tag{30}$$

I next show that $\Xi(A)$ is positive and increasing for all $A \in (\underline{A}, \overline{A})$. This will allow me to sign the expression in equation (30). The derivative of Ξ is given by

$$\begin{aligned}\Xi'(A) &= \frac{r}{r-\mu}\left[1 - B_1 A^{-\theta-1} - B_2 A^{\theta-1}\right] \\ &= \frac{r}{r-\mu}\left[1 - \left(\frac{\overline{A}^{2\theta}\underline{A}^{\theta+1} - \overline{A}^{\theta+1}\underline{A}^{2\theta}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}}\right)A^{-\theta-1} - \left(\frac{\overline{A}^{\theta+1} - \underline{A}^{\theta+1}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}}\right)A^{\theta-1}\right] \\ &= \frac{r}{(r-\mu)A^{\theta+1}}\left[A^{\theta+1} - \frac{\overline{A}^{2\theta}\underline{A}^{\theta+1} + \overline{A}^{\theta+1}\underline{A}^{2\theta} - \overline{A}^{\theta+1}A^{2\theta} + \underline{A}^{\theta+1}A^{2\theta}}{\overline{A}^{2\theta} - A^{2\theta}}\right] \\ &= \frac{r}{(r-\mu)A^{\theta+1}}\left[A^{\theta+1} - \frac{\overline{A}^{2\theta} - A^{2\theta}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}}\underline{A}^{\theta+1} - \frac{A^{2\theta} - \underline{A}^{2\theta}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}}\overline{A}^{\theta+1}\right].\end{aligned}$$

Ξ is increasing in A if and only if the term in brackets is positive. Define $x_1 = \underline{A}^{2\theta}$, $x_2 = \overline{A}^{2\theta}$, and $x = A^{2\theta}$. Then $x = \eta x_1 + (1-\eta)x_2$, where $\eta = (x_2 - x)/(x_2 - x_1) \in [0, 1]$. So the term in brackets can be written as

$$A^{\theta+1} - \eta\underline{A}^{\theta+1} - (1-\eta)\overline{A}^{\theta+1} = (\eta x_1 + (1-\eta)x_2)^{\frac{\theta+1}{2\theta}} - \eta x_1^{\frac{\theta+1}{2\theta}} - (1-\eta)x_2^{\frac{\theta+1}{2\theta}}.$$

Note that since $\theta > 1$, the function $x \mapsto x^{(\theta+1)/2\theta}$ is strictly concave. Therefore, by Jensen's inequality,

$$(\eta x_1 + (1-\eta)x_2)^{\frac{\theta+1}{2\theta}} - \eta x_1^{\frac{\theta+1}{2\theta}} - (1-\eta)x_2^{\frac{\theta+1}{2\theta}} \geq 0,$$

with inequality strict whenever $\underline{A} < A < \overline{A}$. This proves that Ξ is strictly increasing in A over $(\underline{A}, \overline{A})$.

Given that Ξ is increasing, I only need to show that $\Xi\left(\underline{A}\right) > 0$ in order to prove that $\Xi(A)$ is positive for all $A \in \left[\underline{A}, \overline{A}\right]$:

$$\begin{aligned}
\Xi\left(\underline{A}\right) &= \frac{r}{r-\mu}\left[\underline{A} + \frac{1}{\theta}B_1\underline{A}^{-\theta} - \frac{1}{\theta}B_2\underline{A}^{\theta}\right] \\
&= \underline{A} + \frac{r}{(r-\mu)\theta}\left(\frac{\overline{A}^{2\theta}\underline{A}^{\theta+1} - \overline{A}^{\theta+1}\underline{A}^{2\theta}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}}\right)\underline{A}^{-\theta} - \frac{1}{\theta}\left(\frac{\overline{A}^{\theta+1} - \underline{A}^{\theta+1}}{\overline{A}^{2\theta} - \underline{A}^{2\theta}}\right)\underline{A}^{\theta} \\
&= \frac{r}{(r-\mu)\theta\left(\overline{A}^{2\theta} - \underline{A}^{2\theta}\right)}\left[\theta\overline{A}^{2\theta}\underline{A} - \theta\underline{A}^{2\theta+1} + \overline{A}^{2\theta}\underline{A} - \overline{A}^{\theta+1}\underline{A}^{\theta} - \overline{A}^{\theta+1}\underline{A}^{\theta} + \underline{A}^{2\theta+1}\right] \\
&= \frac{r}{(r-\mu)\theta\left(\overline{A}^{2\theta} - \underline{A}^{2\theta}\right)}\left[(\theta+1)\overline{A}^{2\theta}\underline{A} - (\theta-1)\underline{A}^{2\theta+1} - 2\overline{A}^{\theta+1}\underline{A}^{\theta}\right] \\
&= \frac{r}{(r-\mu)\theta\left(\overline{A}^{2\theta} - \underline{A}^{2\theta}\right)}\left[\left((\theta+1)\overline{A}^{\theta-1} - 2\underline{A}^{\theta-1}\right)\overline{A}^{\theta+1}\underline{A} - (\theta-1)\underline{A}^{2\theta+1}\right] \\
&> \frac{r}{(r-\mu)\theta\left(\overline{A}^{2\theta} - \underline{A}^{2\theta}\right)}\left[\left((\theta+1)\underline{A}^{\theta-1} - 2\underline{A}^{\theta-1}\right)\overline{A}^{\theta+1}\underline{A} - (\theta-1)\underline{A}^{2\theta+1}\right] \\
&= \frac{r(\theta-1)}{(r-\mu)\theta\left(\overline{A}^{2\theta} - \underline{A}^{2\theta}\right)}\left[\overline{A}^{\theta+1}\underline{A}^{\theta} - \underline{A}^{2\theta+1}\right] > 0,
\end{aligned}$$

where the inequalities are both consequences of the assumptions that $\theta > 1$ and $\overline{A} > \underline{A}$.

I can now determine the sign of the expression in equation (30). Given that $\Xi'(y) > 0$, the sign of $\partial^2_{yA}\gamma_0(y, A)$ is determined by the sign of the following expression:

$$\alpha\frac{\partial\pi(y,A)}{\partial A}\frac{\partial^2\pi(y,A)}{\partial y\partial A}\Xi(A) + (1+\alpha)\kappa'(y)\frac{\partial\pi(y,A)}{\partial A} - \kappa(y)\frac{\partial^2\pi(y,A)}{\partial y\partial A}.$$

Note that α does not show up in π, κ, or Ξ, so as α gets large, the sign of the above expression is determined by the sign of the following expression

$$\frac{\partial\pi(y,A)}{\partial A}\frac{\partial^2\pi(y,A)}{\partial y\partial A}\Xi(A) - w_0'(y)\frac{\partial\pi(y,A)}{\partial A}.$$

Thus, for large α, if $\partial^2_{yA}\pi(y, A)$ is positive and $w_0'(y)$ is not too positive, then $\partial^2_{yA}\gamma_0(y, A) > 0$, and if $\partial^2_{yA}\pi(y, A)$ is negative and $w_0'(y)$ is not too negative, then $\partial^2_{yA}\gamma_0(y, A) < 0$. □

Proof of proposition 9. I prove the proposition for the case that $\pi(y, A)$ is supermodular. The other case is proved identically. Let $\gamma(y, A) = 1 - H(\Pi(y, A))$ denote the probability of a layoff conditional on a machine breakdown. The separation rate is log-supermodular if the following expression is positive:

$$\frac{\partial^2}{\partial y\partial A}\log s(y,A) = \frac{1}{(\varsigma + \delta\gamma(y,A))^2}\left[\delta(\varsigma + \delta\gamma(y,A))\frac{\partial^2\gamma(y,A)}{\partial y\partial A} - \delta^2\frac{\partial\gamma(y,A)}{\partial y}\frac{\partial\gamma(y,A)}{\partial A}\right].$$

The assumption that $\Pi(y, A)$ and its first two derivatives vary smoothly with δ implies that $\partial^2_{yA}\gamma(y, A)$ is a smooth function of δ. Therefore, by Taylor's theorem,

$$\frac{\partial^2}{\partial y\partial A}\log s(y,A) = \frac{\delta}{\varsigma}\frac{\partial^2\gamma_0(y,A)}{\partial y\partial A} + R(\delta; y, A), \tag{31}$$

where the residual R is second order:

$$\lim_{\delta\to 0}\frac{R(\delta;y,A)}{\delta}=0,$$

and convergence is uniform in y and A since y and A belong to compact sets. Now, by Lemma 4, for any open interval $\mathcal{A}$ contained in $[\underline{A},\overline{A}]$, I can choose a closed set $\mathcal{C}\subset[\underline{y},\overline{y}]\times[\underline{A},\overline{A}]$ that contains $Y\times\mathcal{A}$, a sufficiently large α, and a sufficiently small bound on w_0' such that $\gamma_0(y,A)$ is strictly supermodular over $\mathcal{C}$. By equation (31), then, I can additionally choose δ sufficiently small to guarantee that $\log s(y,A)$ is strictly supermodular over $\mathcal{C}$ and hence over $Y\times\mathcal{A}$. □

B Algebra

B.1 Derivatives with respect to λ

Throughout the proof, I let $\Psi(y,A)=\Pi(y,A)-k$ and let $\partial\mathcal{F}(\lambda)/\partial\lambda(y')$ denote the Fréchet derivative of function $\mathcal{F}$ of λ with respect to λ evaluated at point $y'\in Y$.

The indifference equation that defines $\underline{\omega}(\lambda,A)$, equation (6), can be written as

$$\int\Psi(\tilde{y},A)\ell(\underline{\omega}(\lambda,A)|\tilde{y})\lambda(\tilde{y})d\tilde{y}=0. \tag{32}$$

Differentiating with respect to λ,

$$\Psi(y',A)\ell(\underline{\omega}(\lambda,A)|y')+\frac{\partial}{\partial\omega}\left(\int\Psi(\tilde{y},A)\ell(\omega|\tilde{y})\lambda(\tilde{y})d\tilde{y}\right)\Big|_{\omega=\underline{\omega}(\lambda,A)}\frac{\partial\underline{\omega}(\lambda,A)}{\partial\lambda(y')}=0.$$

Therefore,

$$\frac{\partial\underline{\omega}(\lambda,A)}{\partial\lambda(y')}=\frac{-\Psi(y',A)\ell(\underline{\omega}(\lambda,A)|y')}{C(\lambda,A)},$$

where

$$C(\lambda,A)=\frac{\partial}{\partial\omega}\left(\int\Psi(\tilde{y},A)\ell(\omega|\tilde{y})\lambda(\tilde{y})d\tilde{y}\right)\Big|_{\omega=\underline{\omega}(\lambda,A)}. \tag{33}$$

Next note that $P(y,\lambda,A)=1-L(\underline{\omega}(\lambda,A)|y)$. Therefore,

$$\frac{\partial P(y,\lambda,A)}{\partial\lambda(y')}=-\ell(\underline{\omega}(\lambda,A)|y)\frac{\partial\underline{\omega}(\lambda,A)}{\partial\lambda(y')}=\frac{\ell(\underline{\omega}(\lambda,A)|y)\Psi(y',A)\ell(\underline{\omega}(\lambda,A)|y')}{C(\lambda,A)}.$$

The free entry condition, equation (7), can be rewritten in terms of Ψ as

$$q(\lambda,A)\int P(\tilde{y},\lambda,A)\Psi(\tilde{y},A)\lambda(\tilde{y})d\tilde{y}=\rho k.$$

Differentiating with respect to λ,

$$\frac{\partial q(\lambda,A)}{\partial\lambda(y')}\frac{\rho k}{q(\lambda,A)}+q(\lambda,A)\left[P(y',\lambda,A)\Psi(y',A)+\int\frac{\partial P(\tilde{y},\lambda,A)}{\partial\lambda(y')}\Psi(\tilde{y},A)\lambda(\tilde{y})d\tilde{y}\right]=0.$$

Therefore,

$$\frac{\partial q(\lambda, A)}{\partial \lambda(y')} = -\frac{q(\lambda, A)^2}{\rho k} \left[P(y', \lambda, A)\Psi(y', A) + \int \frac{\partial P(\tilde{y}, \lambda, A)}{\partial \lambda(y')} \Psi(\tilde{y}, A)\lambda(\tilde{y}) d\tilde{y} \right].$$

I can substitute for $\partial P(\tilde{y}, \lambda, A)/\partial \lambda(y')$ in the second term to get,

$$\int \frac{\partial P(\tilde{y}, \lambda, A)}{\partial \lambda(y')} \Psi(\tilde{y}, A)\lambda(\tilde{y}) d\tilde{y} = \frac{\Psi(y', A)\ell(\underline{\omega}(\lambda, A)|y')}{C(\lambda, A)} \int \ell(\underline{\omega}(\lambda, A)|\tilde{y})\Psi(\tilde{y}, A)\lambda(\tilde{y}) d\tilde{y} = 0,$$

where the second equality follows equation (32). Therefore,

$$\frac{\partial q(\lambda, A)}{\partial \lambda(y')} = -\frac{1}{\rho k} q(\lambda, A)^2 P(y', \lambda, A)\Psi(y', A).$$

Next note that

$$p(\lambda, A) = m\left(q(\lambda, A)^{-1}\right).$$

Therefore,

$$\frac{\partial p(\lambda, A)}{\partial \lambda(y')} = \frac{-m'(q(\lambda, A)^{-1})}{q(\lambda, A)^2} \frac{\partial q(\lambda, A)}{\partial \lambda(y')} = \frac{m'(\lambda, A)}{\rho k} P(y', \lambda, A)\Psi(y', A),$$

where with some abuse of notation I have let $m'(\lambda, A) = m'(q(\lambda, A)^{-1})$. Finally,

$$\begin{aligned}\frac{\partial f(y, \lambda, A)}{\partial \lambda(y')} &= p(\lambda, A)\frac{\partial P(y, \lambda, A)}{\partial \lambda(y')} + \frac{\partial p(\lambda, A)}{\partial \lambda(y')} P(y, \lambda, A) \\ &= \left[\frac{p(\lambda, A)}{C(\lambda, A)} \ell(\underline{\omega}(\lambda, A)|y)\ell(\underline{\omega}(\lambda, A)|y') + \frac{m'(\lambda, A)}{\rho k} P(y, \lambda, A)P(y', \lambda, A) \right] \Psi(y', A).\end{aligned}$$

B.2 Derivatives with respect to A

Differentiating equation (32) with respect to A,

$$\int \frac{\partial \Psi(\tilde{y}, A)}{\partial A} \ell(\underline{\omega}(\lambda, A)|\tilde{y})\lambda(\tilde{y}) d\tilde{y} + \frac{\partial}{\partial \omega}\left(\int \Psi(\tilde{y}, A)\ell(\omega|\tilde{y})\lambda(\tilde{y}) d\tilde{y} \right)\bigg|_{\omega = \underline{\omega}(\lambda, A)} \frac{\partial \underline{\omega}(\lambda, A)}{\partial A} = 0.$$

Therefore,

$$\frac{\partial \underline{\omega}(\lambda, A)}{\partial A} = -\frac{1}{C(\lambda, A)} \int \frac{\partial \Psi(\tilde{y}, A)}{\partial A} \ell(\underline{\omega}(\lambda, A)|\tilde{y})\lambda(\tilde{y}) d\tilde{y},$$

where $C(\lambda, A)$ is defined in equation (33). Next recall that $P(y, \lambda, A) = 1 - L(\underline{\omega}(\lambda, A)|y)$, so

$$\begin{aligned}\frac{\partial P(y, \lambda, A)}{\partial A} &= -\ell(\underline{\omega}(\lambda, A)|y)\frac{\partial \underline{\omega}(\lambda, A)}{\partial A} \\ &= \frac{\ell(\underline{\omega}(\lambda, A)|y)}{C(\lambda, A)} \int \frac{\partial \Psi(\tilde{y}, A)}{\partial A} \ell(\underline{\omega}(\lambda, A)|\tilde{y})\lambda(\tilde{y}) d\tilde{y}.\end{aligned}$$

I next differentiate the free entry condition with respect to A to get

$$\frac{\partial q(\lambda, A)}{\partial A} \frac{\rho k}{q(\lambda, A)} + q(\lambda, A) \int \left[\frac{\partial P(\tilde{y}, \lambda, A)}{\partial A} \Psi(\tilde{y}, A) + P(\tilde{y}, \lambda, A)\frac{\partial \Psi(\tilde{y}, A)}{\partial A} \right] \lambda(\tilde{y}) d\tilde{y} = 0.$$

Therefore,

$$\frac{\partial q(\lambda, A)}{\partial A} = -\frac{q(\lambda, A)^2}{\rho k}\left[\int \frac{\partial P(\tilde{y}, \lambda, A)}{\partial A}\Psi(\tilde{y}, A)\lambda(\tilde{y})d\tilde{y} + \int P(\tilde{y}, \lambda, A)\frac{\partial \Psi(\tilde{y}, A)}{\partial A}\lambda(\tilde{y})d\tilde{y}\right].$$

I can substitute for $\partial P(\tilde{y}, \lambda, A)/\partial A$ in the first term in the brackets to get

$$\begin{aligned}\int \frac{\partial P(\tilde{y}, \lambda, A)}{\partial A}&\Psi(\tilde{y}, A)\lambda(\tilde{y})d\tilde{y} \\ &= \frac{1}{C(\lambda, A)}\left(\int \ell(\underline{\omega}(\lambda, A)|\tilde{y})\Psi(\tilde{y}, A)\lambda(\tilde{y})d\tilde{y}\right)\left(\int \frac{\partial \Psi(\tilde{y}, A)}{\partial A}\ell(\underline{\omega}(\lambda, A)|\tilde{y})\lambda(\tilde{y})d\tilde{y}\right) = 0,\end{aligned}$$

where I am using the fact that the first integral is zero by equation (32). So,

$$\frac{\partial q(\lambda, A)}{\partial A} = -\frac{q(\lambda, A)^2}{\rho k}\int P(\tilde{y}, \lambda, A)\frac{\partial \Psi(\tilde{y}, A)}{\partial A}\lambda(\tilde{y})d\tilde{y}.$$

Differentiating $p(\lambda, A) = m(q(\lambda, A)^{-1})$ with respect to A, I get

$$\begin{aligned}\frac{\partial p(\lambda, A)}{\partial A} &= \frac{-m'(q(\lambda, A)^{-1})}{q(\lambda, A)^2}\frac{\partial q(\lambda, A)}{\partial A} \\ &= \frac{m'(\lambda, A)}{\rho k}\int P(\tilde{y}, \lambda, A)\frac{\partial \Psi(\tilde{y}, A)}{\partial A}\lambda(\tilde{y})d\tilde{y},\end{aligned}$$

where $m'(\lambda, A) = m'(q(\lambda, A)^{-1})$. Finally,

$$\begin{aligned}\frac{\partial f(y, \lambda, A)}{\partial A} =& p(\lambda, A)\frac{\partial P(y, \lambda, A)}{\partial A} + \frac{\partial p(\lambda, A)}{\partial A}P(y, \lambda, A) \\ =& p(\lambda, A)\frac{\ell(\underline{\omega}(\lambda, A)|y)}{C(\lambda, A)}\int \frac{\partial \Psi(\tilde{y}, A)}{\partial A}\ell(\underline{\omega}(\lambda, A)|\tilde{y})\lambda(\tilde{y})d\tilde{y} \\ &+ P(y, \lambda, A)\frac{m'(\lambda, A)}{\rho k}\int P(\tilde{y}, \lambda, A)\frac{\partial \Psi(\tilde{y}, A)}{\partial A}\lambda(\tilde{y})d\tilde{y}.\end{aligned}$$

C Robustness Check

In this appendix I modify the model introduced in Section 2 by assuming that the quit rate is procyclical and argue that the predictions of the model are robust to this modification. I do this by simply assuming that the quit rate is given by a smooth and increasing function $\varsigma(A)$.

It is easy to see how this extension can lead to a locally acyclical separation rate. Let λ^* denote the distribution of ability at the steady state equilibrium corresponding to some A^*. Suppose that $\varsigma(A)$ is such that

$$\varsigma'(A^*) = -\delta\frac{\partial}{\partial A}\int \gamma(\tilde{y}, A^*)\,\lambda^*(\tilde{y})\,d\tilde{y}, \tag{34}$$

where $\gamma(y, A) = 1 - H(\Pi(y, A))$ denotes the probability of a layoff conditional on a machine breakdown.[25] Then it is trivially the case that the average separation rate, $\varsigma(A) + \delta\int\gamma(\tilde{y}, A)\lambda(\tilde{y})d\tilde{y}$, is independent of A when the economy is close to the steady state equilibrium.

[25] It is always possible to find a sufficiently small δ and a sufficiently flat $\varsigma(A)$ that come arbitrarily close to satisfying equation (34)—assuming that the endogenous variables of the model vary continuously with its parameters. This is due to the fact that equation (34) is satisfied when $\delta = 0$ and $\varsigma'(A) \equiv 0$.

The following proposition shows that the separation rate is log-supermodular in the modified model—under conditions that are essentially identical to the conditions that give rise to a log-supermodular separation rate in the baseline model.

Proposition 10. *Assume that the quit rate is given by a smooth and increasing function $\varsigma(A)$, the cost of repair is distributed according to a Pareto distribution, the wage rate is given by $w(y, A) = w_0(y) + \beta(y)A$, the flow profit, $\pi(y, A)$, is strictly supermodular in y and A, and $\Pi(y, A)$ and its first two derivatives vary smoothly with δ and ς'. Given any open interval $\mathcal{A} \subset [\underline{A}, \overline{A}]$, if the rate of machine breakdown, δ, the slope of the base wage, $|w_0'|$, and the slope of the quit rate, ς', are sufficiently small and the distribution of the cost of repair has a sufficiently thin tail, then the separation rate, $s(y, A)$, is strictly log-supermodular in (y, A) over $Y \times \mathcal{A}$.*

Proof. The proof is similar to the proof of Proposition 9. The separation rate is log-supermodular if the following expression is positive:

$$\frac{\partial^2 \log s(y, A)}{\partial y \partial A} = \frac{1}{(\varsigma(A) + \delta\gamma(y, A))^2}\left[\delta(\varsigma(A) + \delta\gamma(y, A))\frac{\partial^2\gamma(y, A)}{\partial y \partial A} - \delta^2\frac{\partial\gamma(y, A)}{\partial y}\frac{\partial\gamma(y, A)}{\partial A} - \delta\frac{\partial\gamma(y, A)}{\partial y}\frac{\partial\varsigma(A)}{\partial A}\right].$$

The assumption that the first two derivatives of $\Pi(y, A)$ vary smoothly with δ implies that $\partial^2_{yA}\gamma(y, A)$ is a smooth function of δ. Therefore, by Taylor's theorem,

$$\frac{\partial^2}{\partial y \partial A}\log s(y, A) = \frac{\delta}{\varsigma(A)}\frac{\partial^2\gamma_0(y, A)}{\partial y \partial A} - \frac{\delta}{\varsigma^2(A)}\frac{\partial\gamma_0(y, A)}{\partial y}\frac{\partial\varsigma(A)}{\partial A} + R(\delta; y, A), \tag{35}$$

where R is a term that is second order in δ, $\gamma_0(y, A) = 1 - H(\Pi_0(y, A))$, and Π_0 is the limit of Π as $\delta \to 0$. There are two differences between the above expression and the one obtained in the proof of Proposition 9. First, in the above expression $\Pi_0(y, A)$ is the value of a firm in a model with a productivity-dependent quit rate, whereas in the proof of Proposition 9 the quit rate was constant. But the assumption that $\Pi(y, A)$ varies smoothly with ς' implies that, for sufficiently small values of ς', the two are close. Second, the above expression has an additional term. But since $\gamma_0(y, A)$ is decreasing in y and $\varsigma(A)$ is increasing in A, the additional term is positive. The result then immediately follows an argument along the lines of the proof of Proposition 9. □

Propositions 1–8 only make use of the fact that the value of a firm, $\Pi(y, A)$, is strictly increasing in y and A. Whenever $\partial\Pi(y, A)/\partial A$ varies continuously with ς' and ς' is not too large, $\Pi(y, A)$ will continue to be increasing in its arguments in the modified model. Therefore, Propositions 1–8, too, will continue to hold in the modified model. This observation together with Proposition 10 show that the predictions of the model are robust to the introduction of a procyclical quit rate.

References

Acemoglu, Daron (1996), "A Microfoundation for Social Increasing Returns to Education." *Quarterly Journal of Economics*, 111, 779–804.

Acemoglu, Daron (1999), "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence." *American Economic Review*, 89, 1259–1278.

Acemoglu, Daron (2001), "Good Jobs Versus Bad Jobs." *Journal of Labor Economics*, 19, 1–21.

Acemoglu, Daron and Jörn-Steffen Pischke (1998), "Why Do Firms Train? Theory and Evidence." *Quarterly Journal of Economics*, 113, 79–119.

Barnichon, Regis (2010), "Building a Composite Help-Wanted Index." *Economics Letters*, 109, 175–178.

Berger, David (2015), "Countercyclical Restructuring and Jobless Recoveries."

Blanchard, Olivier Jean and Peter Diamond (1990), "The Cyclical Behavior of the Gross Flows of U.S. Workers." *Brookings Papers on Economic Activity*, 1990, 85–155.

Blanchard, Olivier Jean and Peter Diamond (1994), "Ranking, Unemployment Duration, and Wages." *Review of Economic Studies*, 61, 417–434.

Caballero, Ricardo J. and Mohamad L. Hammour (1994), "The Cleansing Effects of Recessions." *American Economic Review*, 84, 1350–1368.

Caballero, Ricardo J. and Mohamad L. Hammour (1996), "On the Timing and Efficiency of Creative Destruction." *Quarterly Journal of Economics*, 111, 805–852.

Coate, Stephen and Glenn Loury (1993), "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83, 1220–1240.

Davis, Steven J., Jason R. Faberman, and John C. Haltiwanger (2013), "The Establishment-Level Behavior of Vacancies and Hiring." *Quarterly Journal of Economics*, 128, 581–622.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger (2012), "Recruiting Intensity During and After the Great Recession: National and Industry Evidence." *American Economic Review: Papers & Proceedings*, 102, 584–588.

Davis, Steven J. and John Haltiwanger (1990), "Gross Job Creation and Destruction: Microeconomic Evidence and Macroeconomic Implications." *NBER Macroeconomics Annual 1990*, 5, 123 – 186.

Davis, Steven J. and John Haltiwanger (1992), "Gross Job Creation, Gross Job Destruction and Employment Reallocation." *Quarterly Journal of Economics*, 107, 819–863.

Davis, Steven J. and Till von Wachter (2011), "Recessions and the Costs of Job Loss." *Brookings Papers on Economic Activity*, 1–72.

Diamond, Peter A. and Ayşegül Şahin (2015), "Shifts in the Beveridge curve." *Research in Economics*, 69, 18–25.

Dixit, Avinash K. (1993), *The Art of Smooth Pasting*. Taylor & Francis.

Gertler, Mark and Antonella Trigari (2009), "Unemployment Fluctuations with Staggered Nash Wage Bargaining." *Journal of Political Economy*, 117, 38–86.

Gibbons, Robert and Lawrence F. Katz (1991), "Layoffs and Lemons." *Journal of Labor Economics*, 9, 351–380.

Greenwald, Bruce C. (1986), "Adverse Selection in the Labour Market." *The Review of Economic Studies*, 53, 325–347.

Guerrieri, Veronica (2007), "Heterogeneity, Job Creation and Unemployment Volatility." *Scandinavian Journal of Economics*, 109, 667–693.

Guerrieri, Veronica (2008), "Inefficient Unemployment Dynamics under Asymmetric Information." *Journal of Political Economy*, 116, 667–708.

Hall, Robert E. (1991), "Recessions as Reorganizations."

Hall, Robert E. (2005a), "Employment Fluctuations with Equilibrium Wage Stickiness." *The American Economic Review*, 95, 50–65.

Hall, Robert E. (2005b), "Job Loss, Job Finding, and Unemployment in the U.S. Economy Over the Past Fifty Years." *NBER Macroeconomics Annual*, 20, 101–137.

Hall, Robert E. and Paul R. Milgrom (2008), "The Limited Influence of Unemployment on the Wage Bargain." *American Economic Review*, 98, 1653–1674.

Hershbein, Brad and Lisa B. Kahn (2016), "Do Recessions Accelerate Routine-Biased Technological Change?"

Hosios, Arthur J. (1990), "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies*, 57, 279–298.

Jackson, Matthew O. and Pietro Tebaldi (2015), "A Forest Fire Theory of the Duration of a Boom and the Size of a Subsequent Bust."

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan (1993), "Earnings Losses of Displaced Workers." *American Economic Review*, 83, 685–709.

Jaimovich, Nir and Henry E. Siu (2012), "The Trend is the Cycle: Job Polarization and Jobless Recoveries."

Jarosch, Gregor (2014), "Searching for Job Security and the Consequences of Job Loss."

Jarosch, Gregor and Laura Pilossoph (2015), "Statistical Discrimination and Duration Dependence in the Job Finding Rate."

Kahn, Lisa B. (2010), "The Long-Term Labor Market Consequences of Graduating From College in a Bad Economy." *Labour Economics*, 17, 303–316.

Kahn, Lisa B. (2013), "Asymmetric Information Between Employers." *American Economic Journal: Applied Economics*, 5, 165–205.

Kahn, Lisa B. and Fabian Lange (2014), "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures." *Review of Economic Studies*, 81, 1575–1613.

Kennan, John (2010), "Private Information, Wage Bargaining and Employment fluctuations." *Review of Economic Studies*, 77, 633–664.

Koenders, Kathryn and Richard Rogerson (2005), "Organizational Dynamics Over the Business Cycle: A View on Jobless Recoveries." *Federal Reserve Bank of St. Louis Review*, 87, 555–579.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo (2013), "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment." *Quarterly Journal of Economics*, 128, 1123–1167.

Lockwood, Ben (1991), "Information Externalities in the Labour Market and the Duration of Unemployment." *Review of Economic Studies*, 58, 733–753.

Milgrom, Paul R. (1981), "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics*, 12, 380–391.

Modestino, Alicia Sasser, Daniel Shoag, and Joshua Ballance (2015a), "Downskilling: Changes in Employer Skill Requirements Over the Business Cycle."

Modestino, Alicia Sasser, Daniel Shoag, and Joshua Ballance (2015b), "Upskilling: Do Employers Demand Greater Skill When Skilled Workers Are Plentiful?"

Moen, Espen R. and Åsa Rosén (2011), "Incentives in Competitive Search Equilibrium." *Review of Economic Studies*, 78, 733–761.

Mueller, Andreas I. (2015), "Separations, Sorting and Cyclical Unemployment."

Nakamura, Emi (2008), "Layoffs and Lemons Over the Business Cycle." *Economics Letters*, 99, 55–58.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz (2012), "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics*, 4, 1–29.

Philippon, Thomas (2006), "Corporate Governance Over the Business Cycle." *Journal of Economic Dynamics and Control*, 30, 2117–2141.

Pissarides, Christopher A. (2009), "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" *Econometrica*, 77, 1339–1369.

Pries, Michael J. (2008), "Worker Heterogeneity and Labor Market Volatility in Matching Models." *Review of Economic Dynamics*, 11, 664–678.

Restrepo, Pascual (2015), "Skill Mismatch and Structural Unemployment."

Rosén, Åsa (1997), "An Equilibrium Search-Matching Model of Discrimination." *European Economic Review*, 41, 1589–1613.

Shapiro, Carl and Joseph E. Stiglitz (1984), "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review*, 74, 433–444.

Shimer, Robert (2005), "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review*, 95, 25–49.

Shimer, Robert (2012a), "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics*, 15, 127–148.

Shimer, Robert (2012b), "Wage Rigidities and Jobless Recoveries." *Journal of Monetary Economics*, 59, S65–S77.

Solon, Gary, Robert Barsky, and Jonathan A. Parker (1994), "Measuring the Cyclicality of Real Wages: How Important is Composition Bias?" *Quarterly Journal of Economics*, 109, 1–25.

Sullivan, Daniel and Till von Wachter (2009), "Job Displacement and Mortality: An Analysis Using Administrative Data." *Quarterly Journal of Economics*, 124, 1265–1306.