University of Toronto
Department of Economics

Working Paper 597

Learning, On-the-Job Search and Wage-Tenure Contracts

By Kevin Fawcett and Shouyong Shi

March 09, 2018

Learning, On-the-Job Search and Wage-Tenure Contracts

Kevin Fawcett
University of Toronto
(kevin.fawcett@mail.utoronto.ca)

Shouyong Shi
Pennsylvania State University
(sus67@psu.edu)

This version: February 2018 *

Abstract

When workers have incomplete information about their ability, they can learn about this ability by searching for jobs, both while employed and unemployed. Search outcomes yield information for updating the belief about the ability which affects optimal search decisions in the future. Firms respond to updated beliefs by altering vacancy creation and optimal wage contracts. To study equilibrium interactions between learning and search, this paper integrates learning into a search equilibrium with on-the-job search and wage-tenure contracts. The model generates results that shed light on a number of empirical facts, such as wage cuts in job-to-job transition, wage growth over tenure, true duration dependence of unemployment, and frictional wage inequality. We calibrate the model to quantify the extent to which learning and on-the-job search explain these empirical facts.

Keywords: Learning, On-the-job search; Contracts; Inequality.
JEL classifications: E21; E24; J60.

*Fawcett: Department of Economics, University of Toronto, Toronto, Ontario, Canada, M5S 3G7. Shi: Department of Economics, Pennsylvania State University, University Park, Pennsylvania, USA, 16802.

1. Introduction

Workers often have incomplete information about their ability and can learn about this ability by searching for jobs. Success in search improves the belief about the worker's ability and failure reduces the belief, both of which affect search decisions in the future. This process of learning through search can continue throughout a worker's career, both in unemployment and while employed. The objective of this paper is to analyze how learning affects the equilibrium in the presence of on-the-job search and wage-tenure contracts. We calibrate the model to quantify the significance of this mechanism for worker transition, wage dynamics, wage inequality, and the duration of unemployment.

The analysis is motivated theoretically and empirically. Previous theoretical work has emphasized how learning affects search by assuming that only unemployed workers can search (e.g., Gonzalez and Shi, 2010, and Doppelt, 2016). When workers can also search while employed, as in reality, continued learning can have far-reaching effects on the equilibrium. As successive matches produce information, they generate particular patterns of on-the-job search and job-to-job (EE) transition. Firms optimally offer wage-tenure contracts not only to recruit and retain workers, but also to optimally update the belief about the worker. Thus, continued learning affects a worker's wage dynamics both over tenure and between jobs. Moreover, on-the-job search affects not only employed workers but also unemployed workers. When workers can search only when they are unemployed, accepting a job destroys the option values of both search and learning. On-the-job search enables workers to retain these option values, which can change unemployed workers' search strategy and unemployment duration. These effects have not been formally analyzed, but they are essential for understanding worker transition, wage dynamics, and unemployment as parts of the same equilibrium.

Empirically, the integrated framework of learning and search on the job can shed light on a number of empirical facts about the labor market. Among employed workers, evidence suggests that the EE transition is large (Fallick and Fleischman, 2004). A significant fraction of the EE transition results in a wage cut initially (e.g., Postel-Vinay and Robin, 2002). In contrast, typical models of on-the-job search are unable to generate wage cuts in EE transitions without assuming negative shocks to match-specific productivity, and they predict fast, but short, EE transition.[1] Incorporating learning through on-the-job search may help a model account for both facts. Intuitively, workers may transition between jobs for more times in an attempt to learn about their ability and may accept wage cuts due to updated beliefs. In addition, Hornstein et al. (2011) find that, quantitatively, standard search models explain only a small fraction of wage dispersion among similar workers in the U.S. data. Both learning and on-the-job search can widen such frictional wage dispersion, especially when firms offer wage-tenure contracts.

For unemployed workers, the duration of unemployment seems to exhibit negative dependence for short duration and positive dependence for long duration. Job searchers reduce their search targets as unemployment duration increases (Kudlyak et al., 2013), and searchers who end up with long durations have had higher search effort throughout the search process (Faberman and Kudlyak, 2015). These patterns suggest that job seekers learn about their job-finding ability and adjust search targets as a response to the information they gain over time.

To address the theoretical and empirical issues raised above, we construct a model with an infinite horizon where workers are risk averse and firms are risk neutral. A worker's type is either high or low. This permanent characteristic is drawn from a common, time-invariant distribution when the worker first enters the economy, and is unknown to all individuals and firms including the workers themselves. Following Acemoglu and Autor (2011) and Lazear (2009), we interpret a worker's type as a bundle of skills that determine how likely it is the worker can perform the tasks required by a job. Relative to a low-type worker, a high-type worker has a bundle of skills that give the worker a higher probability to be able to perform a job and, hence, turn a meeting with a firm into a productive match. All workers can search independently of their employment status. Search is directed; that is, a worker observes all offers and optimally chooses an offer to apply to. Firms create

[1] Examples are Burdett and Mortensen (1998), Burdett and Coles (2003), Shi (2009) and Postel-Vinay and Robin (2002). Section 5 will provide numerical results of such a model with complete information.

vacancies to enter the market competitively. Search outcomes are publicly observed and information remains symmetric. A search outcome yields information for updating the (common) belief about the worker, but it does not fully reveal the worker's type because a meeting turns into a match probabilistically. In each period, a worker faces an exogenous probability of exiting the economy, and so learning never fully uncovers a worker's type.

An offer is represented by the worker value, which is the expected lifetime utility of the worker conditional on optimal choices in the future. The worker value is implemented by firms through wage-tenure contracts that are formulated recursively. In each period, the firm employing the worker inherits the most recent belief about the worker and the worker value. The firm chooses the wage in the current period and worker value for next period to maximize the value of the match while delivering the previously promised worker value. Firms optimally backload wages not only to increase retention of the worker, as in Shi (2009) and Burdett and Coles (2003), but also to explore learning.

Search and wage-tenure contracts generate endogenous heterogeneity in two dimensions – the belief about the worker and the worker value. Workers who have the same belief can differ in the worker value because different histories of search outcomes and tenure can lead to different worker values. Similarly, workers can reach the same value through different paths of search outcomes that yield different beliefs. Despite the endogenous distribution of workers over the belief and the worker value, the equilibrium is block recursive as defined by Shi (2009) and Menzio and Shi (2010). That is, workers' decisions, firms' decisions, and the matching probabilities are all functions of the above two dimensions of individual states and not of the distribution of workers. Block recursivity enables us to analyze and compute the dynamic equilibrium tractably.

A key feature of the model is that, given the worker value, the firm value of a match can increase in the belief. Although a higher belief does not increase output in a match, it increases the outside value that the worker receives and, thereby, relaxes the firm's promise-keeping constraint and enables the firm to deliver the promised value to the worker more easily. Specifically, if the worker exogenously separates into unemployment, a higher belief increases the value of unemployment by increasing the worker's job-finding probability and, if the worker does not separate exogenously, a higher belief increases the probability that the worker finds a new match through on-the-job search. On the negative side, a higher belief reduces the firm value by increasing the probability that the employee will find another job and leave the firm. For most equilibrium worker values, the positive effect of a higher belief on the firm value through the relaxation of the promise-keeping constraint dominates the negative effect through the endogenous separation probability.

The dependence of the firm value on the belief has novel implications for job creation and wage-tenure contracts. First, given the offer, firms create more vacancies for high belief workers than for low belief workers. Thus, the matching rate for a worker increases in the belief even after controlling for the worker's true type and the search target. Second, a recruiting firm can offer a higher value than a worker's incumbent employer for some beliefs and worker values. If a worker stays with a firm after failing to find a new match, the belief about the worker deteriorates as remaining with the firm is a result of a failed search attempt. Conversely, if the worker finds a new match, the belief about the worker increases and can give the recruiting firm a higher value than the incumbent firm in the new match for the same worker value. For intermediate to high beliefs, this increase in a recruiting firm's value can be large enough to outweigh the cost of posting the vacancy. Third, the extent of wage backloading depends not only on the promised value to the worker, but also on the belief about the worker. For any given worker value, a higher belief increases the extent of wage backloading and can reduce the initial wage in a new match. A firm backloads wages not only to increase retention of the worker, but also for the contingency that the belief about the worker will deteriorate after the worker fails to find a new match. In the latter case, the firm can taper wage growth for the worker in the future. The higher the belief about the worker, the more room there is for the belief to deteriorate in the future, and so the larger is the extent of wage backloading.

To quantify these effects we calibrate the model and simulate the equilibrium. In section 5, we discuss the quantitative results in detail and relate them to the empirical facts described earlier. We conduct counterfactual experiments by shutting down the learning and on-the-job search mechanisms, in turn, to analyze the quantitative roles of each element and the effect of their interaction in equilibrium. For employed workers, the results can be summarized as follows. First, for job stayers, wages grow over tenure more slowly and for a longer time for a high-belief worker than for a low-belief worker. In particular, wage growth is significant for almost ten years for a high-belief worker, because such a worker is motivated to find a new match frequently in an attempt to prevent the belief from falling. Second, high-type workers' EE transition rates are about 80% higher than low-type workers', and this difference remains for most of tenure lengths. Third, approximately 18.2% of EE transitions result in wage cuts. Such wage cuts are more frequent for low-type workers, but are typically larger for high-type workers. The frequency of such wage cuts in EE transitions increases in a worker's tenure. Fourth, frictional wage dispersion is significant. Even after controlling for the true type of the worker (which is unobservable), the mean-min ratio in wages is 2.36. Wage dispersion is smaller when it is measured by the 90-10 wage differential, but it is still larger than in related models.

The interaction between learning and on-the-job search is important for these quantitative results. Eliminating on-the-job search eliminates all the EE transitions. On the other hand, if information is complete, the only cause of wage growth for a stayer is the firm's desire to backload wages to retain the worker. This consideration generates only short-lived wage growth for a job stayer – it is almost completed in less than one year. In addition, eliminating incomplete information reverses the relative EE rate between the two types of workers. In that case, high-type workers' EE rates are about 40% lower than low-type workers' for tenure less than one year and, for tenure greater than or equal to one year, almost all EE transitions are made by low-type workers. Moreover, eliminating on-the-job search or learning reduces the mean-min wage ratio substantially.

For unemployed workers, they reduce their search targets as their unemployment duration lengthens, which is consistent with the finding in Kudlyak et al. (2013). Duration dependence is positive. However, after controlling for the true type of the worker, duration dependence is positive only for workers whose beliefs are close to the low type. For workers entering unemployment with an intermediate or high belief, true duration dependence is negative for short duration and positive for long duration. With a straightforward extension of the model to incorporate search effort, these results on duration dependence imply that search effort is negatively related to the worker belief, which is consistent with the finding in Faberman and Kudlyak (2015). Specifically, workers who have longer unemployment duration are those who enter unemployment with a relatively low belief. These workers have higher search effort, at least in the first seven months of unemployment.

This paper is related to the large literature on directed search originated in Peters (1991) and Montgomery (1991).[2] More specifically, this paper integrates two strands of the search literature. The first is equilibrium models of learning through search that exclude on-the-job search (e.g., Burdett and Vishwanath, 1988, Gonzalez and Shi, 2010, and Doppelt, 2016).[3] The second strand considers models with on-the-job search that assume complete information (e.g., Burdett and Coles, 2003, and Shi, 2009).[4] The integration of learning and on-the-job search produces new features not only of job-to-job transitions, wage dynamics and frictional wage dispersion, but also of duration dependence of unemployment. Additionally, these features shed new light on the empirical facts described earlier. Mathematically, both the belief and the worker value are state variables of workers' and firms' decisions and, thereby, the model generates rich interactions between the two in the equilibrium. Such interactions do not arise in either strand of the literature, because only one of the two variables is required to describe the state of a worker or a firm.

This paper is also related to empirical literatures that examine the facts discussed here. On worker transitions, Postel-Vinay and Robin (2002) document wage changes with job-to-job transitions. Topel and Ward (1992) document that job changes are frequent and that wage growth is driven by a combination of external and internal sources. On duration dependence of unemployment, the empirical literature has emphasized the difficulty in

[2]Other well known papers are Moen (1997), Julien et al. (2000), Burdett et al. (2001), Shi (2001), and Menzio and Shi (2010, 2011).

[3]Gonzalez and Shi (2010) assume that the belief about an employed worker stays constant until the worker separates exogenously into unemployment. Doppelt (2016) allows the belief about a worker to evolve during employment, however, because on-the-job search is not allowed, such evolution only affects the worker's transition in the future after the worker becomes unemployed.

[4]Other examples include Burdett and Mortensen (1998), Delacroix and Shi (2006), Menzio and Shi (2010, 2011), Postel-Vinay and Robin (2002), and Lise and Robin (2017).

distinguishing true duration dependence of unemployment from unobserved worker heterogeneity and dynamic selection (e.g., See Heckman and Borjas, 1980). Kudlyak et al. (2013) and Faberman and Kudlyak (2015) provide evidence to show that search behavior changes endogenously with duration. We emphasize this endogenous response to duration in our analysis. On frictional wage inequality, Hornstein et al. (2011) show that most standard search models are only able to account for only a small fraction of wage inequality measured in the data. Tsuyuhara (2016) shows that introducing moral hazard into a directed search model with on-the-job search can increase frictional wage inequality significantly. Shi (2016) shows that introducing convex vacancy costs and firm investment can account for the majority of frictional wage inequality in the data.

It is useful to distinguish learning in this paper from that in the literature originated in Jovanovic (1979). In this literature, the object to be learned is a worker's productivity with a firm and the information is generated by the stochastic output produced in a match. In contrast, in our setting, the ability to be learned is general and the information comes from searching for new matches in the market. The two learning models have opposite implications on the informational content of tenure and capture different reasons for wage growth over tenure. In Jovanovic (1979), a worker chooses to stay with a firm only if the worker has received sufficiently many high realizations of output. As a result, longer tenure in a firm indicates higher productivity, which is rewarded with wage growth. In our paper, longer tenure indicates lower ability as a worker remains with a firm only if the worker has failed to find a new match. Nevertheless, wages grow over tenure because backloaded wages must eventually be delivered as deteriorating beliefs over tenure reduce the room for backloading wages further. To emphasize such "learning from the market", we abstract from match-specific productivity. Moreover, we integrate learning into a search equilibrium and incorporate on-the-job search. Equilibrium responses of firms to learning in vacancy creation and wage-tenure contracts are critical for generating features of job-to-job transitions and duration dependence in unemployment observed in the economy.

The remainder of this paper is organized as follows. Section 2 describes the model. Section 3 analyzes optimal decisions and defines the equilibrium. After calibrating the model, section 4 illustrates the value and policy functions. Section 5 presents the quantitative results on worker transition, wage dynamics, unemployment duration and frictional wage dispersion. The appendices provide additional analysis and proofs.

2. Model Environment

We consider an economy with discrete time and an infinite horizon. Workers and firms discount the future with the common factor $(1+r)^{-1}$, where $r > 0$. Firms are risk neutral and their measure is determined by competitive entry into the labor market. There is a unit measure of risk-averse workers each of whom can supply one unit of labor inelastically in each period. The utility function in a period is $u(w)$, where w is earnings.[5] The function u satisfies: $0 < u'(w) < \infty$ and $-\infty < u''(w) < 0$ for all $w > 0$. For an employed worker, earnings represent the worker's wage income. Every employed worker produces the same amount y in a period. For an unemployed worker, "earnings" are equal to home production $b \in (0, y)$.[6] In each period with probability $\delta \in (0,1)$, a match is exogenously destroyed and the worker becomes unemployed. In addition, an exit shock occurs with probability $1-\sigma$ that takes a worker out of the economy, where $\sigma \in (0,1)$.

Worker type and incomplete information: Workers are heterogeneous in a permanent type $i \in \{H, L\}$ that affects productivity stochastically, as modeled by Gonzalez and Shi (2010). Upon meeting a firm, a type i worker's productivity with the firm is realized to be $y > b$ with probability a_i and $y' < 0$ with probability $1 - a_i$, where $0 < a_L < a_H \leq 1$. Since $y' < 0$, a firm hires a worker only if the realization is y. Thus, a high-type worker is more likely to be productive than a low-type worker but, conditional on being productive with a given firm, output is the same for all employed workers. We refer to a_i as the productivity of a type i worker. Although this modeling of productivity is stylized, it is meant to capture skills more broadly in the following way. Each worker possesses a vector of skills and jobs consists of a vector of tasks. Upon meeting a worker, a firm determines

[5]For a model with savings and on-the-job search, see Chaumont and Shi (2017).

[6]As is customary in the literature, home production is the sum of the unemployment benefit, the imputed value of additional leisure, home production activity, and any savings on work-related costs.

if the skills possessed by the worker are capable of performing the tasks required by the job. Conditional on having a skill vector that is sufficient to perform all tasks of the job, all workers are equally productive in the match. High ability workers possess a bundle of skills that are capable of performing a greater number of tasks and, as a result, are more likely to be productive when meeting a random firm.[7]

A worker's type is unknown to all market participants, including the worker him-self/herself. When a worker first enters the economy, their type is drawn from a distribution such that $a = a_H$ with probability $p_0 \in (0, 1)$ and $a = a_L$ with probability $1-p_0$. The initial mean belief is $\mu_0 = p_0 a_H + (1 - p_0)\, a_L$. Search outcomes are stochastic and therefore two workers of the same type can have different labor market histories. As a result, the model endogenously generates heterogeneous beliefs about worker ability. In a generic period, let the belief be $a = a_H$ with probability P_H and $a = a_L$ with probability $P_L = 1 - P_H$. The mean belief, $\mu \equiv P_H a_H + P_L a_L$, is part of a worker's resume defined later.

Information remains symmetric over time, as the belief about a worker is common to all participants in the economy. In particular, a worker's search outcome in the labor market is publicly observed. The worker and other participants in the economy update the belief about the worker's ability in the same way. If a worker stays in the market for a sufficiently long time, learning will uncover the true type eventually. The exit shock with probability $1 - \sigma$ is introduced to prevent this uninteresting outcome.

Submarkets and search: Employed and unemployed workers are both able to search. The labor market is organized as a continuum of submarkets indexed by the belief about the applicants, μ, and the expected lifetime utility offered to a worker, z. The cost of creating a vacancy for a period is $k \in (0, y)$. Firms enter the market competitively and make zero expected profit for posting a vacancy. Such competition induces the matching rate as a function of the characteristics of the submarket, (μ, z) (see subsection 3.4). Search is directed. Specifically, a worker chooses a submarket to search knowing that submarkets differ in the offer and the matching probability. As the notation suggests, submarket (μ, z)

[7] This formulation of ability is consistent with the broad interpretation of human capital in the literature, e.g., Lazear (2009) and Acemoglu and Autor (2011). See the Introduction for the discussion.

is restricted for applicants whose resumes support the mean belief μ. This restriction is reasonable since a worker's resume is typically available to the public. A recruiting firm wants to condition the offer on μ as it affects the firm's future value through the employee's search decision in the future and the probability of the worker leaving the firm. The search choice separates the applicants into different submarkets.

Matching inside a submarket is random. A worker may experiment by searching for a job only to gather information about the ability but not to accept the match. If some workers experiment, a firm needs to make an inference on the fraction of applicants who experiment. This inference complicates the analysis significantly by introducing the distribution of workers as a state variable in the decision problems of workers and firms. To eliminate this complexity, we impose restriction (A.2) in Appendix A to make experimentation not optimal for workers and verify that this condition is satisfied in the calibrated equilibrium. This condition and directed search imply that a worker will always accept a match.[8] If a worker fails to match in a period, the worker remains in the current position, whether it be with a firm or unemployment. To preserve symmetric information, we assume that only the search outcome is observed. If a worker fails to match, no information on whether the worker is screened by a firm is revealed or kept.

Contracts and resumes: An offer is delivered by a wage contract conditional on the worker staying with the firm. The wage can vary over tenure because the belief about the worker is updated and the firm has an incentive to backload wages for risk averse workers. Wage contracts are formulated recursively. At the end of a generic period, the expected lifetime utility of a worker is denoted by v and referred to as the *worker value*. Upon matching with a new firm, the worker value is equal to the offer z. To deliver value v to the worker at the end of a period, a firm is assumed to commit to a pair (w_{+1}, v_{+1}), where w_{+1} is the wage to be paid at the beginning of the next period and v_{+1} the continuation value at the end of the next period. See subsection 3.3 for further analyses.

[8] The incumbent firm's commitment to not matching a worker's outside offer is a common assumption in the literature (see Burdett and Mortensen, 1998, Burdett and Coles, 2003, and Shi, 2009). With symmetric firms, the assumption is necessary for generating job-to-job transitions. For a model that allows for heterogeneous firms and matching outside offers, see Postel-Vinay and Robin (2002).

A worker's labor market history can be encoded into the pair (μ, v), which is referred to as the worker's *resume*. In contrast to Gonzalez and Shi (2010) and Doppelt (2016), who assume that only unemployed workers can search, the belief μ alone is insufficient for summarizing an employed worker's resume. With on-the-job search, two employed workers can reach the same belief with different worker values or the same worker value with different beliefs. Since these two workers face different optimal search problems, both μ and v are needed to describe a worker's state. Similarly, a recruiting firm's payoff depends on both the offer z and the belief about the applicant, μ, because workers with different beliefs have different rates of separating from the firm in the future. However, there is no need to add an applicant's value at the current job to the description of submarkets (μ, z). Given μ, applicants who differ in the worker value will optimally choose to search for different offers and, hence, will endogenously separate into different submarkets.

Matching: Employed and unemployed workers can have different parameters of search efficiency, denoted by λ_j for $j \in \{e, u\}$ where $0 \leq \lambda_e \leq \lambda_u$.[9] For a type i worker with employment status j, the worker's efficiency search units are $\lambda_j a_i$. In any arbitrary submarket, let the measure of type i workers be ℓ_i and the total efficiency search units of workers be $\ell_e = \Sigma_{i,j} \lambda_k a_i \ell_i$. Let θ be the ratio of efficiency search units to the measure of vacancies in the submarket, which is referred to as the tightness. The measure of vacancies in the submarket is $\theta \ell_e$ and the measure of (productive) matches in the submarket is given by the function $M(\ell_e, \theta \ell_e)$, where M has constant returns to scale and satisfies the standard assumptions. The matching rate per efficiency search unit of workers is

$$x \equiv \frac{M(\ell_e, \theta \ell_e)}{\ell_e} = M(1, \theta). \tag{2.1}$$

Thus, a type i worker with employment status j in the submarket successfully matches with probability $\lambda_j a_i x$. The corresponding probability for filling a vacancy is $q \equiv \frac{M}{\theta \ell_e} = \frac{x}{\theta}$.

It is convenient to use x instead of θ as the organizing variable. For this purpose, we invert (2.1) to write $\theta = \theta(x)$. Then, the vacancy-filling probability is $q(x) \equiv \frac{x}{\theta(x)}$. Since

[9] The parameter λ_e is convenient because we can nest the model without on-the-job search by setting $\lambda_e = 0$. In the baseline calibration, $\lambda_e = \lambda_u = 1$.

the matching function M is relevant for the analysis only through the implied matching probabilities x and $q(x)$, we impose assumptions directly on the latter. First, we require $0 \leq \lambda_j x \leq 1$ and $x \leq 1$ for $j \in \{e, u\}$. Since $\lambda_e \leq \lambda_u$, the requirement is equivalent to

$$0 \leq x \leq \min\{\frac{1}{\lambda_u}, 1\}. \tag{2.2}$$

Second, the vacancy-filling probability should be bounded in $[0, 1]$. Incorporating the possibility that $\lim_{x\to 0} q(x) < 1$, we express the bounds on q as

$$0 \leq q(x) \leq \bar{q} \equiv \min\{1, \lim_{x\to 0} q(x)\}. \tag{2.3}$$

Third, the vacancy-filling probability depends on x as follows:

$$q'(x) < 0 \text{ and } q''(x) > 0 \text{ for all interior } x \text{ and } q(x).$$

This dependence is intuitive. When the matching probability per efficiency unit of workers' search increases, the matching probability for a vacancy must fall, which leads to $q'(x) < 0$. Moreover, the marginal reduction in the matching probability for a vacancy diminishes as x continues to increase, which leads to $-q''(x) < 0$.

In the equilibrium, competitive entry of vacancies into submarkets determines $x = x(\mu, z)$, which is characterized in subsection 3.4. Firms and workers take the function $x(\mu, z)$ as given. When μ is understood in the context, we sometimes refer to x, instead of (μ, z), as a worker's choice of the submarket.

Belief updating: Consider a type i worker in the employment status j, where $i \in \{H, L\}$ and $j \in \{e, u\}$. Prior to search taking place in the period, the belief is $a = a_H$ with probability P_H and $a = a_L$ with probability P_L, with the mean $\mu = P_H a_H + P_L a_L$. Suppose that the worker searches in a submarket where the matching rate per efficiency search unit is x. Let $s = 1$ indicate a match success and $s = 0$ a match failure. The probability that $s = 1$ is $\lambda_j a_i x$. If $s = 1$, the posterior probability assigned to being a type i worker is:

$$\Pr(a_i|s = 1, x, \lambda_j) = \frac{P_i a_i \lambda_j x}{P_H a_H \lambda_j x + P_L a_L \lambda_j x} = \frac{P_i a_i}{\mu}.$$

Using the definition of μ to write $P_H = \frac{\mu - a_L}{a_H - a_L}$, we can compute the posterior (mean) belief following a match success as

$$\begin{aligned} &\mathbb{E}(a|s=1,x,\mu,\lambda_j) = \Pr(a_H|s=1,x,\lambda_j)a_H + \Pr(a_L|s=1,x,\lambda_j)a_L \\ &= a_H + a_L - \tfrac{a_H a_L}{\mu} \equiv \phi(\mu). \end{aligned} \tag{2.4}$$

If the worker fails to match ($s = 0$), the posterior probability is:

$$\Pr(a_i|s=0,x,\lambda_j) = \frac{P_i(1-a_i\lambda_j x)}{P_H(1-a_H\lambda_j x) + P_L(1-a_L\lambda_j x)} = \frac{P_i(1-a_i\lambda_j x)}{1-\mu\lambda_j x}.$$

In this case, the posterior (mean) belief following a match failure is:

$$\begin{aligned} &\mathbb{E}(a|s=0,x,\mu,\lambda_j) = \Pr(a_H|s=0,x,\lambda_j)a_H + \Pr(a_L|s=0,x,\lambda_j)a_L \\ &= a_H - \tfrac{(a_H-\mu)(1-a_L\lambda_j x)}{1-\mu\lambda_j x} \equiv F(\lambda_j x,\mu). \end{aligned} \tag{2.5}$$

The mean belief is Markovian. The posterior mean belief given each search outcome is only a function of the prior mean belief, the search choice x, and the search efficiency parameter λ. As shown later, the mean belief is a sufficient statistic for the belief that is relevant for decisions, and so we refer to the mean belief simply as the *belief*. Intuitively, the belief improves after a match success and deteriorates after a match failure. That is, $\phi(\mu) > \mu > F(\lambda_j x, \mu)$ for all interior (x, μ). If the prior belief has reached one of the corners, $\mu = a_i$, then the search outcome does not influence the belief anymore, in which case $\phi(a_i) = F(\lambda_j x, a_i) = a_i$. Moreover, conditional on searching in the same submarket and receiving the same search outcome, the higher is the prior belief μ, the higher is the posterior belief. That is, given x, both $\phi(\mu)$ and $F(\lambda_j x, \mu)$ are increasing in μ.

After a match success, the posterior belief $\phi(\mu)$ is independent of (x, λ_j). That is, it is independent of where the worker has just searched and how high the worker's search efficiency parameter is. To understand this result, note that the probability of a match success is multiplicative in (a_i, x, λ_j). Conditional on searching in the same submarket, the likelihood ratio of a match success between the two types is $\frac{a_L}{a_H}$, which is independent of x and λ_j. If a worker succeeds in a match, the submarket where the worker searches and the worker's employment status do not provide additional information for distinguishing between the two types.

In contrast, after a match failure, the submarket where a worker searched contains information about the worker's ability. The posterior belief $F(\lambda_j x, \mu)$ strictly decreases in $\lambda_j x$ for all interior prior beliefs μ. This is because the likelihood ratio of a match failure between the two types, $\frac{1-a_L\lambda_j x}{1-a_H\lambda_j x}$, is strictly increasing in $\lambda_j x$. If a worker searches in a submarket where a match is more easily obtained (i.e., a higher x) or if a worker's employment status gives a higher search efficiency parameter λ, the worker is less likely to fail to match. However, the probability of a match failure falls by less for the low-type worker than for the high-type worker. If the worker fails to find a match in this case, it is rational to attribute the failure more to low ability than to bad luck. Thus, after a match failure, the belief deteriorates more sharply if $\lambda_j x$ is higher. In particular, if a worker searches in a submarket where x is close to 0, failure to match reveals little about the worker's ability as it is very unlikely for either type of worker to find a match in that submarket. Conversely, if the worker searches in a submarket where the matching probability is close to 1 for a high-type worker, failure to match indicates almost certainly that the worker has low ability. Note that since there is no cost to search, it is always feasible for a worker to keep the belief constant by searching in the submarket with $x = 0$.

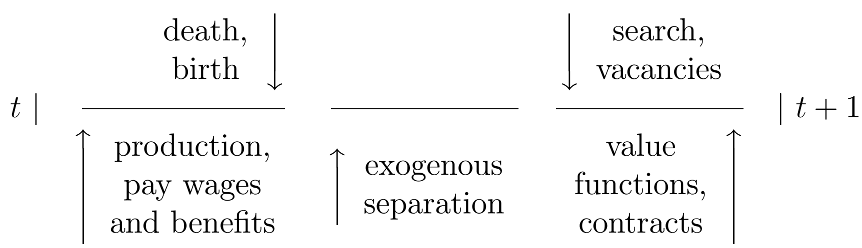


Figure 1. Timing of events in a period

Timing: The timing of events in a period is depicted in Figure 1. At the beginning of each period, production takes place, employed workers are paid wages and unemployed workers receive the unemployment benefit. Second, a proportion $(1 - \sigma)$ of all workers exit the economy and are replaced by new workers who enter the economy through unemployment and whose abilities are drawn from the prior distribution $(p_0, 1 - p_0)$. Third, employed workers who have survived the exit shock separate into unemployment exogenously with probability $\delta \in (0, 1)$. Fourth, vacancies are created, contracts are offered, employed workers who have survived the separation shock and workers who started the period in unemployment search.[10] Employed workers' search generates endogenous separation and job-to-job transitions. Finally, firms with an employee choose a wage to be paid to the worker at the start of the next period, w_{+1}, and a continuation value for the worker at the end of the next period conditional on the worker remaining in the match, v_{+1}. All value functions are measured at the end of the period.

3. Optimal Decisions and the Equilibrium

3.1. Employed Worker's Decision

Consider a worker who is employed at the end of a period in the state (μ, v), where μ is the belief about the worker's ability and v is the worker value promised by the firm. To deliver the promised value v, the firm commits to next period's wage, w_{+1}, and next period's continuation value, v_{+1}. Given the continuation value v_{+1}, the worker chooses an offer z to search for in the next period. For this decision, the worker's value of staying at the job is v_{+1}. Thus, it is convenient to refer to the worker by (μ, v_{+1}) instead of (μ, v). Let $\mathbb{Z}_\mu$ denote the feasible set of offers available for a type μ worker to search for.

In submarket (μ, z), the matching rate per efficiency search unit is $x(\mu, z)$, which will be characterized in subsection 3.4. A worker searching in the submarket will be matched with probability $a_i \lambda_e x(\mu, z)$ if the worker's ability is a_i, and the worker's gain is $(z - v_{+1})$. If the worker fails to match, the continuation value at the current job is v_{+1}. The worker's expected gain of searching in submarket (μ, z) is:

$$\mathbb{E}\left[a_i \lambda_e x(\mu, z)(z - v_{+1})\right] = \mu \lambda_e x(\mu, z)(z - v_{+1}).$$

The equality follows from the result that v_{+1} depends on the belief only through the mean, which we will establish in subsection 3.3. The worker will choose the search target z to maximize the *return on search*:

$$R_e(\mu, v_{+1}) \equiv \max_{z \in \mathbb{Z}_\mu} \{\mu \lambda_e x(\mu, z)(z - v_{+1}) + v_{+1}\}. \tag{3.1}$$

[10] To simplify the computation, workers who have just separated into unemployment in the period cannot search in the same period.

In the above problem, an employed worker makes the tradeoff between the matching probability and the expected gain in lifetime utility in each submarket. For any given μ, a submarket with a higher offer z is associated with a lower matching rate x since the submarket provides a lower expected value to firms and attracts less vacancies as a result. The tradeoff depends on the belief, μ, and the next period's continuation value of staying in the match, v_{+1}. Denote the optimal choice of z in (3.1) by the policy function $g_e(\mu, v_{+1})$. The implied matching rate per efficiency search unit in the targeted submarket is $x_e(\mu, v_{+1}) \equiv x(\mu, g_e(\mu, v_{+1}))$. The worker separates endogenously from the match in the next period with probability $\mu\lambda_e x_e(\mu, v_{+1})$.

The belief about the worker is updated to $\phi(\mu)$ if the worker forms a new match and to $F(\lambda_e x_e, \mu)$ if the worker fails to match. These updated beliefs do not directly appear in the decision problem (3.1), because firms commit to the contracts. When the worker succeeds in a match with an outside firm, the updated belief will affect how the outside firm will deliver the offer z, but this effect does not change the worker's payoff given the firm's commitment to the offer. Similarly, when the worker fails to match with an outside firm, the updated belief will affect how the incumbent firm will deliver the continuation value v_{+1}, but it does not change the worker's payoff given the incumbent firm's commitment to the continuation value. These features imply that an employed worker does not experiment (see Appendix A), and so (3.1) is valid.

3.2. Unemployed Worker's Decision

Consider an unemployed worker with belief μ at the end of a period. Let $U(\mu)$ denote the value of this worker, which is derived from the utility of home production, $u(b)$, and the return on search in the next period denoted by $R_u(\mu)$. That is,

$$(1+r)U(\mu) = u(b) + \sigma R_u(\mu) \tag{3.2}$$

The multiplier $(1+r)$ appears on the left-hand side because U is measured at the end of a period. The parameter σ is the probability that a worker survives the exit shock. The return on search will come from the next period. If the worker searches in submarket (μ, z), the expected matching probability is $\mu\lambda_{\mu}x\,(\mu, z)$. If the worker fails to match, the belief will decrease to $\mu_{+1} = F\,(\lambda_u x, \mu)$, where F is given by (2.5) and $x = x\,(\mu, z)$. In this case, the worker's value at the end of the next period will be $U\,(\mu_{+1})$. Thus, the gain from a match will be $[z - U\,(\mu_{+1})]$. An unemployed worker's optimal search target z solves the following problem which generates the return on search as

$$R_u(\mu) \equiv \max_{z\in \ \ \mu} \{\mu\lambda_u x(\mu, z)\,[z - U(\mu_{+1})] + U(\mu_{+1})\} \tag{3.3}$$

subject to $\mu_{+1} = F\,(\lambda_u x(\mu, z), \mu)$. Denote the optimal target offer z in (3.3) as $g_u\,(\mu)$ and the implied matching rate per efficiency search unit as $x_u\,(\mu) = x\,(\mu, g_u\,(\mu))$. Note that if an unemployed worker finds a match, the updated belief $\phi\,(\mu)$ does not directly appear in (3.3), because the recruiting firm commits to the offer z.

In contrast to an employed worker, an unemployed worker directly considers the effect of search on belief updating after a match failure. If the worker chooses to search next period in a submarket where the matching probability is high, then failure to match will induce the belief to fall sharply. If the value of unemployment is an increasing function of the belief, as it should be intuitively, then the future value for the worker will be reduced sharply. That is, searching in an "easy" submarket quickly destroys the option value of learning if the worker fails to find a match. This consideration induces an unemployed worker to search for high offers initially and reduce the search target over an unemployment spell, as emphasized by Gonzalez and Shi (2010) and examined in subsection 4.3. The consideration may also induce an unemployed worker to experiment, i.e., to search for an offer for the information with the intention of rejecting the match. We impose (A.2) in Appendix A to rule out experimentation, which is satisfied in the calibrated equilibrium.

3.3. Optimal Contracts and Backloaded Wages

For a firm with an employee in the state (μ, v) at the end of a period, the firm value is $J\,(\mu, v)$. To deliver the promised value v to the worker, the firm chooses next period's wage, w_{+1}, and next period's continuation value for the worker, v_{+1}. As shown in (3.1), the continuation value will affect the worker's search decision and the future belief. The firm takes these effects into account. However, the firm is assumed to commit to (w_{+1}, v_{+1}) in each period and, in particular, v_{+1} cannot be contingent on the value searched by the worker. This assumption is made for two reasons.[11] First, if v_{+1} is allowed to depend on the offer searched by the worker, then a firm can effectively match a worker's outside offer z by setting $v_{+1} = z + \varepsilon$ if the worker succeeds in search, and $v_{+1} = U(\mu_{+1}) + \varepsilon$ if the worker fails in search, where $\varepsilon > 0$ is arbitrarily small. This offer matching will prevent the worker from moving to another firm. In the equilibrium, all firms will offer $U(\phi(\mu))$ when recruiting a worker with belief μ and will reduce the continuation value to $U(\mu_{+1})$ afterward. As a result, no job-to-job transitions will occur in the equilibrium. The commitment to a number v_{+1} eliminates this uninteresting outcome. Second, when v_{+1} is independent of the search outcome, an employed worker does not have incentive to experiment, as analyzed in Appendix A. This simplifies the analysis significantly.

To formulate the optimal choices (w_{+1}, v_{+1}), note that the firm's profit at the beginning of the next period will be $y - w_{+1}$. The firm rationally expects that the employee will search for the offer $z = g_e(\mu, v_{+1})$, as analyzed in subsection 3.1. Thus, the employee will separate from the firm endogenously with probability $a_i \lambda_e x_e(\mu, v_{+1})$, where a_i is the worker's ability and $x_e = x(\mu, g_e(\mu, v_{+1}))$. The expected probability of endogenous separation is $\mu \lambda_e x_e(\mu, v_{+1})$. In addition, the employee will separate into unemployment with probability δ and will exit the economy with probability $1 - \sigma$, both of which are exogenous. After such separation or exit, the firm value will become zero. The match will survive at the end of next period with expected probability $\sigma(1-\delta)[1 - \mu\lambda_e x_e(\mu, v_{+1})]$. In this event, the belief about the worker will decrease to $\mu_{+1} = F(\lambda_e x_e, \mu)$, the worker will obtain the continuation value v_{+1}, and the firm will receive the value $J(\mu_{+1}, v_{+1})$. Thus, the optimal contract choices (w_{+1}, v_{+1}) solve the following problem:

$$(1+r)J(\mu, v) = \max_{(w_{+1}, v_{+1})} \{y - w_{+1} + \sigma(1-\delta)[1 - \mu\lambda_e x_e(\mu, v_{+1})] J(\mu_{+1}, v_{+1})\} \tag{3.4}$$

[11] This assumption is meaningful when considering plays out of the equilibrium. If the worker happens to search in a suboptimal submarket, the updated belief will be off the equilibrium path, but the firm is not allowed to change v_{+1} to respond to this unexpected change in the belief.

subject to $\mu_{+1} = F(\lambda_e x_e(\mu, v_{+1}), \mu)$ and

$$(1+r)v = u(w_{+1}) + \sigma\left[(1-\delta)R_e(\mu, v_{+1}) + \delta U(\mu)\right]. \tag{3.5}$$

The constraint (3.5) is the promise-keeping constraint. The promised value v is delivered in the next period through three components: the utility of consumption, the return on search while employed, and the value of unemployment if the match is exogenously destroyed. Note that exogenous separation does not change the belief. We denote the firm's optimal contract that solves (3.4) by the policy functions $w_{+1} = w(\mu, v)$ and $v_{+1} = n(\mu, v)$. The feature that v_{+1} depends on the belief only through the mean was used earlier in subsection 3.1 to compute the return on search to an employed worker.

The optimal contract backloads wages for two reasons. One is the firm's incentive to increase retention of the worker, as analyzed by Burdett and Coles (2003) and Shi (2009). For any promised value, the firm can delay some wage payments to the future to increase the worker's opportunity cost of leaving the firm. The other reason for backloading wages is for the contingency of belief updating, which is absent in the literature. If the worker fails to find a new match and stays with the firm, the belief about the worker will deteriorate. In that case, the firm will want to taper wage growth for the worker in the future. The more backloaded are wages, the more can the firm save on wage payments in this contingency. Balancing these reasons for backloading wages, the incentive to smooth consumption over time for the riske-averse worker prevents wages from being completely backloaded.

A critical feature of the model is that both reasons for backloading wages are stronger if the belief about the worker is higher. A worker with a higher belief is able to find a new match with a higher probability, which increases the firm's incentive to backload wages for retention. In addition, a higher belief leaves a larger room for the belief to deteriorate when the worker fails to find a new match. The firm backloads more wages for this contingency. Mathematically, the belief affects the contracting problem above directly by itself and indirectly through the worker's return on search, R_e, and the value of unemployment, U. In particular, a higher belief increases R_e and U, which make it cheaper for a firm to deliver any promised value v to the worker.

To illustrate the positive dependence of wage backloading on the belief, we depict a likely case of the optimal contract in Figure 2. For any given (μ, v), denote the firm's objective function in (3.4) as $EJ(w_{+1}, v_{+1}; \mu, v)$, where (w_{+1}, v_{+1}) are the contract terms. Denote the right-hand side of the promise-keeping constraint (3.5) as $PK(w_{+1}, v_{+1}; \mu)$. Take two workers with the same promised value v but with different beliefs μ_1 and μ_2, where $\mu_1 < \mu_2$. For each worker, the optimal contract is the tangency point between the firm's indifference curve and the promise-keeping constraint, given by point A for worker μ_1 and point B for worker μ_2. All curves are downward sloping because the firm and the worker both face the tradeoff between the current wage and the continuation value. As depicted, the optimal contract for worker μ_2 has a lower current wage and a higher continuation value than for worker μ_1.

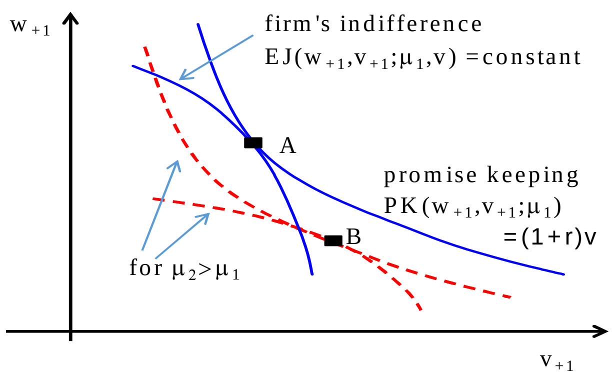


Figure 2. The dependence of the optimal contract on the belief

The continuation value that a firm can offer to an employee is bounded above because the firm value in the future cannot be negative (see Appendix B for a proof). Define $\bar{v}(\mu_{+1})$ as the worker value at which the firm value in the future will be zero with the belief μ_{+1}. That is, $\bar{v}(\mu_{+1})$ solves:

$$J(\mu_{+1}, \bar{v}(\mu_{+1})) = 0. \tag{3.6}$$

It is useful to express this bound as a function $v^*(\mu)$ that depends on the current belief μ instead of the future belief μ_{+1}. Since x_e depends on the continuation value v_{+1} according to $x_e(\mu, v_{+1})$, the bound $v^*(\mu)$ solves the following equation:

$$v^*(\mu) = \bar{v}(F(\lambda_e x_e(\mu, v^*(\mu)), \mu)). \tag{3.7}$$

If this equation has two or more solutions, we choose the largest among them as v^*. At $v_{+1} = v^*(\mu)$, the matching rate is $x_e^*(\mu) \equiv x_e(\mu, v^*(\mu))$ and a match failure will decrease the belief to $\mu_{+1}^*(\mu) \equiv F(\lambda_e x_e^*(\mu), \mu)$. We call $v^*(\mu)$ or equivalently, $\bar{v}\left(\mu_{+1}^*(\mu)\right)$, the *retention upper bound* on the continuation value of a worker whose current belief is μ.

3.4. Optimal Offer and Vacancy Creation

Firms enter the submarkets competitively. Consider a firm that creates a vacancy to enter submarket (μ, z). The vacancy costs k for the period and is filled with probability $q(x(\mu, z))$, where the function $x(\mu, z)$ is taken as given by the firm. If the vacancy is not filled, the value to the firm is zero. If the vacancy is filled, the belief about the newly recruited worker improves to $\phi(\mu)$ and the firm value is $J(\phi(\mu), z)$. Thus, the expected profit of creating a vacancy for submarket (μ, z) is $q(x) J(\phi(\mu), z) - k$. In equilibrium, this expected profit of a vacancy must be non-positive in all submarkets because of competitive entry of vacancies. That is, $q(x) J(\phi(\mu), z) \leq k$ for all (μ, z).[12]

Recall from (2.3) that $q(x) \leq \bar{q} \leq 1$. If $\bar{q} J(\phi(\mu), z) > k$, a positive measure of vacancies enter submarket (μ, z) to drive down the expected profit of a vacancy to zero. That is, $q(x) J(\phi(\mu), z) = k$ and $q(x) < \bar{q}$. In this case, $x > 0$. If $\bar{q} J(\phi(\mu), z) < k$, no vacancies are created in submarket (μ, z), in which case $q(x) = \bar{q}$ and $x = 0$. A borderline case has $\bar{q} J(\phi(z), z) = k$, which we put together with the first case without loss of generality. Thus, competitive entry of vacancies into submarkets yields:

$$q(x) J(\phi(\mu), z) \leq k \text{ and } \left(x \geq 0,\ q(x) \leq \bar{q}\right),\ \forall (\mu, z),$$

where the two sets of inequalities hold with complementary slackness. This condition defines the function $x(\mu, z)$ for every (μ, z). Since $q(x)$ is a strictly decreasing function for all $q(x) < \bar{q}$ and is equal to $\bar{q}$ otherwise, we can rewrite the above condition as

$$x(\mu, z) = \begin{cases} q^{-1}\left(\frac{k}{J(\phi(\mu), z)}\right), & \text{if } J(\phi(\mu), z) \geq k/\bar{q} \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

[12] If expected profit of a vacancy were strictly positive in a submarket, then a firm would create an infinite number of vacancies in the submarket, which would imply a tightness $\theta(x) \to \infty$. In this case, the vacancy-filling probability would be $q(x) = \frac{x}{\theta(x)} \to 0$, which would lead to the contradiction that expected profit of a vacancy is strictly positive in the submarket.

Note that a recruiting firm cares about the belief μ and the offer z, but not directly about the applicant's option value v_{+1} or employment status. The latter two factors affect a worker's optimal search choice. However, once a worker is matched with a firm in submarket (μ, z), the payoffs of the worker and firm depend only on (μ, z). Generally, a worker's state is completely summarized by (μ, v), where $v = z$ if the worker is newly hired.

The offer to an applicant is bounded above because the vacancy cost is positive. Define $\hat{v}(\mu)$ as the *recruiting upper bound* on the worker value by:

$$J(\phi(\mu), \hat{v}(\mu)) = k/\bar{q}. \tag{3.9}$$

Recall that the retention upper bound on a worker μ is $v^*(\mu) = \bar{v}\left(\mu^*_{+1}(\mu)\right)$, where $\bar{v}$ is defined by (3.6), v^* is defined by (3.7), and $\mu^*_{+1}(\mu)$ is the updated belief of a μ worker in equilibrium. In contrast to (3.6), the right-hand side of (3.9) is strictly positive and the first argument in J on the left-hand side is $\phi(\mu)$ instead of μ_{+1}. As a result, it is possible that $\hat{v}(\mu) > v^*(\mu)$ if the benefit to a firm of employing a worker with belief $\phi(\mu)$ outweighs the vacancy cost. In this case, a recruiting firm can offer a worker more than what the incumbent firm can afford. In Appendix B, we compare the functions $(\hat{v}, v^*, \bar{v})$ and explain why the unified upper bound, $\max\{v^*(\mu), \hat{v}(\mu)\}$, is needed for the entire domain of μ.

3.5. Equilibrium Definition

Let Ω_t denote the distribution of workers over (μ, v) at the end of period t. An equilibrium of the economy is defined below:

Definition 3.1. *An equilibrium consists of policy functions, $g_e(\mu, v_{+1})$, $(w, n)(\mu, v)$ and $g_u(\mu)$, return and value functions, $R_e(\mu, v_{+1})$, $R_u(\mu)$, $U(\mu)$ and $J(\mu, v)$, a matching rate function $x(\mu, z)$, and a distribution of workers over (μ, v) that satisfy:*

(i) Worker beliefs are updated according to (2.4) and (2.5).

(ii) Given the function x, the policy function $z = g_e(\mu, v_{+1})$ solves an employed worker's problem in (3.1) and yields the return R_e.

(iii) Given the function x, the policy function $z = g_u(\mu)$ solves an unemployed worker's problem in (3.3) and yields the return R_u while the function U satisfies (3.2).

(iv) Given the functions x, (g_e, g_u) and (R_e, R_u, U), the policy functions (w, n) solve the firm's optimal contracting problem in (3.4) and yield the firm value function J.
(v) The function $x(\mu, z)$ satisfies the condition of competitive entry of vacancies, (3.8).
(vi) Given Ω_t, the distribution Ω_{t+1} is consistent with the flows of workers in period $(t+1)$ induced by optimal job applications, separation, exit and entry of workers.
If Ω_t is constant over time, in addition, then the equilibrium is stationary.

The equilibrium is block recursive and therefore can be tractably computed (see Shi, 2009, Menzio and Shi, 2010). That is, workers' and firms' optimal decisions and the matching rate function are independent of the distribution of workers over states. Specifically, when creating a vacancy in a given (μ, z) submarket, firms have no uncertainty about the type of worker they will employ as only type μ workers can apply to the submarket. As a result, the value function $J(\mu, v)$ is known for each submarket and is independent of the distribution of workers over (μ, v) states. Given the firm value function J, the matching rate function $x(\mu, z)$ is determined through the competitive entry of vacancies (3.8).

4. Calibrated Equilibrium

We compute the baseline model and compare the equilibrium policy functions, value functions and model-generated moments with two nested models. (i) *Model CI* (complete information): this nested model retains on-the-job search and wage-tenure contracts, but assumes complete information. It is a discrete-time version of the model in Shi (2009) augmented with heterogeneity in the worker ability. When a worker enters the economy, the type is drawn from the same distribution as in the baseline but is known. In this model, μ represents the true ability of a worker rather than the belief. (ii) *Model NS* (no on-the-job search): this nested model retains incomplete information but eliminates on-the-job search, which is the model in Gonzalez and Shi (2010). The superscripts CI and NS are given to these two nested models, respectively. In the analysis below, most of the comparisons will be between the baseline and model CI in order to illustrate how learning affects the equilibrium when introduced into a model with on-the-job search. In the supplementary Appendix F, we outline the steps of calibration and computation.

4.1. Calibration

The utility function and the matching function are

$$u(w) = \frac{w^{1-\gamma}}{1-\gamma}, \quad M(\ell_e, \theta\ell_e) = \frac{(\ell_e)(\theta\ell_e)}{[(A\ell_e)^\rho + (\theta\ell_e)^\rho]^{1/\rho}}$$

where $\rho, A > 0$. Using (2.1), we get $x = [\left(\frac{\theta}{A}\right)^{-\rho} + 1]^{-1/\rho}$ as the matching probability per search efficiency unit, which is a Dagum (1975) function.[13] From this expression for x we solve θ as a function of x and then solve $q(x) = \frac{1}{A}(1 - x^\rho)^{1/\rho}$. Incorporating the restrictions on the matching rate for a vacancy in (2.3), we have

$$q(x) = \min\{\frac{1}{A}(1 - x^\rho)^{1/\rho}, \bar{q}\},$$

where $\bar{q} = \min\{\frac{1}{A}, 1\}$. The restrictions on the matching rate per efficiency unit in (2.2) are imposed on the domain of x.

There are fourteen parameters in total which are listed in Table 1 together with their values and calibration targets. Ten of the parameters are calibrated to match standard targets in the literature or are normalized. The remaining four parameters, (A, ρ, a_L, p_0), are calibrated to match moments in the data, which requires computation of the equilibrium. We set the length of a period to one month and normalize $y = 1 = \lambda_u$. The coefficient of relative risk aversion, γ, and the interest rate, r, are chosen according to the convention. The parameter value $a_H = 1$ ensures the matching probability for a worker does not exceed one. The probability of surviving the exit shock, σ, is set to give a worker 40 years as the expected length of a career. The separation rate δ is set to 0.026 to match the evidence in the Current Population Survey (CPS). The vacancy cost, k, is set to 0.20, which is consistent with values for this parameter in the literature, and the unemployment benefit, b, to 0.40 to match the literature on frictional wage dispersion.[14] The purpose of

[13] This function is a cumulative distribution function that Dagum (1975) used to study the income/wealth distribution. Note that when $\rho = 1 = A$, the matching function reduces to $M = \frac{\theta\ell_e}{1+\theta}$.

[14] There is disagreement in the literature over the appropriate value of home production which ranges from 0.3 to 0.9. We set b in this interval, with the consideration that the value of b sets an upper bound on wage dispersion. We use the value $b = 0.40$ as in Hornstein et al. (2011).

keeping the notation λ_e is to nest the model without on-the-job search by setting $\lambda_e = 0$. In the baseline we set $\lambda_e = \lambda_u = 1$. We emphasize that this value of λ_e does not lead to an unrealistically high rate of job-to-job (EE) transition; on the contrary, we will target the EE rate to the one in the data (see below).

Table 1. Parameter calibration (monthly)

Parameter		Value	Calibration target
r		0.0041	annual interest rate = 5%
σ		0.9979	40 year average career
δ		0.026	separation rate in CPS data
b		0.40	typical value in the literature
γ		2	conventional value
k		0.20	$k/y = 0.20$
y		1	normalization
λ_u		1	normalization
λ_e		1	benchmark
a_H		1	worker matching probability ≤ 1
η_1 :	A	0.976	2.2% monthly EE rate
	ρ	0.900	6.0% unemployment rate
η_2 :	a_L	0.551	32% wage loss (initial)
	p_0	0.655	12% wage loss (year 5)

We put the remaining four parameters into two groups. The two parameters in the matching function are $\eta_1 = (A, \rho)$ and the two parameters in the prior distribution of the worker type are $\eta_2 = (a_L, p_0)$. Denote $\eta = (\eta_1, \eta_2)$. The parameters in η are jointly determined using the method described in the supplementary Appendix F by targeting four moments in the data: the unemployment rate (6.0%), the monthly job-to-job transition rate (2.2%), the percentage loss in reemployment wages following displacement (32%), and the percentage wage loss five years after reemployment (12%).[15] The parameters in η_1 are relatively important for matching the unemployment rate and job-to-job matching rate, and the parameters in η_2 are relatively important for generating large and persistent wage losses following displacement. We choose the wage-loss targets to match evidence from Davis and von Wachter (2011), who estimate initial reemployment wages to be 25-39%

[15] An unemployment rate of 6.0% and monthly job-to-job transition rate of 2.2% are consistent with the search literature. See Hornstein et al. (2011). A targeted transition rate of 2.2% is also consistent with recent empirical estimates. See Wolthoff (2016).

following displacement and 10-15% five years after reemployment.[16] Appendix F reports the fit of the baseline model and the two nested models with respect to these targets. In particular, in the baseline model, the calibration yields the unemployment rate as 6.1% and the EE rate as 2.18%, which are very close to the targets.

In the subsections below, we analyze the computed value and policy functions and compare the baseline with model CI. In section 5, we analyze equilibrium features related to worker transition, wage dynamics and unemployment.

4.2. Firm's Value Function, Optimal Contracts and Bounds on Worker Value

The firm value function $J(\mu, v)$ **and the matching rate function** $x(\mu, v)$: The function J determines the function x through competitive entry of vacancies which, in turn, is central to workers' search decisions. In Figure 3, the right panel depicts $J(\mu, \bullet)$ for five values of μ, where $\mu_i < \mu_{i+1}$, $\mu_1 = a_L + \epsilon$ and $\mu_5 = a_H - \epsilon$ for some arbitrarily small $\epsilon > 0$. The higher the worker value v, the lower is the value that the firm gets for any given belief; that is, $J(\mu, v)$ decreases in v.[17] The left panel in Figure 3 shows that $J(\bullet, v)$ is increasing in μ for five given values of v. We label the values of v such that $v_i < v_{i+1}$.[18] The firm value increases more steeply in μ when the worker value is high.

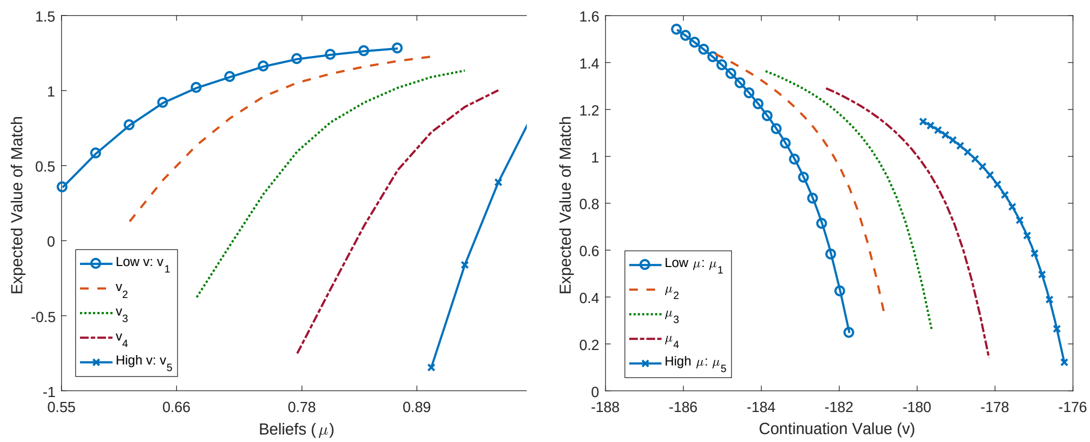
It may be puzzling why the firm value increases in μ. A higher μ does not affect output but increases the expected probability of endogenous separation, which should have a negative effect on the firm value. However, this negative effect of a higher μ on the firm value can be dominated by a positive effect through the promise-keeping constraint. Specifically, since a higher belief increases a worker's expected matching probability, it increases the value of unemployment, $U(\mu)$, and the return on search for an employed worker, $R_e(\mu, v_{+1})$

[16] We choose these as targets for wage loss noting that there are not clear targets in the search literature for wage losses and epmirical estimates of losses can vary significantly by data sets and methodology. See Couch and Placzek (2010) for a review of the empirical literature on wage losses following displacement. Our parameter estimates are not sensitive to the values we choose in these ranges.

[17] See Proposition E.1 in Appendix E for a proof.

[18] In the left panel of Figure 3, the five worker values $v_1, ..., v_5$ are all relatively "high" compared to the values observed in equilibrium across all beliefs. These values of v are found in the domains of workers with relatively high beliefs, which explains why the function J in the left panel of Figure 3 is negative for low belief workers.

(see subsections 4.3 and 4.4). For any given contract terms (w_{+1}, v_{+1}), these higher values and returns to a worker outside the match relax the firm's promise-keeping constraint (3.5) and, thereby, increases the value of the match to the firm. When the promised value v is high, this positive effect of a higher μ on the firm value through the promise-keeping constraint dominates the negative effect through the endogenous separation probability, because the separation probability is low when v is high. This is the case in the left panel in Figure 3 because most equilibrium values of v are relatively high.



$J(\bullet, v)$ $\quad$ $J(\mu, \bullet)$

Figure 3. Firm value function $J(\mu, v)$

In the equilibrium, the properties of J imply the properties of the matching rate per efficiency search unit, $x(\mu, z)$. Since $x(\mu, z)$ must satisfy (3.8) for competitive entry of vacancies, it inherits the features of $J(\phi(\mu), z)$. In particular, $x(\mu, z)$ strictly decreases in z for all $z < \hat{v}(\mu)$ where $\hat{v}$ is defined by (3.9). For any given μ, a match formed in a submarket with a higher offer yields a lower value to a firm. A smaller number of vacancies enter such a submarket, and so the matching rate for an applicant in the submarket is lower. Also, given the offer z, the matching rate function $x(\mu, z)$ increases in μ because $J(\mu, z)$ is increasing in μ and firms create more vacancies as a result.

The retention upper bound $v^*(\mu)$ and the recruiting upper bound $\hat{v}(\mu)$ on the worker value: $v^*(\mu)$ is defined by (3.7) and $\hat{v}(\mu)$ by (3.9). These bounds are important for understanding worker transitions and wage dynamics in equilibrium. In fact, section 5 will show that workers reach the retention upper bound $v^*(\mu)$ within a few periods of being matched with a firm. However, if $\hat{v}(\mu) > v^*(\mu)$, they may continue to search and match with a positive probability. This causes the worker's belief to deteriorate when they stay with their incumbent firm. Most interactions between learning and on-the-job search occur for such workers. Figure 4 depicts $(\hat{v}, v^*)$. These bounds increase in μ because they are defined by setting J to some constants and because $J(\mu, v)$ increases in μ. The two bounds intersect at $\mu = \mu_L$ and $\mu = \mu_H$, where $a_L < \mu_L < \mu_H < a_H$. Furthermore, $\hat{v}(\mu) > v^*(\mu)$ for $\mu \in (\mu_L, \mu_H)$, and $v^*(\mu) > \hat{v}(\mu)$ for $\mu \in (a_L, \mu_L) \cup (\mu_H, a_H)$.

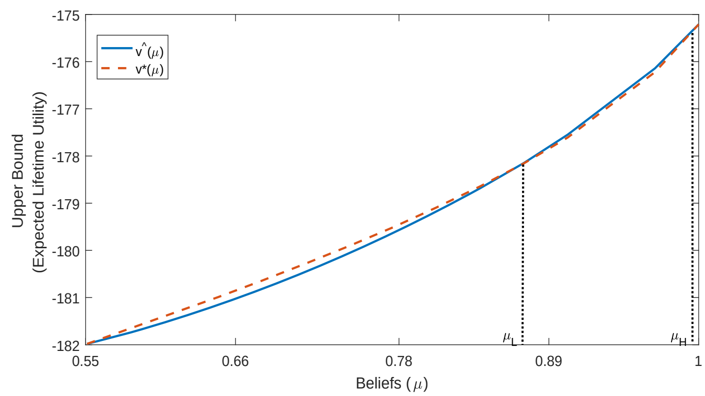
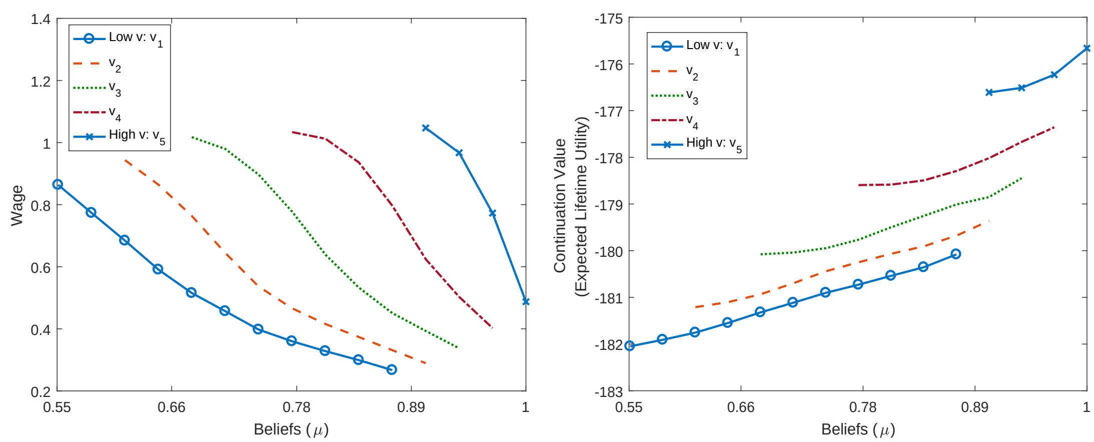


Figure 4. Bounds on recruiting, $\hat{v}(\mu)$, and retention, $v^*(\mu) = \bar{v}\left(\mu^*_{+1}\right)$

Learning is critical for the feature that the recruiting upper bound can be higher than the retention upper bound for some beliefs. The explanation, given in Appendix B, can be summarized here. When a worker stays with a firm, the belief decreases, but if the worker moves to a new match, the belief increases. This gives a recruiting firm an advantage over the incumbent firm since the value of a match to a firm is increasing in the worker's belief as discussed above. Conversely, a recruiting firm has a disadvantage because they must incur the vacancy cost to recruit a worker. When the vacancy cost is small, the advantage exceeds the disadvantage, in which case a recruiting firm can offer a higher value to a worker than what the incumbent firm can (i.e., $\hat{v}(\mu) > v^*(\mu)$). In section 5, we illustrate that this feature is important for the baseline to generate wage dispersion in the equilibrium and wage cuts in job-to-job transitions.

The optimal contract: $w_{+1} = w(\mu, v)$ **and** $v_{+1} = n(\mu, v)$. For any given μ, the wage and the continuation value are increasing in the value promised to the worker, because a firm wants to use both to deliver the higher promised value. This dependence is not depicted. Instead, Figure 5 shows how $w(\bullet, v)$ (in the left panel) and $n(\bullet, v)$ (in the right panel) depend on μ for five given values of v. Given v, the current wage decreases in μ. Also, the continuation value increases slowly in μ for low values of v, and increases more sharply in μ for high values of v where the firm is more constrained in its contract choices. These features of the policy functions confirm the analysis in subsection 3.3. That is, for any promised value, a higher belief increases the firm's capacity to backload wages. To backload wages by more, the firm reduces the initial wage, as shown in the left panel in Figure 5. In addition, when the belief is higher, there is more room for the belief to deteriorate after a match failure. Anticipating this possibility, the firm keeps the continuation relatively flat with respect to μ, as shown in the right panel of Figure 5. However, when μ is close to a_H, the belief will deteriorate very slowly even if the worker stays with the firm. Since backloaded wages have to be delivered, the continuation value increases more sharply in μ when μ is close to a_H.



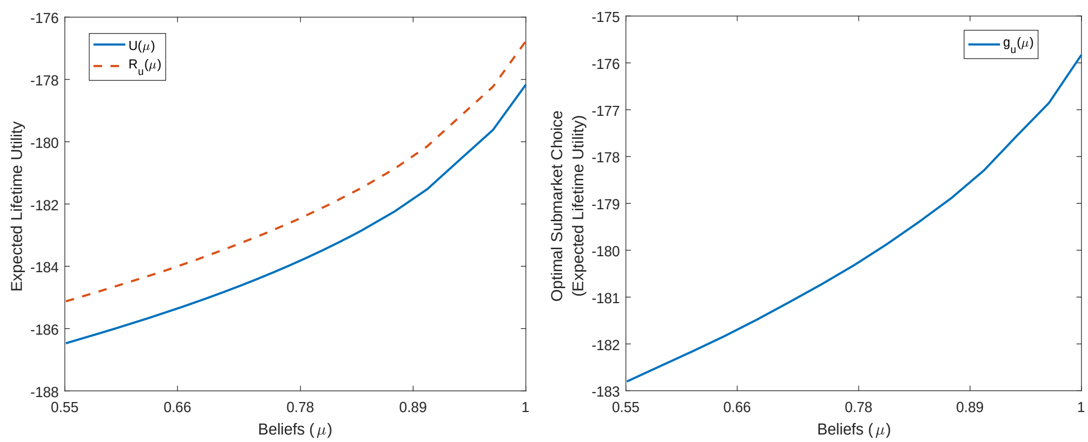
$w(\bullet, v)$ $\qquad\qquad$ $n(\bullet, v)$

Figure 5. Optimal wage $w(\mu, v)$ and continuation value $n(\mu, v)$

To summarize, given the same worker value, a higher belief leads to a higher firm value, more backloading of wages, and more vacancies created for workers. Relative to a low-belief worker, a high-belief worker's current wage is *lower* and future wages rise *more slowly* over tenure given the worker value v.

4.3. Unemployed Workers

The left panel in Figure 6 depicts the value function $U(\mu)$ and the return on search $R_u(\mu)$ for unemployed workers. The two functions are increasing because, the higher the belief, the higher the expected probability of the worker finding employment. Furthermore, these functions are convex in the belief, which reflects a fundamental role of learning, as emphasized by Gonzalez and Shi (2010). Search outcomes generate information by inducing dispersion in the posterior beliefs about a worker's ability. This dispersion is valuable to a worker only if the value function $U(\mu)$ is strictly convex in the belief. The right panel in Figure 6 depicts an unemployed worker's optimal search target, $g_u(\mu)$. Not surprisingly, an unemployed worker with a higher belief searches for a higher offer, because the option value of remaining unemployed is higher.



$U(\mu)$ $g_u(\mu)$

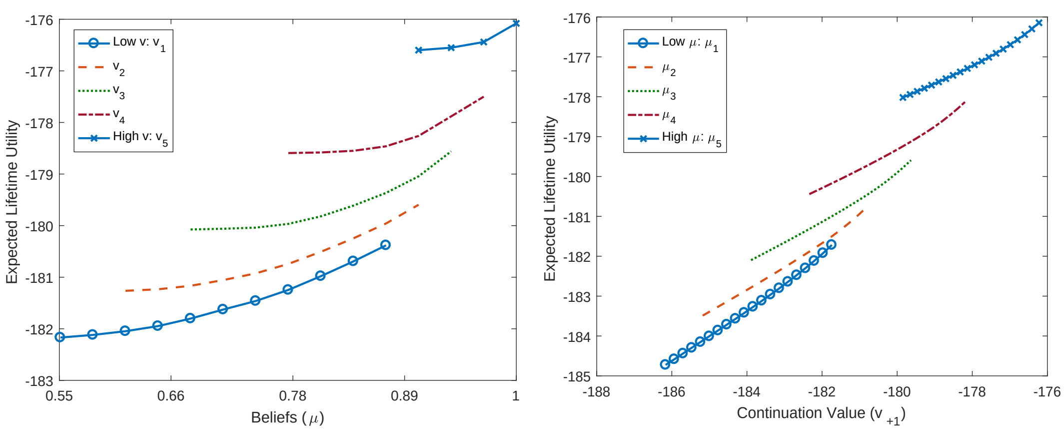
Figure 6. Unemployed workers' value and policy functions

Similar to the learning model without on-the-job search studied by Gonzalez and Shi (2010), the informational content of search depends on the search target. When a worker searches in a submarket with a higher matching rate per efficiency search unit, x, the worker is supposed to be able to find a match relatively easily. If the worker fails to find a match in such a submarket, the belief about the worker will deteriorate precipitously. That is, $F(\lambda_u x, \mu)$ is a decreasing function of x. Taking into account this dependence of belief updating on the submarket searched, an unemployed worker searches for a relatively high target offer initially and reduces the search target over time as he/she stays unemployed.

We will illustrate this role of learning on the optimal search target in Appendix C by comparing the g_u policy function in the baseline with model CI.

4.4. Employed Workers

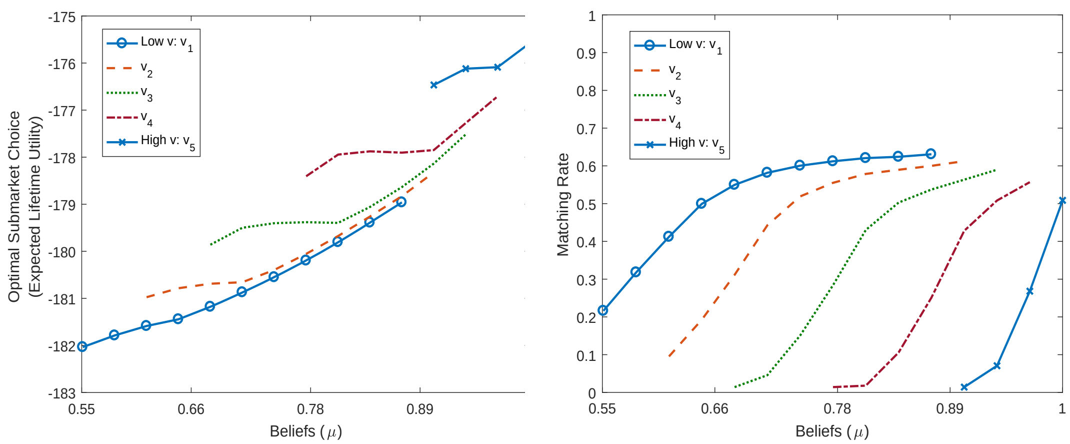
An employed worker's return on search is $R_e(\mu, v_{+1})$, given by (3.1). Figure 7 shows $R_e(\bullet, v_{+1})$ in the left panel for five values of v_{+1} and $R_e(\mu, \bullet)$ in the right panel for five values of μ. Given v_{+1}, the return $R_e\left(\mu, v_{+1}\right)$ is increasing and convex in μ. These properties are similar to those of an unemployed worker's value function. In addition, $R_e\left(\mu, v_{+1}\right)$ is increasing in the continuation value v_{+1}, since v_{+1} serves as an "insurance" against a match failure. However, the surplus of search, $\left[R_e\left(\mu, v_{+1}\right) - v_{+1}\right]$, is decreasing and convex in the continuation value v_{+1} (not plotted). As the continuation value approaches the retention upper bound $v^*\left(\mu\right)$, the surplus of search becomes zero for high or low beliefs. For intermediate beliefs, the surplus of search can still be positive because a match with an outside firm can improve the belief about the worker (see subsection 4.2).



$R_e\left(\bullet, v_{+1}\right)$ $R_e\left(\mu, \bullet\right)$

Figure 7. An employed worker's return on search $R_e(\mu, v_{+1})$

An employed worker's optimal search target is $g_e\left(\mu, v_{+1}\right)$ and the implied matching rate per efficiency search unit is $x_e\left(\mu, v_{+1}\right) = x\left(\mu, g_e\left(\mu, v_{+1}\right)\right)$. Since a type i worker's job-to-job transition probability is $a_i \lambda_e x_e\left(\mu, v_{+1}\right)$, the rate x_e captures the difference in the transition rate across $\left(\mu, v_{+1}\right)$ conditional on the worker's true type. Given μ, if a worker has a higher continuation value, the worker will optimally search in a higher offer submarket. Since the match surplus is lower for a firm that recruits in such a submarket, a smaller number of vacancies exist in the submarket, which leads to a lower matching rate for a worker. Thus, for given μ, $g_e(\mu, v_{+1})$ increases in v_{+1} and $x_e(\mu, v_{+1})$ decreases in v_{+1} (not plotted).[19] For the dependence on μ, Figure 8 depicts $g_e(\mu, v_{+1})$ in the left panel and $x_e(\mu, v_{+1})$ in the right panel for five levels of v_{+1}. The function $g_e(\mu, v_{+1})$ increases in μ; that is, a higher belief induces an employed worker to search for a higher offer. This higher search target has a negative effect on the worker's matching rate x_e because a smaller number of vacancies are created for the higher offer. However, this negative effect is dominated by the equilibrium effect that more vacancies are created for a high-belief worker than for a low-belief worker, as analyzed in subsection 4.2. As a result, the matching rate $x_e(\mu, v_{+1})$ is increasing in μ, especially at high values of μ.



$g_e(\bullet, v_{+1})$ $\quad$ $x_e(\bullet, v_{+1})$

Figure 8. An employed worker's optimal search target, $g_e(\mu, v_{+1})$, and the matching rate per efficiency search unit, $x_e(\mu, v_{+1})$

4.5. Comparisons with Model CI

In order to further understand the effect of learning, we compare the dependence on μ between the baseline model and model CI (i.e., the model with complete information). For the comparisons in this section, we compute the two models for the same parameter estimates from subsection 4.1. In model CI, μ is the true type of a worker instead of the belief. For this to be meaningful, a worker's type in model CI is allowed to be any value

[19]See Proposition E.1 in Appendix E for a proof.

in $[a_L, a_H]$ instead of $\{a_L, a_H\}$. We focus on a partial list of comparisons between the two models and delegate additional comparisons to Appendix C.

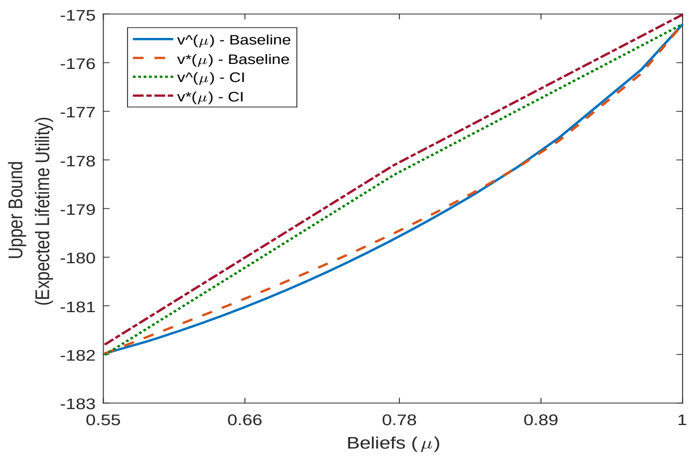
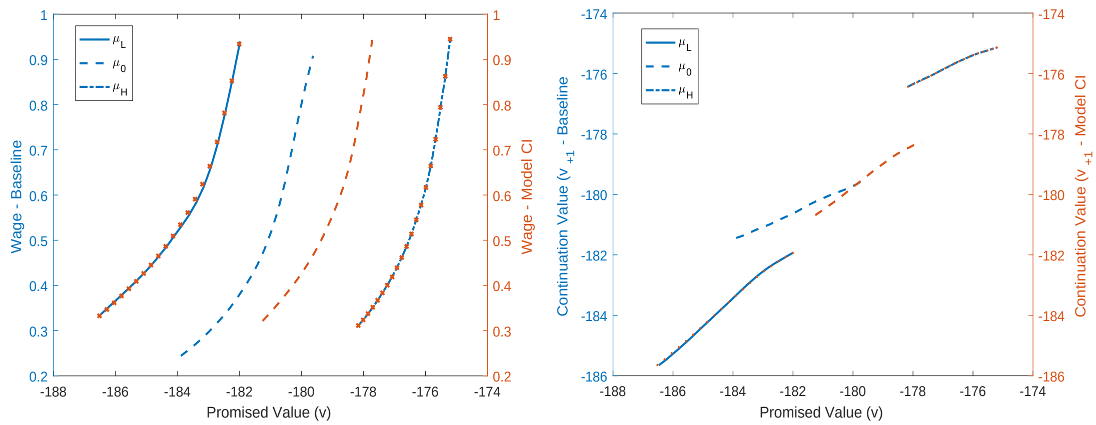


Figure 9. The baseline and model CI: $v^*(\mu)$ and $\hat{v}(\mu)$

The upper bounds on the worker value: $v^*(\mu)$ **and** $\hat{v}(\mu)$. In model CI, the bound $v^*(\mu)$ needs to be redefined as $v^*(\mu) = \bar{v}(\mu)$ because a worker's μ does not change over time. Figure 9 depicts the retention upper bound $v^*(\mu)$ and the recruiting upper bound $\hat{v}(\mu)$ in the two models. Model CI has $v^*(\mu) > \hat{v}(\mu)$ for all μ, but the baseline has $v^*(\mu) < \hat{v}(\mu)$ for $\mu \in (\mu_L, \mu_H)$ (see Figure 4). When learning is irrelevant as in model CI, a recruiting firm has no advantage over a retention firm in improving the belief. With only the disadvantage of having to incur the vacancy cost, a recruiting firm cannot offer as much as an incumbent firm can. Moreover, the bounds in model CI are higher than in the baseline for all interior beliefs, and this difference increases when the belief is further away from the two endpoints. In the baseline, as tenure increases, the firm inherits a worker with a lower belief and can rely less on backloading wages to deliver expected lifetime utility that it had promised in the contract. For this reason, the firm optimally offers a worker a lower continuation value in the baseline than in model CI given the belief and worker value. As the belief becomes more interior, it has more room to be updated following a search outcome, and so the larger is the difference between the two models in the upper bounds on the worker value increases. This difference vanishes when the belief is at either endpoint because the belief ceases to update at the endpoints.

The optimal contract: For three levels of μ, Figure 10 depicts the wage $w(\mu, v)$ in the left panel and the continuation value $n(\mu, v)$ in the right panel. For beliefs close to the true ability types, the functions $w(\mu, \bullet)$ and $n(\mu, \bullet)$ are almost indistinguishable between the two models, because belief updating is small in these cases. However, for intermediate beliefs, these policy functions are significantly different between the two models. For a given wage, the worker value is lower in the baseline than in model CI. Also, for a given worker value, the continuation value is higher in the baseline, which reflects an increased ability to backload wages. These differences between the two models are primarily driven by the difference depicted in Figure 9. That is, when the belief is intermediate, an employed worker whose value is close to the retention upper bound in the baseline continue to search for new matches to prevent the belief from falling, but such a worker in model CI stops searching because μ does not change and there is no possibility of finding a better match.



$w(\mu, \bullet)$ $\quad$ $n(\mu, \bullet)$

Figure 10. The baseline and model CI: $w(\mu, v)$ and $n(\mu, v)$

Unemployed workers' value and policy functions: Figure 11 depicts $U(\mu)$ and $g_u(\mu)$, which are increasing in μ in both models. $U(\mu)$ is higher in model CI than in the baseline for all interior μ, because a worker's μ in model CI does not fall during unemployment. Furthermore, $U(\mu)$ in model CI is concave rather than convex as learning is irrelevant and search outcomes do not contain valuable information as in the baseline. Notice that the optimal search target, $g_u(\mu)$, is lower in the baseline than in model CI for all interior μ. However, in subsection 4.3, we have explained that learning motivates an unemployed worker to search for higher offers initially. In Appendix C, we resolve this puzzle by showing that the value of an unemployed worker and the amount of vacancy creation are smaller in the baseline than in model CI. These equilibrium effects reduce an unemployed worker's optimal search target. They dominate the direct effect of learning and lead to a lower $g_u(\mu)$ in the baseline than in model CI.

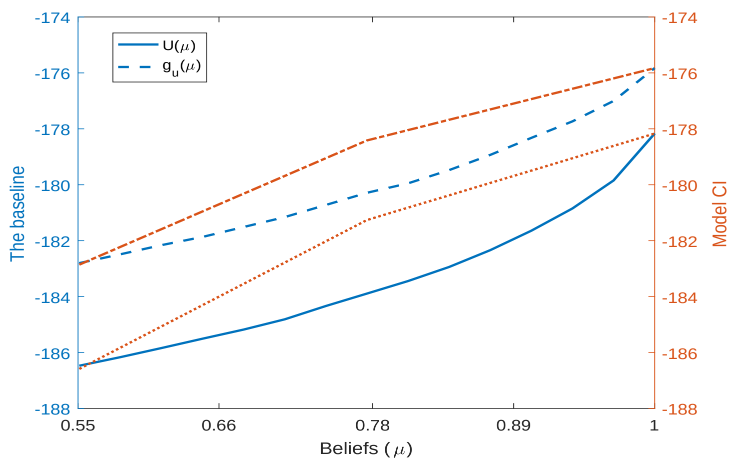


Figure 11. The baseline and model CI: $U(\mu)$ and $g_u(\mu)$

5. Worker Transition, Wage Dynamics and Inequality

Using the computed policy functions, we simulate the model with a large number of workers until the distribution of workers over states converges (see Appendix F). We report the features of this stationary equilibrium and relate them to empirical facts on worker transition, wage dynamics, unemployment duration, and frictional wage inequality.[20]

5.1. Wage Dynamics and Job-to-Job Transition

Subsections 4.2 and 4.4 explained how learning and search interact with wage-tenure contracts in equilibrium. Specifically, if the belief about a worker is higher, there is a larger capacity for backloading wages. When a worker stays with a firm, the belief falls, in which case the firm can scale back the continuation value for the worker. To illustrate this interaction, the left panel in Figure 12 depicts wages over tenure for three belief levels and the right panel depicts the continuation value over tenure, where tenure is the duration,

[20] The model also generates significant wage losses due to unemployment, since these losses were used as part of the calibration targets.

in months, of employment in the same firm. The two panels depict both the baseline and model CI. Consider the right panel first. In both models, the worker value becomes close to the retention upper bound v^* within a few periods. In model CI, the worker value remains at this level until the worker exogenously separates into unemployment or exits the economy. In the baseline, however, the worker value starts to decline because the retention upper bound falls as the belief deteriorates.

Now consider the left panel in Figure 12. For all belief levels, wages in model CI grow rapidly initially and then become almost constant after less than one year. Such a fast but short wage growth indicates that job-to-job transitions take place in a short time in model CI. As workers stop moving up in the value, the room for backloading wages diminishes quickly, and previously backloaded wages materialize. A similar pattern of wage growth emerges in the baseline model for workers with low beliefs, because there is not much room for backloading wages for such workers. However, for high-belief workers, the initial wage is lower, wages grow with tenure more slowly and for a longer time. The baseline model predicts that wage growth is significant for almost ten years for high belief workers. A firm postpones much of the wage payment to the future in anticipation of the reduction in the belief when the worker stays with the firm. However, when the latter event occurs, the firm will reduce the continuation value and will not need to increase wages much.

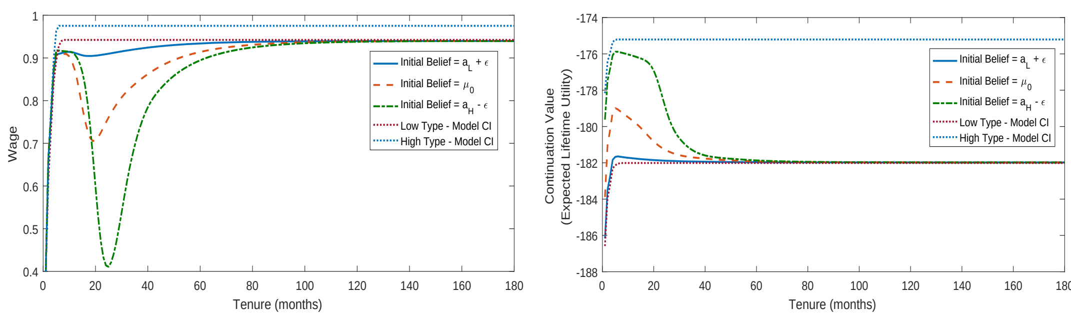


Figure 12. Wages and continuation values over tenure

Given the above description of a firm's incentive, one might conjecture that wages may fall eventually after a worker has stayed with the firm for a sufficiently long time. However, this does not happen in Figure 12. Under a set of sufficient conditions, we prove in the supplementary Appendix E that wages increase over tenure as long as the continuation value is below the retention upper bound. Moreover, when the continuation value reaches the retention upper bound, wages remain constant at y while the continuation value declines along the retention upper bound with the belief. These sufficient conditions seem to be satisfied in Figure 12 once tenure exceeds a threshold.[21]

Table 2. Monthly EE transition rates with tenure (%)

	Baseline			Model CI		
tenure (years)	Type L	Type H	Average	Type L	Type H	Average
avg	1.51	2.82	2.18	1.95	0.57	1.24
<1	2.74	4.74	3.82	3.06	1.86	2.53
1 - 2	0.81	1.44	1.14	1.22	0.00	0.64
2 - 3	0.79	1.40	1.09	1.23	0.00	0.61
3 - 4	0.82	1.39	1.09	1.18	0.00	0.53
4 - 5	0.75	1.46	1.07	1.20	0.00	0.50
>5	0.81	1.45	1.07	1.18	0.00	0.38

The wage-tenure profile in Figure 12 is conditional on a worker remaining with a firm. As such, it should not be taken as evidence that a high-belief worker is worse off than a low-belief worker. While wages grow with tenure less quickly for a high-belief worker than for a low-belief worker, the former obtains a higher value on average because a high-belief worker is more likely to make the transition to another job with a higher value. Table 2 lists job-to-job (EE) transition rates conditional on the true type of workers.[22] In both the baseline and model CI, the EE rate falls with tenure. However, the two models have opposite predictions about the relative EE rate between high- and low-type workers. In the baseline, high-type workers' EE rates are about 80% higher than low-type workers', and this difference remains for most of tenure lengths. In model CI, in contrast, high-type workers' EE rates are about 40% lower than low-type workers' for tenure shorter than one year. For tenure longer than one year, almost all EE transitions in model CI are made

[21]Past this tenure threshold, declining wages are not optimal for the firm because they conflict with backloading. If a wage profile had a decreasing section in the future, the firm could bring such low wages to the present by depressing the current wage. For the same value promised to the worker, this change in the wage profile would increase the firm value by increasing current profit and retention.

[22]To make the comparison between the baseline and model CI in Table 2 more appropriate, we re-calibrate model CI to match the empirical moments described in section 4.1. This calibration results in parameter estimates for model CI of $\eta^{CI} = (\rho^{CI}, A^{CI}, a_L^{CI}, p_0^{CI}) = (0.90, 1.004, 0.535, 0.50)$. When using the paramters estimated for the baseline, model CI generates an average EE rate of 1.16%.

by low-type workers. These differences illustrate the quantitative importance of belief updating for EE transitions in the baseline. Namely, a high-type worker is motivated to find a new match more frequently because, if the worker stays with a firm, the belief will deteriorate more quickly than for a low-type worker.[23]

A striking implication of the baseline is that wages can fall when a worker moves from one job to another, although the worker value increases in the transition. This is particularly the case for a worker whose continuation value is close to the retention upper bound. If such a worker stays with the incumbent firm, the belief deteriorates and the continuation value falls. The firm will become increasingly constrained in its ability to postpone wage payments, and so wages will be high for such workers. If the worker moves to another firm, the new match improves the belief about the worker, which increases the attracting firm's ability to backload wages. This can result in a wage decrease relative to the wage in the previous match. The longer a worker's tenure with a firm, the closer the worker value will be to the retention upper bound, and therefore the more likely it is that a job-to-job transition results in a wage cut. It should be emphasized that this interaction between leaning and wage backloading is the cause of wage cuts in EE transitions. If wages were fixed over tenure, an EE transition would have to be compensated by an immediate increase in the wage. If there were no learning, as in model CI, an EE transition would cause no improvement in the belief about the worker. In this case, wages would still be backloaded, but the higher promised value in the new match of the EE transition would have to be delivered by a new wage-tenure contract whose initial wage is higher than the wage the worker has just left (see Shi, 2009).

Table 3 lists the percentage of EE transitions with wage changes less than or equal to a given percentage. The baseline model produces frequent wage reductions associated with EE transitions. Overall, about 18.22% of EE transitions experience wage cuts and about 2.13% of EE transitions have wage cuts greater than 5%. On average, an EE transition is more likely to be associated with a wage cut for low-type workers than for high-type

[23] The effect of the belief on the EE transition rate suggests that job-to-job transitions may have positive occurrence dependence in the baseline. Workers who just made a job-to-job transition in the previous period are likely the ones with high beliefs who will make such transitions again.

workers, but the size of such wage cuts are larger for high-type workers. For high-type workers, 16.44% of EE transitions result in a wage cut and approximately 2.43% of such cuts are greater than 5%. In comparison, for low-type workers, 21.65% of EE transitions result in a wage cut, and about 1.55% of those cuts are greater than 5%. Also, confirming the above analysis, the frequency of EE transitions with wage cuts increases with a worker's tenure in the previous match. This frequency is 11.4% for workers with tenure less than one year and increases to 32.07% for workers with tenure between 2 to 3 years. However, as tenure increases, the magnitude of the wage cut associated with an EE transition falls.

Table 3. Changes in wages with EE transitions

		Percentage of EE transitions with wage changes $< x\%$									
tenure (years)	type	-20	-15	-10	-5	0	5	1	15	20	100
All	L	0.60	1.10	1.34	1.55	21.65	55.80	63.42	69.52	77.48	100
	H	0.83	1.67	2.02	2.43	16.44	51.92	60.87	69.22	77.93	100
	avg.	0.75	1.48	1.79	2.13	18.22	53.25	61.74	69.32	77.78	100
< 1	L	0.90	1.65	2.01	2.31	10.62	35.36	45.11	54.27	66.21	100
	H	1.18	2.37	2.86	3.38	11.78	34.07	44.54	56.37	68.72	100
	avg.	1.09	2.13	2.58	3.03	11.40	34.50	44.73	55.68	67.89	100
$1-2$	L	0	0	0	0.03	37.07	95.28	100	100	100	100
	H	0	0	0	0.19	22.60	94.11	100	100	100	100
	avg.	0	0	0	0.13	27.56	94.51	100	100	100	100
$2-3$	L	0	0	0	0.05	43.11	97.06	100	100	100	100
	H	0	0	0	0.14	25.82	94.25	100	100	100	100
	avg.	0	0	0	0.10	32.07	95.26	100	100	100	100

Frequent wage cuts associated with EE transitions exist in the data but have been difficult for theories to explain without the loss of a match-specific component or firm heterogeneity. With administrative French data, Postel-Vinay and Robin (2002) find that the fraction of annual EE transitions with wage cuts greater than or equal to 5% ranges from 28.5% to 42.9%, depending on the category of workers. A typical explanation for wage cuts is the loss of skill in the transition between jobs. Our theory, based on learning and on-the-job search, provides an alternative explanation. The theory yields the additional prediction that the frequency of wage cuts in EE transitions increases with tenure. It is interesting to check whether this is supported by the data.

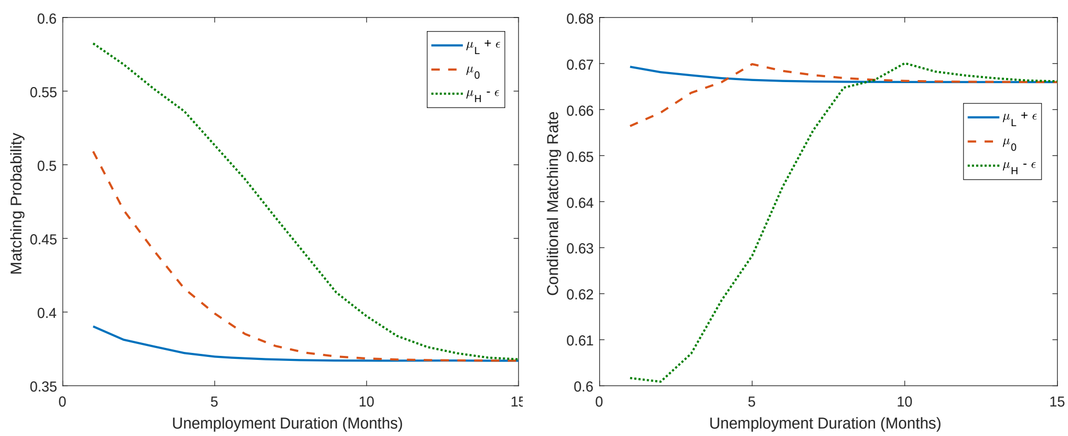
5.2. Duration Dependence in Unemployment

For workers who enter unemployment with a given belief μ, we calculate how the expected matching rate, $\mu\lambda_u x_u$, and the matching rate per efficiency unit, x_u, change with unemployment duration. Figure 13 depicts these matching rates for three initial belief levels. As shown in the left panel, the expected matching rate falls over the duration for all three initial beliefs. Thus, the longer a worker has been unemployed, the more likely the worker will remain unemployed. There is positive duration dependence in unemployment. Moreover, the higher the belief of the worker when the worker enters unemployment, the more steeply the expected matching rate falls over unemployment duration.

An important cause of positive duration dependence is dynamic selection by unemployed workers. High-type workers are more likely to find jobs and exit unemployment, while low-type workers are more likely to remain unemployed. For any cohort of workers entering unemployment, the fraction of low-type workers who stay unemployed in the cohort increases over time. As the belief deteriorates, the average job-finding probability in the cohort declines over unemployment duration. Controlling for this selection effect yields so-called "true" duration dependence in unemployment, which measures how the future employment probability depends on unemployment duration among identical workers. In our model, the selection effect is captured by μ that multiplies $\lambda_u x_u$ in the expected matching rate. After controlling for the worker type, the matching rate varies with unemployment duration entirely through x_u, which is shown in the right panel of Figure 13.

True duration dependence can be a non-monotonic function of the belief entering unemployment. For workers who enter unemployment with a low belief, the matching rate per efficiency search unit is relatively flat and decreases slightly in unemployment duration. That is, true duration dependence is positive for low-belief workers. For high-belief workers, as unemployment duration increases, the matching rate per efficiency search unit increases first and then decreases. That is, true duration dependence is negative for short duration and positive for long duration. This turning point is about ten months of unemployment for workers who enter unemployment with a belief close to a_H.

To explain the results on true duration dependence, recall that $x_u(\mu) = x(\mu, g_u(\mu))$, where $g_u(\mu)$ is the optimal target value searched by an unemployed worker with the belief μ. The function $x(\mu, g_u)$ increases in μ and decreases in g_u. By reducing the belief μ, an increase in unemployment duration reduces x_u through the direct effect of μ but increases x_u through g_u. The direct effect of μ on x_u arises from the response of firms. Since the firm value of hiring a worker increases in μ (see Figure 3), deteriorating beliefs induce firms to reduce vacancies which reduces x_u for the workers. Conversely, the effect of g_u on x_u arises from the response of workers. Deteriorating beliefs induce a worker to reduce the search target, which increases x_u. For a worker who enters unemployment with a low belief, both effects are weak, although the reduction in vacancies dominates to produce true positive duration dependence. For a worker who enters unemployment with a high belief, the reduction in the search target dominates initially to produce true negative duration dependence. After the worker has stayed in unemployment for a sufficiently long time, the belief has deteriorated to a sufficiently low level so that the relative strength of the two effects is reversed and true duration dependence becomes positive.[24]



$\mu\lambda_u x_u$ $\lambda_u x_u$

Figure 13. Dependence of matching rates on unemployment duration

The above analysis on true duration dependence illustrates the importance of incomplete information and on-the-job search. If a worker's type is known, as in model CI, the matching rate should be the same for all workers of the same type and be independent of

[24]Expressed as functions of μ instead of unemployment duration, the unconditional rate $\mu\lambda_u x_u(\mu)$ is decreasing and $\lambda_u x_u(\mu)$ is hump-shaped.

unemployment duration. On the other hand, if on-the-job search is rule out, as in model NS, firms have no incentive to backload wages, which eliminates the benefit of a higher belief in relaxing the promise-keeping constraint. In this case, the firm value depends on the belief entirely through the promised wage/value, which implies that x_u depends on μ entirely through the search target g_u. As unemployment duration increases, the optimal target falls, and so x_u increases. True duration dependence is negative in model NS, although dynamic selection can still make the matching rate unconditional on the worker type decrease in unemployment duration.

The contrast between the two panels in Figure 13 illustrates that unobserved heterogeneity among workers is important for understanding duration dependence. The right panel also shows that understanding workers' and firms' responses to duration is critical for understanding true duration dependence. Recent empirical evidence supports this analysis. Using high-frequency panel data on individuals' job applications from a job posting website, Kudlyak et al. (2013) find that the types of jobs applied to are highly correlated with the applicants' education level when the applicants first enter the website. This correlation drops by 33% from week 2 to week 26 of search, with half of the reduction happening by week 5. As unemployment duration increases, job seekers who initially applied to high-education jobs start to apply to low-education jobs. Thus, job applicants lower their search target over the duration, as our theory suggests.[25]

With similar data, Faberman and Kudlyak (2017) find that longer-duration job seekers send relatively more applications per week throughout their entire search. That is, job seekers who end up with longer duration have had higher search effort throughout the search process. To relate this finding to our results, note first that workers who enter unemployment with a lower belief will have longer unemployment duration on average, as shown in the left panel of Figure 13. We next note that search effort can be introduced into the baseline by endogenizing λ_u. Increasing search effort has similar effects to reducing the search target, as both increase $\lambda_u x_u$. As a result, the finding in Faberman and Kudlyak

[25] The dataset explored by Kudlyak et al. (2013) does not contain information on wages. However, because wages are likely to increase in the education level, it is reasonable to infer that wages implicit in the jobs applied to also fall with the duration of search.

(2017) on search effort suggests that the lower the belief of the worker when they enter unemployment, the higher is $\lambda_u x_u$. As the right panel of Figure 13 shows, this suggested effect is consistent with the baseline model's prediction for unemployment duration shorter than 7 months. As unemployment duration becomes sufficiently long, beliefs converge to a_L, in which case the effect of the initial belief on search effort declines to zero.[26]

5.3. Frictional Wage Dispersion

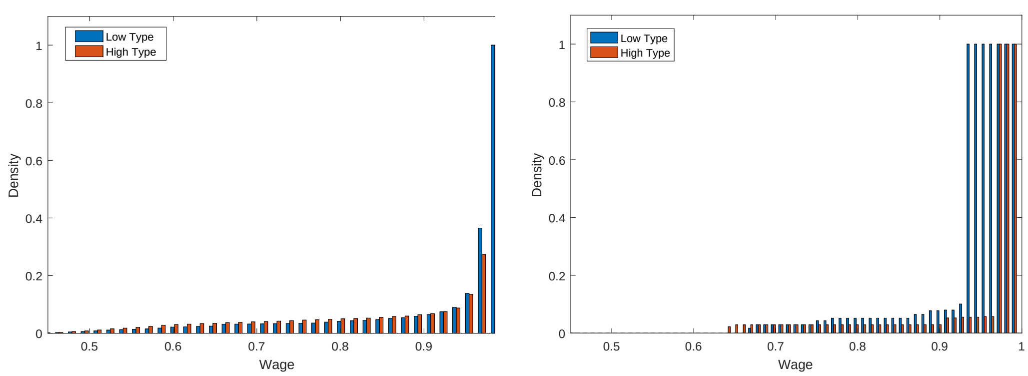
Frictional wage dispersion refers to dispersion in wages after controlling for observable heterogeneity and among workers. Hornstein et al. (2011) propose to measure frictional dispersion by the ratio of the mean to the minimum wage, denoted as Mm. They estimate that the mean-min ratio is approximately 2 in the data but only approximately 1.05 in a number of standard search models. The baseline model in this paper introduces on-the-job search and learning to a standard model of directed search. On-the-job search endogenously generates a wage ladder, and learning endogenously generates heterogeneity in beliefs. Analytically, both elements can widen frictional wage dispersion.

To see how much these mechanisms increase wage dispersion *quantitatively*, we report statistics on frictional wage dispersion for the two types of workers separately to contrast between the baseline and model CI. However, it should be emphasized that all wage dispersion in the baseline, including dispersion between the two types of workers, is frictional because an observer cannot tell the two types apart. All workers draw their types from the same distribution when they enter the economy. While beliefs are updated over time, the true type of a worker is never revealed.

Figure 14 displays the wage density functions of the two types of workers in the equilibrium. The left panel gives the wage density functions for low- and high-type workers in the baseline and the right panel for model CI. In both models, wages are heavily concentrated

[26]Duration dependence is also studied with the audit approach that sends ficticious resumes to employers to determine how callback rates vary with unemployment duration. These studies have obtained conflicting results (e.g. Farber et al., 2015, versus Kroft et al., 2013), and have all ignored workers' endogenous responses to duration. It is difficult to map these field experiments into aspects of the actual data where endogenous responses to duration potentially play an important role.

at the highest level for each type of worker. This shows that on-the-job search takes a worker to the highest wage quickly. However, wages are less concentrated at the highest level in the baseline than in model CI as learning slows down the process of rising wages, particularly for high-type workers. As beliefs deteriorate over tenure within a firm in the baseline, the continuation value eventually falls, which limits the extent to which wages grow. Thus, learning is critical to generate dispersion in equilibrium wages. Referring to the right panel of Figure 14, in model CI there is a difference in the highest wage that can be earned for low- and high-type workers, which limits the amount of wage dispersion generated by the model, especially for low-type workers. Conversely, in the baseline, wages are determined by the worker's belief and not by the true type. As a result, both types can earn the same highest wage in equilibrium. Furthermore, the highest wage earned in the baseline is higher than the highest wage earned for either type in model CI. This results from the fact that in the baseline, recruiting firms can continue to attract workers from their match through an improvement in the belief, which causes incumbent firms to increase the highest wage they are willing to pay a worker, a force that is absent in model CI since the worker's type is known.



Baseline Model CI

Figure 14. Wage density functions

Finally, in model CI the lowest wage earned by any worker is approximately 0.65, which is the reemployment wage that workers earn after leaving unemployment. In the baseline, workers with beliefs close to the true ability types earn similar wages to these when they become unemployed. However, workers with intermediate beliefs have much lower reemployment wages. Workers with beliefs close to μ_0 earn a reemployment wage between 0.45 and 0.5 when exiting unemployment (see Figure 19 of Appendix C). Referring to both panels of Figure 14, the difference in the minimum wage between the baseline and model CI is significant. Learning is critical for this lower reemployment wage through the evolution of beliefs within a match, as previously discussed. Overall, this effect of learning on the highest and lowest equilibrium wages should significantly increase Mm as wages are heavily concentrated at the highest level and the minimum wage is lowered for each type.

Table 4. Equilibrium wage dispersion in three models

	Baseline			Model CI			Model NS		
	L	H	avg.	L	H	avg.	L	H	avg.
Mm	2.35	2.36	2.36	1.38	1.50	1.46	1.033	1.034	1.034
$\frac{90}{10}$	1.043	1.041	1.042	1.016	1.000	1.000	1.002	1.008	1.003
$\frac{50}{10}$	1.038	1.040	1.039	1.016	1.000	1.000	1.000	1.000	1.000
st.dev.	0.078	0.088	0.083	0.008	0.009	0.008	0.008	0.009	0.008

Four measures of wage dispersion are listed in Table 4. In addition to the baseline and model CI, we report wage dispersion in model NS that preserves incomplete information but eliminates on-the-job search. The mean-min ratio (Mm) in the baseline model is about 2.36, even after controlling for a worker's observable type. This is not surprising given that wages are determined by the workers belief and not the type. The mean-min ratio generated by the baseline is much larger than in the literature and similar to what is measured in the data.[27] In model CI, the mean-min ratio is also large but significantly smaller than in the baseline.[28] In all three models, wage dispersion is much smaller by the other three measures: the 90-10 percentile ratio, the 50-10 percentile ratio, and the standard deviation. The 90-10 and 50-10 ratios in model CI indicate almost no dispersion. This is not surprising given the large concentration of workers at the highest wage (see Figure 14). In the baseline, in contrast, the average 90-10 ratio between the two types of workers is about 1.042 and the average 50-10 ratio is about 1.039. Such dispersion is small but significant relative to model CI, suggesting that learning plays a significant role in generating wage dispersion.

[27]Shi (2016) proposes an alternative model that generates a mean-min ratio in wages as 1.8. In this model, firms can undertake investment that increases a job's productivity.

[28]For model CI, we use the same parameter estimates that were found when calibrating the baseline model. If model CI is recalibrated to match the empirical targets, it generates similar dispersion numbers to those presented in Table 4.

To assess frictional wage dispersion, it is important to mention two calibration targets. One is the unemployment rate (6%) and the other is the ratio of home to market production ($b/y = 0.4$). Hornstein et al. (2011) explain why these two targets critically restrict frictional wage dispersion. A low unemployment rate implies a high job-finding probability for unemployed workers. For unemployed workers to search and accept jobs quickly, the option value of continuing to search while unemployed must be low. Furthermore, since the ratio of home to market production is significant, the low option value of search implies that wage dispersion must be small. As a result of these two targets, most search models generate negligible frictional wage dispersion, with $Mm = 1.046$. In fact, Hornstein et al. (2011) show that for the canonical search model to generate a sizable mean-min ratio in wages while meeting the targeted unemployment rate, home production must be negative.

The interaction between learning and on-the-job search is critical for increasing frictional wage dispersion. Eliminating learning, as in model CI, reduces the average mean-min ratio from 2.36 to 1.46 and the average 90-10 ratio from 1.04 to 1.00. Also, eliminating on-the-job search but retaining incomplete information as in model NS reduces the average mean-min ratio significantly to 1.034 and the average 90-10 ratio to 1.00. Again, the explanation is the high job-finding probability implied by the low unemployment rate. When on-the-job search is absent, the fast transition of unemployed workers into employment terminates the learning process quickly. As a result, there is not much dispersion in beliefs or wages. Furthermore, in the absence of on-the-job search, unemployed workers significantly increase their reservation wage as the opportunity cost of accepting a low wage is much higher. This increase in the reemployment wage of workers significantly limits the amount of frictional wage dispersion in model NS by increasing the minimum wage in equilibrium.

Finally, we note that directed search has a tendency to compress wage dispersion relative to undirected search. When search is undirected, a worker accepts any offer above the reservation wage. When search is directed, however, a worker searches only for the offer that maximizes the expected surplus of search. This optimal target of search exceeds the reservation wage because the latter yields a suboptimally low expected surplus to the worker. Thus, it is remarkable that model CI can generate a sizable mean-min ratio.

6. Conclusion

We have analyzed an equilibrium model where workers learn about their ability by searching for jobs, both on the job and during unemployment, and firms offer wage-tenure contracts. A worker's state is described by the belief about the worker's ability and the present value promised to the worker. While the belief and the worker value jointly affect search decisions and contracts, search outcomes yield information for updating the belief about the worker. For a given worker value, the firm value increases in the belief. Thus, the higher the belief, the more vacancies are created for the workers, and the higher the matching probability for the workers. In addition, a higher belief induces a firm to backload wages by more, which may result in the initial wage in the match to be lower. The calibrated equilibrium yields a number of interesting results about worker transition, wage dynamics, unemployment duration, and frictional wage inequality. These results are useful for interpreting empirical regularities in the data.

There are several opportunities for future research. One is to restrict the extent to which the belief is contractible. In this model, the belief and the worker value summarize a worker's history of search, employment, wages, and unemployment. Not all of this history may be contractible. One may find it interesting to restrict the contractible aspects to the most recent wage, the most recent spell of unemployment and the number of unemployment spells. While this restriction enables one to obtain results on how worker transition and wage dynamics depend on the specific aspects, it comes with the cost of increasing the dimension of the state space significantly. Another direction of research is to relax the assumption that all matches produce the same level of output. By allowing output to be positively correlated with a worker's ability, the extension allows a worker to learn both from searching in the market, as in this paper, and from producing in the current match, as in Jovanovic (1979). In such an extension, realized output in a match is a state variable in addition to the belief and the worker value. Finally, one may want to use the model to structurally estimate the parameters, especially those related to the worker type.

Appendix

A. Ruling out Experimentation

Experimentation by workers can occur only when the belief is interior. If μ has reached either a_L or a_H, the belief does not change any further (see (2.4) and (2.5)). Thus, we only consider interior values of μ in this appendix. Also, if the offer searched by a worker is higher than the value of staying in the current position, it is optimal for the worker to accept the match rather than reject the match. Thus, experimentation can occur only when a worker searches for an offer less than or equal to the continuation value in the current position.[29]

Employed workers have no incentive to experiment. When an employed worker stays with the incumbent firm after search, the continuation value v_{+1} promised by the firm is independent of the current search outcome. Experimentation is dominated by the choice of searching for the offer $z > v_{+1}$ that solves the problem (3.1). Since the search target gives a higher value than the continuation value in the current match, the worker always moves to the new match if search is successful.

An unemployed worker may have an incentive to experiment, as the updated belief has a direct effect on the worker value. Consider an unemployed worker whose belief is μ before search in a period. If the worker fails to match, the continuation value will be $U(\mu_{+1})$, where $\mu_{+1} = F(\lambda_u x, \mu)$ is given by (2.5). After getting an offer and rejecting it, the worker's belief will be updated to $\phi(\mu)$ and the continuation value will be $U(\phi(\mu))$. If the worker experiments, the worker will search for an offer $z \leq U(\phi(\mu))$, as explained above. The optimal choice of experimentation solves the following problem that generates the return on experimentation as:

$$R_u^e(\mu) \equiv \max_{z \leq U(\phi(\mu))} \left\{ \mu \lambda_u x(\mu, z) \left[U(\phi(\mu)) - U(\mu_{+1}) \right] + U(\mu_{+1}) \right\}, \tag{A.1}$$

subject to $\mu_{+1} = F(\lambda_u x(\mu, z), \mu)$.

[29] An alternative way to rule out experimentation is to impose a direct cost of search, as in Burdett and Vishwanath (1988). We do not follow this approach because it changes the model.

It is not optimal for an unemployed worker with belief μ to experiment if and only if

$$R_u(\mu) \geq R_u^e(\mu). \tag{A.2}$$

Experimentation is not optimal if and only (A.2) holds for all μ.

To express (A.2) in detail, we assume that the function $U(\mu)$ is strictly increasing and convex in μ for all interior μ (see Figure 6). Note that $F(\lambda x, \mu)$ defined by (2.5) is strictly increasing and concave in x for all interior μ, with the additional property that $(1-\mu\lambda x) F_{11}'' - 2\mu F_1' = 0$. As a result, the above assumption on $U(\mu)$ implies that $(1-\mu\lambda x) U(F(\lambda x, \mu))$ is strictly convex in x for all interior μ.

Since $x(\mu, z)$ is decreasing in z (see subsection 4.2), we rewrite the constraint $z \leq U(\phi(\mu))$ in (A.1) as $x \geq x_L \equiv x(\mu, U(\phi(\mu)))$. Denote the highest equilibrium value of x as x_H. We can therefore express optimal experimentation as a choice of x instead of z:

$$R_u^e(\mu) \equiv \max_{x \in [x_L, x_H]} \left[\mu\lambda_u x U(\phi(\mu)) + (1-\mu\lambda_u x) U(F(\lambda_u x, \mu))\right]. \tag{A.3}$$

Under the assumptions on U, the objective function in (A.3) is strictly convex in x for all interior μ. If the worker experiments, the optimal choice is either $x = x_L$ or $x = x_H$. However, experimenting with $x = x_L$ cannot be optimal. To see this, suppose that $x = x_L$ achieves the maximum in (A.3). A match success yields $U(\phi(\mu))$ with the experimentation. Since the objective function of experimentation is convex in x, it must be decreasing at x_L. Compare the maximum return on experimentation with the return on not experimenting while searching for the offer $U(\phi(\mu)) + \varepsilon$, where $\varepsilon > 0$ is arbitrarily small. The level of x associated with this offer is $x(\mu, U(\phi(\mu)) + \varepsilon) = x_L - \varepsilon'$, where $\varepsilon' > 0$ is arbitrarily small. Since the objective function of experimentation is decreasing in x at x_L, it is higher at $x_L - \varepsilon'$ than at x_L. In addition, if the worker accepts the match, the worker value increases by $\varepsilon > 0$ relative to the value of rejecting the match. Thus, experimenting with the offer $U(\phi(\mu))$ is dominated by searching for $U(\phi(\mu)) + \varepsilon$ and accepting the match.

The above analysis implies that, if $U(\mu)$ is strictly increasing and convex in μ for all interior μ, then (A.2) can be rewritten as

$$R_u(\mu) \geq \mu\lambda_u x_H U(\phi(\mu)) + (1-\mu\lambda_u x_H) U(F(\lambda_u x_H, \mu)). \tag{A.4}$$

This result is intuitive. The more widely dispersed is the posterior belief, the more information the search outcome contains. If a worker intends to experiment, it is optimal to search for an offer whose outcome will widen the posterior distribution of the belief by the most since U is strictly convex. The belief updated after a match success is independent of where the worker searched, but the belief updated after a match failure falls by more if the offer searched by the worker is lower. Thus, searching for the lowest offer induces the widest dispersion in the posterior belief between the two outcomes of search.

We characterize x_H, the highest value of x in the equilibrium. Let z_L be the offer associated with x_H. For experimentation in submarket z_L to be meaningful, there must be firms that offer z_L in equilibrium. However, if all workers applying for z_L are experimenting, a vacancy created for z_L makes a loss. Thus, a necessary condition for experimentation to occur is that some other workers apply for z_L and do not experiment. These are the workers with the lowest belief $\mu = a_L$. Note that $\phi(a_L) = a_L$ and $F(\lambda_u x, a_L) = a_L$ for all x. Thus, z_L is the solution to the following search problem:

$$R_u(a_L) = \max_z \; a_L \lambda_u x(\lambda_u a_L, z)(z - U_L) + U_L,$$

where U_L solves $(1+r)U_L = u(b) + \sigma R_u(a_L)$. The matching rate x implied by z_L is $x_H = x\left(\lambda_u a_L, z_L\right)$. Note that x_H is the same for all μ.

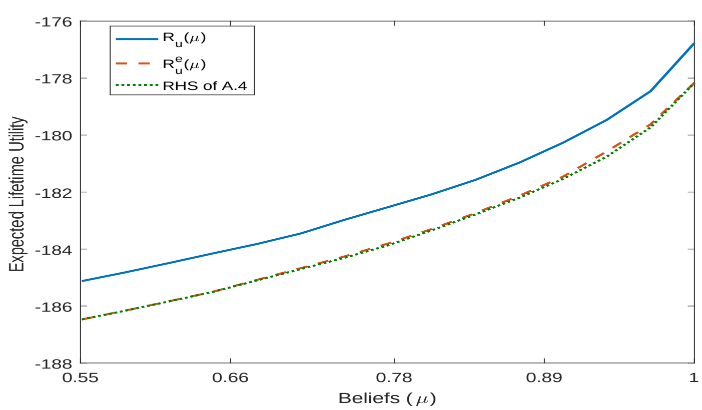


Figure 15. Experimentation is not optimal

With the parameters calibrated in subsection 4.1, Figure 15 depicts $R_u(\mu)$, $R_u^e(\mu)$ and the right-hand side of (A.4). All three functions are convex. Confirming the above analysis,

$R_u^e(\mu)$ is equal to the right-hand side of (A.4). Both functions are lower than $R_u(\mu)$. Thus, experimentation is not optimal under the calibrated baseline parameters.

B. Upper Bounds on the Worker Value

The retention upper bound on the worker value is v^* defined by (3.7) and the recruiting upper bound on the worker value is $\hat{v}$ defined by (3.9). For these upper bounds to be well-defined, we prove the following lemma:

Lemma B.1. *$J(\mu, v)$ decreases in v for any given μ. In the equilibrium, $J(\mu, v) \geq 0$.*

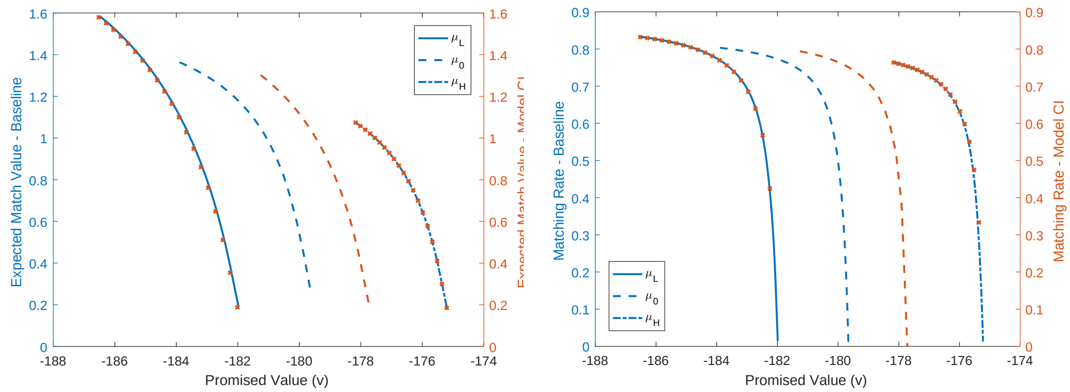
Proof. If $J(\mu, v)$ is strictly increasing in v in some interval $[v_1, v_2]$, then a firm promising a value lower than v_2 in this interval can increase both expected profit and the worker value by offering v_2 instead. This contradicts the optimality of the contract that offers v_2. We prove $J(\mu, v) \geq 0$ by induction. Note first that every submarket with $x > 0$ has $J \geq k/\bar{q} > 0$. That is, a firm's value at the beginning of employing a worker is strictly positive. Suppose that the firm value has been non-negative up to a period at end of which the worker's state is (μ_0, v_0). In any future period $i \geq 1$, if the worker has stayed with the firm up to i, the worker's state will be $\left(\mu_{+(i-1)}, v_{+(i-1)}\right)$. Set $w_{+i} = y$ and let v_{+i} satisfy the promise-keeping constraint: $(1+r)\, v_{+(i-1)} = u(y) + \sigma\left[(1-\delta)\, R_e\left(\mu_{+(i-1)}, v_{+i}\right) + \delta U\left(\mu_{+(i-1)}\right)\right]$. For any given (μ_0, v_0), this path of contracts, $\{(w_{+i}, v_{+i})\}_{i\geq 1}$, is feasible and yields zero value to the firm. Since $J(\mu_0, v_0)$ is the value of the firm under optimal contracts, then $J(\mu_0, v_0) \geq 0$. **QED**

The firm value may reach zero as the belief deteriorates while a worker stays with a firm. Thus, $\bar{v}(\mu_{+1})$ defined by (3.6), is the maximum continuation value that a firm can offer to a job stayer whose belief will be μ_{+1} after failing to match with another firm. By (3.7), we express this retention upper bound as a function of the current belief, $v^*(\mu) = \bar{v}\left(\mu_{+1}^*(\mu)\right)$, where $\mu_{+1}^*(\mu)$ is the future belief consistent with the continuation value $\bar{v}(\mu_{+1})$. However, $v^*(\mu)$ is not the maximum value that a recruiting firm can offer to an applicant with belief μ. Relative to a firm that already has a worker, a recruiting firm has the disadvantage of having to incur the vacancy cost but also has the advantage of employing a worker with an improved belief. When an applicant with belief μ is recruited by a new firm, the belief improves to $\phi(\mu)$, but if the worker stays with the incumbent firm, the belief deteriorates to $\mu_{+1} = F(\lambda_e x_e, \mu)$. The highest value that a recruiting firm can offer to a worker of belief μ is $\hat{v}(\mu)$ defined by (3.9). If k is sufficiently small, then $\hat{v}(\mu) > v^*(\mu)$; otherwise, $\hat{v}(\mu) \leq v^*(\mu)$. Figure 4 depicts $\hat{v}(\mu)$ and $v^*(\mu)$ with the baseline parameters in subsection 4.1. It shows that $\hat{v}(\mu) > v^*(\mu)$ if and only if $\mu_L < \mu < \mu_H$.

The unified upper bound on the worker value of belief μ is $\max\{v^*(\mu), \hat{v}(\mu)\}$ in the equilibrium. If $\hat{v}(\mu) > v^*(\mu)$, clearly $\hat{v}(\mu)$ is the upper bound because some recruiting firms offer $\hat{v}(\mu)$. If $\hat{v}(\mu) < v^*(\mu)$, the retention upper bound $v^*(\mu)$ can still be reached because of the incentive to backload wages. For example, consider an employed worker whose current belief is μ' and let the belief after failing to find a new match be updated to $\mu = \mu^*_{+1}(\mu')$, where μ^*_{+1} is defined in the text following (3.7). To backload wages, the firm may offer a feasible continuation value $v^*(\mu') = \bar{v}(\mu)$ even if $\bar{v}(\mu) > \hat{v}(\mu)$. This continuation value is binding when the worker stays with the firm.

C. Additional Comparisons between the Baseline and Model CI

The firm value function and the matching rate function: For the two models, Figure 16 depicts $J(\mu, v)$ in the left panel and $x(\mu, z)$ in the right panel for three given values of μ. The firm value is lower in the baseline than in model CI for all interior (μ, v). In the baseline, a firm expects the belief to deteriorate when the worker stays with the firm. This reduction in the belief increases the cost of delivering the promised value to the worker, as explained in subsection 4.2, and reduces the firm's present value. In model CI, this effect of μ does not exist. This difference between the two models vanishes when μ is at either end, a_L or a_H, at which point the belief ceases to update. The function $x(\mu, z)$ inherits the features of $J(\mu, z)$ because x is determined by J through competitive entry of vacancies (see (3.8)). For any interior (μ, z), the incentive to create vacancies in submarket (μ, z) is weaker in the baseline than in model CI, which causes $x(\mu, z)$ to be lower in the baseline. Note that for intermediate values of μ, the equilibrium support of v is lower in the baseline than in model CI.



$J(\mu, \bullet)$ $x(\mu, \bullet)$

Figure 16. The baseline and model CI: $J(\mu, v)$ and $x(\mu, z)$

Employed workers' matching rate per efficiency search unit: $x_e(\mu, v_{+1}) = x(\mu, g_e(\mu, v_{+1}))$, where $g_e(\mu, v_{+1})$ is the optimal search target. Figure 17 depicts $x_e(\mu, v_{+1})$ for three given levels of μ. Again, the main difference between the baseline and model CI occurs when the belief is intermediate. In this case, $x_e(\mu, \bullet)$ has the same shape in the two models but the support is lower in the baseline than in model CI.[30]

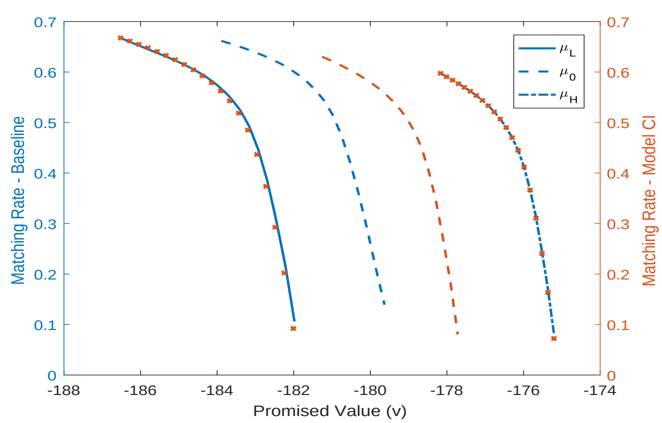


Figure 17. The baseline and model CI: $x_e(\mu, \bullet)$

Unemployed workers' optimal search targets: $g_u(\mu)$. Subsection 4.5 (Figure 11) shows that the policy function $g_u(\mu)$ is lower in the baseline than in model CI for all interior μ, which seems to contradict the analysis in subsection 4.3 that learning motivates

[30] To economize on space, we do not depict the expected matching probability, $\mu\lambda_e x_e$, which has similar properties to x_e.

a worker to search for higher offers initially. To resolve this puzzle, we note that $U(\mu)$ and $x(\mu,z)$ are different in equilibrium between the two models. The function $U(\mu)$ determines the option value of remaining unemployed and the function $x(\mu,z)$ affects the job-finding probability, and therefore both affect the optimal search target of an unemployed worker. To separate these equilibrium effects from the effect of learning, we decompose the change from the baseline to model CI in the following steps. First, we isolate the effect of learning by keeping $U(\mu)$ and $x(\mu,z)$ as in the baseline but assuming that the belief does not update after a match failure. That is, suppose that the belief μ of an unemployed worker will not fall to $F(\lambda_u x,\mu)$ as in (3.3). This counterfactual policy function solves:

$$g_u^{CF_1}(\mu) \equiv \arg\max_{z\in \;\mu} \left\{\mu\lambda_u x(\mu,z)\left[z-U(\mu)\right]+U(\mu)\right\}.$$

Second, we capture the additional effect of the function U by assuming that the future value function for unemployed workers is the one in model CI, $U^{CI}(\mu)$. This second counterfactual policy function solves:

$$g_u^{CF_2}(\mu) \equiv \arg\max_{z\in \;\mu} \left\{\mu\lambda_u x(\mu,z)\left[z-U^{CI}(\mu)\right]+U^{CI}(\mu)\right\}.$$

Finally, by changing $x(\mu,z)$ to $x^{CI}(\mu,z)$, we capture the change from $g_u^{CF_2}(\mu)$ to $g_u^{CI}(\mu)$ resulting from the additional equilibrium effect through the matching rate function $x(\mu,z)$.

The counterfactual policy functions $g_u^{CF_1}(\mu)$ and $g_u^{CF_2}(\mu)$ are depicted in Figure 18. For all interior μ, $g_u^{CF_1}(\mu) < g_u(\mu)$, although this difference is small. Thus, learning indeed motivates an unemployed worker to increase the search target. However, this effect of learning on the search target is reversed by the equilibrium effect through the future value function of unemployed workers; i.e., $g_u^{CF_2}(\mu) > g_u(\mu)$ for all interior μ. Since $U^{CI}(\mu) > U(\mu)$ for all interior μ, an unemployed worker has a higher option value in model CI than in the baseline, which motivates the worker to search for a higher offer in model CI. Moreover, $g_u^{CI}(\mu) - g_u^{CF_2}(\mu) > 0$ for all interior μ, and this difference is larger than $[g_u^{CF_2}(\mu) - g_u^{CF_1}(\mu)]$ and $[g_u(\mu) - g_u^{CF_1}(\mu)]$. Thus, the equilibrium effect through the matching rate function $x(\mu,z)$ is larger than the effect of learning and the effect through the function U. As depicted in Figure 16, $x^{CI}(\mu,z) > x(\mu,z)$ for all interior (μ,z). With the higher matching rate in model CI, the optimal tradeoff for an unemployed worker is to search for a higher offer than in the baseline. Therefore, the equilibrium effect of competitive entry of vacancies is critical for the quantitative analysis.

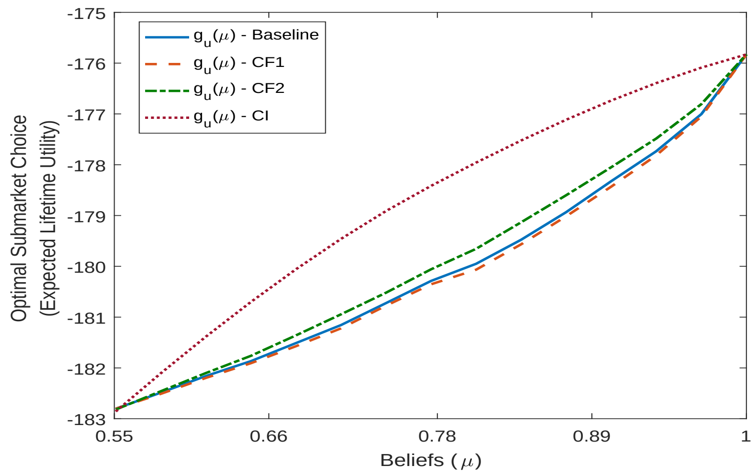


Figure 18. Unemployed workers' optimal search and decomposition

Unemployed workers' reemployment wages: $w(\phi(\mu), g_u(\mu))$, where $\phi(\mu)$ is the updated belief of the worker at reemployment and $g_u(\mu)$ is the optimal search target. The belief $\phi(\mu)$ and the optimal search target $g_u(\mu)$ exert opposite effects on the reemployment wage. By increasing the posterior, a higher belief μ increases the room for a firm to backload wages for the reemployed worker, which reduces the starting wage. On the other hand, by increasing the search target, a higher belief increases the starting wage that is required for delivering the promised value. Figure 19 depicts the reemployment wage. In the baseline, when the belief increases, the reemployment wage first decreases and then increases. That is, for low values of μ, the effect of a higher μ on backloading wages dominates but, for high values of μ, the effect through the search target dominates. In model CI, in contrast, the effect of a higher μ on backloading wages dominates for all values of μ, and so the reemployment wage increases in μ.

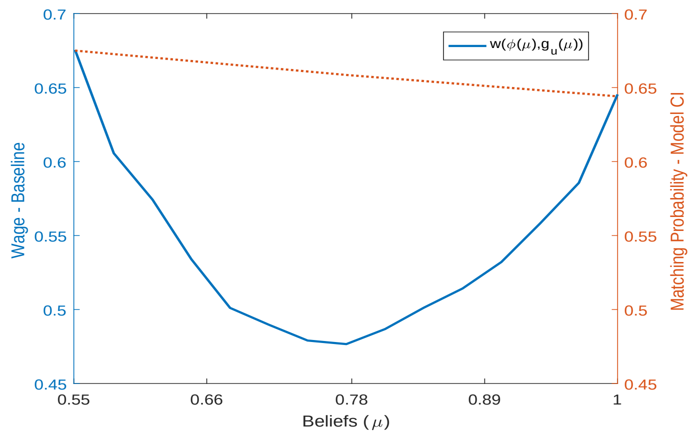


Figure 19. The baseline and model CI: reemployment wages

Employed workers' return and policy functions: For any given v_{+1}, the contrasts between the two models in $R_e(\mu, v_{+1})$ are similar to those in $U_u(\mu)$, and the contrasts in g_e are similar to those in $g_u(\mu)$ analyzed above.

References

[1] Acemoglu, D. and D. Autor, 2011, "Skills, tasks and technologies: implications for employment and earnings," Handbook of Labor Economics 4, 1043-1171.

[2] Burdett, K. and M. Coles, 2003, "Equilibrium wage-tenure contracts," Econometrica 71, 1377-1404.

[3] Burdett, K. and D.T. Mortensen, 1998, "Wage differentials, employer size and unemployment," International Economic Review 39, 257-273.

[4] Burdett, K. and T. Vishwanath, 1988, "Declining reservation wages and learning," Review of Economic Studies 55, 655-665.

[5] Burdett, K., Shi, S. and R. Wright, 2001, "Pricing and Matching with Frictions," Journal of Political Economy 109, 1060-1085.

[6] Chaumont, G. and S. Shi, 2017, "Wealth accumulation, on the job search and inequality," SSRN working paper (http://ssrn.com/abstract=2893777).

[7] Couch, K.A. and D.W. Placzek, 2010, "Earnings losses of displaced workers revisited," American Economic Review 100, 572-589.

[8] Dagum, C., 1975, "A model of income distribution and the conditions of existence of moments of finite order, " Bulletin of the International Statistical Institute 46, 199-205.

[9] Davis, S.J. and T. von Wachter, 2011, "Recessions and the costs of job loss," Brookings Papers on Economic Activity.

[10] Delacroix, A. and S. Shi, 2006, "Directed search on the job and the wage ladder," International Economic Review 47, 651-699.

[11] Doppelt, R., 2015, "The hazards of unemployment: A macroeconomic model of job search and resume dynamics," manuscript, Penn State University.

[12] Faberman, J.R. and M. Kudlyak, 2017, "The intensity of job search and search duration," working paper, Federal Reserve Bank of San Francisco.

[13] Fallick, B. and C. Fleischman, 2004, "Employer-to-employer flows in the U.S. labor market: The complete picture of gross worker flows," Manuscript, Federal Reserve Board.

[14] Farber, H.S., Silverman, D. and T. von Wachter, 2015, "Factors determining callbacks to job applications by the unemployed: an audit study," NBER Working Paper 21689.

[15] Gonzalez, F.M. and S. Shi, 2010, "An equilibrium theory of learning, search and wages," Econometrica 78, 468-510.

[16] Heckman, J.T. and G.J. Borjas, 1980, "Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence," Economica, 47, 247-283.

[17] Hornstein, A., Krusell, P. and G.L. Violante, 2011, "Frictional wage dispersion in search models: A quantitative assessment," American Economic Review 101, 2873-2898.

[18] Jovanovic, B., 1979, "Job matching and the theory of turnover," Journal of Political Economy 87, 972-990.

[19] Julien, B., Kennes, J. and I. King, 2000, "Bidding for Labor," Review of Economic Dynamics 3, 619-649.

[20] Kroft, K., Lange, F. and M.J. Notowidigdo, 2013, "Duration dependence and labor market conditions: evidence from a field experiment," Quarterly Journal of Economics 129, 1123-1167.

[21] Kudlyak, M., Lkhagvasuren, D. and R. Sysuyev, 2013, "Systematic job search: new evidence from individual job application data," working paper 12-03R, Federal Reserve Bank of Richmond.

[22] Lazear, E., 2009, "Firm-specific human capital: a skill-weights approach," Journal of Political Economy 117, 914-40.

[23] Lise, J. and J-M. Robin, 2017, "The macrodynamics of sorting between workers and firms," American Economic Review 107, 1104-1135.

[24] Marinescu, I. and R. Wolthoff, 2015, "Opening the black box of the matching function: the power of words," manuscript, University of Toronto.

[25] Menzio, G. and S. Shi, 2010, "Block recursive equilibria for stochastic models of search on the job," Journal of Economic Theory 145, 1453-1494.

[26] Menzio, G. and S. Shi, 2011, "Efficient search on the job and the business cycle," Journal of Political Economy 119, 468-510.

[27] Moen, E.R., 1997, "Competitive search equilibrium," Journal of Political Economy 105, 385-411.

[28] Montgomery, J.D., 1991, "Equilibrium wage dispersion and interindustry wage differentials," Quarterly Journal of Economics 106, 163-179.

[29] Mortensen, D.T. and C.A. Pissarides, 1994, "Job creation and job destruction in the theory of unemployment," Review of Economic Studies 61, 397-415.

[30] Peters, M., 1991, "Ex Ante Price Offers in Matching Games: Non-Steady State," Econometrica 59, 1425-1454.

[31] Postel-Vinay, F. and J-M Robin, 2002, "Equilibrium wage dispersion with worker and employer heterogeneity," Econometrica 70, 2295-2350.

[32] Shi, S., 2001, "Frictional assignment i: efficiency," Journal of Economic Theory 98, 232-260.

[33] Shi, S., 2009, "Directed search for equilibrium wage-tenure contracts," Econometrica 77, 561-584.

[34] Shi, S., 2016, "Efficient job upgrading, search on the job and output dispersion," SSRN working paper at http://ssrn.com/abstract=2729019.

[35] Topel, R.H. and M.P. Ward, 1992, "Job mobility and the careers of young men," Quarterly Journal of Economics 107, 439-479.

[36] Tsuyuhara, K. 2016, "Dynamic contracts with worker mobility via directed on-the-job search," International Economic Review 57, 1405-1424.

[37] Wolthoff, R., 2017, "Applications and interviews: firms' recruiting decisions in a frictional labor market," The Review of Economic Studies, forthcoming.

Supplementary Appendix of

"Learning, On-the-Job Search
and Wage-Tenure Contracts"

by Kevin Fawcett and Shouyong Shi

2018

This supplementary appendix provides additional analyses and proofs. The sections, figures and tables are numbered sequentially from the appendix to the paper.

E. Optimal Search and Optimal Contracts

The following proposition characterizes some features of an employed worker's optimal search and a firm's optimal contracts:

Proposition E.1. *(i) $x_e(\mu, v_{+1})$ is decreasing in v_{+1}. (ii) Assuming that $x_e(\mu, v_{+1})$ is single valued for each v_{+1}, then $R_e(\mu, v_{+1})$ is differentiable with respect to v_{+1} when $x_e(\mu, v_{+1})$ is interior, and the derivative is $\frac{dR_e(\mu,v_{+1})}{dv_{+1}} = 1 - \mu\lambda_e x_e(\mu, v_{+1}) \geq 0$. (iii) Assuming that the envelope conditions hold for (μ, v) in the firm's problem, (3.4), then $J'_v(\mu, v) = \frac{-1}{u'(w_{+1})}$.*

In addition to the assumptions in (ii) and (iii), assume that $(1 - \mu\lambda_e x_e) J(F(\lambda_e x_e, \mu), v_{+1})$ is decreasing in x_e for any given (μ, v_{+1}). Then, (iv) and (v) hold:
(iv) $w_{+2} \geq w_{+1}$ whenever the optimal choice v_{+1} is interior, i.e., if $v_{+1} < v^*(\mu)$ where $v^*(\mu) = \bar{v}(\mu^*_{+1}(\mu))$ is defined by (3.7). (v) If $v = \bar{v}(\mu)$ in a period, then $w_{+i} = y$, $v_{+i} = v^*(\mu_{+(i-1)})$ and $J(\mu_{+i}, v_{+i}) = 0$ for all $i \geq 1$.

Part (i) is depicted in the right panel of Figure 16 and (iii) in the right panel of Figure 3. Part (ii) is not depicted but intuitive: the return on search increases in the worker's option value v_{+1}. Parts (iv) and (v) require the additional assumption that, for given (μ, v_{+1}), the expected future value of the firm decreases in the employee's matching rate with an outside firm. This intuitive assumption is satisfied in the calibrated model after tenure passes a threshold. Under this additional assumption, part (iv) states that wages are increasing with tenure when the continuation value is below the retention upper bound.

Part (v) states that once the promised value has reached the retention upper bound, it will change along this upper bound in the future as beliefs update, while wages are equal to y and the firm value remains zero. Parts (iv) and (v) together imply that, under the additional assumption, wages do not fall with tenure.

Proof of Proposition E.1:

(i) Consider an employed worker's search decision in (3.1). Inverting the function $x(\mu, z)$ to write z as $z(\mu, x)$, we can express (3.1) as

$$R_e(\mu, v_{+1}) \equiv \max_{x \in X} \{\mu \lambda_e x [z(\mu, x) - v_{+1}] + v_{+1}\},$$

where X is a closed interval. Denote the objective function in the above problem by the temporary notation $f(x, \mu, v_{+1})$. It is easy to verify that $f(x, \mu, v_{+1})$ has strict decreasing differences in (x, v_{+1}). Since (x, v_{+1}) lie in a rectangle, which is a lattice, then f is strictly submodular in (x, v_{+1}) and the optimal choice of x decreases in v_{+1} (see Topkis, 1998).

(ii) If $x_e(\mu, v_{+1})$ is single valued for each v_{+1}, then it is continuous in v_{+1} by the Theorem of the Maximum (see Stokey and Lucas, 1989). For any v_{+1} such that $x_e(\mu, v_{+1})$ is interior, $x_e(\mu, v_{+1} \pm \varepsilon)$ is a feasible choice for x at v_{+1} and $x_e(\mu, v_{+1})$ is feasible for x at $(v_{+1} \pm \varepsilon)$, where $\varepsilon > 0$ is arbitrarily small. Using these features and continuity of $x_e(\mu, v_{+1})$ in v_{+1}, we can compute the one-sided derivatives of $R_e(\mu, v_{+1})$ with respect to v_{+1} and verify that they are both equal to $[1 - \mu \lambda_e x_e(\mu, v_{+1})]$. See Amir et al. (1991) and Gonzalez and Shi (2010, Appendix C).

(iii) From (3.5) we can solve the wage as:

$$w_{+1} = \tilde{w}(\mu, v_{+1}, v) \equiv u^{-1}\left((1+r)v - \sigma\left[(1-\delta)R_e(\mu, v_{+1}) + \delta U(\mu)\right]\right).$$

Substituting this and $\mu_{+1} = F(\lambda_e x_e, \mu)$ into (3.4) yields:

$$\begin{aligned}&(1+r)J(\mu, v) \\ &= \max_{v_{+1}} \left[y - \tilde{w}(\mu, v_{+1}, v) + \sigma(1-\delta)(1 - \mu\lambda_e x_e) J(F(\lambda_e x_e, \mu), v_{+1})\right],\end{aligned} \tag{E.1}$$

where $x_e = x_e(\mu, v_{+1})$. If the envelope condition for v holds, then

$$J'_v(\mu, v) = \frac{-\tilde{w}'_3(\mu, v_{+1}, v)}{1+r} = \frac{-1}{u'(w_{+1})}.$$

(iv) A change in v_{+1} affects the objective function in (E.1) in two ways. One is the direct effect through $\tilde{w}(\mu, v_{+1}, v)$ and $J(F, v_{+1})$. The other is the indirect effect through x_e in the term $(1 - \mu\lambda_e x_e) J(F(\lambda_e x_e, \mu), v_{+1})$. Under the hypothesis in (ii), $\tilde{w}_2'(\mu, v_{+1}, v) = -\sigma(1-\delta)(1-\mu\lambda_e x_e)$. Under the hypothesis in (iii), $J(F, v_{+1})$ is differentiable in both arguments. In particular, $J_2'(F, v_{+1}) = \frac{-1}{u'(w_{+2})}$. Thus, the objective function in (E.1) is differentiable with respect to v_{+1} for any given x_e, and the derivative is:

$$\sigma(1-\delta)(1-\mu\lambda_e x_e)\left[\frac{1}{u'(w_{+1})} - \frac{1}{u'(w_{+2})}\right].$$

This derivative is strictly positive if and only if $w_{+1} > w_{+2}$. Under the assumption that the term $(1 - \mu\lambda_e x_e) J(F(\lambda_e x_e, \mu), v_{+1})$ is decreasing in x_e, the term is increasing in v_{+1}, because $x_e(\mu, v_{+1})$ is decreasing in v_{+1} (see (i)). If $w_{+1} > w_{+2}$, an increase in v_{+1} strictly increases the objective function in (E.1). In this case, the choice v_{+1} is optimal only if it is at the upper corner, $v^*(\mu) = \bar{v}\left(\mu_{+1}^*\right)$, where $\mu_{+1}^* = F(\lambda_e x_e, \mu)$. If the optimal choice is $v_{+1} < v^*(\mu)$, then the benefit of increasing v_{+1} must be zero, which requires $w_{+2} \geq w_{+1}$. The strict inequality $w_{+2} > w_{+1}$ holds if $(1 - \mu\lambda_e x_e) J(F(\lambda_e x_e, \mu), v_{+1})$ is strictly decreasing in x_e and $x_e(\mu, x_{+1})$ is strictly decreasing in v_{+1}.

(v) Suppose $v = \bar{v}(\mu)$ in a period. Note first that $J(\mu, v) \geq 0$ for all equilibrium (μ, v) (see Lemma B.1). We prove the result that, for all integers $i \geq 1$, either $(J_{+i} = 0,$ $v_{+i} = \bar{v}(\mu_{+i}),\ w_{+(i+1)} \geq y)$ or $\left(J_{+i} > 0,\ v_{+i} < \bar{v}(\mu_{+i}),\ w_{+(i+1)} > w_{+i} \geq y\right)$. The proof is by induction. First, we prove that the result holds for $i = 1$. Since $J_{+1} \geq 0$, the Bellman equation for $J(\mu, \bar{v}(\mu)) = 0$ implies $w_{+1} \geq y$, where the inequality is strict if $J_{+1} > 0$. If $J_{+1} = 0$, then $v_{+1} = \bar{v}(\mu_{+1})$ by the definition of $\bar{v}$. Since $J_{+2} \geq 0$, the Bellman equation for $J_{+1}\,(=0)$ implies $w_{+2} \geq y$. If $J_{+1} > 0$, then $v_{+1} < \bar{v}(\mu)$. For this interior v_{+1} to be optimal, it must be the case that $w_{+2} > w_{+1}\ (\geq y)$ (see (iv)). Next, supposing that the stated result holds for an arbitrary $i \geq 1$, we prove that the result holds for $(i+1)$. If $J_{+(i+1)} = 0$, the definition of $\bar{v}$ implies $v_{+(i+1)} = \bar{v}\left(\mu_{+(i+1)}\right)$. Since $J_{+(i+2)} \geq 0$, the Bellman equation for $J_{+(i+1)}\ (=0)$ implies $w_{+(i+2)} \geq y$. If $J_{+(i+1)} > 0$, then $v_{+(i+1)} < \bar{v}\left(\mu_{+(i+1)}\right)$. For this interior choice of $v_{+(i+1)}$ to be optimal, it must be the case that $w_{+(i+2)} > w_{+(i+1)} \geq y$.

The above result implies that if $v = \bar{v}(\mu)$, then $w_{+i} \geq y$ for all $i \geq 0$. Because a firm's value is the sum of discounted profits, then $J(\mu, \bar{v}(\mu)) \leq 0$. For $J(\mu, \bar{v}(\mu)) = 0$ to hold as is defined for $\bar{v}(\mu)$, it must be true that $w_{+i} = y$ for all i. This requires $J_{+i} = 0$ for all i and, hence, $v_{+i} = \bar{v}(\mu_{+i})$ for all i. By the definition of v^* in (3.7), $v_{+i} = v^*\left(\mu_{+(i-1)}\right)$ in this case for all $i \geq 1$. **QED**

F. Estimation of Parameters and Computation Algorithm

We denote the set of parameters to be calibrated by $\eta = (A, \rho, a_L, p_0)$ and denote the empirical moments to be targeted in calibration by $m(\eta)$. We separate η into two sets of parameters $\eta = (\eta_1, \eta_2)$ and the moments $m(\eta)$ into two sets of moments $m(\eta) = (m_1(\eta), m_2(\eta))$ where the parameters η_1 are critical for determining the moments $m_1(\eta)$ and the parameters η_2 are critical for determining the moments $m_2(\eta)$. We choose $\eta_1 = (A, \rho)$, $\eta_2 = (a_L, p_0)$, the moments $m_1(\eta)$ to be the unemployment rate (6.0%) and the monthly job-to-job transition rate (2.2%), and the moments $m_2(\eta)$ to be the average wage loss upon reemployment relative to pre-displacement wages (32%), and the average wage loss five years after reemployment (12%).

We choose the parameters in η to solve the following two-stage minimization problem, where we first solve:

$$\begin{aligned} \hat{\eta}_1(\tilde{\eta}_2) &= \arg\min_{\tilde{\eta}_1} SSE_1(\tilde{\eta}_1, \tilde{\eta}_2) \\ &= \arg\min_{\tilde{\eta}_1} \left[m_1(\tilde{\eta}_1, \tilde{\eta}_2) - m_1(\eta_1, \eta_2)\right]^2 \end{aligned} \tag{F.1}$$

for all $\tilde{\eta}_2$. Given the correspondence $\hat{\eta}_1(\tilde{\eta}_2)$, we then solve the following problem:

$$\begin{aligned} \hat{\eta}_2 &= \arg\min_{\tilde{\eta}_2} SSE_2(\hat{\eta}_1(\tilde{\eta}_2), \tilde{\eta}_2) \\ &= \arg\min_{\tilde{\eta}_2} \alpha \left[m_1(\hat{\eta}_1(\tilde{\eta}_2), \tilde{\eta}_2) - m_1(\eta_1, \eta_2)\right]^2 \\ &+ (1-\alpha)\left[m_2(\hat{\eta}_1(\tilde{\eta}_2), \tilde{\eta}_2) - m_2(\eta_1, \eta_2)\right]^2 \end{aligned} \tag{F.2}$$

where $\alpha_j \in [0, 1]$ is the weight placed on moments $m_j(\eta)$ for $j = \{1, 2\}$ in the second stage of the calibration procedure where $\alpha_1 + \alpha_2 = 1$.[31] In (F.1), for each $\tilde{\eta}_2$, we choose the parameters $\tilde{\eta}_1$ that minimize the sum of squared errors between the model predicted moments $m_1(\tilde{\eta}_1, \tilde{\eta}_2)$ and the target moments $m_1(\eta_1, \eta_2)$. We denote the solutions to (F.1) by $\hat{\eta}_1(\tilde{\eta}_2)$, and given these solutions we choose the set of parameters $\tilde{\eta}_2$ to solve (F.2)

[31] We choose $\alpha_1 = \alpha_2 = 0.5$ and find that parameter estimates are not overly sensitive to this choice given the two stage proceduere we employ.

by minimizing the weighted sum of squared errors between the model predicted moments $m_j(\hat{\eta}_1(\tilde{\eta}_2), \tilde{\eta}_2)$ and the corresponding moments in the data $m_j(\eta_1, \eta_2)$.[32]

The two-step minimization problem places additional importance on matching the moments in $m_1(\eta)$, and as a result, places additional importance on estimating the parameters in η_1. We do this for three reasons. First, the moments $m_1(\eta)$ have clear targets in the data and we choose these moments to be consistent with the majority of the search literature. Conversely, there is some disagreement over estimates of the wage losses at different points in time following displacement. We use Davis and von Wachter (2010) as a guide and choose values that are consistent with the majority of the empirical literature. However, since there is disagreement in the literature with respect to the magnitude of estimates, we de-emphasize the model's ability to match the moments $m_2(\eta)$.[33] Second, we believe it is critical for the model to be able to match the moments in $m_1(\eta)$, as these are fundamental features of the labor market and which have a significant effect on the equilibrium policy and value functions in the model, especially as they relate to worker transition and wage dynamics. Furthermore, these are standard moments targeted in the search literature and we feel it is critical to match these targets to make comparisons with other work. Finally, while the model is able to match the moments in $m_1(\eta)$ quite well, but it is unable to match the moments in $m_2(\eta)$ as well. As a result, if the model is calibrated to minimize $[m(\tilde{\eta}) - m(\eta)]^2$ instead of using the two-step process described above, sufficient emphasis is placed on the model's ability to match $m_2(\eta)$, which results in a significant loss in the model's ability to match $m_1(\eta)$. However, since $[m_2(\tilde{\eta}) - m_2(\eta)]^2$ is relatively large for all $\tilde{\eta}$, we believe we should match $m_1(\eta)$ as well as possible.

The two-step minimization problem described by (F.1) and (F.2) is solved over a grid of η. The solution is $\eta = (0.976, 0.90, 0.551, 0.655)$. With these parameters, the baseline model generates an equilibrium unemployment rate of 6.10%, a monthly job-to-job transition rate of 2.18%, average initial wage losses equal to approximately 44%, and average wage losses after 5 years equal to approximately 2.31%.

[32] We can set $\alpha = 0$ so that all weight is placed on the second set of moments that were not targeted in the first step.

[33] See Couch and Placzek (2010) for a review of the empirical literature on earnings losses from displacement.

Using the calibrated parameters, we compute the unemployment rate, the job-to-job transition rate, and average wage losses following displacement in the baseline and the two nested models provided that these moments are applicable. The results are listed in Table 5. In Model CI, the unemployment rate is approximately equal to the unemployment generated by the baseline, while the job-to-job transition rate is approximately one half of that generated in the baseline. Model NS has no job-to-job transition by construction. The baseline's better match with the data on the job-to-job transitions rate than in the two nested models is an important prediction that drives many of the results. While for the optimal parameters model CI is able to match initial wage losses well, it is unable to generate persistence in wage losses as workers transition to the highest wage quickly as previously discussed.[34]

Table 5. Simulated moments

Moment	Target	Baseline model	Complete information (CI)	No on-the-job search (NS)
u	6.00%	6.10%	6.10%	5.37%
EE rate	2.20%	2.18%	1.16%	-
Wage $\text{Loss}_{t=0}$	32%	44%	32.5%	1.24%
Wage $\text{Loss}_{t=5}$	12%	2.31%	0.00%	1.24%

Computation of the model relies on the block-recursive formulation of the equilibrium in definition 3.1. The primary difficulty in computing the baseline model is to identify the upper-bound functions $\overline{v}$ and $\hat{v}$, which determine the set of equilibrium offers for each belief. The equilibrium set of offers varies with beliefs and depends on the equilibrium policy and value functions. As a result, the upper-bound functions must be determined endogenously. We start by noting that for beliefs close to the true types, equilibrium policy and value functions are similar for the baseline model and model CI as beliefs no longer evolve when they are close to the bounds in the baseline. As a result, we start by computing model CI to identify the values of $\overline{v}$ and $\hat{v}$ at a_L and a_H. We then use these values to make an initial guess at the functions $\overline{v}$ and $\hat{v}$ for the baseline, which determine

[34]When model CI is recalibrated to match the target unemloyment and job-to-job transition rates, it is only able to generate an EE rate of 1.24, which is still significantly smaller than that in the baseline. This results from the fact that workers quickly reach the highest worker value at which point transitions no longer take place. This emphasizes the main result of our paper, that workers continue to transition near the highest worker value through an evolution of the belief.

upper-bound of the the domain of offers in equilibrium.

We describe the computation algorithm in detail below. We compute model CI by initially guessing the worker's value function $U(\mu)$ and iterate over this value function until convergence. Given that only unemployed workers that started the period in unemployment are able to search, the upper-bound of the offer space can be calculated given our choice of the value function $U(\mu)$. In the description of the computation procedure below, we use the superscripts CI and BL to differentiate between model CI and the baseline model respectively.

A. Create a grid of parameters $\eta = (\rho, A, a_L, p_0)$. We start with a rough grid of parameters to determine the general region where the solution to the problem given by (F.1) and (F.2) lies. We note that for extreme values of a_L and p_0 the learning process is extreme and the model takes significantly longer to compute. To avoid this, we focus on values of these parameters that are intermediate. Starting with the first set of parameters in the grid, we proceed to step B.

B. Computation of Model CI

1. Discretize belief space. Set the discretized belief space: $M^{CI} = \{a_L, \mu_0, a_H\}$.

2. Choose initial values for the unemployed value function $U^{CI}(\mu)$ defined on the grid M^{CI}. We start by choosing values that are increasing in μ. Set the lower-bound of offer space $\underline{z}_{\mu}^{CI} = U^{CI}(\mu)$ for every $\mu \in M^{CI}$.

a. Calculate the upper-bound function $\overline{v}^{CI}(\mu) = \frac{u(y,\gamma)+\sigma\delta U^{CI}(\mu)}{(1+r)-\sigma(1-\delta)}$ from the unemployed Bellman equation given by (3.2). Set the upper-bound of the offer space $\overline{z}_{\mu}^{CI} = \overline{v}^{CI}(\mu)$ for every $\mu \in M^{CI}$.

b. Discretize offer space. For every $\mu \in M^{CI}$, create a grid of $n_z > 0$ points in the range $\left[\underline{z}_{\mu}^{CI}, \overline{z}_{\mu}^{CI}\right]$ to obtain the set of offers Z_{μ}^{CI}. For every $\mu \in M^{CI}$, create a grid of $n_v > 0$ points in the range $\left[\underline{z}_{\mu}^{CI}, \overline{z}_{\mu}^{CI}\right]$ to obtain the set of states V_{μ}^{CI} where $n_z >> n_v$.

c. Choose initial values for the firm's value function $J^{CI}(\mu, v)$ defined on the grid $M^{CI} \times V^{CI}$. Initially use choose values that are increasing in μ and decreasing in v.

d. Interpolate the value function $J^{CI}(\mu, v)$ defined on the grid $M^{CI} \times V^{CI}$ onto $M^{CI} \times Z^{CI}$ to get $J^{CI}(\mu, z)$. Given the function $J^{CI}(\mu, z)$, compute the matching rate function $x^{CI}(\mu, z)$ from condition (3.8) for vacancy creation.

i. Given $x^{CI}(\mu, z)$, solve the employed worker's problem given by (3.1) to get the worker's policy functions $x_e^{CI}(\mu, v_{+1})$ and $g_e^{CI}(u, v_{+1})$ and the return on search $R_e^{CI}(\mu, v_{+1})$ defined on $M^{CI} \times V^{CI}$.

ii. Given the solutions to the employed worker's problem and the value function $U^{CI}(\mu)$, solve the firm's optimal contracting problem given by (3.4) and (3.5). Compute the value function implied by this solution, $TJ^{CI}(\mu, v)$.

iii. If the value function implied by the solution in step (ii) above is not sufficiently close to the initial value of $J^{CI}(\mu, v)$, then repeat step (d) using $TJ^{CI}(\mu, v)$ as the firm's value function to compute $x^{CI}(\mu, v)$. If the value function implied by step (ii) is sufficiently close to the value function $J^{CI}(\mu, v)$, then proceed to step (e). Specifically, for some norm $\|\cdot\|$ and some arbitrarily small $\epsilon_J > 0$, if $\left\|TJ^{CI} - J^{CI}\right\| > \epsilon_J$ then repeat step (e) using TJ^{CI} as the firm's value function, and if $\left\|TJ^{CI} - J^{CI}\right\| \leq \epsilon_J$ then proceed to step (e) using the matching rate function computed from TJ^{CI}.

e. Given the matching rate function computed in (d), solve the unemployed worker's problem given by (3.3) to get the return on search $R_u^{CI}(\mu)$.

f. Compute the updated unemployed worker's value function $TU^{CI}(\mu)$ using the solution from step (e) and the unemployed Bellman equation given by (3.2). If this value function is not sufficiently close to the initial value function $U^{CI}(\mu)$, then repeat step (2) using $TU^{CI}(\mu)$ as the new unemployed value function. If this value function is sufficiently close to the initial guess of U^{CI}, then proceed to step (3). Specifically, for some for some norm $\|\cdot\|$ and some arbitrarily small $\epsilon_\mu > 0$, if $\left\|TU^{CI} - U^{CI}\right\| > \epsilon_\mu$ then repeat step (2) using TU^{CI} as the unemployed worker's value function, and if $\left\|TU^{CI} - U^{CI}\right\| \leq \epsilon_\mu$ then proceed to step (3).

3. Store relevant policy and value functions for model CI.

C. Computation of the Baseline Model:

1. Discretize belief space. For some arbitrarily small $\epsilon > 0$, create an even grid of $n_\mu > 3$ points in the range $[a_L + \epsilon, a_H - \epsilon]$ to obtain the set of beliefs M^{BL}.

2. Choose bounds of offer space and worker value space. We must first choose the lower and upper bounds of the offer space for each μ, denoted by $\underline{z}_\mu^{BL}$ and $\overline{z}_\mu^{BL}$ respectively.

a. Choose upper-bound of offer space $\overline{z}_\mu^{BL}$. In model CI, $\overline{v}^{CI}(\mu) > \hat{v}^{CI}(\mu)$ for all μ. We use $\overline{v}^{CI}$ to determine the upper-bound of the offer-space for the baseline. We interpolate the function $\overline{v}^{CI}(\mu)$ defined on M^{CI} for all beliefs in M^{BL} to obtain the function $\overline{v}^{CI^*}(\mu)$. We use this interpolated function as the initial values for the upper-bound of the offer space $\overline{z}_\mu^{BL} = \overline{v}^{CI^*}(\mu)$. In subsequent steps, $\overline{z}_\mu^{BL}$ is determined endogenously to be $\overline{z}_\mu^{BL} = \max\{\hat{v}^{BL}(\mu), v^{*BL}(\mu)\}$. In the equilibrium, $\overline{v}^{CI^*}(\mu) > \max\{\hat{v}^{BL}(\mu), v^{*BL}(\mu)\}$ for all $\mu \in M^{BL}$, and as a result, this choice of the upper-bound of the offer space is suitable for the first iteration as it does not exclude any equilibrium offers in the baseline.

b. Choose lower-bound of offer space $\underline{z}_\mu^{BL}$. Interpolate $U^{CI}(\mu)$ defined on M^{CI} for all beliefs in M^{BL} to obtain the function $U^{CI^*}(\mu)$. Choose the lower-bound of the offer-space $\underline{z}_\mu^{BL} = U^{CI^*}(\mu)$.

3. Discretize offer space. For every $\mu \in M^{BL}$, create a grid of $n_z > 0$ points in the range $\left[\underline{z}_\mu^{BL}, \overline{z}_\mu^{BL}\right]$ to obtain the set of offers Z_μ^{BL}. For every $\mu \in M^{BL}$, create a grid of $n_v > 0$ points in the range $\left[\underline{z}_\mu^{BL}, \overline{z}_\mu^{BL}\right]$ to obtain the set of states V_μ^{BL} where $n_z >> n_v$

4. Choose initial values for the firm's value function $J^{BL}(\mu, v)$. Interpolate the value function $J^{CI}(\mu, v)$ obtained in part (1) onto the $M^{BL} \times V^{BL}$ grid obtained in the previous steps to obtain the function $J^{CI^*}(\mu, v)$. We use this interpolated function as the initial guess of the firm's value function $J^{BL} = J^{CI^*}$. Compute the upper-bound functions $v^{*BL}(\mu)$ and $\hat{v}^{BL}(\mu)$ implied by the value function $J^{BL}(\mu, v)$ using (3.7) and (3.9) respectively.

a. Interpolate the value function $J^{BL}(\mu, v)$ defined over $M^{BL} \times V^{BL}$ onto $M^{BL} \times Z^{BL}$ to get the value function $J^{BL}(\mu, z)$. Given the value function $J^{BL}(\mu, z)$, compute the matching rate function $x^{BL}(\mu, z)$ from (3.8) for vacancy creation.

b. Choose initial values for the unemployed value function $U^{BL}(\mu)$. Use the interpolated value function $U^{CI^*}(\mu)$ defined on M^{BL}.

  i. Given the value function $U^{BL}(\mu)$ and matching rate function $x^{BL}(\mu, z)$, solve the unemployed worker's problem given by (3.3) to obtain $R_u^{BL}(\mu)$.

  ii. Compute the updated unemployed value function $TU^{BL}(\mu)$ from the solutions in (b) above and (3.2). If this updated value function is not sufficiently close to the initial guess, then repeat step (b) using $TU^{BL}(\mu)$ as the initial guess of the unemployed value function. If the updated value function is sufficiently close to the initial guess, then proceed to the next step. Specifically, for some norm $\|\cdot\|$ and some arbitrarily small $\epsilon_\mu > 0$, if $\left\|TU^{BL} - U^{BL}\right\| > \epsilon_\mu$ then repeat step (b) using $TU^{BL}(\mu)$ as the initial guess, and if $\left\|TU^{BL} - U^{BL}\right\| \leq \epsilon_\mu$ then proceed to step (c).

c. Given $x^{BL}(\mu, z)$, solve the employed worker's problem given by (3.1) to get the worker's policy functions $x_e^{BL}(\mu, v_{+1})$ and $g_e^{BL}(u, v_{+1})$, and the employed return on search $R_e^{BL}(\mu, v_{+1})$.

d. Given the solutions to the employed worker's problem and the value function $U^{BL}(\mu)$ from step (b), solve the firm's optimal contracting problem given by (3.4) and (3.5). Compute the updated firm value function implied by this solution, $TJ^{BL}(\mu, v)$. If the value function implied by the solution to the firm's problem in the previous step is not sufficiently close to the initial guess of $J^{BL}(\mu, v)$, then repeat steps (3)-(4) using $\underline{z}_\mu^{BL} = TU^{BL}(\mu)$ and $\overline{z}_\mu^{BL} = \max\{v^{*BL}(\mu), \hat{v}^{BL}(\mu)\}$ for every $\mu \in M^{BL}$. Use the updated value function $TJ^{BL}(\mu, v)$ as the initial firm's value function. Specifically, for some norm $\|\cdot\|$ and some arbitrarily small $\epsilon_J > 0$, if $\left\|TJ^{BL} - J^{BL}\right\| > \epsilon_J$

then repeat steps (3)-(4) using TJ^{BL} as the firm's value function, and if $\left\|TJ^{BL} - J^{BL}\right\| \leq \epsilon_J$ then proceed to step (6).

6. Store all policy and value functions for the baseline.

D. Simulate the baseline model:

1. Create a grid of states, tenure, and ability, indexed by (μ, v, t, a), where the v dimension includes an additional grid point for unemployed workers.

a. Create a grid of beliefs with $n_{\mu_2} > n_\mu > 0$ points denoted by M^{BL_2}.

b. For each $\mu \in M^{BL_2}$ create a grid of $n_{z_2} > n_z > 0$ points over the equilibrium range of offers Z_μ^{BL}. For values of μ that are in M^{BL_2} but not in M^{BL}, interpolate $\underline{z}_\mu^{BL}$ and $\overline{z}_\mu^{BL}$ to get the bounds of the offer space to create $Z_\mu^{BL_2}$.

c. Include an additional point for an unemployed worker for each belief in the v dimension of the grid.

d. For each state (μ, v) and unemployment, create a grid of $T \geq 1$ points to store the tenure of each worker in the simulation.

e. For each (μ, v, t) and each unemployed pair (μ, t), create two points: one for low-type workers and one for high-type workers.

f. Each point in the grid is indexed by the state, tenure and ability of the worker (μ, v, t, a) where there are $n_{v_2} + 1$ points in the v dimension to keep track of unemployed workers. Call this grid G.

2. Create an initial distribution Ω_0 of workers over the points in G for some very large number of workers, N. Given N and p_0, place $p_0 N$ and $(1 - p_0)N$ workers in the high and low type grids respectively.

3. For each point in G, determine where a worker will move for each stochastic event in the model using the policy functions and assumptions of the model. That is, create a mapping $\Gamma^{BL} : G \to G'$ where G' is the position of each worker in the following period conditional on the outcome of stochastic events.

a. For any (μ, v, t, a) worker, if the worker is removed from the economy by the exit shock with probability $1-\sigma$, they are replaced in the next period by an unemployed worker with $(\mu_0, t=0, a)$. Note that since the number of workers N is very large, a sufficiently large number of workers exit the economy in each period and we simply replace workers with a worker of the same type for consistency in the number of workers of each type in each iteration.

b. For an employed worker with (μ, v, t, a) that survived the exit shock with probability σ, if they become unemployed with probability δ they begin next period in unemployment with $(\mu, t=0, a)$.

c. For an employed worker with (μ, v, t, a) that survived the exit and separation shocks with probability $\sigma(1-\delta)$, they are optimally offered the continuation value $n^{BL}(u, v)$. The worker optimally searches for the offer $g_e^{BL}(\mu, n^{BL}(u, v))$ with associated matching rate $x_e^{BL}(\mu, n^{BL}(u, v))$.

i. If successful with probability $a\lambda_e x_e^{BL}(\mu, n^{BL}(u, v))$, the (μ, v, t, a) worker begins the next period as a $(\phi(\mu), g_e^{BL}(\mu, n^{BL}(u, v)), t=0, a)$ worker.

ii. If unsuccessful with probability $1 - a\lambda_e x_e^{BL}(\mu, n^{BL}(u, v))$, the (μ, v, t, a) worker begins the next period as a $(F(\lambda_e x_e^{BL}(\mu, n^{BL}(u, v)), \mu), n^{BL}(u, v), t+1, a)$ worker.

d. For an unemployed worker with (μ, t, a) that survived the exit shock and started the period in unemployment, they optimally search for the offer $g_u^{BL}(\mu)$ with associated matching rate $x_u^{BL}(\mu)$.

i. If successful with probability $a\lambda_u x_u^{BL}(\mu)$, the (μ, t, a) unemployed worker begins the next period as a $(\phi(\mu), g_u^{BL}(\mu), t=0, a)$ employed worker.

ii. If unsuccessful with probability $1 - a\lambda_u x_u^{BL}(\mu)$, the (μ, t, a) unemployed worker begins the next period as a $(F(\lambda_u x_u^{BL}(\mu), \mu), t+1, a)$ unemployed worker.

4. Simulate the economy by drawing stochastic events from the appropriate distributions and placing individuals from the distribution Ω_t into the new distribution

Ω_{t+1} according to the mapping Γ^{BL} defined in step (3). Store all transitions and wage changes from Ω_t to Ω_{t+1}.

5. For some norm $\|\cdot\|$ and some arbitrarily small $\epsilon > 0$, if $\|\Omega_{t+1} - \Omega_t\| > \epsilon$ then repeat step 4 using Ω_{t+1} as the initial distribution of workers. If $\|\Omega_{t+1} - \Omega_t\| \leq \epsilon$ then proceed to the next step.

6. Store the final distribution and compute features of the equilibrium including the unemployment rate, job-to-job transition rates by ability and tenure, equilibrium wage distribution by types, characteristics of unemployment, and the incidence of wage decreases in transitions.

E. Simulate model CI: Store the final distribution of workers and compute features of the equilibrium. Do so by repeating C using the equilibrium policy functions and state space for model CI and note that with perfect information $\phi(\mu) = F(\lambda_i x, \mu) = \mu$ to create the mapping Γ^{CI}.

F. Calibrate model. Repeat steps (B) to (E) for each set of parameters from (A). Once complete, choose the optimal set of parameters that solves the problem described by (F.1) and (F.2).

Additional references used in this supplementary appendix:

Amir, R., Mirman, L.J. and W.R. Perkins, 1991, "One-sector nonclassical optimal growth: optimality conditions and comparative dynamics," International Economic Review 32, 625-644.

Stokey, N., R.E. Lucas, Jr. with E. Prescott, 1989, Recursive Methods in Economic Dynamics. Cambridge, MA: Harvard University Press.

Topkis, D.M., 1998, Supermodularity and Complementarity. Princeton, NJ: Princeton University Press.