



Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
<http://economics.sas.upenn.edu/pier>

PIER Working Paper 11-040

“Efficient Learning, Job Turnover and Wage Dispersion”

by

Fei Li

<http://ssrn.com/abstract=1967503>

Efficient Learning, Job Turnover and Wage Dispersion*

Fei Li[†]

University of Pennsylvania

November 28, 2011

Abstract

This paper studies the aggregate consequences of individual learning in the labor market. Specifically, I examine this issue in a model of directed search on the job. Once matched, a firm-worker pair gradually learns the match-specific quality, taking the history of realized production as signals. Heterogeneity in beliefs about the match quality and in the job search behavior of workers naturally occurs, resulting from a variety of individual histories. I describe the efficient learning and searching strategy and implement the efficient allocations through a market mechanism in which the labor contract depends deterministically on tenure. Consistent with the stylized facts, the model successfully predicts the tenure effect on both the job separation rate and the probability of on-the-job search, and when search frictions are small, the model generates a dispersed wage distribution with a flat tail, along the lines of observations.

Keywords: Learning, Directed Search, Block Recursive Equilibrium, Turnover, Wage Dispersion

JEL Classification Codes: D83, J31

*I am grateful to my adviser, George Mailath, for insightful instruction and encouragement. I thank Hanming Fang, Guido Menzio and Andrew Postlewaite for their invaluable guidance and excellent comments that improved the paper. I also thank Naoki Aizawa, Kenneth Burdett, Chao Fu, Nils Gornemann, Can Tian, Xi Weng, and participants at the 2011 Royal Economic Society Conference, and the Upenn theory lunch for insightful comments. All remaining errors are mine.

[†]Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104; Email: fei@econ.upenn.edu

Contents

1	Introduction	2
2	The Model	5
2.1	Physical Environment	5
2.2	Timing	6
2.3	Learning Process	7
3	Efficient Allocation	7
3.1	Formulation of the Planner's Problem	7
3.2	Solution to the Planner's Problem	11
3.3	Continuous Time Limit	13
4	Decentralization	17
4.1	Market Equilibrium	17
4.2	Empirical Implications of Continuous Time Limit	21
4.2.1	Job Turnover	22
4.2.2	Stationary Wage Distribution	24
4.2.3	Distribution Convergence	27
5	Extensions	28
6	Conclusion	30
A	Appendix: Proofs of Results on the Planner's Problem	31
A.1	Separability of Planner's Problem	31
A.2	The Planner's Individual Problem in Continuous Time Limit	33
A.3	Solution Characterization of Planner's Problem in Continuous Time Limit	36
B	Appendix: Proofs of Results on the Market Mechanism	38
B.1	Decentralization	38
B.2	Stationary Distribution	39

1 Introduction

It is well established that learning about the match quality of firm-worker pairs plays an important role in explaining job turnover at the individual level.¹ However, few studies have shed light on the aggregate consequences resulting from individual learning. This paper analyzes the impact of individual learning on wage dispersion, the interplay between learning and search frictions, and how the socially efficient learning allocation can be implemented through a market mechanism.

I develop a directed on-the-job search model to address the aforementioned issues. In particular, I introduce the learning behavior of matched firm-worker pairs into a frictional labor market. A firm and a worker may meet in the labor market. Their pair-specific match quality is initially unknown (with the same prior on both sides) and revealed gradually over time. Employed workers can search for a new job and so can unemployed workers, and the optimal search strategy depends on a worker's subjective evaluation of his current job status. Over time, a worker adjusts his evaluation of current job match quality based on his past job performance. The diversity of individual histories results in ex post heterogeneity in subjective evaluation and therefore in the job search behavior among workers. Optimistic workers are attracted by well-paying jobs only. Pessimistic workers, who are or will be unemployed, are less selective and are willing to take offers with lower pay. This job status contingent search strategy can generate interesting empirical implications for both job turnover and wage dispersion.

In general, the equilibrium is hard to characterize. In a standard search model, as Burdett and Mortenson (1998) and Shi (2009) show, firms may post different wage schemes, which induces two dimensions of ex post heterogeneity among employed workers: (1) their subjective evaluation of the current match quality, and (2) the wage scheme promised by their current employer. A worker's job search behavior depends on both of them, and therefore it is hard to analyze, especially in an equilibrium model. To avoid the curse of dimensionality, I follow Menzio and Shi (2011) and consider an environment in which the employment contract is complete and the allocation is socially efficient. Focusing on this particular environment not only simplifies the analysis but also generates interesting and economically relevant dynamics.

To characterize the socially efficient allocation, I start with the social planner's problem. A social planner decides (1) the separation rule of existing matches and (2) search strategy for each worker. The efficient separation rule is given as a cutoff belief about match quality. When the belief about match quality is higher than the cutoff level, the planner keeps the underlying

¹For example, Bergemann and Valimaki (1994), Eeckhout and Weng (2010), Felli and Harris (1996), Jovanovic (1979), Miller (1984), Nagypal (2007).

match. Otherwise, the planner destroys the match and naturally stops learning about its quality. Following the literature of directed search (Acemoglu and Shimer 1999, Moen 1997), I assume there are numerous locations in the economy. A match forms only if a worker meets a firm at the same location. To workers, locations differ from each other in the probability of finding a new job and promised pay. The efficient choice of a searching location is determined by the current state of a worker. If an employed worker is in a good match, he does not search for a new job. If an employed worker's current match quality is uncertain, he is sent to a specific location to find a new job, and the probability of getting a new job is non-increasing in the belief about his current match quality. An unemployed worker searches for a job at the location with the highest job-finding probability. Thanks to the block recursive structure of the model, the social planner's solution does not depend on the distribution of workers' states, and therefore, the analysis works whether it is at the steady state or not.

Next, for tractability, I take the original discrete time model to its continuous time limit to characterize its approximated empirical implication. The discrete time model has a well-defined continuous time limit, the solution of which is uniquely characterized by a set of equations. The dynamics of the continuous time limit of the social planner's problem is governed by an ordinary differential equation. This implies that when the time interval is small, the solution to a social planner's problem in discrete time can be approximated by the solution to a simple ODE.

I then turn to the decentralization of the planner's solution in both discrete time and continuous time settings and suggest possible empirical implications of the market equilibrium. The planner's solution can be implemented if a firm and a worker sign a bilaterally efficient contract upon forming their match. In particular, a bilaterally efficient contract may consist of a contractible tenure contingent on-the-job search strategy, a tenure independent wage agenda and a probationary period. Under this specific contract, the market equilibrium can implement the socially efficient allocation.

The main results are equilibrium predictions at both the individual and aggregate levels. First, this model can preserve the qualitative correlations between tenure, separation rates, and on-the-job search probability that are observed in the individual-level data. The separation rate as a function of tenure first increases at low tenure levels, then decreases, and eventually becomes constant. The on-the-job search probability as a function of tenure is decreasing first and constant in the end. Second, at the aggregate level, this model generates a stationary wage distribution whose curvature depends on learning speed and search frictions. With small search frictions, one can expect a wage distribution with a flat tail. Intuitively, when search frictions are small, it is easy to find a job quickly, and therefore it is socially inefficient to create many vacancies for workers whose match quality is believed to be not low enough. Since those workers are aiming for a high

wage, the new match formation with a high wage has low mass. Hence, the probability densities of high wage offers are small in the stationary wage distribution. Finally, a simple comparative statics of this model reveals that the wage dispersion increases as a result of an improvement in aggregate productivity. These empirical implications illustrate that the coexistence of individual learning and search frictions has important aggregate consequences.

These theoretical results are roughly consistent with a variety of stylized facts emerging from the data at both the micro and macro levels. For example, Farber (1994) and Nagypal (2007) find that the separation rate increases early in tenure and decreases later, but in the end, the separation rate becomes constant. In addition, Pissarides and Wadsworth (1994) show that on-the-job search intensity is decreasing in tenure. Finally, many empirical papers, such as Postel-Vinay and Robin (2002), and Mortenson (2005), find that within-group wage dispersion has a flat tail, and its inequality rises from the 1970s to the 1990s as productivity improves. See Kambourov and Manovskii (2009) for details.

My paper is closely related to Moscarini (2005), who also studies the aggregate consequences of pair-specific match quality learning in a Mortenson-Pissarides model. In his paper, search is random, and a matched worker and firm keep bargaining over wages over time. His model implies that a low-quality match bears a higher probability of being destroyed than a high-quality match does, and therefore, the longer a worker stays, the higher are the match quality he expects to have. This *ex post* selection mechanism generates a tenure-wage effect and a wage dispersion with a flat tail. Although learning is the common driving force of both Moscarini's model and mine, the mechanisms are very different and lead to different implications. In my paper, search is directed and wage is determined by a contract posting mechanism instead of Nash bargaining. Moreover, in Moscarini's model, the shape of the wage distribution results from the combination of an *ex post* selection mechanism and a tenure effect; in my model, the wage dispersion is generated by the combination of learning and on-the-job search. Burdett and Vishwanath (1988) consider a random search model in which workers learn about the unknown distribution of wages from the random arrival of wage offers and showed that learning from search can induce reservation wages to decline with the duration of unemployment. Gonzalez and Shi (2010) also employ a directed search framework to study learning issues. Yet, they focus on the learning of workers' job search ability.

The rest of this paper is organized as follows. Section 2 presents the basic environment, individual payoff and learning process. The social planner's problem is characterized in Section 3. Section 4 considers a simple contract to implement the social planner's allocation in a frictional labor market and discusses the empirical implications. A number of extensions are discussed in Section 5. Section 6 concludes. All technical proofs can be found in the Appendices.

2 The Model

2.1 Physical Environment

Time is discrete and lasts forever with a time interval of $\Delta > 0$. Denote $T = 0, \Delta, 2\Delta, \dots$ as the calendar time of the economy. The economy is populated by a continuum of workers of measure one and by a continuum of firms of measure greater than 1. Each worker has the utility function $\sum_{T=0}^{\infty} e^{-rT} C_T \Delta$, where $C_T \Delta \in \mathbb{R}$ is the worker's consumption in $[T, T + \Delta)$ and r is his discount rate. Each firm has the payoff function $\sum_{T=0}^{\infty} e^{-rT} \Pi_T \Delta$, where $\Pi_T \Delta \in \mathbb{R}$ is the firm's profit in $[T, T + \Delta)$. Each firm has one vacancy and can hire at most one worker. Vacant firms or unemployed workers are unproductive.

The match between a firm and a worker is either good or bad. If the match is good, a matched firm receives 1 unit of payoff in each period with probability $\lambda \Delta$; if the match is bad, a matched firm receives nothing. Initially, a matched worker-firm pair shares symmetric information about the match quality with a common prior $\alpha_0 \in (0, 1)$ that the current match is good. They observe the outcomes and hold common posterior beliefs α_t throughout time, where α_t denotes the subjective probability that they assign to the match being good at t , where t denotes the worker's tenure in his current job with $t = 0, \Delta, 2\Delta, 3\Delta, \dots$. For simplicity, no extra flow payoff is generated by a match. A match is destroyed exogenously with probability $\delta \Delta$ in each period. Unemployed workers enjoy a flow payoff of $b \Delta > 0$, which can be interpreted as his home production. To avoid a trivial case, assume $\alpha_0 \lambda > b > 0$, that is, a new match is better than no match, but no match is better than a bad match.

There is a continuum of locations indexed by a real number $l \in [0, 1]$. A vacant firm and a worker can match only if they are searching in the same location. In each period, both firms and workers decide which location to enter. A location is interpreted as a submarket if there are firms and workers there. Different submarkets can be indexed by the promised value to the worker, $x \in \mathbb{R}$, posted by firms in that market. I denote mapping $\Gamma : [0, 1] \rightarrow \mathbb{R} \cup \emptyset$ as the submarket assignment function. In other words, $x = \Gamma(l)$ is the promised value to the worker specified by the contract offered at location l , while $\Gamma(l) = \emptyset$ means there is no submarket at location l . At location l with $\Gamma(l) \neq \emptyset$, the ratio between the number of jobs that are vacant and the number of searching workers is denoted by $\tilde{\theta}(l) \in \mathbb{R}^+$. I refer to $\theta(x)$ as the tightness of the submarket at location l such that $x = \Gamma(l)$. In other words, I do not distinguish between the two markets $l \neq l'$ with the same x .

All submarkets are frictional. In particular, workers and firms that are searching in the same location are brought into contact by a meeting technology with constant returns to scale that can be described in terms of the market tightness $\theta \in \mathbb{R}^+$. In particular, each period a worker

finds a vacant job with probability $p(\theta(l))\Delta$ at location l , where $\theta(l)$ is the market tightness at location l and function $p : \mathbb{R}^+ \rightarrow [0, 1]$ is twice continuously differentiable, strictly increasing, strictly concave, which satisfies (i) $p(0) = 0$, (ii) $\lim_{\theta \rightarrow 0} p'(\theta) = \infty$, and (iii) $\lim_{\theta \rightarrow \infty} p(\theta)$ is bounded by a finite number. Similarly, a vacancy meets a worker with probability $q(\theta(l))\Delta$ in location l where $q : \mathbb{R}^+ \rightarrow [0, 1]$ is a twice continuously differentiable, strictly decreasing function such that $q(\theta) = p(\theta)/\theta$, $q(0) = 1$ and $\lim_{\theta \rightarrow \infty} q(\theta) = 0$. When a firm and a worker meet, a new match is formed, and the worker's old match, if any, is destroyed.

Each firm chooses to enter at most one submarket by paying a maintaining cost $k\Delta$ in each period and posts an employment contract x , which is the promised value to the worker. All workers, whether employed or unemployed, observe all available offers in the labor market and choose one submarket to enter and search for a new job. Different wage dynamics are allowed given the identical initial expected promise. In general, a worker's individual wage dynamics can depend on both the aggregate market variables and the match-specific payoff history.

2.2 Timing

Each period is divided into three stages: 1) separation, 2) search and matching, and 3) production. At the separation stage, an existing match is destroyed with probability $z \in [\delta\Delta, 1]^2$, which is chosen by the matched firm, where $\delta > 0$ measures the exogenous separation rate.³

At the search and matching stage, each potential vacant firm chooses whether to create a vacancy and which submarket it enters. While the workers who lose their jobs in the separation stage are not allowed to search immediately, the rest, both employed and unemployed, can search. They observe all available offers in every submarket and choose whether and where to search for a new job. When a worker and a vacancy meet, they match and sign the contract proposed by the firm.

At the production stage, an unemployed worker produces $b\Delta$ units of output. An employed worker with a good match produces the reward 1 with probability $\lambda\Delta$, and those with a bad match produce zero. Put differently, no news is bad news. A matched firm and a worker learn the quality of their match commonly by observing their past production outcome.

²Since I focus on the case where Δ is small, $[\delta\Delta, 1]$ is nonempty. Similarly, $p(\theta)\Delta$, $q(\theta)\Delta$ and $\lambda\alpha\Delta$ are strictly less than 1 for any θ and α .

³Alternatively, one can denote $z_e \in [0, 1]$ as the probability that a match is endogenously destroyed conditional on its not being destroyed by Nature, and define the ex ante probability of being destroyed as $z = \delta\Delta + (1 - \delta\Delta)z_e$.

2.3 Learning Process

The learning process is summarized as follows. Denote \mathcal{H} as the collection of all signal histories for an existing match. $\{\mathcal{F}_t\}_{t=0}^\infty$ is the filtration generated by all possible individual histories. The subjective belief α_t is a \mathcal{F}_t -adapted stochastic process. As stated before, each new match starts with the same initial belief α_0 .

For a match with $\alpha = 1$, no belief adjustment happens regardless of its current period output. For a match with $\alpha \leq \alpha_0$, if the unit of payoff is received in tenure period t , $\alpha_{t+\Delta}$ jumps to 1; otherwise, by standard Bayes' rule updating, the evolvement of α_t in each period follows

$$\alpha_{t+\Delta} - \alpha_t = \frac{-\lambda\alpha_t(1 - \alpha_t)\Delta}{\alpha_t(1 - \lambda\Delta) + (1 - \alpha_t)}.$$

When Δ goes to zero, one obtains the well-known limit $\dot{\alpha}_t = -\lambda(1 - \alpha_t)\alpha_t$, which is 0 if α_t equals either 0 or 1.

Denote B_Δ as the set of all possible values α_t based on the production history. For any $\alpha \in B_\Delta$, define α^- such that $\alpha = \alpha^- + \frac{-\lambda\alpha^-(1-\alpha^-)\Delta}{\alpha^-(1-\lambda\Delta)+(1-\alpha^-)}$, and α^+ such that $\alpha^+ = \alpha + \frac{-\lambda\alpha(1-\alpha)\Delta}{\alpha(1-\lambda\Delta)+(1-\alpha)}$. In other words, α^- (α^+) is the last (next) period belief about the match with current belief α . When Δ goes to zero, the Hausdorff distance between B_Δ and $[0, \alpha_0] \cup \{1\}$ goes to zero.

3 Efficient Allocation

To characterize the efficient allocation, I solve the social planner's problem first. The social planner's search strategy and separation strategy depend both on the aggregate employment state and on a worker's individual employment status. The former includes the unemployment rate and the distribution of the current match quality. The latter includes: (1) whether the worker is employed, and (2) the belief about the worker's match quality if he is employed. Formally, define $\Omega = [0, \alpha_0] \cup \{u\} \cup \{1\}$ as a worker's state space. A worker's state $\omega \in \Omega$ can be interpreted as follows. For an uncertain matched worker, his type $\omega \in [0, \alpha_0]$ is the belief about the current match quality. For a matched worker who has sent a good signal before, $\omega = 1$. For an unemployed worker, $\omega = u$. Denote the probability measure μ_T as the state of the economy. Let $\Xi = \Delta(\Omega)$ denote the set to which μ_T belongs for all T .

3.1 Formulation of the Planner's Problem

At each stage, the social planner decides on job creation, match formation and separation, and workers' job search strategy. Formally, at the beginning of a period, the planner observes the

aggregate state μ . At the separation stage, the planner chooses the probability $z : [0, \alpha_0] \cup \{1\} \rightarrow [\delta\Delta, 1]$ of destroying a match for each belief α . At the search and matching stage, the planner observes the adjusted aggregate state and chooses how many vacancies to open at each location and which location a worker should visit, and therefore he can determine $\theta(\omega)$, the tightness at the location where workers with state ω look for new matches, $\theta : \Omega \rightarrow \mathbb{R}^+$. As is standard in models of directed search, the planner will find it optimal to send workers with different states to different locations but will have no incentive to send workers with the same states to different locations. Thus, there is no loss in generality in indexing the active location by the state of workers ω who are searching for new jobs there.

After the first two stages, the economy's state change to $\tilde{\mu}$ due to the job reallocations. To compute the updated state of the economy $\tilde{\mu}$, it is useful to derive the transition probabilities for an individual worker in the first two stages. Suppose the planner's optimal separation probability is z in the separation stage, then a matched worker with state α becomes unemployed with probability $z(\alpha)$; thus the increment of the measure of unemployed workers is $\int_{\alpha \in [0, \alpha_0] \cup \{1\}} z d\mu$, and the measure of a match that survives with state α is $\mu(\alpha)(1 - z(\alpha))$, assuming that the law of large numbers holds in this environment.

At the search and matching stage, workers' state may change due to both unemployment-to-employment and employment-to-employment transition. Suppose the planner's optimal search policy is $\theta(\omega)$. In other words, he sends workers with state ω to a location where the tightness is $\theta(\omega)$. First, consider a worker who enters this period unemployed. With probability $1 - p(\theta(u))\Delta$, the worker does not meet a firm at this stage. In this case, the worker remains unemployed. With probability $p(\theta(u))\Delta$, the worker meets a firm and becomes employed. Next, consider an employed worker who enters this period with state α and survives the separation stage. With probability $1 - p(\theta(\alpha))\Delta$, the worker does not meet a firm at this stage and remains α . With probability $p(\theta(\alpha))\Delta$, the worker meets a firm and forms a new match. In this case, his state changes to α_0 .

After aggregating the transition probabilities of individual workers, the measure of workers who are unemployed at the production stage is given by

$$\tilde{\mu}(u) = \mu(u)(1 - p(\theta(u))\Delta) + \int_{\alpha \in [0, \alpha_0] \cup \{1\}} z d\mu, \quad (1)$$

while the measure of workers who are employed with state $\omega \in B_\Delta / \{\alpha_0\}$ is given by

$$\tilde{\mu}(\alpha) = \mu(\alpha)[1 - z(\alpha)](1 - p(\theta(\alpha))\Delta). \quad (2)$$

Similarly, the measure of workers with state α_0 is given by

$$\begin{aligned} \tilde{\mu}(\alpha_0) &= \mu(\alpha_0)[1 - z(\alpha_0)] + p(\theta(u))\Delta\mu(u) \\ &\quad + \int_{\alpha \in [0, \alpha_0] \cup \{1\}} p(\theta(\alpha))\Delta[1 - z(\alpha)]d\mu, \end{aligned} \quad (3)$$

where the first term of the right-hand side is the measure of workers who enter this period with state α_0 and survive the separation stage, the second term is the measure of workers who enter this period unemployed and find a job in the search stage, and the last term is the measure of workers who enter this period employed, survive the separation stage and change their job in the search stage.

Clearly, there is no reason to separate a good match or replace it with another one, since there is no way to generate a strictly positive surplus by replacing a good match; hence, the separation and search policies for good matches are $z(1, \mu) = \delta\Delta$ and $\theta(1, \mu) = 0$ in any state of the economy μ . Therefore, one can focus on the choice of $(z(\alpha), \theta(\alpha), \theta(u))$. Given the state of the economy μ and a choice function $(z(\alpha), \theta(\omega))$, the social planner's ex ante flow payoff at the production stage is given by

$$\begin{aligned} F(z, \theta|\mu) = & \int_{\alpha \in [0, \alpha_0]} \lambda\alpha\Delta\tilde{\mu} + \tilde{\mu}(u)b\Delta + \tilde{\mu}(1)\lambda\Delta \\ & - k\Delta \left[\int_{\alpha \in [0, \alpha_0]} \theta(\alpha)[1 - z(\alpha)]d\mu + \mu(u)\theta(u) \right], \end{aligned} \quad (4)$$

where the first and third terms are the expected flow payoff created by existing matches, the second term is unemployed workers' benefits, the last term is the sum of new contract posting costs in all locations, and $\tilde{\mu}$ can be calculated by (1,2,3).

At the production stage, the economy's state changes to $\hat{\mu}$ due to learning of uncertain matches. For a worker whose state is $\omega \in \{1, u\}$, there is nothing to learn; thus his state will be unchanged regardless of his output in the production stage. For a worker whose state is $\alpha \in B_\Delta/\{1\}$, his state will change at the production stage for sure. With probability $\lambda\alpha\Delta$, his state becomes $\omega = 1$ if a breakthrough is realized; with complementary probability, his state becomes $\omega = \alpha^+$ if no breakthrough is realized. After aggregating the transition probabilities of individual workers at the production stage, the economy's state becomes

$$\begin{aligned} \hat{\mu}(\alpha) &= \tilde{\mu}(\alpha^-)(1 - \lambda\alpha^-\Delta) \text{ when } \alpha^- \in B_\Delta, \\ \hat{\mu}(u) &= \tilde{\mu}(u), \quad \hat{\mu}(\alpha_0) = 0, \text{ and} \\ \hat{\mu}(1) &= \tilde{\mu}(1) + \int_{\alpha \in [0, \alpha_0]} \lambda\alpha\Delta d\mu', \end{aligned}$$

where $\tilde{\mu}$ is given by (1,2,3). In sum, the planner's decisions imply that next period's distribution of workers' state $\hat{\mu}(\alpha)$ is given by

$$\hat{\mu}(u) = \mu(u)(1 - p(\theta(u))\Delta) + \int_{\alpha \in [0, \alpha_0] \cup \{1\}} z d\mu, \quad (5)$$

$$\hat{\mu}(\alpha) = \begin{cases} \mu(\alpha)[1 - z(\alpha)](1 - p(\theta(\alpha))\Delta) & \text{if } \alpha^- \in B_\Delta / (\{1\} \cup \{\alpha_0\}), \\ (1 - \lambda\alpha_0\Delta)\{\mu(\alpha_0)[1 - z(\alpha_0)] + p(\theta(u))\Delta\mu(u) \\ + \int_{\alpha \in [0, \alpha_0]} p(\theta(\alpha))\Delta[1 - z(\alpha)]d\mu\} & \text{if } \alpha^- = \alpha_0, \end{cases} \quad (6)$$

$$\hat{\mu}(\alpha_0) = 0, \text{ and} \quad (7)$$

$$\begin{aligned} \hat{\mu}(1) &= \mu(1)[1 - \delta\Delta] + \int_{\alpha \in [0, \alpha_0]} \lambda\alpha\Delta[1 - z(\alpha)](1 - p(\theta(\alpha))\Delta)d\mu(\alpha) \\ &\quad + \lambda\alpha_0\Delta\{\mu(\alpha_0)[1 - z(\alpha_0)] + p(\theta(u))\Delta\mu(u) + \int_{\alpha \in [0, \alpha_0]} p(\theta(\alpha))\Delta[1 - z(\alpha)]d\mu\}. \end{aligned} \quad (8)$$

Notice that $\mu(\Omega / (B_\Delta \cup \{u\})) = 0$ in any period.

The planner maximizes the sum of the present value of the discounted flow payoff at rate r . Hence, the planner's society problem is formulated as

$$\mathbf{U}^\Delta(\mu) = \max_{\theta, z} F(z, \theta | \mu) + e^{-r\Delta} \mathbf{U}^\Delta(\hat{\mu}), \quad (9)$$

subject to (5-8), and $z : \Omega \rightarrow [\delta\Delta, 1], \theta : \Omega \rightarrow \mathbb{R}^+$.

In general, the optimal policies depend on both μ and ω . Following Menzio and Shi (2011), Lemma 1 shows that the social planner's society problem is separable, that is, the search and learning strategy of a worker depend only on his own state instead of the distribution μ .

Lemma 1. *The planner's value function \mathbf{U}^Δ is the unique solution to (9) such that (5-8), and satisfies the following properties:*

1. \mathbf{U}^Δ is linear in μ . That is, $\mathbf{U}^\Delta(\mu) = \mu(u)S^\Delta(u) + \mu(1)S^\Delta(1) + \int_{\alpha \in [0, \alpha_0]} S^\Delta(\alpha)d\mu$, where the component value function $S^\Delta(u)$ is given by

$$\begin{aligned} S^\Delta(u) &= \max_{\theta} \{-k\theta\Delta + [1 - p(\theta)\Delta][b\Delta + e^{-r\Delta}S^\Delta(u)] \\ &\quad + p(\theta)\Delta\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S^\Delta(1)] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}S^\Delta(\alpha_0^+)\}\}. \end{aligned} \quad (10)$$

The component value function $S^\Delta(\alpha)$ for $\alpha \in [0, \alpha_0]$ is given by

$$\begin{aligned} S^\Delta(\alpha) &= \max_{z, \theta} z[b\Delta + e^{-r\Delta}S^\Delta(u)] + (1 - z)\{[1 - p(\theta)\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}S^\Delta(1)] \\ &\quad + [1 - p(\theta)\Delta](1 - \lambda\alpha\Delta)e^{-r\Delta}S^\Delta(\alpha^+)] \\ &\quad + p(\theta)\Delta\lambda\alpha_0\Delta[1 + e^{-r\Delta}S^\Delta(1)] + p(\theta)\Delta(1 - \lambda\alpha_0\Delta)e^{-r\Delta}S^\Delta(\alpha_0^+)\}. \end{aligned} \quad (11)$$

The component value function $S^\Delta(1)$ is given by

$$S^\Delta(1) = \delta\Delta[b\Delta + e^{-r\Delta}S^\Delta(u)] + (1 - \delta\Delta)\{\lambda\Delta + e^{-r\Delta}S^\Delta(1)\}. \quad (12)$$

2. The policy functions $\theta_*^\Delta, z_*^\Delta$ do not depend on the distribution μ .
3. $S^\Delta(\alpha)$ is strictly increasing in α .

Thanks to Lemma 1, one can solve the planner's society problem (9) by finding the solution of the planner's individual worker problem, $S^\Delta(\omega)$ for $\omega \in B_\Delta \cup \{u\}$, with $U^\Delta(\mu)$ being given by their μ -weighted sum. In the planner's individual worker problem associated with an unemployed worker, the planner chooses $\theta(u)$ to maximize the present value of the output generated by this worker, net of the cost of vacancies assigned to him. Similarly, in the planner's individual worker problem associated with a worker with state α , the planner chooses $z(\alpha), \theta(\alpha)$ to maximize the present value of the output generated by this worker, net of the cost of the vacancies assigned to him. That the planner's society problem can be decomposed into countably many individual worker problems results from the fact that the search process is directed rather than random. Under random search, the planner has to choose the same tightness for workers with different states because all workers search in the same location; thus, the planner's society problem cannot be decomposed into worker-state-specific individual problems, and the solution will depend on the current state of the economy, μ .

3.2 Solution to the Planner's Problem

As I argued, it is suboptimal to separate a good match, and the efficient choice of $z_*^\Delta(1) = \delta\Delta$, $\theta_*^\Delta(1) = 0$. The efficient search choice for an unemployed worker is $\theta_*^\Delta(u)$ such that

$$k = p'(\theta_*^\Delta(u))\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S^\Delta(1)] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}S^\Delta(\alpha_0^+) - b\Delta - e^{-r\Delta}S^\Delta(u)\}. \quad (13)$$

The left-hand side of (13) is the marginal cost of increasing the tightness at the location for unemployed workers. The right-hand side is the marginal benefit, which is given by the product of two terms. The first term is the marginal increase in the probability of workers' state transition from unemployed to α_0 . The second term is the state transition surplus, which is given by the expected value difference between the new state and the current state. The existence of an interior solution depends on (1) the Inada conditions of $p(\cdot)$, and (2) the second term of the right-hand side is strictly positive. Hence $\theta_*^\Delta(u) > 0$.

Similarly, the efficient search choice for employed workers with state $\alpha \in B_\Delta/\{1\}$ is $\theta_*^\Delta(\alpha)$ such that

$$\begin{aligned} k = & p'(\theta_*^\Delta(\alpha))\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S^\Delta(1)] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}S^\Delta(\alpha_0^+) \\ & - \lambda\alpha\Delta[1 + e^{-r\Delta}S^\Delta(1)] - (1 - \lambda\alpha\Delta)e^{-r\Delta}S^\Delta(\alpha^+)\}. \end{aligned} \quad (14)$$

Notice that the second term of the RHS of (14) is strictly increasing in α ; thus $\theta_*^\Delta(\alpha)$ is strictly decreasing in α . Since $p(\cdot)$ satisfies the Inada conditions, $\theta_*^\Delta(\alpha) > 0$ when $\alpha < \alpha_0$, and $\theta_*^\Delta(\alpha) = 0$ when $\alpha \in \{\alpha_0, 1\}$.

Finally, since the social planner's α -specific problem is linear in z , the efficient choice for the separation probability is $z_*^\Delta(\alpha) = 1$ if

$$\begin{aligned} b\Delta + e^{-r\Delta}S^\Delta(u) \geq & [1 - p(\theta)\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}S^\Delta(1)] \\ & + [1 - p(\theta)\Delta](1 - \lambda\alpha\Delta)e^{-r\Delta}S^\Delta(\alpha^+)p(\theta)\Delta\lambda\alpha_0\Delta[1 + e^{-r\Delta}S^\Delta(1)] \\ & + p(\theta)\Delta(1 - \lambda\alpha_0\Delta)e^{-r\Delta}S^\Delta(\alpha_0^+), \end{aligned} \quad (15)$$

and $z_*^\Delta(\alpha) = \delta\Delta$ otherwise. The left-hand side of (15) is the value induced by separating the current match. The worker becomes unemployed but does not have the opportunity to search for a new match in the current period. The right-hand side is the value induced by keeping the current match. When the left-hand side is greater than the right-hand side, the planner destroys the current match for sure. Otherwise, Nature destroys the match with probability $\delta\Delta$. Notice that the left-hand side does not depend on α , but the right-hand side is strictly increasing in α ; thus the solution can be summarized by a stopping belief $\alpha_*^\Delta \in B_\Delta$ such that

$$z_*^\Delta(\alpha) = \begin{cases} 1 & \text{if } \alpha \leq \alpha_*^\Delta, \\ \delta\Delta & \text{if } \alpha > \alpha_*^\Delta, \end{cases}$$

for all $\alpha \in B_\Delta$. Combining (13), (14) and $S^\Delta(u) \leq S^\Delta(\alpha)$ for any $\alpha \in [\alpha_*^\Delta, \alpha_0] \cup \{1\}$ yields that $\theta_*^\Delta(u) \geq \theta_*^\Delta(\alpha)$ for all $\alpha \in (\alpha_*^\Delta, \alpha_0]$.

The properties of the socially efficient search and separation rule are summarized as follows.

Lemma 2. *The policy correspondences $(z_*^\Delta, \theta_*^\Delta)$ satisfy the following properties:*

1. *They are singleton valued, i.e., they are functions.*
2. *There is α_*^Δ such that, for any $\alpha \in B_\Delta$, $z_*^\Delta(\alpha) = 1$ if $\alpha \leq \alpha_*^\Delta$ and $z_*^\Delta(\alpha) = \delta\Delta$ if $\alpha > \alpha_*^\Delta$.*
3. *$S^\Delta(u) \leq S^\Delta(\alpha)$ for any $\alpha \in [\alpha_*^\Delta, \alpha_0] \cup \{1\}$.*
4. *The search strategy $\theta_*^\Delta(1) = \theta_*^\Delta(\alpha_0) = 0$, $\theta_*^\Delta(u) > 0$, $\theta_*^\Delta(\alpha) > 0$ for any $\alpha \in B_\Delta / (\{1\} \cup \{\alpha_0\})$, and $\theta_*^\Delta(\alpha)$ is decreasing and smaller than $\theta_*^\Delta(u)$ for all $\alpha \in (\alpha_*^\Delta, \alpha_0]$.*

It is worth noting that the block recursivity of the planner's problem does not imply that the social planner's value is irrelevant to the distribution. First, it is clear that the social planner's choice affects the workers' state transition and therefore $\hat{\mu}$. Second, the measure of job creation depends on μ . For example, the measure of vacancy created for unemployed workers is $v(\mu) = \theta_*^\Delta(u)\mu(u)$. When the unemployment rate $\mu(u)$ is high, the measure of vacancy in the submarket for unemployed workers $v(\mu)$ would also be high even though $\theta_*^\Delta(u)$ does not depend on $\mu(u)$.

3.3 Continuous Time Limit

The problem (10,11,12) has no closed-form solution, and therefore, its empirical implications are difficult to characterize. I examine a continuous time limit of the model obtained by taking the length of each period to zero. In the limit, for an unemployed worker, a new job arrives as a Poisson process whose rate depends on the tightness in the market where the worker is searching. For an existing match, a good signal arrives with a rate either λ or 0, which depends on the match quality, and the exogenous separation arrives with a rate δ . The planner chooses an optimal stopping belief such that an existing match is destroyed when its match quality is lower than this cutoff belief. For each $\omega \in \Omega$, one can solve the associated value $S(\omega)$ and policy function $\theta(\omega)$, which is the analogue of $S^\Delta(\omega)$ and $\theta_*^\Delta(\omega)$ in the discrete time model, and solve the planner's aggregate payoff by simply calculating an integral, $\int_{\omega \in \Omega} S(\omega) d\mu$. As I will show, both $S(\alpha)$ and $\theta(\alpha)$ are differentiable; thus, the empirical implications of the continuous time model can be easily derived.

Note that the value and strategy, $S(\omega)$ and $\theta(\omega)$, are the solution of the continuous time planner's individual worker problem, whose formal expression is relegated to Appendix A. The continuous time version of planner's society problem involves consideration of infinite-dimensional states, μ_T , and complicated resulting dynamics. The proof of the separability of the planner's problem is difficult to work with, which is different from that in the discrete time model in Menzio and Shi (2011). The goal of taking the limit is to approximate the equilibrium dynamics in the original discrete time model when Δ is small and to obtain tractability of the equilibrium strategy, $\theta(\omega)$, and its implications. I characterize the solution of the planner's individual worker problem in continuous time in the following Lemma.

Lemma 3. *The solution of the planner's problem in continuous time satisfies the following Hamilton-Jacobi-Bellman (HJB) equations:*

$$rS(u) = b + p(\theta(u))[S(\alpha_0) - S(u)] - k\theta(u), \quad (16)$$

where

$$k = p'(\theta(u))[S(\alpha_0) - S(u)], \quad (17)$$

when $\omega = u$,

$$rS(1) = \lambda + \delta[S(u) - S(1)], \quad (18)$$

when $\omega = 1$,

$$S(\alpha) = S(u), \quad (19)$$

when $\omega \in [0, \alpha_*)$,

$$\begin{aligned} rS(\alpha) &= \alpha\lambda + \lambda\alpha(S(1) - S(\alpha)) - \lambda\alpha(1 - \alpha)S'(\alpha) \\ &\quad + \delta(S(u) - S(\alpha)) + p(\theta(\alpha))[S(\alpha_0) - S(\alpha)] - k\theta(\alpha), \end{aligned} \quad (20)$$

where

$$k = p'(\theta(\alpha))[S(\alpha_0) - S(\alpha)] \text{ for } \omega \in (\alpha_*, \alpha_0), \text{ and } \theta(\alpha_0) = 0, \quad (21)$$

with boundary conditions

$$S(u) = S(\alpha_*), S'(\alpha_*) = 0, \quad (22)$$

when $\omega \in [\alpha_*, \alpha_0]$.

The proof is omitted since it follows the standard verification argument, which can be found in chapter 4 of Oksendal and Sulem (2005). The following Proposition shows that the solution of the social planner's discrete time problem can be approximated by equations (16–21) when the time interval between consecutive periods is small enough.

Proposition 1. *As Δ goes to zero, the planner's value function $S^\Delta(\omega)$ in the discrete time model converges uniformly to $S(\omega)$ on Ω , where $S(\omega)$ is the solution of (16–21). The policy function $\theta_*^\Delta(\omega)$ converges uniformly to the policy function $\theta(\omega)$ when $\omega \neq 1$, and 0 when $\omega = 1$, the optimal stopping belief α_*^Δ converges uniformly to α_* .*

I present an intuitive argument here, and relegate the proof to the Appendix A. To show that $S^\Delta(\omega) \rightarrow S(\omega)$, one needs to ensure the existence and uniqueness of the limit solution, $S(\omega)$. The following Lemma characterizes the properties of the solution of problem (16–21).

Lemma 4. *The solution of (16–21) exists and satisfies the following properties:*

1. *It is unique.*
2. *$S(\alpha)$ is increasing for all $\alpha \in [\alpha_*, \alpha_0]$, $\theta(\alpha)$ is decreasing for all $\alpha \in [\alpha_*, \alpha_0]$.*
3. *$S(\alpha)$ is convex for all $\alpha \in [\alpha_*, \alpha_0]$.*

After characterizing the properties of the limit $S(\omega)$, one can show the convergence step by step.

Policy function. First, look at the policy function $\theta_*^\Delta(\omega)$ associated with $S^\Delta(\omega)$ for $\omega \in (\alpha_*^\Delta, \alpha_0] \cup \{u\}$. For an unemployed worker, notice that the second term of the right-hand side of (13) can be represented as

$$S^\Delta(\alpha_0^+) - S^\Delta(u) - \underbrace{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S^\Delta(1)] - (\lambda\alpha_0\Delta + r\Delta)S^\Delta(\alpha_0^+) - b\Delta + r\Delta S^\Delta(u)}_{O(\Delta)},$$

where the summation of the last four terms is $O(\Delta)$. The Taylor expansion implies that $S^\Delta(\alpha_0^+) = S^\Delta(\alpha_0) + \frac{S^\Delta(\alpha_0^+) - S^\Delta(\alpha_0)}{\alpha_0^+ - \alpha_0}(\alpha_0^+ - \alpha_0)$ where $\alpha_0^+ - \alpha_0$ is $O(\Delta)$. Suppose $\lim_{\Delta \rightarrow 0} \frac{S^\Delta(\alpha_0^+) - S^\Delta(\alpha_0)}{\alpha_0^+ - \alpha_0}$ is finite, when Δ goes to zero, the right-hand side of (13) goes to $p'(\theta_*^\Delta(u))[S^\Delta(\alpha_0) - S^\Delta(u)]$. Since the left-hand side of (13) does not depend on Δ , (13) can be approximated by

$$k = p'(\theta_*^0(u))[S^\Delta(\alpha_0) - S^\Delta(u)], \quad (23)$$

when Δ is close to zero.

Similarly, for an employed worker with belief $\alpha \in (\alpha_*^\Delta, \alpha_0)$, (14) can be approximated by

$$k = p'(\theta_*^0(\alpha))[S^\Delta(\alpha_0) - S^\Delta(\alpha)], \quad (24)$$

when Δ is close to zero. For α_0 -worker, $\theta_*^\Delta(\alpha_0) = 0$, since there is no surplus from job replacement.

Value function. Then consider the value function $S^\Delta(\omega)$. For an unemployed worker, (10) can be reformulated as

$$S^\Delta(u) = b\Delta + e^{-r\Delta}S^\Delta(u) + e^{-r\Delta}\Delta \max_{\theta} \{p(\theta)[S^\Delta(\alpha_0) - S^\Delta(u) + O(\Delta)] - k\theta\},$$

where the optimal $\theta_*^\Delta(u)$ can be approximated by $\theta_*^0(u)$, which satisfies (23). Hence, together with $1 - r\Delta$ as an approximation to $e^{-r\Delta}$, $S^\Delta(u)$ can be approximated by

$$b\Delta - k\theta\Delta + (1 - r\Delta)S^\Delta(u) + (1 - r\Delta)p(\theta_*^0(u))\Delta[S^\Delta(\alpha_0) - S^\Delta(u)] + O(\Delta^2).$$

By taking Δ to zero and ignoring the $O(\Delta)$ term, one obtains

$$rS^0(u) = b - k\theta_*^0(u) + p(\theta_*^0(u))[S^0(\alpha_0) - S^0(u)].$$

Similarly, for an employed worker with belief $\alpha \in (\alpha_*^\Delta, \alpha_0]$, $S^\Delta(\alpha)$ can be approximated by

$$\begin{aligned} & \delta b\Delta^2 + \delta\Delta(1 - r\Delta)S^\Delta(u) + (1 - \delta\Delta)\{\lambda\alpha\Delta + \lambda\alpha\Delta(1 - r\Delta)S^\Delta(1) \\ & + (1 - \lambda\alpha\Delta)(1 - r\Delta)S^\Delta(\alpha^+) + (1 - r\Delta)\{p(\theta_*^0(\alpha))\Delta[S^\Delta(\alpha_0) - S^\Delta(u)] - k\theta_*^0(\alpha)\Delta\} + O(\Delta^2), \end{aligned}$$

and $\lim_{\Delta \rightarrow 0} S(\alpha^+) = S(\alpha) + \lim_{\Delta \rightarrow 0} \frac{S(\alpha^+) - S(\alpha)}{\alpha^+ - \alpha}(\alpha^+ - \alpha) = S(\alpha) - \lambda\alpha(1 - \alpha)S'(\alpha)$. Hence, by taking Δ to zero, one obtains

$$\begin{aligned} rS^0(\alpha) &= \alpha\lambda + \lambda\alpha(S^0(1) - S^0(\alpha)) - \lambda\alpha(1 - \alpha)\frac{dS^0(\alpha)}{d\alpha} \\ &\quad + \delta(S^0(u) - S^0(\alpha)) + p(\theta_*^0)[S^0(\alpha_0) - S^0(\alpha)] - k\theta_*^0. \end{aligned}$$

By the same logic, $rS^0(1) = \lambda + \delta[S^0(u) - S^0(1)]$. It is clear that $S^0(\omega)$ for any $\omega \in \Omega$ satisfying (16,18,19,20) and $\theta_*^0(\omega)$ for any $\omega \in [\alpha_*, \alpha_0] \cup \{u\}$ satisfying (17,21).

Boundary condition. Finally, consider the limit of the optimal stopping belief. α_*^Δ such that $S^\Delta(\alpha) \geq S^\Delta(u)$ when $\alpha \geq \alpha_*^\Delta$ and $S^\Delta(\alpha) < S^\Delta(u)$ when $\alpha < \alpha_*^\Delta$. Hence, $S^\Delta(\alpha_*^\Delta) = S^\Delta(u)$ due to the continuity of the value function. By taking Δ to zero, one obtains $S^0(\alpha_*^0) = S^0(u)$, where α_*^0 is the accumulated point of α_*^Δ . In the optimal stopping literature, it is called the value-matching condition.

Another boundary condition, smooth-pasting conditions, is also necessary to determine the optimal stopping belief when Δ goes to zero. To see why, consider a deviation by stopping at $\alpha_*^{\Delta+}$ instead of α_*^Δ . Let $S(\alpha|\alpha_*)$ be the value at α by adopting stopping belief α_* . Then the value difference induced by two stopping beliefs, $\alpha_*^{\Delta+}$ and α_*^Δ is

$$\begin{aligned} & S(\alpha_*^{\Delta+}|\alpha_*^{\Delta+}) - S(\alpha_*^{\Delta+}|\alpha_*^\Delta) \\ = & (1 - \delta\Delta)\lambda(\alpha_*^{\Delta+} - \alpha_*^\Delta)\Delta[1 + e^{-r\Delta}S^\Delta(1) - e^{-r\Delta}S^\Delta(u)] \\ & + (1 - \delta\Delta)e^{-r\Delta}\{[p(\theta_*^\Delta(\alpha_*^{\Delta+})) - p(\theta_*^\Delta(\alpha_*^\Delta))]\Delta[S^\Delta(\alpha_0) - S^\Delta(u)] \\ & - k[\theta_*^\Delta(\alpha_*^{\Delta+}) - \theta_*^\Delta(\alpha_*^\Delta)]\Delta\} + O(\Delta^2), \end{aligned}$$

which is $O(\Delta^2)$, and therefore $S'(\alpha_*^{0+}) = \lim_{\Delta \rightarrow 0} \frac{S(\alpha_*^{\Delta+}|\alpha_*^{\Delta+}) - S(\alpha_*^{\Delta+}|\alpha_*^\Delta)}{\alpha_*^{\Delta+} - \alpha_*^\Delta} = 0$. Similarly, one can verify that $S'(\alpha_*^{0-}) = 0$ by considering the value difference induced by stopping at $\alpha_*^{\Delta-}$ and α_*^Δ . Hence, when Δ goes to zero, the value-matching and smooth-pasting conditions hold at $\alpha = \alpha_*^0$.

In sum, $S^0(\omega)$ satisfies (16,18,19,20), $\theta_*^0(\omega)$ satisfies (17,21), and α_*^0 satisfies (22). It is worth mentioning here the possibility that on-the-job search improves the social value of learning because $\theta(\alpha) = 0$ is always a (suboptimal) choice for any $\omega \in \Omega$.

Remark 1. When $p(\theta) = \min\{\theta, 1\}$, there are no search frictions in the market, and the on-the-job search decision problem is a linear programming problem with a corner solution: $\theta(\alpha) = 1$ if $S(\alpha_0) - S(\alpha) > k$, $\theta = 0$ otherwise.

Remark 2. When the match quality is known, $\alpha_0 = 1$,⁴ the absence of learning implies that matched workers' and firms' values are constant over time and on-the-job search is not efficient. Hence, there is only one active submarket in the social planner's solution.

The intuition of the two remarks above is straightforward. The social planner's fundamental trade-off is between the replacement premium of an existing match and cost of creating a vacancy. When match quality is common knowledge, neither learning nor on-the-job search has value; hence, the optimal allocation is a corner solution. When the market is non-linearly frictionless, the fundamental trade-off becomes a linear programming problem, and distinguishing a match with a different belief is not necessary. Hence, the optimal allocation is a corner solution as well.

⁴In order to avoid a trivial result, assume the publicly known match quality is greater than b .

4 Decentralization

In this section, I consider the implementation of the social planner's solution. I first describe the structure of the labor market and the nature of the employment contracts. I then derive the conditions on the firms and workers' value and policy functions that need to be satisfied in the market equilibrium. I also establish that there exists a unique equilibrium for the market economy and that this equilibrium is efficient, in the sense that it decentralizes the solution to the planners' problem, and block recursive, in the sense that the individual's value and policy functions depend only on his own state. Finally, I derive some empirical implications of the market economy.

I assume that the contracts offered by firms to workers are bilaterally efficient in the sense that they maximize the joint value of the match, that is, the sum of the worker's expected lifetime utility and the firm's expected lifetime profits. I make this assumption because there are a variety of specifications of the contract space under which the contracts that maximize the firm's profits are, in fact, bilaterally efficient. As I will show, the profit-maximizing contracts are bilaterally efficient if the contract space is complete in the sense that a contract can specify the tenure independent wage, w , the separation probability z and the worker's on-the-job search strategy. This result is intuitive. The firm maximizes its profits by choosing the contingencies z, x so as to maximize the joint value of the match and by choosing the contingencies for w so as to deliver the promised value x .⁵

4.1 Market Equilibrium

In order to decentralize the social planner's optimal allocation, I first define the joint surplus of an existing match $M(\omega, \mu)$ for $\omega \in B_\Delta \cup \{u\}$. Note that $M(\omega, \mu)$ is not the social surplus generated by the match since the matched worker and firm do not take into account the wage posting cost paid by the worker's potential new employer.

Labor market supply side. First, consider an employed worker at the beginning of the search and matching stage. Since the contract is bilaterally efficient, given the equilibrium market tightness function $\theta(x, \mu)$, the worker chooses to search in the submarket with promised value x to maximize

⁵Moreover, one can prove that the profit-maximizing contracts are bilaterally efficient if they can specify the wage only as a function of tenure and productivity (while the separation and search decisions are made by the worker). This result is also intuitive. The firm maximizes its profits by choosing the wage when it meets a worker so as to deliver the promised value x and by choosing the wage as a function of the belief about the match so as to induce the worker to maximize the joint value of the match (by setting the wage equal to the product of the match). Alternatively, profit-maximizing contracts are bilaterally efficient if they can specify severance transfers that induce the worker to internalize the effect of his separation and search decisions on the profits of the firm. See Moen and Rosen (2004), and Menzio and Shi (2009, 2011) for more examples.

the continuation value of his current match, which is given by

$$p(\theta(x, \mu))\Delta x + [1 - p(\theta(x, \mu))\Delta]\{\lambda\alpha\Delta[1 + e^{-r\Delta}M^\Delta(1, \mu)] + (1 - \lambda\alpha\Delta)e^{-r\Delta}M^\Delta(\alpha^+, \mu)\}.$$

With probability $p(\theta)\Delta$, the worker may find a new job and obtains a value x and the firm obtain zero, while with a complementary probability the current match survives and it may send a good signal with a subjective probability $\lambda\alpha\Delta$.

At the separation stage, the employer predicts the worker's choice and chooses the separation probability $z_d \in [\delta\Delta, 1]$ to maximize the value of their match. Hence, their joint value is given by

$$\begin{aligned} M^\Delta(\alpha, \mu) = \max_{z, x} & z[b\Delta + e^{-r\Delta}M^\Delta(u, \mu)] \\ & + (1 - z)\{[1 - p(\theta(x, \mu))\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}M^\Delta(1, \mu)] \\ & + [1 - p(\theta(x, \mu))\Delta](1 - \lambda\alpha\Delta)e^{-r\Delta}M^\Delta(\alpha^+, \mu)] + p(\theta(x, \mu))\Delta x\}. \end{aligned} \quad (25)$$

If the match is separated following the contract, the firm obtains zero and the worker becomes unemployed. Otherwise, the worker follows the contract to search and maximize the continuation value of his current match.

By the same logic, an unemployed worker chooses to search in the submarket with promised value x to maximize his value, which is given by

$$M^\Delta(u, \mu) = \max_x \{[1 - p(\theta(x, \mu))\Delta][b\Delta + e^{-r\Delta}M^\Delta(u, \mu)] + p(\theta(x, \mu))\Delta x\}, \quad (26)$$

where the interpretation of the Bellman equation is similar to the one above.

Clearly, given a bilaterally efficient contract, an employed worker with state 1 will choose not to search in any submarket; thus his value is given by

$$M^\Delta(1, \mu) = \delta\Delta[b\Delta + e^{-r\Delta}M^\Delta(u, \mu)] + (1 - \delta\Delta)\{\lambda\Delta + e^{-r\Delta}M^\Delta(1, \mu)\}, \quad (27)$$

and $x(1, \mu) = \emptyset$ and $z(1, \mu) = \delta\Delta$.

Labor market demand side. Firms without a match are on the demand side of the labor market. They choose whether to enter the labor market and which submarket to enter. The competition in the labor market implies that firms' expected discounted profit is zero, and there is no difference between any of the submarkets for any firm. A firm may form a new match with probability $q(\theta)\Delta$, which depends on the tightness of the market the firm is in, and its expected profit is given by

$$\lambda\alpha_0\Delta(1 + e^{-r\Delta}M^\Delta(1, \mu)) + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}M^\Delta(\alpha_0^+, \mu) - x.$$

By posting a new job, a firm needs to pay a flow cost $k\Delta$ in each period. Hence, firms' free-entry condition implies that

$$\lambda\alpha_0\Delta(1 + e^{-r\Delta}M^\Delta(1, \mu)) + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}M^\Delta(\alpha_0^+, \mu) - x = k/q(\theta) = k\theta/p(\theta), \quad (28)$$

which pins down the equilibrium market tightness function $\theta(x, \mu)$.

The solution concept I used is the bilaterally efficient recursive competitive equilibrium (RCE) in which a matched firm and worker maximize their joint value.

Definition 1. A (bilaterally efficient) recursive competitive equilibrium (RCE, henceforth) consists of a market tightness function $\theta_d^\Delta : \mathbb{R} \times \Xi \rightarrow \mathbb{R}^+$, where $\theta_d^\Delta(x, \mu)$ is the market tightness of a submarket with promised value x in aggregate state μ , a value for the unemployed worker $M^\Delta(u, \mu)$, and an optimal submarket choice $x_d^\Delta(u, \mu) \in \mathbb{R} \cup \emptyset$, a joint value for the firm-worker match $M^\Delta : B_\Delta \times \Xi \rightarrow \mathbb{R}^+$, and policy functions for the firm-worker match $z_d^\Delta : B_\Delta \times \Xi \rightarrow [\delta\Delta, 1]$ and $x_d^\Delta : B_\Delta \times \Xi \rightarrow \mathbb{R} \cup \emptyset$. These functions satisfy the following conditions: (i) θ_d^Δ satisfies (28) for all $(x, \mu) \in \mathbb{R} \times \Xi$; (ii) $M^\Delta(u, \mu)$ satisfies (26) for all $\mu \in \Xi$, and $x_d^\Delta(u, \mu) \in \mathbb{R}$ is the associated policy function, (iii) $M^\Delta(1, \mu)$ satisfies (27) for all μ , and $M^\Delta(\alpha, \mu)$ satisfies (25) for all $(\alpha, \mu) \in (0, \alpha_0] \times \Xi$, and $z_d^\Delta(\omega, \mu)$ and $x_d^\Delta(\omega, \mu)$ are the associated policy functions.

Condition (i) guarantees that the market tightness function θ is consistent with the firm's incentives to create vacancies. Condition (ii) guarantees that the search strategy of an unemployed worker maximizes his lifetime utility, given the market tightness function. Condition (iii) guarantees that the matched firms and workers maximize their joint value. In general, workers' and firms' equilibrium strategies depend on both their individual states and the state of the economy, which is a distribution over Ω . Fortunately, following Menzio and Shi (2011), it can be shown that all bilaterally efficient recursive competitive equilibria satisfy the block recursive property in the sense that all firms' and workers' equilibrium strategy depend on their individual states only.

Definition 2. A (bilaterally efficient) block recursive equilibrium (BRE, henceforth) consists of a market tightness function $\theta_d^\Delta : \mathbb{R} \rightarrow \mathbb{R}^+$, where $\theta_d^\Delta(x)$ is the market tightness of a submarket with promised value x , a value for the unemployed worker M_u^Δ , and an optimal submarket choice $x_d^\Delta(u) \in \mathbb{R} \cup \emptyset$, a joint value for the firm-worker match $M^\Delta : B_\Delta \rightarrow \mathbb{R}^+$, and policy functions for the firm-worker match $z_d^\Delta : B_\Delta \rightarrow [\delta\Delta, 1]$ and $x_d^\Delta : B_\Delta \rightarrow \mathbb{R} \cup \emptyset$. These functions satisfy the following conditions: (i) θ_d^Δ satisfies (28) for all $(x, \mu) \in \mathbb{R} \times \Xi$; (ii) $M^\Delta(u)$ satisfies (26) for all $\mu \in \Xi$, and $x_d^\Delta(u) \in \mathbb{R}$ is the associate policy. (iii) $M^\Delta(1)$ satisfies (27) for all μ , and $M^\Delta(\alpha)$ satisfies (25) for all $(\alpha, \mu) \in (0, \alpha_0] \times \Xi$, and $z_d^\Delta(\alpha)$ and $x_d^\Delta(\alpha)$ are the associated policy functions.

The interpretation of three equilibrium conditions are similar to those in RCE. Taken together, they ensure that in a BRE, the strategies of each agent are optimal given the strategies of the other agents. The only difference between a BRE and a standard RCE is that the individual decision does not depend on market distribution μ .

Proposition 2. Suppose firms and workers can sign bilaterally efficient contracts. Then

1. All bilaterally efficient RCEs are block recursive.
2. There exists a unique BRE.
3. The BRE is socially efficient.

The proof of the first statement is omitted since it is similar to the argument of theorem 2 in Menzio and Shi (2011). Since the problems (25,26,27,28) are equivalent to the problems (10,11,12), the existence, uniqueness and efficiency of the equilibrium can be ensured. Since $M^\Delta(\omega) = S^\Delta(\omega)$ for any $\Delta > 0$, $M^\Delta(\omega)$ converges to $M(\omega) = S(\omega)$ as $S^\Delta(\omega)$ does.

Corollary 1. *As Δ goes to zero,*

1. The value function $M^\Delta(\omega)$ converges uniformly to the unique solution $S(\omega)$ on Ω , where $S(\omega)$ is given by (16-20).
2. The equilibrium $\theta_d^\Delta \circ x_d^\Delta(\omega)$ uniformly converges to $\theta(\omega)$ defined in Lemma 3 for $\omega \in \Omega$.
3. The equilibrium $\theta_d^\Delta(x)$ uniformly converges to a strictly decreasing and differentiable function $\theta_d(x)$ such that

$$k = q(\theta_d(x))[S(\alpha_0) - x]$$

4. The equilibrium $x_d^\Delta(\omega)$ uniformly converges to a strictly decreasing and differentiable function $x^*(\omega)$ such that

$$x^*(\omega) = \arg \max_x p(\theta(x))[S(\alpha_0) - S(\omega)] - k\theta(x)$$

5. The equilibrium stopping belief α_d^Δ induced by z_d^Δ uniformly converges to α_* .

The block recursivity of the competitive equilibrium crucially depends on the fact that search is directed. If one replaced the assumption of directed search with random search, the equilibrium could not be block recursive. Under random search, workers with different beliefs all have to search in the same market. When this is the case, the firms' expected value from meeting a worker depends on how employed workers are distributed across different beliefs about their current match quality, as this distribution determines the probability that the employment contract offered by the firm will be accepted by a randomly selected worker. In turn, the free-entry condition implies that the probability that a firm meets a worker must also depend on the distribution of workers. Since the meeting probability between firms and workers depends on the distribution, so do all the agents' value and policy functions.

4.2 Empirical Implications of Continuous Time Limit

In section 4.1, I show that when matched workers and firms make the decision to maximize their joint value $M(\alpha)$, the market equilibrium is equivalent to the social planner's problem. However, how can the allocation above be implemented by a contract? In this section, I will provide one contract that can induce the bilateral efficient allocation and characterize its empirical implications.

Consider that workers and firms can contract $\mathcal{C} = (w_c, \alpha_c, x_c) \in \mathbb{R} \times [0, \alpha_0] \times (\mathbb{R} \cup \emptyset)$, where there is (1) a tenure independent flow wage w_c , (2) the separation cutoff belief α_c , (3) and the worker's on-the-job search strategy $x_c(\alpha)$. Denote $J(\alpha|\mathcal{C})$, $V(\alpha|\mathcal{C})$ as the matched firm's and worker's value when the belief about their match is α , and the wage is w_c .

Since utility is perfectly transferable between firm and worker, the firm's optimal contract in market (θ, x) maximizes its expected life time profit such that the worker's expected life time utility is not less than x .

Proposition 3. *In equilibrium, the profit-maximizing contract \mathcal{C}^* is given by $(w, \alpha_*, x^*(\alpha))$, where*

1. *the wage clause, $w = W(x)$, specifies the worker's flow wage during the current match, where W is strictly increasing in the promised value x and such that $V(\alpha_0, W(x)) = x$.*
2. *the stopping belief clause, α_* , specifies the match separation criteria, which satisfies the requirement in Proposition 1, and*
3. *the on-job-search clause, $x^*(\alpha)$, specifies the submarket the α -worker should target, and satisfies $\theta_d \circ x^*(\alpha) = \theta(\alpha)$ where $\theta_d(x)$ and $\theta(\alpha)$ are given by Corollary 1 and Proposition 1.*

A tenure-contingent on-the-job search strategy will internalize the externality of on-the-job search costs and firing decisions at each belief level; thus, it is not surprising that the contract implemented is bilaterally efficient. Since utility is transferable, the different constant wage distributes the joint value between matched firms and workers differently. Since there is a one-to-one map between x and wage, multiple wages appear in equilibrium in the discrete time model. In the continuous time limit, the belief updates continuously and so does the target submarket wage. Hence, the equilibrium wage set is connected.

Corollary 2. *There exists a strictly increasing and differentiable function $w^{new} : (\alpha_*, \alpha_0) \cup \{u\} \rightarrow \mathbb{R}$ such that*

$$w^{new} = W \circ x^*.$$

This result is immediate given Corollary 1 and Proposition 3, and the strictly monotonicity and differentiability of $w^{new}(\cdot)$ results from those of W and x^* . It says that if a worker with state ω forms a new match, he gets a unique $w^{new}(\omega)$ as his new wage level. In particular, consider an employed worker with a wage w and an employment state, $\omega = \alpha \in (\alpha_*, \alpha_0)$, it must be that he actively searches for other jobs. If this worker with (w, α) gets a new job, then his new wage level will be given by $w^{new}(\alpha)$, which depends solely on α , but not on his previous wage level w . This will be useful when examining the stationary wage distribution.

4.2.1 Job Turnover

At the micro-level, it has been well established that (1) on-the-job search probability is decreasing in tenure, and (2) the separation rate is a function of tenure, which is increasing at first, then decreasing, and eventually becomes constant. The relevant literature can be traced back to Parsons (1972), who finds a negative correlation between match quality and the separation rate. Later, Farber (1994) finds the initially positive and soon negative relation between a worker's tenure and the separation rate. Parsons (1991) and Pissarides and Wadsworth (1994) find a negative relation between tenure and the propensity to search on the job. See a survey by Farber (1999) for more details. A more recent work, Nagypal (2007), finds that learning on match quality is the main driving force of the dynamics of separation rates. My model has predictions that are in line with these stylized facts at the individual level, especially the on-the-job search (OJS) probability, the employment-to-employment transition (EE) rate, the employment to unemployment transition rate (EU) and the separation rate. The macro-level predictions on wage distribution will be derived in section 4.2.2.

The driving force of the employment dynamics in this mechanism is learning and updating beliefs about match qualities. A specific match is either good for sure or still has unknown quality. For a good match, there is no on-the-job search, and job destruction happens only exogenously. If the match quality is uncertain, the belief is less favorable as t increases, and consequently, search intensity while on the job is higher, as well as the EE rate. Endogenous job destruction finally takes place when the belief becomes "unbearable."

OJS Probability. Consider a randomly picked worker with tenure t . Without knowing his signal history, one does not know for sure whether this match is good or whether this worker is searching on the job. Nonetheless, it is possible to find the ex ante probability that a randomly chosen worker is searching on the job, as a function of t , which is

$$\sigma_t \equiv \begin{cases} \alpha_0 \Pr(\tau > t) + (1 - \alpha_0) & t \leq t^*, \\ 0 & t > t^*, \end{cases}$$

where τ is a random variable representing the time at which the good signal occurs, so the probability that it has not happened by t is $\Pr(\tau > t) = \exp(-\lambda t)$, which is decreasing in t . The critical time cutoff is defined as $t^* = \inf\{t > 0 | \alpha_0 - \alpha_* = \int_0^{t^*} \lambda \alpha_s (1 - \alpha_s) ds\}$, at which point the belief hits α_* and the firm optimally destroys the current uncertain match, so the worker becomes unemployed. Before this point, a match is bad with ex ante probability $1 - \alpha_0$, and the worker always searches for a new job. With complementary probability α_0 , the match is good, and the worker searches only when a good signal has not arrived; hence, the quality remains uncertain. Therefore, the model predicts that the on-the-job search intensity measured as the ex ante probability is at first decreasing in workers' tenure but eventually becomes constant.

EE Transition. The employment-to-employment transition rate is defined as the probability that an employed worker actively searches for and gets a new job, which is a function of tenure t given by

$$\xi_t^{ee} \equiv \sigma_t p(\theta(\alpha_t)).$$

The EE transition happens only if a worker is looking for another job. Conditional on that, the probability that he actually finds a new job is $p(\theta(\alpha_t))$. The tenure effect on ξ_t^{ee} is driven by two forces in opposite directions,

$$\dot{\xi}_t^{ee} = \underbrace{p' \theta' \dot{\alpha}_t}_{(+)} - \underbrace{\alpha_0 \lambda \exp(-\lambda t) p(\theta(\alpha))}_{(+)}.$$

The first one is that conditional on the match not having sent a good signal before, the matched worker becomes more pessimistic over time, and therefore, his on-the-job search becomes more aggressive. Thus the probability of getting a new job becomes greater, and this raises the EE rate. The second one that lowers the EE rate is simply the decreasing probability of a good match not having sent a good signal. When t is small, by assumption of the matching function, $p'(\theta)$ is large, so the first force dominates the second. As t approaches t^* , α goes to α_* , and θ' goes to zero because of the convexity of $S(\alpha)$, which implies that the effect of the first force goes to zero. Hence, the second one becomes dominant. Yet, if a random match's tenure is greater than t^* , only the good match can survive, in which case the EE transition rate is zero. To summarize, the expected EE transition rate is initially increasing, then decreasing, and becomes negligible in the long run.

EU Transition. The employment-to-unemployment transition rate as a function of tenure is $\xi_t^{eu} = \delta$ for any $t \neq t^*$, when the EU transition happens only as a result of exogenous separation. At $t = t^*$, in addition to exogenous separations, all matches that did not send a good signal will be endogenously destroyed, the measure of which is positive. The atom of EU rate results from the assumption of a precise and uniform learning process. If the learning process is heterogeneous as a

consequence of either different priors or noisy observations, such an atom can be eliminated. The mass point in EU hazard showing at a particular tenure point is considered empirically irrelevant. However, it fits the observation in the academic job market, where learning is based on relatively uniform and precise information on the quality and quantity of research publications.

Separation Rate. Separation of an existing match may result from either EE or EU transition, hence the separation rate of an existing match with tenure t , ξ_t must be $\xi_t = \xi_t^{ee} + \xi_t^{eu}$. When $t < t^*$, it is given by

$$\xi_t \equiv \alpha_0 \{ \Pr(\tau > t) [p(\theta(\alpha_t)) + \delta] + \Pr(\tau \leq t) \delta \} + (1 - \alpha_0) [p(\theta(\alpha_t)) + \delta].$$

For a match with tenure $t < t^*$, the quality can be either good or bad, and tenure has two opposing effects on the separation probability, which are,

$$\dot{\xi}_t = \underbrace{p' \theta' \dot{\alpha} \sigma_t}_{(+)} - \underbrace{\alpha_0 \lambda \exp(-\lambda t) p(\theta(\alpha))}_{(+)}.$$

Generically, $\xi_t = \xi_t^{ee} + \delta$, the economic intuition behind $\dot{\xi}_t$ is almost identical to that of the EE rate.

Consider a match with $t > t^*$, then this match must be good for sure and separation occurs only as exogenous destruction at rate δ . Accordingly, the tenure effect $\dot{\xi}_t$ is zero. At $t = t^*$, as discussed before, a positive measure of matches will be separated. Just as that of EU hazard, this mass point of separation hazard at t^* is also empirically irrelevant. In summary, the expected separation rate has a shape similar to that of the EE rate, which is also increasing at first, then decreasing, and eventually becomes constant.

4.2.2 Stationary Wage Distribution

Now I turn to the model's macro-level predictions on employment dynamics, which are in line with another set of empirical findings. Specifically, I look at the wage dispersion. The equilibrium search and learning strategy, (θ, α_*) , implies that a worker's state $\omega(T)$ follows a continuous time Markov process, which is right continuous with respect to calendar time. Let $\Omega^* = \{u\} \cup \{1\} \cup [\alpha_*, \alpha_0]$ denote the state space of $\omega(T)$ in the equilibrium. It is clear that the process is aperiodic, strongly recurrent, and ergodic. In a large labor market with a continuum population of workers, by assuming "the law of large number" holds, the invariant distribution, if it exists, can be interpreted as the stationary cross-sectional distribution of workers' state. In particular, the following Proposition shows that the market equilibrium has the unique stationary wage distribution μ^* : There are two mass points at $\omega = 1$ and $\omega = u$. For $\omega \in [\alpha_*, \alpha_0]$, the probability density function is well defined. Denote the p.d.f. of stationary belief distribution as

$\phi(\alpha)$ for $\alpha \in [\alpha_*, \alpha_0]$, $\beta = \mu^*(1)$ the probability mass at $\omega = 1$, and $v = \mu^*(u)$ at $\omega = u$. Since the policy functions of matched workers and firms do not depend on wages, the evolution and hence the stationary distribution of beliefs are identical for different wages. These findings can be summarized in the following Proposition.

Proposition 4. *The stationary distribution of workers' state, μ^* , is characterized by (v, β, ϕ) where*

$$v = \mu^*(u), \beta = \mu^*(1), \phi(\alpha) = \frac{d\mu^*(\alpha)}{d\alpha} \text{ for } \alpha \in [\alpha_*, \alpha_0],$$

the probability density function $\phi(\alpha)$ is given by

$$\phi(\alpha) = \kappa(\alpha)/A \text{ for } \alpha \in (\alpha_*, \alpha_0], \quad (29)$$

$$\phi(\alpha_*) = 0, \quad (30)$$

and β, v such that

$$\beta = \frac{1}{\delta A} \int_{\alpha_*}^{\alpha_0} \lambda \alpha \kappa(\alpha) d\alpha, \text{ and } v = \frac{\int_{\alpha_*}^{\alpha_0} \lambda \alpha \kappa(\alpha) d\alpha + \int_{\alpha_*}^{\alpha_0} \delta \kappa(\alpha) d\alpha + 1}{Ap(\theta(u))}, \quad (31)$$

where

$$A = \int_{\alpha_*}^{\alpha_0} \kappa(\alpha) d\alpha + \frac{\int_{\alpha_*}^{\alpha_0} (\lambda \alpha + \delta) \kappa(\alpha) d\alpha + 1}{\delta p(\theta(u))} \int_{\alpha_*}^{\alpha_0} \lambda \alpha \kappa(\alpha) d\alpha,$$

and

$$\kappa(\alpha) = \exp \left[\int_{\alpha_*}^{\alpha} \frac{\lambda s + \delta + p(\theta(s))}{\lambda s(1-s)} ds \right].$$

In equilibrium, all matched pairs with the same α exhibit the same behavior in job searching, separation and new match formation, regardless of different current wages. At each time, the inflow at each w is the rate at which new matches are formed with contracted wage $w = w^{new}(\omega)$, $\omega \in \Omega$, and with the same prior belief α_0 . By Corollary 2, new matches formed with w come from two sources, the unemployed workers if $w = w^{new}(u)$, and employed workers with $\alpha = (w^{new})^{-1}(w)$. Meanwhile, the outflow is determined by integrating all rates at which matches with wage w and different beliefs $\alpha \in (\alpha_*, \alpha_0) \cup \{1\}$ are destroyed. Stationarity requires inflow and outflow to be equal at all wages w . And because of an identical belief distribution for every w , the outflow rates are in fact the same for all wages. Consequently, the shape of the stationary wage distribution can be captured by the wage distribution of newly formed matches with (w, α_0) as there is a one-to-one mapping between the new wage and the belief about the quality of the previous match, again by Corollary 2. The measure of new matches is $\phi(\alpha)p(\theta(\alpha))$, where α is the belief about the quality of the previous match, the denominator is the total measure of new matches, and the numerator is the probability that a worker with α gets a new job. The following Proposition summarizes the intuition above.

Proposition 5. *The stationary wage measure, $G_w^*(w)$, satisfies the following properties*

1. *It has a bounded and connected support $[\underline{w}, \bar{w})$, where $\underline{w} = w^{new}(u) = \lim_{\alpha \searrow \alpha_*} w^{new}(\alpha)$, and $\bar{w} = \lim_{\alpha \nearrow \alpha_0} w^{new}(\alpha)$,*
2. *$G_w^*(\underline{w}) = g_w^u > 0$, and $G_w^*(w) = 0$ for all $w \in (\underline{w}, \bar{w})$, and*
3. *For $w \in (\underline{w}, \bar{w})$, the p.d.f. $g_w(w) = \frac{dG_w^*(w)}{dw}$ is well defined and given by*

$$g_w(w) = \frac{\phi((w^{new})^{-1}(w))p(\theta(w^{new})^{-1}(w))}{\int_{\alpha_*}^{\alpha_0} \phi(s)[p(\theta(s)) + \delta]ds + 1/A}.$$

For $w = \underline{w}$, there is a mass point with measure $g_w^u > 0$, such that

$$g_w^u = \frac{vp(\theta(u))}{\int_{\alpha_*}^{\alpha_0} \phi(s)[p(\theta(s)) + \delta]ds + 1/A}.$$

The first part directly results from Proposition 4 and Corollary 2. The second part is obtained by making inflow equal outflow for all $w \in [\underline{w}, \bar{w})$.

Since $\int_{\alpha_*}^{\alpha_0} \phi(s)[p(\theta(s)) + \delta]ds + 1/A$ is a constant number in the steady state, I focus on the curvature of $\phi(\alpha)p(\theta(\alpha))$,

$$\frac{d}{d\alpha}[\phi(\alpha)p(\theta(\alpha))] = \underbrace{\phi'p}_{(+)} + \underbrace{\phi p'\theta'(\alpha)}_{(-)},$$

with $\phi' > 0$ and $\theta' < 0$. The first term $\phi'p$ is the *separation effect* due to both endogenous and exogenous separation of current matches. When $\alpha \rightarrow \alpha_0$, the separation rate is small, and therefore the separation effect is bounded. The second term $\phi p'\theta'(\alpha)$ is the *search effect* due to the change in search behavior. In particular, employed workers adjust their target submarkets of on-the-job search according to the change in α : the smaller an α , the bigger a probability of EE transition. Hence, the second term is non-increasing in α . When $\alpha \rightarrow \alpha_0$, we have $\theta(\alpha) \rightarrow 0$, and $p' \rightarrow \infty$, and therefore the second term goes to infinity, while the first effect remains finite. Therefore, the search effect may dominate the separation effect when α is close to α_0 , and the whole term becomes negative.

When the friction in the labor market is small enough, one can obtain a wage distribution with a flat tail, which is consistent with robust empirical findings in the literature.⁶ For intuition, consider a labor market with a Cobb-Douglas matching function $p(\theta) = \theta^\gamma$ where $\gamma \in [0, 1]$. The planner maximizes the social gain from on-the-job search of employed workers, $\max_\theta \theta^\gamma [S(\alpha_0) - S(\alpha)] - k\theta$. The marginal cost of job creation is k , and the marginal benefit is $\gamma\theta^{\gamma-1}[S(\alpha_0) - S(\alpha)]$

⁶See Mortensen (2005).

which depends on both γ and $S(\alpha_0) - S(\alpha)$. Higher α leads to a smaller surplus of replacement $S(\alpha_0) - S(\alpha)$, and therefore, fewer vacancies are created. Smaller search frictions (large γ) also lead to fewer vacancies being created. In the decentralized labor market, a low job-finding rate for high α workers implies a small measure of high wage offers, and therefore, the mass of high-wage jobs is relatively small in the stationary wage distribution. When search frictions are small (γ close to 1), the planner does not create many vacancies unless α is small enough, $\theta(\alpha)$ is small for large α . By Proposition 4 and Corollary 2, $w = w^{new}(\theta(\alpha))$ has a small mass when α is large, but a large mass when α is small. Hence, the stationary wage distribution has a flat tail.

Related stylized facts in the macro-labor literature are summarized in Mortenson (2005). On the theoretical side, this flat tail has been attributed to multiple-job applications (Galenianos and Kircher, 2009), on-the-job training (Fu, 2011), a combination of ex post selection and tenure effect (Moscarini, 2005), etc. However, in this paper, it is the coexistence of match quality learning and on-the-job search frictions that generates the wage dispersion. On the one hand, without search frictions, there is only one submarket; hence, the market wage distribution is degenerated even though there is a non-trivial learning process about match quality. On the other hand, if the match quality is deterministic, then there is no need for learning, and naturally no ex post heterogeneity among workers, so the wage dispersion disappears as well. Therefore, it is the interaction between the two attributes that leads to the difference in wages. It would be interesting to empirically decompose the effect of search frictions and the effect of learning in a directed search framework, which is left to future work.

4.2.3 Distribution Convergence

A natural question is whether the stationary distribution μ^* is also the limit distribution of the process. In other words, from an arbitrary initial state distribution $\mu_0 \in \Xi$, whether μ_T will eventually converge to the unique stationary distribution μ^* on the path of equilibrium. In a discrete time model with time length Δ , since a worker's state $\omega(T)$ follows a Markov process with finite state space Ω_Δ^* , μ_T converges to its invariant distribution μ_Δ^* by theorem 11.4 in Stokey, Lucas and Prescott (1989). Yet in the continuous time model, the state space Ω^* is a continuum and the Markov process can not be described by a transition matrix. Hence, the contraction mapping argument in Stokey, Lucas and Prescott (1989) can not be directly applied here. The following Proposition shows that from any initial distribution $\mu_0 \in \Xi$, the market equilibrium leads μ_T to the unique stationary wage distribution μ^* .

Proposition 6. *For any $\Delta > 0$, let $k \in \mathbb{N}$ and $T = k\Delta$, then μ_T converge to μ^* in the total variation norm as k goes to infinity.*

Since the Proposition 5 holds for any $\mu_T \in \Delta(\Omega^*)$, the wage distribution $G_w^{k\Delta}$ also converges to the stationary wage distribution G_w^* in the total variation norm.

5 Extensions

Additional Flow Payoff. If a match can produce a small and quality-independent flow payoff $y > 0$ as well as the quality-contingent reward, the logic of our model still works. The positive flow payoff reduces firms' experimentation cost and therefore extends the duration of experimentation. This observation implies that the equilibrium within the group wage dispersion may be enhanced as productivity grows, which is consistent with empirical observations about wage inequality trends.⁷ Furthermore, one can allow the flow payoff $y(t)$ to be a stochastic process. The block recursive properties still hold, but the solution of the social planner's learning problem is described as a partial differential equation. One can choose a particular y process such as a deterministic increasing trend to study the effect of technology improvement on the labor market, and an Itô process to study the relation between real business cycles and unemployment dynamics.

Bad News Cases. Assume the match quality independent flow payoff is $y > 0$, and $y - (1 - \alpha_0)\lambda > b > y - \lambda$ to avoid a trivial case. Suppose the payoff structure and learning process are given as follows: if a match is good, it never generates a negative one payoff; if it is bad, it may generate a unit of loss at a random time with a Poisson arrival rate λ . In this case, when bad news is realized, the firm learns that the match is bad and therefore fires the worker immediately. By observing a history with no bad news, a matched firm and worker become more and more optimistic about their match quality. In this economy, any existing match has a belief α higher than α_0 in equilibrium; thus on-the-job search is not valuable. The equilibrium has only one labor market with market tightness $\theta(u)$. In the decentralized market, there is only one equilibrium wage.

Poisson Bandit Cases. In the benchmark model, I assume that a bad match can not generate any profit, which seems restrictive. What if it can generate one unit of reward with a lower rate $\lambda_b \in (0, \lambda)$? To avoid a trivial case, I assume $\lambda_b < b < \alpha_0\lambda + (1 - \alpha_0)\lambda_b$. In other words, a new match is better than no match, but no match is better than a bad match. Then, given no reward arriving in $[t, t + \Delta)$, the belief at the end of that time period is $\alpha_{t+\Delta} = \frac{\alpha_t \exp(-\lambda\Delta)}{\alpha_t \exp(-\lambda\Delta) + (1 - \alpha_t) \exp(-\lambda_b\Delta)}$ by Bayes' rule. Yet, if one reward is realized in $[t, t + \Delta)$, the belief about the match quality jumps up from α_t to

$$\alpha_{t+\Delta} = \frac{\alpha_t[1 - \exp(-\lambda\Delta)]}{\alpha_t[1 - \exp(-\lambda\Delta)] + (1 - \alpha_t)[1 - \exp(-\lambda_b\Delta)]},$$

⁷Such as Bound and Johnson (1992), Autor, Katz and Kearney (2009), Kambourov and Manovskii (2009), etc.

by Bayes' rule. When Δ goes to zero, the updating can be approximated by

$$\lim_{\Delta \rightarrow 0} \frac{\alpha_{t+\Delta} - \alpha_t}{\Delta} = \begin{cases} -(\lambda - \lambda_b)\alpha_t(1 - \alpha_t) & \text{no reward at } t, \\ \frac{\lambda\alpha_t}{\lambda\alpha_t + \lambda_b(1 - \alpha_t)} & \text{one reward at } t, \end{cases}$$

and the probability that more than one reward is realized is $O(\Delta^2)$, which is negligible when Δ is small. By the same logic, one can solve the social planner's optimal stopping belief, and on-the-job search strategy. Over time, good matches can survive with higher probability than bad ones due to the dynamics of endogenous separation driven by learning; thus the empirical implications for job transitions still hold qualitatively. When the search friction is small, the wage distribution has a flat tail as well. However, the implications are slightly different from those in the benchmark model in the following sense: (1) No match is believed to be good for sure, and therefore the endogenous separation will not disappear even for the match with long tenure. (2) The arrival of a reward can increase the belief about match quality; thus it is possible that a belief $\alpha_t \in (\alpha_0, 1)$ appears in equilibrium. Clearly, it is inefficient to destroy a match with belief higher than α_0 , and therefore, employed workers with belief $\alpha > \alpha_0$ will not search on-the-job under a bilaterally efficient contract.

Informative Signal. In the benchmark model, the match is modeled as an experience good whose quality needs to be slowly learned over time. Yet, in some situations, the employer can extract non-trivial information about the match quality through an interview. Suppose a firm can draw an informative signal of the match quality and update its belief about the match quality through an interview before the match is formed. The signal is drawn from a match quality dependent distribution that satisfies monotone likelihood ratio property (MLRP), and the updated posterior $\tilde{\alpha}_0 \in [\underline{\alpha}_0, \bar{\alpha}_0]$, where $0 < \underline{\alpha}_0 < \bar{\alpha}_0 < 1$. In this extension, the social planner will form a new match only if the updated posterior $\tilde{\alpha}_0$ is higher than a cutoff level that depends on the worker's current state. For an unemployed worker, this cutoff is the stopping time belief α_* . For an employed worker, this cutoff is the belief α_t about the worker's current match quality. Let $\Pr(\tilde{\alpha}_0 > \alpha_t)$ be the ex ante probability that the posterior is larger than the worker's current belief. Hence, the on-the-job search policy is determined by

$$\max_{\theta} p(\theta) \Pr(\tilde{\alpha}_0 > \alpha_t) \{ \mathbb{E}[S(\tilde{\alpha}_0) | \tilde{\alpha}_0 > \alpha_t] - S(\alpha_t) \} - k\theta.$$

It is clear that both $\Pr(\tilde{\alpha}_0 > \alpha_t)$ and $\mathbb{E}[S(\tilde{\alpha}_0) | \tilde{\alpha}_0 > \alpha_t] - S(\alpha_t)$ are non-increasing in α_t ; thus the optimal policy $\theta(\alpha)$ is non-increasing in α , which is similar to that in the benchmark model. Hence the empirical implications for workers' turnover and the stationary wage distribution predicted by the benchmark model are qualitatively preserved.

Costly On-the-Job Search. Suppose workers' on-the-job search requires a flow cost $\epsilon\Delta$. To avoid a trivial case where on-the-job search is always suboptimal, I assume ϵ is small enough. Since

the gain from on-the-job search $\max_{\theta}\{p(\theta)[S(\alpha_0) - S(\alpha)] - k\theta\}$ is increasing in the job replacement premium, $[S(\alpha_0) - S(\alpha)]$, for small enough ϵ , there exists a cut-off belief α^ϵ such that

$$\begin{aligned}\max_{\theta} p(\theta)[S(\alpha_0) - S(\alpha)] - k\theta &\leq \epsilon \text{ if } \alpha \geq \alpha^\epsilon, \\ \max_{\theta} p(\theta)[S(\alpha_0) - S(\alpha)] - k\theta &> \epsilon \text{ if } \alpha < \alpha^\epsilon.\end{aligned}$$

In other words, matched workers would search on-the-job only if they believed the match quality is low enough. When $\alpha < \alpha^\epsilon$, the social planner's problem is unchanged, and therefore it is obvious that introducing costly on-the-job search does not change the main result but reduces the social surplus $S(\alpha)$ for each α and therefore shortens the duration of experimentation.

6 Conclusion

This paper presents a model of learning in a frictional labor market and studies the macroeconomic consequences of individual learning behavior between matched firms and workers. The match quality of a worker-firm pair is ex ante homogeneous but unknown. Over time, a matched worker gradually learns the match quality of his current job. Because of the diversity of individual histories, a heterogeneity arises among employed workers, which results in different on-the-job search behavior. I show that this learning can produce (1) interesting dynamics of workers' job transition and (2) wage dispersion among seemingly identical workers.

The implications are consistent with the stylized empirical evidence. The analysis shows that (1) the separation rate as a function of tenure is increasing first, decreasing later, and constant in the end; (2) the on-the-job search probability as a function of tenure is decreasing first and constant in the end; (3) when search frictions are small, the wage distribution has a flat tail as predicted by the empirical literature; and (4) in an extension, the model can generate a rise in wage inequality with an improvement in aggregate productivity. The contribution of this paper is to shed light on the aggregate consequences of individual learning in a tractable model and identify another source of wage dispersion with a flat tail. For future research, I believe it would be interesting to empirically evaluate the impact of learning on wage dispersion.

A Appendix: Proofs of Results on the Planner's Problem

A.1 Separability of Planner's Problem

Proof of Lemma 1

The proof is basically similar to that of theorem 1 in Menzio and Shi (2011). First, given $\Delta > 0$, let $C(\Xi)$ be the space of bounded continuous functions $R : \Xi \rightarrow \mathbb{R}$ with the sup norm, $\|R\| = \sup_{\mu \in \Xi} R(\mu)$. Define bounded function $f : \Omega \times \mathbb{R}^+ \times [\delta\Delta, 1] \rightarrow \mathbb{R}$, where given θ, z , $f(\theta, z, \omega)$ is the expected flow output produced by a worker with state ω . Hence

$$f(\theta, z, \omega) = \begin{cases} (1 - p(\theta)\Delta)b\Delta + p(\theta)\Delta\lambda\alpha_0\Delta & \text{if } \omega = u, \\ (1 - z)\lambda\Delta + zb\Delta & \text{if } \omega = 1, \\ (1 - z)[(1 - p(\theta)\Delta)\lambda\alpha\Delta + p(\theta)\Delta\lambda\alpha_0\Delta] + zb\Delta & \text{if } \omega \in B_\Delta/\{1\}. \end{cases}$$

For $\omega \in \Omega/(B_\Delta \cup \{u\})$, $f(\theta, z, \omega) = 0$ for any (θ, z) . Then the aggregated flow output $F(\theta, z, \mu) = \int_{\omega \in \Omega} f(\theta, z, \omega) d\mu$.

Define the operator \mathbf{T}^Δ on $C(\Xi)$. The planner's problem is reformulated as

$$\begin{aligned} \mathbf{T}^\Delta R(\mu) &= \max_{z, \theta} \left[\int_{\omega \in \Omega} f(\theta, z, \omega) d\mu + e^{-r\Delta} R(\hat{\mu}) \right] \\ \text{s.t. } & \text{(5 -- 8.) and } z : \Omega \rightarrow [\delta\Delta, 1], \theta : \Omega \rightarrow \mathbb{R}^+. \end{aligned} \tag{A.1}$$

Since the maximand is bounded in $[0, \lambda/r]$, $\mathbf{T}^\Delta R$ is bounded; Define $\bar{\theta}$ such that $k = p'(\bar{\theta})\lambda/r$. The job replacement benefit is bounded by λ/r . Then (13) and (14) imply that the optimal $\theta(\omega) \leq \bar{\theta}$ for any ω , and therefore one can replace the constraint that $\theta : \Omega \rightarrow \mathbb{R}^+$ in (A.1) by $\theta : \Omega \rightarrow [0, \bar{\theta}]$ without loss of any generality.

For each $R \in C(\Xi)$ and $\mu \in \Xi$, the maximand in (A.1) is continuous in (z, θ) and the set of feasible choices for $(z, \theta) \in [\delta\Delta, 1] \times [0, \bar{\theta}]$ is compact. Hence, the maximum is attained. Since the maximand is continuous, it follows from the Theorem of Maximum that $\mathbf{T}^\Delta R$ is continuous. Hence, $\mathbf{T}^\Delta : C(\Xi) \rightarrow C(\Xi)$. By Blackwell sufficient conditions for a contraction mapping, the solution of (A.1) is essentially unique. Since $\lim_{n \rightarrow \infty} e^{-n(r\Delta)} R^* = 0$, $R = \mathbf{U}^\Delta$ solves the social planner's problem for each $\Delta > 0$.

Given a particular $\mu \in \Xi$, let $C'(\Xi) \subset C(\Xi)$ be the set of functions $R : \Xi \rightarrow \mathbb{R}$ that are bounded, continuous and linear in the measure of unemployed workers, $\mu(u)$, the measure of successful matches, $\mu(1)$, and the pdf of existing uncertain matches $\mu(\alpha)$. In other words, if $R \in C'(\Xi)$, there exists $R_u, R_1 \in \mathbb{R}$ and $R_e : (0, \alpha_0] \rightarrow \mathbb{R}$ such that

$$R(\mu) = \mu(u)R_u + \mu(1)R_1 + \int_{\alpha \in [0, \alpha_0]} R_e(\alpha) d\mu.$$

Consider an arbitrary function $R \in C'(\Xi)$. After substituting the constraints into the maximand of (A.1) and changing the order of integrals and maximization,⁸ I have

$$(\mathbf{T}^\Delta R)(\Xi) = \mu(u)\tilde{R}_u + \mu(1)\tilde{R}_1 + \int_{\alpha \in [0, \alpha_0]} \tilde{R}_e(\alpha) d\mu, \quad (\text{A.2})$$

where

$$\tilde{R}_e(\alpha) = \max_{(\theta, z) \in [0, \bar{\theta}] \times [\delta\Delta, 1]} z[b\Delta + e^{-r\Delta}R_u] + (1-z)\{[1 - p(\theta)\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}R_1] \quad (\text{A.3})$$

$$+ [1 - p(\theta)\Delta](1 - \lambda\alpha\Delta)e^{-r\Delta}R(\alpha^+)] \\ + p(\theta)\Delta\lambda\alpha_0\Delta[1 + e^{-r\Delta}R_1] + p(\theta)\Delta(1 - \lambda\alpha_0\Delta)e^{-r\Delta}R(\alpha_0^+)\},$$

$$\tilde{R}_u = \max_{\theta \in [0, \bar{\theta}]} \{-k\theta\Delta + [1 - p(\theta)\Delta][b\Delta + e^{-r\Delta}R_u] \quad (\text{A.4})$$

$$+ p(\theta)\Delta\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}R_1] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}R(\alpha_0^+)\}\}, \text{ and}$$

$$\tilde{R}_1 = \max_{(\theta, z) \in [0, \bar{\theta}] \times [\delta\Delta, 1]} z[b\Delta + e^{-r\Delta}R_u] + (1-z)\{[1 - p(\theta)\Delta]\lambda\Delta[1 + e^{-r\Delta}R_1] \quad (\text{A.5})$$

$$+ [1 - p(\theta)\Delta](1 - \lambda\Delta)e^{-r\Delta}R_1] \quad (\text{A.6})$$

$$+ p(\theta)\Delta\lambda\alpha_0\Delta[1 + e^{-r\Delta}R_1] + p(\theta)\Delta(1 - \lambda\alpha_0\Delta)e^{-r\Delta}R(\alpha_0^+)\}. \quad (\text{A.7})$$

The definition of R and (A.2) imply that $\mathbf{T}^\Delta : C'(\Xi) \rightarrow C'(\Xi)$. Moreover, $C'(\Xi)$ is a closed subset of $C(\Xi)$, thus $\mathbf{U}^\Delta \in C'(\Xi)$.

Let $C''(\Xi) \subset C'(\Xi)$ be the set of functions $R : \Xi \rightarrow \mathbb{R}$ such that R_e is non-decreasing in α . Since $TR \in C'(\Xi)$, the maximand in (A.4) is non-decreasing in α . Hence $T : C''(\Xi) \rightarrow C''(\Xi)$. Since $C''(\Xi)$ is a closed subset of $C'(\Xi)$, $\mathbf{U}^\Delta \in C''(\Xi)$. Let $C'''(\Xi)$ be the set of functions $R : \Xi \rightarrow \mathbb{R}$ such that R_e is strictly increasing in α . I need to show that $R_e(\alpha)$ is strictly increasing in α . The difficulty is that it is hard to apply a fixed point theorem in $C'''(\Xi)$ which is not a closed set. Yet, I will show that any fixed point in $C''(\Xi)$ must belong to $C'''(\Xi)$ in this problem. Let R_e be non-decreasing in α , then the maximand in (A.4) is strictly increasing due to the fact that $\lambda\Delta[1 + e^{-r\Delta}R(1) - R(\alpha^+)] > 0$. Hence, if $\mathbf{U}^\Delta \in C''(\Xi)$, it must be the case that $\mathbf{U}^\Delta \in C'''(\Xi)$.

It follows that the policy correspondences $(z_*^\Delta, \theta_*^\Delta)$ solve the maximization problem (A.3-A.5) for $(R_e, R_u, R_1) = (S(\alpha), S(u), S(1))$. Since the maximand and the constraints in (A.3-A.5) do not depend on the economy state μ , $(\theta_*^\Delta, z_*^\Delta)$ depend on the individual state ω only.

Hence, I have

$$\mathbf{U}^\Delta(\mu) = \int_{\alpha \in B_\Delta / \{1\}} S^\Delta(\alpha) d\mu(\alpha) + \mu(u)S^\Delta(u) + \mu(1)S_1^\Delta, \quad (\text{A.8})$$

where $S^\Delta(\alpha), S^\Delta(u), S_1^\Delta$ are given by (10-12). ■

⁸A pointwise optimization result solves the intergration optimization with measure 1.

A.2 The Planner's Individual Problem in Continuous Time Limit

The Planner's Individual Control Problem in Continuous Time

A history for a particular worker is defined as $\{(y_T, m_T)\}_{T \geq 0} \in (\{0, 1, b\}^{\mathbb{R}^+} \times \mathbb{R}^+)$, where the stochastic process $\{y_T\}$ documents the history of realized production signal, where $y_T \in \{b, 1, 0\}$ for any $T \in \mathbb{N}$, and m_T denotes the starting time of the current match to trace both EE and UE transitions. Denote $\{\mathcal{F}_T\}_{T \geq 0}$ as the filtration generated by $\{(y_T, m_T)\}_{T \geq 0}$. Let θ_T, τ^d denote the control, and stopping time, which are both \mathcal{F}_T -adapted. Let $h^T = (y_{T'}, m_{T'})_{T'=0}^T$ be a possible history of a particular worker until calendar time T .

Given a control $\{\theta_T\}$ and a stopping time τ^d , the performance function is given by

$$S(\{h^T, \theta_T, \tau^d\}) = \mathbb{E} \left[\int_T^{T+\tau} e^{-rs} (y_{T+s} - k\theta_{T+s}) ds + e^{-r\tau} S(h^{T+\tau}, \theta, \tau^d) \right],$$

where $\tau = \tau_1 \wedge \tau_2 \wedge \tau_3 \wedge \tau^d$, τ_1 stands for the stopping time at which a good signal of the current match is sent, τ_2 is the stopping time when the worker finds a new job, τ_3 is the stopping time of exogenous separation, and τ^d stands for the stopping time when the current match is destroyed endogenously.

Without loss of any generality, we can focus on Markov control, where the state variable $\omega(T)$ is \mathcal{F}_T -adapted. And $\omega(T)$ evolves as follows: (1) $\omega(T) \in \Omega$, (2) ω jumps to α_0 with rate $p(\theta_T)$, (3) it jumps to 1 with rate λ only if $y \neq b$ and the current match is good, (4) it jumps to u with rate δ when $y_T \neq b$, and (5) if $y_T \neq b$, $d\omega(T) = -\lambda\omega(T)(1 - \omega(T))dT$. Since θ_T is \mathcal{F}_T -adapted, $\omega(T)$ is a well-defined stochastic process, and \mathcal{F}_T -adapted.

Given a control $\theta(\omega)$ and a cutoff belief to destroy the current match $\alpha_d \in [0, 1]$ such that $\theta_T = \theta(\omega(T))$, $\alpha_d = \sup\{\alpha_s | s = \tau^d\}$, the performance function conditional on ω is

$$S(\omega | \theta, \tau^d) = \mathbb{E}^\omega \left[\int_0^\tau e^{-rs} (y_s - k\theta_s) ds + e^{-r\tau} S(\tau, \omega_\tau | \theta, \tau^d) \right].$$

The problem is, for each $\omega \in \Omega$, to find the value function $S(\omega)$, an optimal separation belief $\alpha_* \in [0, 1]$, and an optimal control $\theta^* \in \mathcal{A}$ such that

$$S(\omega) \equiv \sup_{\theta(\omega), \alpha_d} S(\omega | \theta, \alpha_d), \tag{A.9}$$

where \mathcal{A} denotes all admissible controls such that \mathcal{F}_T -adapted processes, positive value and $\theta(\omega) \in L^2(\Omega, \mathbb{R}^+)$. ■

Proof of Proposition 1

For each $\Delta > 0$, Lemma 1 implies that one can define an operator $\hat{\mathbf{T}}^\Delta$ on $C(\Omega)$, the space of bounded continuous functions $S : \Omega \rightarrow \mathbb{R}$ with the sup norm, as (A.3,A.4,A.5), and it is equivalent to an optimal control combined stopping time problem (A.9). By definition of a fixed point, $(\hat{\mathbf{T}}^\Delta S^\Delta)(\omega) = S^\Delta(\omega)$.

I want to show that $[\theta^\Delta(\alpha), \theta^\Delta(u), \alpha_*^\Delta] \rightarrow [\theta(\alpha), \theta(u), \alpha_*]$ and $S^\Delta(\omega)$ uniformly converge to $S(\omega)$ as Δ goes to zero, where $S(\omega), \theta(\alpha), \theta(u), \alpha_*$ satisfying Proposition 1. In particular,

$$\begin{aligned} \|S^\Delta - S\| &= \|\hat{\mathbf{T}}^\Delta S^\Delta - \hat{\mathbf{T}}^\Delta S + \hat{\mathbf{T}}^\Delta S - S\| \\ &\leq (1 - r\Delta)\|S^\Delta - S\| + \|\hat{\mathbf{T}}^\Delta S - S\|, \end{aligned}$$

where $\|\cdot\|$ is the sup-norm on $B_\Delta \cup \{u\}$, and the inequality comes from the Blackwell discount theorem and triangle inequality. Hence

$$\|S^\Delta(\omega) - S(\omega)\| \leq \frac{1}{r\Delta} \|\hat{\mathbf{T}}^\Delta S - S\|.$$

I want to show that

$$\lim_{\Delta \rightarrow 0} \frac{1}{r\Delta} \|\hat{\mathbf{T}}^\Delta S - S\| = 0,$$

where

$$\begin{aligned} (\hat{\mathbf{T}}^\Delta S)(u) &= \max_{\theta} \{-k\theta\Delta + [1 - p(\theta)\Delta][b\Delta + e^{-r\Delta}S(u)] \\ &\quad + p(\theta)\Delta\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S(1)] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}S(\alpha_0^+)\}\}, \end{aligned}$$

$$\begin{aligned} (\hat{\mathbf{T}}^\Delta S)(\alpha) &= \max_{z, \theta} z[b\Delta + e^{-r\Delta}S(u)] + (1 - z)\{-k\theta\Delta + [1 - p(\theta)\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}S(1)] \\ &\quad + [1 - p(\theta)\Delta](1 - \lambda\alpha\Delta)e^{-r\Delta}S(\alpha^+)] \\ &\quad + p(\theta)\Delta\lambda\alpha_0\Delta[1 + e^{-r\Delta}S(1)] + p(\theta)\Delta(1 - \lambda\alpha_0\Delta)e^{-r\Delta}S(\alpha_0^+)\}, \end{aligned}$$

and

$$(\hat{\mathbf{T}}^\Delta S)(1) = \delta\Delta[b\Delta + e^{-r\Delta}S(u)] + (1 - \delta\Delta)\{\lambda\Delta + e^{-r\Delta}S(1)\},$$

with optimal control $(\hat{z}^\Delta, \hat{\theta}^\Delta)$ such that

$$\hat{z}^\Delta = \begin{cases} 1 & \text{if } \alpha \in B_\Delta, \text{ and } \alpha < \hat{\alpha}_*^\Delta, \\ \delta\Delta & \text{if } \alpha \in B_\Delta, \text{ and } \alpha \geq \hat{\alpha}_*^\Delta, \end{cases}$$

where $\hat{\alpha}^\Delta \in (0, \alpha_0]$, and $\hat{\theta}^\Delta$ such that

$$k = p'(\hat{\theta}^\Delta(u))\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S(1)] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}S(\alpha_0^+) - b\Delta - e^{-r\Delta}S(u)\}, \quad (\text{A.10})$$

$$\begin{aligned}
k &= p'(\hat{\theta}^\Delta(\alpha))\{\lambda\alpha_0\Delta[1 + e^{-r\Delta}S(1)] + (1 - \lambda\alpha_0\Delta)e^{-r\Delta}S(\alpha_0^+) \\
&\quad - \lambda\alpha\Delta[1 + e^{-r\Delta}S(1)] - (1 - \lambda\alpha\Delta)e^{-r\Delta}S(\alpha^+)\},
\end{aligned} \tag{A.11}$$

when $\alpha \in B_\Delta/\{1\}$, and $\hat{\theta}^\Delta(1) = 0$.

The differentiability of $S(\alpha)$ implies that $S(\alpha^+) = S(\alpha) + S'(\alpha)[\alpha^+ - \alpha] + O(\Delta^2)$ for all $\alpha \in [\alpha_*, \alpha_0]$. Multiplying both sides of equations (16,18,19,20) by Δ , and adding $(1 - r\Delta)S(\omega)$ for each $\omega \in B_\Delta \cup u$, one obtains

$$\begin{aligned}
S(u) &= b\Delta + p(\theta(u))\Delta[S(\alpha_0) - S(u)] - k\theta(u)\Delta + (1 - r\Delta)S(u), \\
S(\alpha) &= \alpha\lambda\Delta + \lambda\alpha\Delta(S(1) - S(\alpha)) - \lambda\alpha\Delta(1 - \alpha)S' \\
&\quad + \delta\Delta(S(u) - S(\alpha)) + p(\theta(\alpha))\Delta[S(\alpha_0) - S(\alpha)] - k\theta(\alpha)\Delta \\
&\quad (1 - r\Delta)S(\alpha), \\
S(1) &= \lambda\Delta + \delta\Delta[S(u) - S(1)] + (1 - r\Delta)S(1).
\end{aligned}$$

Hence, on $B_\Delta \cup \{u\}$, $\hat{\mathbf{T}}^\Delta S - S$ is well-defined.

Since (A.10,A.11) are continuous in Δ , $e^{-r\Delta} - (1 - r\Delta) = O(\Delta)$, and $p(\cdot)$ is single-value and continuous, $\lim_{\Delta \rightarrow 0} \|\hat{\theta}^\Delta - \theta\| = 0$. And therefore

$$\begin{aligned}
&[1 - p(\theta)\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}S(1)] + [1 - p(\theta)\Delta](1 - \lambda\alpha\Delta)e^{-r\Delta}S(\alpha^+) \\
&+ p(\theta)\Delta\lambda\alpha_0\Delta[1 + e^{-r\Delta}S(1)] + p(\theta)\Delta(1 - \lambda\alpha_0\Delta)e^{-r\Delta}S(\alpha_0^+),
\end{aligned}$$

goes to $S(\alpha)$, and $b\Delta + e^{-r\Delta}S(u)$ goes to $S(u)$ when Δ goes to zero. Hence $\lim_{\Delta \rightarrow 0} \|\hat{\alpha}_*^\Delta - \alpha_*\| = 0$, and therefore for $\omega \in (\alpha_*^\Delta, \alpha_0] \cup \{u\}$

$$\begin{aligned}
\frac{(\hat{\mathbf{T}}^\Delta S)(\omega) - S(\omega)}{\Delta} &= [p(\hat{\theta}^\Delta(\omega)) - p(\theta)][S(\alpha_0) - S(\omega)] - k(\hat{\theta}^\Delta(\omega) - \theta(\omega)) + O(\Delta) \\
&= p'(\theta(\omega))[S(\alpha_0) - S(\omega) - k](\hat{\theta}^\Delta(\omega) - \theta(\omega)) + O(\Delta).
\end{aligned}$$

For $\omega = 1$, $\frac{(\hat{\mathbf{T}}^\Delta S)(1) - S(1)}{\Delta} = 0$.

Hence, $\frac{\hat{\mathbf{T}}^\Delta S(\omega) - S(\omega)}{\Delta} = O(\Delta)$ for $\omega \in \Omega$. Since the convergence is independent of the path, the S^Δ uniformly converges to S , and therefore $(\theta^\Delta, z_*^\Delta)$ uniformly converges to (θ, z_*) . ■

A.3 Solution Characterization of Planner's Problem in Continuous Time Limit

Proof of Lemma 4

Existence and Uniqueness. Given any $S(u) \in [\frac{b}{r}, \frac{\lambda}{r}]$, from (16,18,19,20), there exists a unique solution $S(\alpha)$ such that

$$S(\alpha) = S(u) + \int_{\alpha_*}^{\alpha} h(x, S(x))dx, \quad (\text{A.12})$$

where $h(\alpha, S(\alpha))$ is the right-hand side of (20) divided by r .

And $\theta(u) = \arg \max \{p(\theta)[S(\alpha_0) - S(u)] - k\theta\}$. For an unemployed worker,

$$rS(u) = b + \max_{\theta} \{p(\theta)[S(\alpha_0) - S(u)] - k\theta\},$$

or

$$S(u) = \frac{b + p(\theta(u))S(\alpha_0) - k\theta(u)}{r + p(\theta(u))}. \quad (\text{A.13})$$

The envelop theorem implies that

$$\frac{dS(u|S(\alpha_0))}{dS(\alpha_0)} = \frac{p(\theta(u|S(\alpha_0)))}{r + p(\theta(u|S(\alpha_0)))} \in (0, 1). \quad (\text{A.14})$$

Combining (A.12) and (A.13) yields

$$S(\alpha) = \frac{b + p(\theta(u))S(\alpha_0) - k\theta(u)}{r + p(\theta(u))} + \int_{\alpha_*}^{\alpha} h(x, S(x))dx.$$

Define an operator $\mathbf{T}_p : C[0, \alpha_0] \rightarrow C[0, \alpha_0]$ where $S(\alpha) \in C[0, \alpha_0]$ is any bounded continuous differentiable function and $\mathbf{T}_p S = \frac{b - p(\theta(u))S(\alpha_0) - k\theta(u)}{r + p(\theta(u))} + \int_{\alpha_*}^{\alpha} h(x, S(x))dx$ where S such that (A.12) and (A.13). To prove the uniqueness of $S(\omega)$, one needs to verify whether \mathbf{T}_p is a contraction mapping. For the second part of $\mathbf{T}_p S$, by the standard contraction mapping argument of the existence and uniqueness of the solution in the problem of an ordinary differential equation with initial value, $\mathbf{T}_p^2 S = \int_{\alpha_*}^{\alpha} h(x, S(x))dx$ satisfies the Blackwell sufficient condition. Moving to the first part, \mathbf{T}_p^1 , one needs to check whether the Blackwell sufficient condition holds.

Let $S_1 > S_2$, then $S(u|S_1(\alpha_0)) > S(u|S_2(\alpha_0))$ following (A.14); thus the monotonicity of \mathbf{T}_p^1 is proved. Move to the discounting property. Let $n \in \mathbb{R}^+$, and $S_3 = S_1 + n$. Following (A.14), $\frac{dS(u|S(\alpha_0))}{dS(\alpha_0)} < 1$ for all $S(\alpha_0) \in [S_1(\alpha_0), S_1(\alpha_0) + n]$; thus the discounting property of \mathbf{T}_p^1 holds. Hence $\mathbf{T}_p = \mathbf{T}_p^1 + \mathbf{T}_p^2$ satisfies the Blackwell sufficient condition, and therefore it is a contraction mapping on a complete functional space, $C[0, \alpha_0]$. There exists a unique solution $S(\alpha)$ such that $S(\alpha) = \mathbf{T}_p S(\alpha)$, and $S(u)$ is also determined uniquely!

Convexity. Consider two α_1, α_2 such that (1) $\alpha_1, \alpha_2 \in B_\Delta(\alpha_0)$, (2) $\alpha_* < \alpha_2 < \alpha_1 \leq \alpha_0$. Let $\alpha^\zeta = \zeta\alpha_2 + (1 - \zeta)\alpha_1$ where $\zeta \in (0, 1)$. I want to show that for all α_1, α_2 and α^ζ , $S(\alpha^\zeta) \leq \zeta S(\alpha_2) + (1 - \zeta)S(\alpha_1)$. Denote $\theta^\zeta(\alpha)$ as the path of optimal on-the-job search starting from α^ζ during the current match. Denote by $S^g(\theta)$ the expected surplus from an arbitrary path of on-the-job search θ conditional on the true match quality being good and similarly for $S^b(\theta)$. Then $S(\alpha^\zeta) = \alpha^\zeta S^g(\theta^\zeta) + (1 - \alpha^\zeta)S^b(\theta^\zeta)$. And it holds that

$$\begin{aligned} S(\alpha_1) &\geq \alpha_1 S^g(\theta^\zeta) + (1 - \alpha_1)S^b(\theta^\zeta), \\ S(\alpha_2) &\geq \alpha_2 S^g(\theta^\zeta) + (1 - \alpha_2)S^b(\theta^\zeta), \end{aligned}$$

since θ^ζ is a feasible price path. Hence

$$\begin{aligned} \zeta S(\alpha_2) + (1 - \zeta)S(\alpha_1) &\geq \left\{ \begin{array}{l} \zeta[\alpha_1 S^g(\theta^\zeta) + (1 - \alpha_1)S^b(\theta^\zeta)] \\ +(1 - \zeta)[\alpha_2 S^g(\theta^\zeta) + (1 - \alpha_2)S^b(\theta^\zeta)] \end{array} \right\} \\ &= (\zeta\alpha_1 + (1 - \zeta)\alpha_2)S^g(\theta^\zeta) \\ &\quad + (1 - \zeta\alpha_1 - (1 - \zeta)\alpha_2)S^b(\theta^\zeta) \\ &= \alpha^\zeta S^g(\theta^\zeta) + (1 - \alpha^\zeta)S^b(\theta^\zeta) = S(\alpha^\zeta), \end{aligned}$$

which contradicts the fact that the solution of the HJB function maximizes the planner's individual worker problem. This proves the claim.

Monotonicity. Since $S'' \geq 0$, and $S'(\alpha_*) = 0$, thus $S' \geq 0$ for all $\alpha \in [\alpha_*, \alpha_0]$. There are two cases. The first one is $S' = 0$ for all $\alpha \in [\alpha_*, \alpha_0]$, which implies that $\alpha_* = \alpha_0$. The second one is that $S' > 0$ for all $\alpha \in (\alpha_*, \alpha_0]$. The optimal on-the-job search decision satisfies $p'(\theta(\alpha))[S(\alpha_0) - S(\alpha)] = k$. Differentiating it yields

$$p''(\theta(\alpha))[S(\alpha_0) - S(\alpha)]\theta'(\alpha) = p'(\theta)S'(\alpha),$$

since $S(\alpha_0) - S(\alpha) > 0$ for any $\alpha < \alpha_0$, $p'' < 0$, I have $\theta'(\alpha) < 0$. p', p'' and $S(\alpha_0) - S(\alpha)$ is bounded. When α goes to α_* , $S'(\alpha)$ goes to zero, and therefore θ' goes to zero. ■

B Appendix: Proofs of Results on the Market Mechanism

B.1 Decentralization

Proof of Proposition 2

Combining (25), (26) and (28) yields

$$\begin{aligned}
 M^\Delta(\alpha, \mu) &= \max_{z, \theta} \{ z[b\Delta + e^{-r\Delta}M^\Delta(u, \mu)] + (1-z)\{[1-p(\theta)\Delta]\lambda\alpha\Delta[1 + e^{-r\Delta}M^\Delta(1, \mu)] \\
 &\quad + [1-p(\theta)\Delta](1-\lambda\alpha\Delta)e^{-r\Delta}M^\Delta(\alpha^+, \mu)] + p(\theta)\Delta[\lambda\alpha_0\Delta(1 + e^{-r\Delta}M^\Delta(1, \mu)) \\
 &\quad + (1-\lambda\alpha_0\Delta)e^{-r\Delta}M^\Delta(\alpha_0^+, \mu)] - k\theta\Delta\}, \\
 M^\Delta(u, \mu) &= \max_{\theta} \{-k\theta\Delta + [1-p(\theta)\Delta][b\Delta + e^{-r\Delta}M^\Delta(u, \mu)] \\
 &\quad + p(\theta)\Delta[\lambda\alpha_0\Delta(1 + e^{-r\Delta}M^\Delta(1, \mu)) + (1-\lambda\alpha_0\Delta)e^{-r\Delta}M^\Delta(\alpha_0^+, \mu)]\},
 \end{aligned} \tag{B.1}$$

$$\tag{B.2}$$

where $\alpha^+ = \alpha + \lambda\alpha(1-\alpha)\Delta$.

One can find that $S^\Delta(\alpha)$ and $S^\Delta(u)$ uniquely solves the above problem since (B.1, B.2) are equivalent to (10, 11), and therefore $S^\Delta(1) = M^\Delta(1, \mu)$. Hence, the equilibrium joint value $M^\Delta(\omega, \mu) = S^\Delta(\omega)$, and the equilibrium job search and job separation strategy equals θ_*^Δ and z_*^Δ , and therefore the equilibrium is efficient.

Proof of Proposition 3

Since the joint value equals $J(\alpha|\mathcal{C}) + V(\alpha|\mathcal{C})$, then the firm's profit maximization problem is given by

$$\max_{w_c, x_c(\alpha), \alpha_c} M(\alpha_0|\mathcal{C}) - x \text{ s.t. } V(\alpha_0|\mathcal{C}) \geq x,$$

where $M(\alpha|\mathcal{C})$ is the joint value of the match given the contract \mathcal{C} . Given the binding constraint $V(\alpha_0|\mathcal{C}) = x$, the profit-maximizing contract \mathcal{C}^* will achieve the joint value $M(\alpha_0) = M(\alpha_0|\mathcal{C}^*)$, by setting an optimal separation rule, on-the-job search strategy given by Proposition 1 and wage $w = W(x)$ such that $V(\alpha_0|\mathcal{C}^*) = x$.

The monotonicity of $W(x)$ can be proved by contradiction. Suppose in equilibrium $w_1 > w_2$ such that $V(\alpha_0|w_1, \alpha_*, \theta) = V(\alpha_0|w_2, \alpha_*, \theta)$. Since the bilaterally efficient contract ensures both workers' on-the-job search and firms' stopping time decision are socially optimal. Two contracts induce the same history and stopping time. Yet, the expected cost of experimentation is different since $\mathbb{E} \int_0^\tau e^{-rt} w_1 dt - \mathbb{E} \int_0^\tau e^{-rt} w_2 dt > 0$. Thus no firm will choose w_1 . ■

B.2 Stationary Distribution

Proof of Proposition 4

The stationary distribution can be calculated by making the inflow equal outflow at any $\omega \in \Omega^*$. The density at α_* is zero since the Markov process is right continuous with respect to calendar time. For interior point α , the only inflow comes from match with belief $\alpha' > \alpha$ that survives but has not sent an good signal, while the outflow is $\mu(\alpha)$. In the steady state, $\mu_{T_1}(\alpha) = \mu_{T_2}(\alpha) = \phi(\alpha)$ for any $T_1, T_2 \geq 0$.

$$\lambda\alpha(1-\alpha)\frac{d}{d\alpha}\phi(\alpha) = [\lambda\alpha + \delta + p(\theta(\alpha))]\phi(\alpha), \quad (\text{B.3})$$

where $\phi(\alpha)$ is the probability density at α .

At α_0 , the inflow comes from matched and unemployed workers who successfully find a new job; outflow is $\phi(\alpha_0)$, in the steady state, $\mu_{T_1}(\alpha_0) = \mu_{T_2}(\alpha_0)$ for any $T_1, T_2 \geq 0$, and thus I have

$$\int_{\alpha_*}^{\alpha_0} p(\theta(\alpha))\phi(\alpha)d\alpha + vp(\theta(u)) = \phi(\alpha_0), \quad (\text{B.4})$$

where v is the measure of u -workers. Both the left-hand side and right-hand side of (B.4) are finite, and therefore there is no mass point at α_0 .

For unemployed workers, the inflow comes from the separation of an existing match, while the outflow is the measure of unemployed workers who find a job. Letting inflow equal outflow, I have

$$vp(\theta(u)) = \beta\delta + \int_{\alpha_*}^{\alpha_0} \delta\phi(\alpha)d\alpha + \phi_u, \quad (\text{B.5})$$

where ϕ_u is the density of workers who have just been fired, β is the measure of 1-workers.

Combining (B.4) and (B.5) yields

$$\phi(\alpha_0) = \int_{\alpha_*}^{\alpha_0} [p(\theta(\alpha)) + \delta]\phi(\alpha)d\alpha + \phi_u + \beta\delta. \quad (\text{B.6})$$

For a successful match, the inflow comes from a new good signal sent by an existing uncertain match, and the outflow comes from the exogenous separation. Inflow equals outflow implies that

$$\beta = \frac{1}{\delta} \int_{\alpha_*}^{\alpha_0} \lambda\alpha\phi(\alpha)d\alpha. \quad (\text{B.7})$$

Given the equilibrium $\theta(\alpha)$, α_* , and matching function $p(\cdot)$, one can obtain a general solution of the ODE (B.3), which is given by

$$\phi_{\bar{A}}(\alpha) = \frac{1}{\bar{A}} \exp\left[\int_{\alpha_*}^{\alpha} \frac{\lambda s + \delta + p(\theta(s))}{\lambda s(1-s)} ds\right], \quad (\text{B.8})$$

where $1/\tilde{A}$ is a constant positive number to ensure $\phi > 0$. To fix A , one needs to use a boundary condition implied by the fact that ϕ is a density function and ν, β are the probability. The condition is given by

$$\int_{\alpha_*}^{\alpha_0} \phi(\alpha) d\alpha = 1 - \nu - \beta. \quad (\text{B.9})$$

Plugging (B.5) and (B.7) into (B.9) yields

$$\int_{\alpha_*}^{\alpha_0} \phi(\alpha) d\alpha = 1 - \frac{\int_{\alpha_*}^{\alpha_0} (\lambda\alpha + \delta)\phi(\alpha) d\alpha + \phi(\alpha_*)}{p(\theta(u))} - \frac{1}{\delta} \int_{\alpha_*}^{\alpha_0} \lambda\alpha\phi(\alpha) d\alpha.$$

Let $\kappa(\alpha) = \exp[\int_{\alpha_*}^{\alpha} \frac{\lambda s + \delta + p(\theta(s))}{\lambda s(1-s)} ds]$, and $\phi_{\tilde{A}}(\alpha) = \kappa(\alpha)/\tilde{A}$. Then the $\tilde{A} = A$ satisfying the boundary condition (B.9) is given by

$$A = \int_{\alpha_*}^{\alpha_0} \kappa(\alpha) d\alpha + \frac{\int_{\alpha_*}^{\alpha_0} (\lambda\alpha + \delta)\kappa(\alpha) d\alpha + 1}{\delta p(\theta(u))} \int_{\alpha_*}^{\alpha_0} \lambda\alpha\kappa(\alpha) d\alpha.$$

Given the solution $\phi(\alpha)$, $\phi(\alpha_0)$, $\phi_u = \lim_{\alpha \rightarrow \alpha_*} \phi(\alpha) = 1/A$, and ν, β can be solved by (B.5) and (B.7). Since (29) and (31) uniquely pin down μ^* , the stationary distribution is unique. ■

Proof of Proposition 5

For any $w \in (\underline{w}, \bar{w})$, suppose the density is $g_w(w)$. The outflow is $g_w(w)[\int_{\alpha_*}^{\alpha_0} \phi(s)[p(\theta(s)) + \delta]ds + \phi_u]$, where $\int_{\alpha_*}^{\alpha_0} \phi(s)p(\theta(s))ds$ is the mass of leaving for a new job, $\int_{\alpha_*}^{\alpha_0} \phi(s)\delta ds$ is the mass of unemployed workers due to a job destruction shock, and ϕ_u is the mass of leaving due to firing. The inflow is the measure of new jobs formed with wage w , which is $\phi((w^{new})^{-1}(w))p(\theta(w^{new})^{-1}(w))$. A worker with state $\omega = (w^{new})^{-1}(w)$ is looking for a new job in the submarket where the promised wage is w and market tightness $\theta = \theta(w^{new})^{-1}(w)$. In the steady state, the inflow equals the outflow, $g_w(w)[\int_{\alpha_*}^{\alpha_0} \phi(s)[p(\theta(s)) + \delta]ds + \phi_u] = \phi((w^{new})^{-1}(w))p(\theta(w^{new})^{-1}(w))$. Since $\phi_u = 1/A$, $g_w(w) = \frac{\phi((w^{new})^{-1}(w))p(\theta(w^{new})^{-1}(w))}{\int_{\alpha_*}^{\alpha_0} \phi(s)[p(\theta(s)) + \delta]ds + 1/A}$. For $w = \underline{w}$, only unemployed workers are looking for a job with this wage. A similar logic can derive the expression of g_w^u . Since w^{new} is strictly monotone and continuous, (\underline{w}, \bar{w}) is connected. Since $S(u) = S(\alpha_*)$, $w^{new}(u) = \lim_{\alpha \rightarrow \alpha_*} w^{new}(\alpha)$. ■

Proof of Proposition 6

For a fixed $\Delta > 0$, an artificial discrete time Markov process is constructed. The equilibrium strategy implies that, from any initial prior $\mu_0 \in \Xi$, $\mu_T(\Omega/\Omega^*) = 0$ after time t^* , and no mass point for $\omega \in [\alpha_*, \alpha_0]$. It is easy to verify that the process satisfies Doeblin's condition (see Stokey, Lucas and Prescott, 1989, chapter 11) after time t^* . By Theorem 11.10 in Stokey, Lucas and Prescott (1989), the invariant distribution μ_Δ is unique, and $\mu_{k\Delta}$ converges to μ_Δ in the total variation

norm as $k \in \mathbb{N}$ goes to infinity. Clearly, the invariant distribution of the continuous time Markov process, μ^* , is also the invariant distribution of the artificial discrete time Markov process; thus, $\mu^* = \mu_\Delta$ and therefore $\mu_{k\Delta}$ strongly converges to μ^* as $k \in \mathbb{N}$ goes to infinity. ■

References

- [1] Acemoglu, D., and R. Shimer (1999): “Efficient Unemployment Insurance,” *Journal of Political Economy*, 107, 893-927.
- [2] Autor, D., L. Katz, and M. Kearney (2009): “Trends in U.S. Wage Inequality: Revising the Revisionists,” *Review of Economics and Statistics*, 90, 300-323.
- [3] Bergemann, D., and J. Valimaki (1994): “Learning and Strategic Pricing,” *Econometrica*, 64: 1125-1150.
- [4] Bound, J., and G. Johnson (1992): “Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations,” *American Economic Review*, 82, 371-392.
- [5] Burdett, K., and D. Mortenson (1998): “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review*, 39, 257-273.
- [6] Burdett, K., and T. Vishwanath (1988): “Declining Reservation Wage and Learning,” *Review of Economic Studies*, 55, 655-665.
- [7] Eeckhout, J., and X. Weng (2010): “Assortative Learning,” unpublished manuscript, University of Pennsylvania.
- [8] Farber, H. (1994): “The Analysis of Inter-firm Worker Mobility,” *Journal of Labor Economics*, 12, 554-593.
- [9] Farber, H. (1999): “Mobility and Stability: The Dynamics of Job Change in Labor Markets,” in *Handbook of Labor Economics*, Vol. 3B, ed. by O. Ashenfelter and D. Card. Amsterdam: Elsevier, 2439-2484.
- [10] Felli, L., and C. Harris (1996): “Learning, Wage Dynamics, and Firm-Specific Human Capital,” *Journal of Political Economy*, 104, 838-869.
- [11] Fu, C. (2011): “Training, Search and Wage Dynamics,” *Review of Economic Dynamics*, 14, 650-666.
- [12] Galenianos, M., and P. Kircher (2009): “Directed Search with Multiple Job Applications,” *Journal of Economic Theory*, 133, 445-471.
- [13] Gonzalez F., and S. Shi (2010): “An Equilibrium Theory of Learning, Search and Wages,” *Econometrica*, 78, 509-537.

- [14] Jovanovic, B. (1979): “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87, 972-989.
- [15] Kambourov, G., and I. Manovskii (2009): “Occupational Mobility and Wage Inequality,” *Review of Economic Studies*, 76, 731-759.
- [16] Menzio, G., and S. Shi (2009): “Block Recursive Equilibria for Stochastic Models of Search on the Job,” *Journal of Economic Theory*, 145, 1453-1494.
- [17] Menzio, G., and S. Shi (2011): “Efficient Search on the Job and the Business Cycle,” *Journal of Political Economy*, 119, 468-510.
- [18] Miller, R. (1984): “Job Matching and Occupational Choice,” *Journal of Political Economy*, 92, 1086-1120.
- [19] Moen, E. (1997): “Competitive Search Equilibrium,” *Journal of Political Economy*, 105, 694-723.
- [20] Moen, E., and Rosen, S. (2004): “Does Poaching Distort Training?” *Review of Economic Studies*, 71, 1143-1163.
- [21] Mortensen, D. (2005): “Wage Dispersion: Why Are Similar Workers Paid Differently?” Cambridge, Massachusetts: MIT Press.
- [22] Moscarini G. (2005): “Job Matching and the Wage Distribution,” *Econometrica*, 73, 481-516.
- [23] Nagypal E. (2007): “Learning by Doing vs. Learning About Match Quality: Can We Tell Them Apart?” *Review of Economic Studies*, 74, 537-566.
- [24] Oksendal, B., and A. Sulem (2005): *Applied Stochastic Control of Jump Diffusions*, Springer.
- [25] Parsons, D.O. (1972): “Specific Human Capital: An Application to Quit Rates and Layoff Rates,” *Journal of Political Economy*, 80, 1120-1143.
- [26] Parsons, D.O. (1991): “The Job Search Behavior of Employed Youth,” *Review of Economics and Statistics*, 73, 597-604.
- [27] Pissarides, C., and J. Wadsworth (1994): “On-the-job Search: Some Empirical Evidence from Britain,” *European Economic Review*, 38, 385-401
- [28] Postel-Vinay, F., and J. Robin (2002): “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, 70, 2295-2350.

- [29] Shi, S. (2001): “Frictional Assignment I: Efficiency,” *Journal of Economy Theory*, 98, 232-260.
- [30] Shi, S. (2009): “Directed Search for Equilibrium Wage-Tenure Contracts,” *Econometrica*, 77, 561–584.
- [31] Stokey N., R. Lucas, and E. Prescott (1989): *Recursive Methods in Economic Dynamics*, Cambridge: Harvard University Press.