

Multidimensional Learning, Optimal Contract, and On-the-job Search

Ji-Woong Moon*, Seokjong Ryu[†]
Shanghai University of Finance and Economics

September 13, 2022

Abstract

We study a model of on-the-job search and optimal wage contracts under uncertainty about both worker ability and match quality. Firms and workers update beliefs through random output realizations, and the optimal contracts and search decisions respond differently to the belief about ability and match quality. We show that the optimal wage has a continuous path even when beliefs jump. The optimal contract exhibits wage cuts over tenure within a match if the worker is valuable only to outside firms. A random promotion can be a result of the optimal contract. We numerically calculate the optimal wage policy as a function of each belief, identifying how the implications of learning can differ if uncertainty is about a worker or match. The multidimensional aspect of learning can explain empirical phenomena which are not easily understood in standard models: the adverse wage effects of previous job-to-job transitions and the higher likelihood of wage cuts for high-wage/ability workers in firms.

*School of Economics, Shanghai University of Finance and Economics, 777 Guoding Rd, Shanghai 200433, China.
E-mail: jiwoong@mail.shufe.edu.cn

[†]School of Economics, Shanghai University of Finance and Economics, 777 Guoding Rd, Shanghai 200433, China.
E-mail: seokjong@mail.shufe.edu.cn

1 Introduction

When a person begins a career, she is uncertain about her talent “and” whether the firm is a good fit for her. Both uncertainties resolve over time jointly, leading to different career choices and compensations. While this story is intuitively appealing, few papers consider the learning and contract problem when both uncertainties exist. For instance, previous papers with uncertain worker ability abstract uncertainty in match quality (Blumen et al., 1955; Gonzalez and Shi, 2010; Papageorgiou, 2014; Carrillo-Tudela and Kaas, 2015), while papers with uncertain match quality assume observable worker ability (Jovanovic, 1979, 1984; Moscarini, 2005; Menzio and Shi, 2010). Considering only one aspect of learning can lead to different interpretations of the same labor market moments. For instance, a longer tenure may suggest a better fit, but it can imply poor outside opportunities due to a bad ability. Alternatively, it may associate with a good ability if the firm provides enough performance-based payments. As such, what the observed wage and job transition patterns imply depends on the underlying uncertainty structure and the payment scheme, but the lack of a unified framework limits our understanding of how learning affects wages and job transitions. This paper fills this gap by building a model of on-the-job search and optimal wage contracts with uncertainty regarding worker ability and match quality.

In this paper, we incorporate Jovanovic (1979) type symmetric learning into the model of on-the-job search and optimal contract (Burdett and Coles, 2003; Shi, 2009). A match’s output is a stochastic process of which law depends on hidden match quality, either good or bad (Moscarini, 2005). We introduce general ability by assuming that hidden worker ability, either high or low, affects the initial match quality draw: a high-quality worker is more likely to be a good fit for every firm (Gonzalez and Shi, 2010). Thus, a history of output realizations is informative also for worker ability. The labor market frictions are modeled by directed search. Both employed and unemployed workers search for the next job, and their optimal search target changes when beliefs update. Firms post a vacancy that commits a wage contract, which specifies wage payment contingent upon public history. The optimal wage contract is well-defined between risk-neutral firms and risk-averse workers in that the firm can only indirectly control the worker’s search decisions through

future wage promises.

The first contribution of this paper is the theoretical characterization of the optimal wage contract. We build a continuous-time model and heuristically derive an HJB-type equation with uncertainty, contrary to previous literature that uses discrete-time approaches in similar environments (Menzio and Shi, 2011; Balke and Lamadon, 2020). We derive three qualitative features of the optimal contract over tenure. First, the optimal wage is continuous in tenure as in Burdett and Coles (2003) and Shi (2009), even if beliefs discontinuously evolve. Firms compensate for a positive shock by a higher wage growth rather than an immediate wage promotion because promising a discontinuous wage path is sub-optimal to risk-averse workers. Second, the optimal wage is not always increasing in tenure, in contrast to Burdett and Coles (2003) and Shi (2009). The wage falls if the job value is negative for the current employer but positive for outside firms. In this case, the current employer provides an incentive to leave the job through an on-the-job search, which can happen when the worker's ability is likely good, but the match quality is likely bad. Third, the optimal contract may exhibit a random promotion during periods of flat wages. It occurs when the firm needs to raise the expected promised value, but it is costly to raise the current value locally.

The second contribution of this paper is to represent the shape of the optimal contract as a function of each belief separately. When underlying uncertainty is multidimensional, the beliefs are also multidimensional and are challenging to interpret, even if the number of states is finite. The specific information structure in this paper allows us to overcome this difficulty and represent the optimal wage contract as a function of each belief dimension separately. We calibrate the model and show that the optimal wage contract is increasing in the promised value while decreasing in the ability and match-quality belief. It does not mean better workers are paid less because they enjoy higher promised values on average.

The last contribution of this paper is to apply the model to understand empirical regularities. We simulate the model and confirm that the model is consistent with well-known empirical regularities such as declining job-to-job transitions and concave wage profiles over tenure. On top of that, we answer two empirical questions that standard models do not easily answer. The first question

is whether better workers stay at a job longer or leave for a better job quicker. Using the simulated economy, we find that workers with a higher ability belief are paid better and stay at a job longer. It confirms a negative performance-quit relationship observed in data (Salamin and Hom, 2005; Hom et al., 2008). Note that the relationship is the opposite if the wage contract is fixed-wage, which highlights the importance of performance-based payment to retain good workers (Lazear, 2000; Franceschelli et al., 2010). This negative ability-quit relationship generates a negative correlation between the past job-to-job transitions and wages in data (Carrillo-Tudela and Kaas, 2015). The ability uncertainty is essential for the negative correlation in that it becomes positive once the ability uncertainty is eliminated, as in standard search models (Burdett and Mortensen, 1998; Delacroix and Shi, 2006).

The second question is the characteristics of job stayers whose wages fall. Note that the optimal wage contract does not exhibit wage cuts within a job if a match's output depends on ability but not on match quality, suggesting that it is essential to distinguish between ability and match quality. The simulation reveals that the incidence of wage cuts is increasing in worker ability belief and wage and decreasing in match quality belief. It is intuitive because these workers have higher outside options but lower value to the current employer, which is what the theory predicts about the necessary condition for wage cuts. It is consistent with recent empirical findings that many firms cut the wages of a subset of workers, and they are highly paid on average (Park and Shin, 2017; Jardim et al., 2019).

Related literature: This paper is related to a large literature on labor market learning (Jovanovic, 1979, 1984; Moscarini, 2005; Gonzalez and Shi, 2010; Papageorgiou, 2014; Carrillo-Tudela and Kaas, 2015; Doppelt, 2016; Carrillo-Tudela and Kaas, 2015). This paper contributes to this strand of literature by building a learning model with uncertainty about both general and specific human capital, and labor market frictions. Beyond realism, the unified model allows us to explain whether better workers stay at a job longer and the characteristics of workers who are more likely to experience wage cuts.

This paper also contributes to a strand of literature that has studied the relationship between

performance and worker turnover. While a negative correlation between performance and quits is intuitive, theoretically (Jovanovic, 1984) and empirically (Salamin and Hom, 2005; Hom et al., 2008), it is less clear how a worker’s innate ability is associated with quits. We show that the negative correlation between ability and quits exists, but only when wages endogenously respond to the belief, which is consistent with the view that performance payment is crucial to retaining good workers (Lazear, 2000; Franceschelli et al., 2010). This explains a negative correlation between job-to-job transitions and wages in data (Light and McGarry, 1998; Carrillo-Tudela and Kaas, 2015).

Lastly, this paper is close related to the literature on the optimal wage contract (Burdett and Coles, 2003; Shi, 2009; Menzio and Shi, 2010; Fawcett and Shi, 2018; Balke and Lamadon, 2020). To our knowledge, this is the first paper that incorporates learning, on-the-job search, and wage contract with risk-averse workers in a unified framework. Li and Weng (2017) studies match-specific uncertainty and on-the-job search, but they solve the social planner’s problem so that there is no wage. In terms of modeling, Balke and Lamadon (2020) is close to our paper. Because of the similar modeling choice, the optimal contract also has several similar features. The difference is that in Balke and Lamadon (2020), the multi-dimensional feature comes from exogenous productivity shocks, while in our paper, the multi-dimensional feature comes through endogenous learning. Also, we develop a continuous-time model that allows us to show that the wage path is continuous regardless of the underlying uncertainty structure.

2 Model

2.1 Environment

Time is continuous. The labor market has a continuum of risk-averse workers with a geometrically distributed lifespan. Workers exit the labor market at a fixed rate, and the same measure of newly-born workers enter the labor market. We normalize the measure of workers in the labor market to

one. The other side of the labor market consists of risk-neutral firms with free entry. Workers and firms discount the future at the same rate, and we denote the effective discount rate by r , which is the sum of the discount and labor market exit rates.

At each point in time, workers are either employed or unemployed. The employed workers are paid according to their wage contracts, which will be explained later. They are separated from the job voluntarily through on-the-job search and involuntarily into unemployment at a fixed rate of δ . The unemployed workers receive a flow payoff b .

2.2 Production

A match between a worker and a firm produces output. The output process is stochastic, and its process depends on the match-specific quality α , which is drawn at the beginning of a match. The match quality is either one of two values $\alpha \in \{\alpha_G, \alpha_B\}$, unobserved, and remains fixed within a match (Jovanovic, 1979; Moscarini, 2005). We assume that $0 < \alpha_B < \alpha_G < 1$ without loss of generality.

A match produces output only occasionally. Specifically, a match receives a production opportunity at a Poisson rate of λ , independently of α . Denote the cumulative production opportunity process by $Y_\tau \in \mathbb{N}$. Upon an arrival of production opportunity at τ ($dY_\tau = 1$), a match succeeds to produce the stock of output $y_\tau = y > 0$ with probability α , and fails so that produces nothing ($y_\tau = 0$) with probability $1 - \alpha$. This output process implies that a match produces the output stock y at a Poisson rate of $\lambda\alpha$, while not producing any output flow. Because $\alpha_G > \alpha_B$, a good match produces the output more frequently than a bad match. Therefore, a successful production raises the likelihood of a good match. However, the arrival of the production opportunity itself is not informative because λ is independent of α .

A departure from the previous literature is that the initial draw of α depends on workers' innate ability $\gamma \in \{\gamma_H, \gamma_L\}$. Workers' ability is unobserved and remains fixed during their lifetime. The ability is drawn from the initial distribution $P(\gamma = \gamma_H) = 1/2$ when entering the labor market and

affects the initial match quality draws as follows.

$$P(\alpha = \alpha_G | \gamma = \gamma_H) = P(\alpha = \alpha_B | \gamma = \gamma_L) = \rho \geq 1/2 \quad (1)$$

In words, an H-type (L-type) worker is more likely to be a good (bad) fit for every match. ρ governs the degree of correlation between the ability and match quality. $\rho = 1/2$ means that worker ability is homogeneous.

2.3 Information structure

Denote the history of production opportunities and their outcomes until tenure τ by $Y^\tau = \{Y_t, y_t\}_{0 \leq t \leq \tau}$. We assume that Y^τ is public information so that even outside firms can observe Y^τ . This assumption eliminates complicated issues arising from asymmetric information between inside and outside firms. It is not a restrictive assumption because the belief about the worker's ability is a sufficient statistic for outside firms, which can be inferred through various methods in reality, such as reference checks and resumes. It is a common assumption in the previous literature (Gonzalez and Shi, 2010; Doppelt, 2016; Fawcett and Shi, 2018).

The concept of the worker's ability in this paper is similar to Gonzalez and Shi (2010) in that it affects the distribution of the initial match quality draws. The difference is that the match quality is immediately observed after meeting in Gonzalez and Shi (2010), while in this paper, workers and firms slowly learn through output realization. This learning shapes how the belief evolves within and across matches over the employment history. The belief process determines wage growth and job transitions endogenously under the optimal contract setting, which is one of the key questions this paper addresses.

2.4 Learning

No learning occurs during unemployment because b is constant, and a match is an experience good.¹ Thus, we only focus on learning within a match. Denote $\mu_\tau = P(\alpha = \alpha_G | h^\tau)$, where h^τ includes Y^τ for the current match, and belief about the worker's ability at $\tau = 0$. Given the output process, μ_τ changes only when a production opportunity arises. Upon the arrival of a production opportunity, the belief jumps to a new value depending on the production outcome. Denote the belief after a successful production at τ by μ_τ^S , and the belief after a failure by μ_τ^F . By the Bayes' rule,

$$\mu_\tau^S = \frac{\alpha_G \mu_{\tau-}}{\alpha_G \mu_{\tau-} + \alpha_B (1 - \mu_{\tau-})}, \quad \mu_\tau^F = \frac{(1 - \alpha_G) \mu_{\tau-}}{(1 - \alpha_G) \mu_{\tau-} + (1 - \alpha_B) (1 - \mu_{\tau-})} \quad (2)$$

where $\mu_{\tau-} = \lim_{t_0 \uparrow \tau} \mu_{t_0}$. A successful production raises the belief about match quality $\mu_\tau^S > \mu_{\tau-}$, while a failure reduces the belief $\mu_\tau^F < \mu_{\tau-}$. Once the belief jumps to a new level, it stays at the level until the next production opportunity arrives. The initial belief μ_0 is determined by the initial belief about the worker's ability.

The worker and the firm learn about the worker's ability as well. Denote $\eta_\tau = P(\gamma = \gamma_H | h^\tau)$. Note that Y^τ is independent of γ conditional on α , meaning that the belief about the worker's ability is constant within a match conditional on α . Therefore,

$$P(\gamma = \gamma_H | \alpha = \alpha_G, h^\tau) = P(\gamma = \gamma_H | \alpha = \alpha_G, \eta_0) = \frac{\eta_0 \rho}{\eta_0 \rho + (1 - \eta_0)(1 - \rho)} \equiv R_G(\eta_0) \quad (3)$$

$$P(\gamma = \gamma_H | \alpha = \alpha_B, h^\tau) = P(\gamma = \gamma_H | \alpha = \alpha_B, \eta_0) = \frac{\eta_0(1 - \rho)}{\eta_0(1 - \rho) + (1 - \eta_0)\rho} \equiv R_B(\eta_0) \quad (4)$$

From Equation (3) and (4), the following relationship between η_t and μ_t can be derived.

$$P(\gamma = \gamma_H, \alpha = \alpha_G | h^\tau) = P(\alpha = \alpha_G | h^\tau) \cdot R_G(\eta_0) \quad (5)$$

$$P(\gamma = \gamma_H, \alpha = \alpha_B | h^\tau) = P(\alpha = \alpha_B | h^\tau) \cdot R_B(\eta_0) \quad (6)$$

$$\Rightarrow \eta_\tau = \mu_\tau R_G(\eta_0) + (1 - \mu_\tau) R_B(\eta_0) \quad (7)$$

¹If firms observe an initial signal before match creation and make hiring decisions based on it, unemployed workers also learn about their ability through job search failures while unemployed.

It shows that the change in η_τ is proportional to the change in μ_τ within a match, and the responsiveness is a function of the initial belief η_0 .

$$d\eta_\tau = L(\eta_0)d\mu_\tau, \quad L(\eta) \equiv R_G(\eta) - R_B(\eta) \quad (8)$$

Equation (8) implies that η_τ is bounded away from 0 and 1 within a match. Intuitively, each match is a match quality draw so that η_τ is bounded above (below) by the case where a good (bad) match quality draw is directly observed. It implies that Equation (9) holds.

$$\frac{\eta_0}{\eta_0 + (1 - \eta_0)^{\frac{\rho}{1-\rho}}} \leq \eta_\tau \leq \frac{\eta_0}{\eta_0 + (1 - \eta_0)^{\frac{1-\rho}{\rho}}} \quad (9)$$

Because the learning about ability is bounded, workers need to experience several matches to learn about their abilities. It generates an additional incentive for workers to search for the next job provided that workers prefer the resolution of uncertainty.

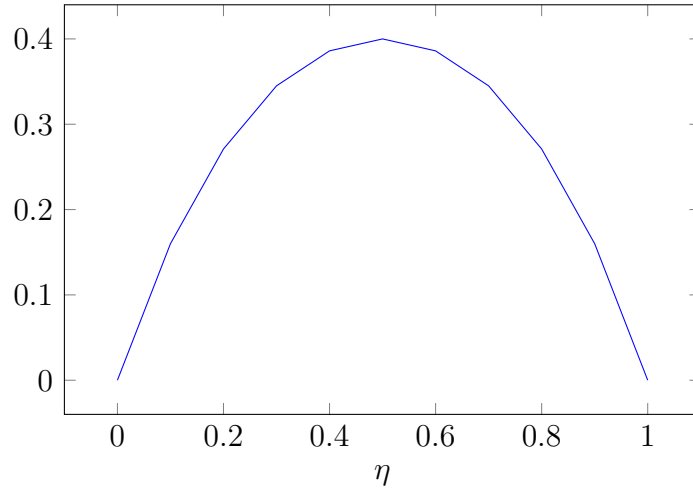


Figure 1: Learning slope $L(\eta)$

Figure 1 illustrates $L(\eta)$. It is maximized at $\eta = 1/2$, meaning that the learning about the worker's ability is faster in a match with a larger initial uncertainty. Given that γ is initially drawn from $P(\gamma = H) = 1/2$, workers with a shorter labor market experience learn about their ability faster, and the learning slows down over the employment history on average.

2.5 Search and matching

Search is directed. Firms post a vacancy with a wage contract, which will be explained in Section 2.6, and workers apply to a vacancy while observing all the wage contracts. The wage contracts can be indexed by the lifetime value to the worker V and belief about worker ability η , implying that a continuum of submarkets indexed by (V, η) exist. A market tightness $\theta(V, \eta)$ is specific to each submarket, determining the job-finding rate and vacancy-filling rate. We denote the job-finding rate function by $p(\theta(V, \eta))$ and vacancy-filling rate function by $q(\theta(V, \eta)) = p(\theta(V, \eta))/\theta(V, \eta)$. We impose standard conditions: $p(\theta) \in [0, \bar{p}]$, $p'(\theta) > 0$, $p''(\theta) < 0$, $q(\theta) \in [0, \bar{q}]$, $q'(\theta) < 0$, $q''(\theta) < 0$.

Consider a worker whose current continuation value is V , and belief about ability is η . The optimal search decision for the worker is described by

$$R(V, \eta) \equiv \max_{V'} p(\theta(V', \eta))(V' - V) \quad (10)$$

Let $F(V, \eta)$ denote the optimal search policy, and $p_F(V, \eta) \equiv p(\theta(F(V, \eta), \eta))$ denote the endogenous job-finding rate induced by the optimal search. An unemployed worker solves the same search problem where the continuation value is the value of the unemployed.

2.6 Contract

A vacancy in a submarket (V, η) commits a wage contract that delivers the lifetime value $V \in [\underline{V}, \bar{V}]$ to a worker with the ability belief η , where $-\infty < \underline{V} < \bar{V} < \infty$. To make the contract space convex, we allow the firm can randomize wages using a public randomization device $X(t)$. To be specific, the wage at tenure τ is a function of the public history $h^\tau = \{X^\tau, Y^\tau, \eta_0\}$, where X^τ is the history of $X(t)$ realizations until τ .² The worker earns the promised wage only if the

²Previous literature introduces a two-point lottery to make the firm's value function concave (Menzio and Shi, 2011; Balke and Lamadon, 2020). There is another reason for us to introduce the public randomization device $X(t)$. As Staudigl and Steg (2017) mentioned, for the continuous-time limit model to be consistent with the discrete-time approximations with frequent actions, we should consider the *correlated equilibrium* in the continuous-time. Moreover, in literature of continuous-time games including Sannikov (2007) and Bernard (2016), the public randomization device plays an important role in making the set of equilibrium payoffs to be convex.

worker stays at the match. Because the output process determines beliefs, the wage fully depends on the beliefs about match quality and ability.

The optimal contract is the contract that maximizes the firm's expected profit subject to the promised lifetime value to the worker V , and given belief (μ, η) . While firms can commit any contract, only the optimal contract can exist in the market because of the free entry condition.

Several points are worth mentioning regarding the contractual environment. First, firms cannot renegotiate or fire workers after hiring, while workers do not commit to staying at the job. Second, the contract cannot depend on workers' search outcomes or targets, meaning that firms cannot counteroffer like in [Postel-Vinay and Robin \(2002\)](#). The two points imply that workers leave their current job whenever a job search is successful, and firms can only indirectly affect workers' search decisions through promised value. This specific contractual environment is studied in previous literature, such as [Burdett and Coles \(2003\)](#); [Shi \(2009\)](#); [Menzio and Shi \(2011\)](#) and [Balke and Lamadon \(2020\)](#).

2.7 Value functions

Define the firms' value function $J(V, \mu, \eta)$ as

$$J(V, \mu, \eta) = \sup_{\tilde{w}=\{\tilde{w}_t\}_{t \geq 0}^{\infty}} \mathbb{E} \int_0^{\infty} e^{-\int_0^t (r+\delta+s_e p_F(V_s, \eta_s)) ds} \left[e^{-\lambda t} \lambda \alpha(\mu_t) y - \tilde{w}_t \right] dt \quad (11)$$

with

$$V_s = \mathbb{E} \int_s^{\infty} e^{-\int_s^t (r+\delta+s_e p_F(V_\tau, \eta_\tau)) d\tau} [u(\tilde{w}_t) + s_e p_F(V_t, \eta_t) F(V_t, \eta_t) + \delta U(\eta_t)] dt, \quad V = V_0 \quad (12)$$

where $\alpha(\mu) = \mu \alpha_G + (1 - \mu) \alpha_B$, $\tilde{w}_t \in \Delta[\underline{w}, \bar{w}]$ with $-\infty < \underline{w} < \bar{w} < \infty$, and $\mathbb{E}$ is an expectation with respect to the probability distribution induced by $w = \{\tilde{w}_t\}_{t \geq 0}^{\infty}$.

While Equation (11) and (12) are defined on the general contract space, it is difficult to consider general dynamic contracts. Therefore, we focus on wage contracts that are recursively described

by the state variable (V, μ, η) following the approach in previous literature (Menzio and Shi, 2011; Fawcett and Shi, 2018; Balke and Lamadon, 2020).

In the recursive formulation, the value function $J(V, \mu, \eta)$ and optimal control $\pi \in \Delta[\underline{V}, \bar{V}]$, $(w, m, V^P, V^F, V^C) \in [w, \bar{w}] \times [0, \infty) \times [\underline{V}, \bar{V}]^3$ satisfy the following HJB-type equation (ref Sannikov), where (π, m) is the randomization of contract over (V^C, V^P) , and (V^S, V^F) is the promised value contingent on the realization of production shock.

$$\begin{aligned} J(V, \mu, \eta) &= \max_{\pi} E_{\pi} \tilde{J}(V^C, \mu, \eta) \quad \text{s.t.} \quad E_{\pi}[V^C] = V \\ \tilde{J}(V^C, \mu, \eta) &= \max_{w, m, V^P, V^S, V^F} \left[(\lambda \alpha(\mu) y - w) + m(J(V^P, \mu, \eta) - J(V^C, \mu, \eta)) \right. \\ &\quad \left. + \lambda \left(\alpha(\mu) J(V^S, \mu^S, \eta^S) + (1 - \alpha(\mu)) J(V^F, \mu^F, \eta^F) - J(V^C, \mu, \eta) \right) \right. \\ &\quad \left. - (\delta + s_e p_F(V^C, \eta)) J(V^C, \mu, \eta) + \mathcal{A}^o J(V^C, \mu, \eta) \right] \end{aligned} \quad (13)$$

where $\mathcal{A}^o$ is the operator defined by

$$\mathcal{A}^o W(t) = \lim_{h \downarrow 0} \frac{\mathbb{E}[W(t+h) - W(t) | \text{No production shock during } (0, h)]}{h} \quad (14)$$

for a stochastic process $W(t)$. The operator $\mathcal{A}^o$ calculates the expected growth of a stochastic process along the path without production shocks. When randomization is ignored, the stochastic process is deterministic conditional on the event that no production shock occurs.

From now on, we will assume that $J(V, \mu, \eta)$ is sufficiently smooth with respect to V . In other words, for a given (V, μ, η) , it is twice-differentiable with respect to V , and it has a continuous second-derivative with respect to V .

Given the definition, $\mathcal{A}^o J(V, \mu, \eta)$ is given by

$$\mathcal{A}^o J(V, \mu, \eta) = J_V(V, \mu, \eta) \mathcal{A}^o V \quad (15)$$

$$\mathcal{A}^o V = (rV - [u(w) + \lambda(\bar{V} - V) + \delta(U(\eta) - V) + m(V^P - V) + s_e R(V, \eta)]) \quad (16)$$

provided that the firm value J is differentiable with respect to V . In the equation, $(\mu^S, \eta^S), (\mu^F, \eta^F)$ are updated beliefs after a production opportunity, δ is the exogenous separation rate, s_e is the exogenous search rate for the employed worker, and $U(\eta)$ is the value of unemployed.

The value of unemployed $U(\eta)$ is the solution of the implicit equation (17).

$$rU(\eta) = u(b) + s_u R(U(\eta), \eta) \quad (17)$$

where s_u is the exogenous search rate for the unemployed. There is no $\mathcal{A}U(\eta)$ term in Equation (17) since η remains constant while unemployed.

Regarding the firm value function, there are several properties that are useful for future arguments.

Lemma 2.1. *The firm value function $J(V, \mu, \eta)$ is continuous, weakly decreasing and weakly concave in V for all (μ, η) .*

Lemma 2.2. *If π is a solution of (13), then $J_V(V, \mu, \eta) = J_V(V^C, \mu, \eta)$ for all V^C that $\pi(V^C) > 0$.*

Lemma 2.1 holds because the contract space is convex. The firm can offer a randomized contract between any two contracts in the beginning. It guarantees that the value function is weakly concave in V . Because $J(V, \mu, \eta)$ is weakly concave, any V^C in the domain of randomization must satisfy $J_V(V, \mu, \eta) = J_V(V^C, \mu, \eta)$, otherwise it cannot satisfy both $V = E_\pi[V^C]$ and $J(V, \mu, \eta) = E_\pi[J(V^C, \mu, \eta)]$.

To understand the contract problem intuitively, we present a discrete approximation of the model. At each point in time, the firm chooses the current wage w and the evolution of the promised value tomorrow. Because the firm fully controls the promised value process, the firm chooses the promised value contingent on the shock realizations: the promised value when the production opportunity arrives and succeeds V^S , the promised value when production opportunity arrives and fails V^F , and the promised value when production opportunity does not arrive V^C . The public randomization implies that the firm chooses a distribution π on (w, V^S, V^F, V^C) . In short, the

value function satisfies the maximization problem (18) subject to the promise-keeping constraint (19)

$$J(V, \mu, \eta) = \max_{\pi} E_{\pi} \left[(\lambda \alpha(\mu) y - w) \Delta + \frac{1}{1 + r \Delta} (\lambda \Delta \bar{J}(V^S, V^F, \mu, \eta) + Q(V^C, \eta) J(V^C, \mu, \eta)) \right] \quad (18)$$

$$\text{where } \bar{J}(V^S, V^F, \mu, \eta) = \alpha(\mu) J(V^S, \mu^S, \eta^S) + (1 - \alpha(\mu)) J(V^F, \mu^F, \eta^F)$$

$$Q(V^C, \eta) = 1 - (\lambda + \delta + s_e p_F(V^C, \eta)) \Delta$$

subject to

$$V = E_{\pi} \left[u(w) \Delta + \frac{1}{1 + r \Delta} (\lambda \Delta \bar{V} + \delta \Delta U(\eta) + s_e p_F(V^C, \eta) \Delta F(V^C, \eta) + Q(V^C, \eta) V^C) \right] \quad (19)$$

$$\text{where } \bar{V} = \alpha(\mu) V^S + (1 - \alpha(\mu)) V^F$$

Intuitive interpretation of the maximization problem is the following. In a short period of time Δ , three types of shocks can occur: the production opportunity arrival with probability $\lambda \Delta$, exogenous separation with probability $\delta \Delta$, and endogenous separation through a job-to-job transition with probability $s_e p_F(V^C, \eta) \Delta$. Upon the arrival of the two separation shocks, the match is destroyed and the firm value becomes 0. When the production opportunity arrives, the match produces expected output $\alpha(\mu) y$, and the expected firm value after Δ becomes $\bar{J}(V^S, V^F, \mu, \eta)$, which is the average between $J(V^S, \mu^S, \eta^S)$ and $J(V^F, \mu^F, \eta^F)$. With probability $Q(V^C, \eta)$, no shock occurs and the firm value after this event is $J(V^C, \mu, \eta)$. Taking into accounts these possibilities, the firm chooses a distribution π over (w, V^S, V^F, V^C) to maximize the firm value. The probability distribution π must deliver the expected lifetime value V to the worker, which is the sum of flow utility $u(w) \Delta$ and the discounted expected future value $(\lambda \Delta \bar{V} + \delta \Delta U(\eta) + s_e p_F(V^C, \eta) \Delta F(V^C, \eta) + Q(V^C, \eta) V^C) / (1 + r \Delta)$. Note that the effects of the belief changes are already incorporated into V^S and V^F . In the appendix, we show that Equation (18) and equation (19) converge to the HJB-type equation (13) as $\Delta \rightarrow 0$.

2.8 Market tightness

Firms post a vacancy with a flow cost $k > 0$. Denote the firm value immediately after matching with a worker by $J_0(V, \eta)$, which satisfies

$$J_0(V, \eta) \equiv J(V, \eta\rho + (1 - \eta)(1 - \rho), \eta) \quad (20)$$

Another interpretation of $J_0(V, \eta)$ is the value of an employed worker to outside firms when outside firms match the current contract value V . The free entry condition requires that the market tightness satisfies (21)

$$q(\theta(V, \eta))J_0(V, \eta) = k \quad (21)$$

if $\bar{q}J_0(V, \eta) \geq k$, and $\theta(V, \eta) = 0$ otherwise. Note that firms have a higher initial belief about match quality $\mu_0 = \eta\rho + (1 - \eta)(1 - \rho)$ in a market with higher η . Given that the increase in the expected output is the first-order effect, $J_0(V, \eta)$ is increasing in η .³ It means that the market for a higher η is tighter, so workers with a higher η find a job more easily conditional on V .

3 Equilibrium

Definition 3.0.1. An equilibrium consists of the firm value function $J(V, \mu, \eta)$, the optimal policy $\pi(V, \mu, \eta)$, the market tightness $\theta(V, \eta)$, the optimal search policy and return to search $F(V, \eta)$, $R(V, \eta)$, and the value of unemployed $U(\eta)$ such that

1. The firm value function $J(V, \mu, \eta)$ is the solution of (11) subject to (12), and the optimal policy $\pi(V, \mu, \eta)$ solves the (13) together with the value function $J(V, \mu, \eta)$ given $F(V, \eta)$, $R(V, \eta)$, $U(\eta)$ and the evolution of beliefs (2) and (7).

³In equilibrium, there are several effects of η on J beyond its effects on μ_0 . First, workers with different η find the next job at different rates. Second, the wage stream required to deliver V differs by η since the return to search is different. Given that $J_0(V, \eta)$ is increasing in η , then the first effect makes this relationship weaker while the second effect makes this relationship stronger. We numerically confirm that not only J_0 , but also J are increasing in η .

2. Workers' search decisions and return to search $F(V, \eta)$ and $R(V, \eta)$ are consistent with the equilibrium tightness $\theta(V, \eta)$.
3. $\theta(V, \eta)$ satisfies the free entry condition (21) given $J(V, \mu, \eta)$.
4. The value of unemployed $U(\eta)$ satisfies (17) given $R(V, \eta)$.

The equilibrium definition is self-explanatory. The equilibrium in this paper is Block-recursive in the sense that the value function and policy functions are independent of aggregate distributions (Shi, 2009; Menzio and Shi, 2011). It is because workers' search is directed by the wage contract. It reduces computational burdens when we calculate the optimal contract, and more substantially when we simulate our model.

4 Characterization of the optimal wage contract

In this subsection, we will derive qualitative features of the optimal contract.

4.1 Continuous wage-path

Recall the firm's problem Equation (13). By substituting $\tilde{J}(V, \mu, \eta)$, one can get the following.

$$\begin{aligned}
rJ(V, \mu, \eta) = & \max_{\pi} E_{\pi} \left[(\lambda \alpha(\mu) y - w) + m(J(V^P, \mu, \eta) - J(V^C, \mu, \eta)) \right. \\
& + \lambda \left(\alpha(\mu) J(V^S, \mu^S, \eta^S) + (1 - \alpha(\mu)) J(V^F, \mu^F, \eta^F) - J(V^C, \mu, \eta) \right) \\
& \left. - (\delta + s_e p_F(V^C, \eta)) J(V^C, \mu, \eta) + J_V(V^C, \mu, \eta) \mathcal{A}^o V^C \right] \\
\text{subject to } & J(V, \mu, \eta) = E_{\pi}[J(V^C, \mu, \eta)], \quad V = E_{\pi}[V^C]
\end{aligned}$$

where the choice variable is a probability distribution π over $(w, m, V^C, V^P, V^S, V^F, V^C)$. From Lemma 2.2, $J_V(V^C, \mu, \eta) = J_V(V, \mu, \eta)$, which means that the firm does not randomize over

(w, V^S, V^F) because the following terms are (weakly) concave with respect to each control.

$$w : -w - J_V(V, \mu, \eta)u(w) \quad (22)$$

$$V^S : \lambda\alpha(\mu)(J(V^S, \mu^S, \eta^S) - J_V(V, \mu, \eta) \cdot V^S) \quad (23)$$

$$V^F : \lambda(1 - \alpha(\mu))(J(V^F, \mu^F, \eta^F) - J_V(V, \mu, \eta) \cdot V^F) \quad (24)$$

Therefore, the randomization is only required for values without production shocks.

Given this formulation, the first order conditions are as the following:

$$\partial_w : -1 - J_V(V, \mu, \eta)u'(w) = 0 \quad (25)$$

$$\partial_V : J_V(V^S, \mu^S, \eta^S) = J_V(V^F, \mu^F, \eta^F) = J_V(V^P, \mu, \eta) = J_V(V^C, \mu, \eta) = J_V(V, \mu, \eta) \quad (26)$$

$$\partial_m : J_V(V^P, \mu, \eta) = J_V(V^C, \mu, \eta) = J_V(V, \mu, \eta) \quad (27)$$

The first order condition with respect to w shows that the optimal wage contract should optimally split the revenue as in previous literature (Burdett and Coles, 2003; Shi, 2009). At the optimal wage, the marginal benefit of reducing wage today must be equal to the marginal cost of doing so, which comes from a higher promised value in the future. By reducing wage marginally, the firm must promise a higher value instantaneously after by $u'(w)$ of which marginal cost is $J_V(V, \mu, \eta)$. Thus, $|J_V \cdot u'| = 1$ at the optimal contract. The first order conditions with respect to V^S and V^F show that J_V is equalized before and after the production shock. Otherwise, the firm can adjust how the value is delivered to the worker so that increase J while fixing the lifetime value of the worker.

Since $u'(w)$ is a function of J_V , the optimal wage is equalized before and after the production shock, regardless of whether it succeeds or not. When there is no production shock, the promised value V evolves according to the optimal π , which is either a degenerate or a two-point lottery. From Lemma 2.2 and first-order condition, $J_V(V^C, \mu, \eta)$ and $J_V(V^P, \mu, \eta)$ are equal to $J_V(V, \mu, \eta)$, meaning that the optimal wage does not vary by the randomization. Combining all

these results, the realized wage-path is continuous with respect to tenure.

Proposition 4.1.1. *The optimal wage-path is continuous in tenure.*

The continuous wage-path is known from previous literature when there is no uncertainty (Burdett and Coles, 2003; Shi, 2009). It is also a natural consequence if the belief and the worker value V continuously evolve. In this paper, however, the beliefs are jump processes. Still, the firm adjusts the promised value so that the optimal wage-path is continuous in tenure. The intuition behind Proposition 4.1.1 is similar to the previous literature. Since promising a discontinuous wage-path is expensive for firms due to risk-aversion, firms provide a continuous wage-path for every possible state realizations. It does not mean that firms do not compensate for a positive shock. Firms compensate for a positive shock not by immediate wage promotion but by future wage growth.

4.2 Wage and value growth

For a given V , the optimal contract may require mixing between two values $V_1 < V < V_2$ with a strictly positive probability π . Because π realization is immediate, the worker's promised value can never stay at such V . At the end point of mixing, such as V_1 and V_2 , the firm does not randomize through π , but may use the promotion rate m to randomize the contract.

Consider V such that $J_{VV}(V', \mu, \eta) < 0$ for all V' in a neighborhood of V . At (V, μ, η) , the optimal contract does not involve mixing, therefore, along the optimal path, one can get the following,⁴

$$J_{VV}(V, \mu, \eta) \mathcal{A}^o V = \frac{u''(w)}{u'(w)^2} \mathcal{A}^o w \quad (28)$$

Note that there is no uncertainty in the optimal wage growth because the optimal wage at t does not depend on the shock realizations at t , therefore, $dw/dt = \mathcal{A}^o w$. Equation (28) shows that $\mathcal{A}^o V$ and the wage growth rate must have the same sign if $J_{VV} < 0$.

⁴Since J is a concave function of V , J is almost everywhere twice-differentiable with respect to V .

By taking derivative of (13) with respect to V ,

$$J_{VV}(V, \mu, \eta) \mathcal{A}^o V = s_e \frac{\partial p_F(V, \eta)}{\partial V} J(V, \mu, \eta) \quad (29)$$

Therefore, the following relationships about the wage growth and firm value holds:

$$\frac{u''(w)}{u'(w)^2} \frac{dw}{dt} = s_e \frac{\partial p_F(V, \eta)}{\partial V} J(V, \mu, \eta) \quad (30)$$

Note that $\frac{\partial p_F(V, \eta)}{\partial V} \leq 0$, and strictly less than 0 if $p_F(V, \eta) > 0$, or equivalently $J_0(V, \eta) > k/\bar{q}$. It describes how the optimal wage evolves over tenure when production shocks do not occur.

Proposition 4.2.1. *Let (V, μ, η) be given. Suppose $J_{VV} < 0$ in a small neighborhood of V . The optimal wage growth rate is positive if $J_0(V, \eta) > k/\bar{q}$ and $J(V, \mu, \eta) > 0$. The optimal wage growth rate is negative if $J_0(V, \eta) > k/\bar{q}$ and $J(V, \mu, \eta) < 0$. The optimal wage growth rate is zero if $J_0(V, \eta) \leq k/\bar{q}$.*

Proposition 4.2.1 intuitively extends the result in the previous literature without uncertainty (Burdett and Coles, 2003; Shi, 2009). When there is no uncertainty, firms provide an increasing wage profile to retain the worker longer in a match. In this model, the same retention incentive appears when the worker is valuable for both the current and outside firms.

This retention incentive works in the opposite direction if the worker is valuable for outside firms, that is $J_0(V, \eta) > k/\bar{q}$, but not for the current firm, that is $J(V, \mu, \eta) < 0$. In this case, the current firm has an incentive to destroy the match, which is infeasible. Instead, the firm provides a decreasing wage profile which encourages the worker's on-the-job search. This wage cuts cannot occur if match quality uncertainty is absent because all firms have an identical value for a worker.

Note that the firm can control the duration of a match through the worker's on-the-job search decision. Therefore, if the worker cannot search on-the-job because the current promised value is already too high, that is $J_0(V, \eta) \leq k/\bar{q}$, the wage profile does not affect the expected duration of the match. Then, the only consideration is the consumption smoothing motive so that the wage is

flat over tenure.

Proposition 4.2.1 provides predictions about wage cuts. To be specific, there is a lower bound $\underline{\eta}$ such that if the initial belief η_0 starts from below $\underline{\eta}$, the optimal wage never declines over tenure in any possible history realization within the match. In general, Proposition 4.2.1 suggests that wage cuts are more likely to occur for workers with a successful career (high η) who performed badly in the current match (low μ). There are three intuitive reasons: first, workers with higher η have better outside opportunities so that their search decisions respond to wages more elastically. Second, these workers are on average highly paid (high V), so that $J(V, \mu, \eta)$ can become negative easily once μ falls. Third, when η_0 is high, η does not vary much with respect to μ . We confirm that it is the case using the simulation.

Proposition 4.2.1 also shows why wage cuts are rare although they exist in the optimal contract. Because outside firms need to pay vacancy costs, wage cuts can occur only if the gap between the value of a worker to the current and outside firms is sufficiently large: $J_0(V, \eta) > k/\bar{q} > 0 > J(V, \mu, \eta)$. The condition is more difficult to be satisfied when $k/\bar{q}$ is large, which suggests that the wage cuts are less prevalent in a labor market with larger frictions.

Lastly, the model suggests that the optimal contract may exhibit a period of flat wage followed by a steeper wage growth. It occurs when the optimal contract requires a randomization, or equivalently, $J_{VV} = 0$ at the current promised value. When the optimal contract randomizes over V_1 and V_2 , and the current value is V_1 , then the value jumps to V_2 at a rate of $\frac{\mathcal{A}^\circ V_1}{V_2 - V_1}$ while remains at V_1 otherwise.⁵ Therefore, occasional promotions with a flat wage in between, commonly observed in reality, can appear in the optimal contract.

5 Calibration

For numerical analysis, we use the following functional forms:

$$u(c) = \log(c), \quad p(\theta) = \frac{\theta}{1 + \theta} \quad (31)$$

⁵This is because the expected value must increase by $\mathcal{A}^\circ V_1$ without shocks.

We calibrate the model in a monthly base. We set the discount rate $r = 0.0052$, which reflect 5% annual discount rate and average labor market duration of 30 years. The exogenous separation rate δ is set to 0.015.

We normalize both the production amount y and the expected output per unit time when a worker newly enters $\lambda\alpha(0.5)$ to one. Then, we assume $\alpha_G - \alpha_B = 10\%$, which generates about 20% productivity difference between the good and bad match. These assumptions imply that $\lambda = 2, \alpha_G = 0.55, \alpha_B = 0.45$. Given the average productivity is normalized to one, we assume the home production b is 0.4 and the vacancy cost k is 0.3. The search rate of the unemployed s_u and the employed s_e are assumed to be 0.5 and 0.45, which generate reasonable unemployment and E–E transition rates.

There is few direct estimate for ρ , which governs the correlation between the worker’s type and match quality. Thus, we use two values $\rho = 3/4$ representing worker heterogeneity, and $\rho = 1/2$ representing homogeneous worker. We denote the economy with $\rho = 3/4$ as the benchmark, and $\rho = 1/2$ as a counterfactual economy where workers are homogeneous. The following table summarizes the choice of parameters.

Parameter	Meaning	Value
r	Discount rate	0.0052
δ	Separation rate	0.015
b	Home production	0.4
y	Output	1
λ	Production opportunity rate	2
α_G	Good match production probability	0.55
α_B	Bad match production probability	0.45
ρ	P(Good match High ability)	3/4, 1/2
s_u	Search rate for the unemployed	0.5
s_e	Search rate for the employed	0.45
k	Vacancy cost	0.3

Table 1: Parameter values

Using the parameters above, we simulate the economy consists of 10,000 workers for 50 years. Table 2 reports some key macro moments from the simulation. The unemployment rate and E–E

transition rates are relatively close to the data counterpart in both the benchmark and counterfactual economies. The model generates a significant wage dispersion in terms of Mean-min ratio through the learning and on-the-job search, but the dispersion is much higher when workers are heterogeneous. In both economies, the median is shorter than the average tenure, indicating that the tenure distribution is right-skewed. The difference between the mean and median tenure is much larger in the benchmark economy.

Statistics	Benchmark ($\rho = 3/4$)	Counterfactual ($\rho = 1/2$)
Average wage	1.0470	1.0672
Standard deviation of wage	0.1753	0.1167
Mean-min ratio	1.8514	1.4275
Unemployment rate	0.0622	0.0619
Average match quality (μ)	0.5407	0.5348
Median tenure (year)	2.1875	2.2500
Average tenure (year)	6.4708	3.3606
E-E transition rate	1.29%	1.56%

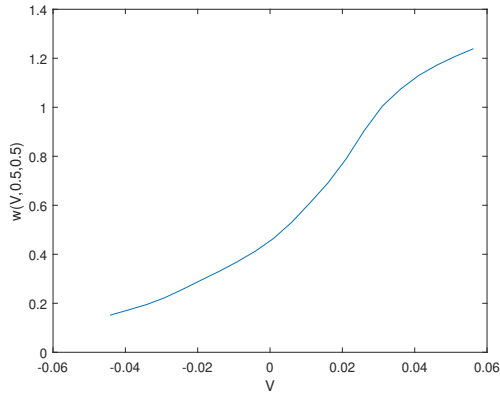
Table 2: Moments from the simulation

5.1 Numerical results

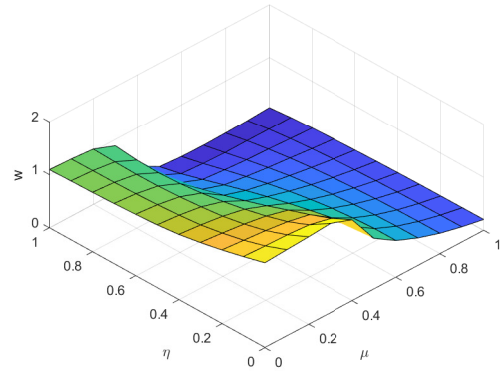
5.1.1 Optimal wage

We first examine the optimal wage policy as a function of (V, μ, η) . For the promised value, it is straightforward that the optimal wage is increasing in the promised value (Figure 2(a)). As a function of beliefs, the optimal wage is decreasing in both μ and η . It is decreasing in μ because the firm has a higher incentive to retain the worker when μ is higher. The firm can retain the worker longer by promising a higher value tomorrow, which leads to a lower current wage given the current promised value (Balke and Lamadon, 2020; Fawcett, 2021).

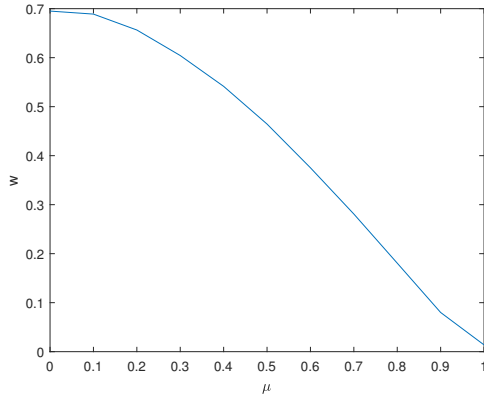
The optimal wage is also a decreasing function of η . There are two opposite effects of η on wage. On the one hand, a higher η raises the job-finding probability of the worker. Because the job-finding probability is concave in the search value, it tends to increase the optimal wage. On the



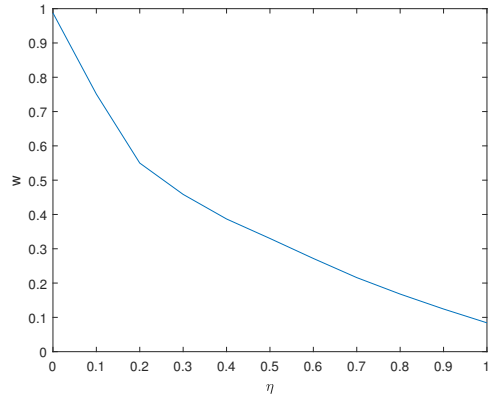
(a) Optimal wage over V



(b) Optimal wage over (μ, η)



(c) Optimal wage over μ



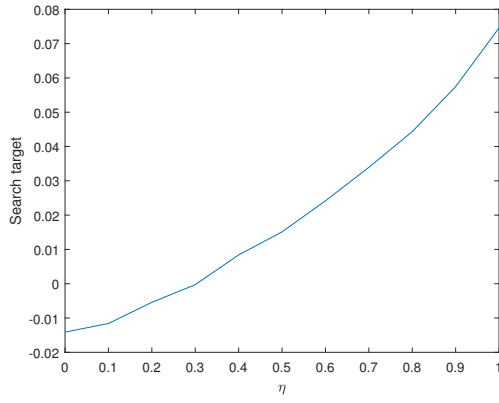
(d) Optimal wage over η

Figure 2: Optimal wage policy

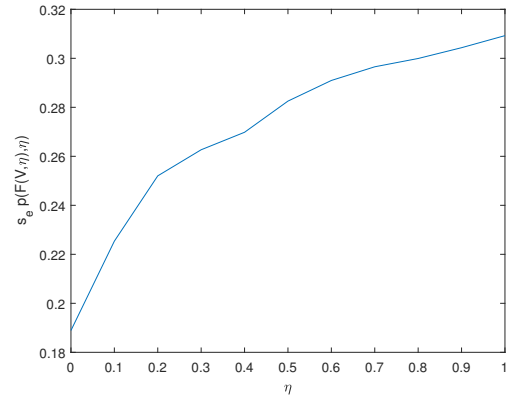
other hand, a higher η increases the value of search. It means that the firm can compensate less to deliver the same promised value. The numerical calculation shows that the latter effect dominates the former one.

Note that Figure 2(d) does not mean that good workers are paid less because good workers have on average higher promised values. There are two reasons. First, the promised value increases when η increases within a match. Second, a higher η worker searches for a higher V . The relationship between the ability and wage are results of all these effects in equilibrium. Especially, it crucially depends on the wage and value distribution. As expected, a strong positive relationship between η and w exists in equilibrium.

5.1.2 Worker's search decision

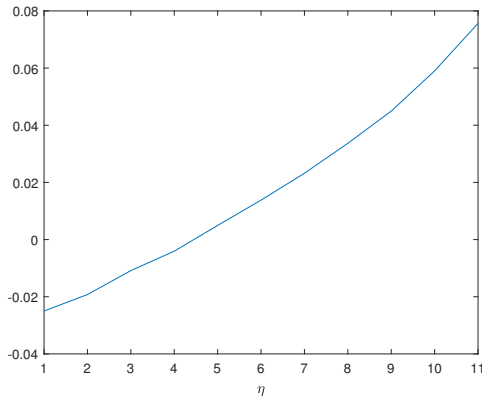


(a) Search target of the employed (given V)

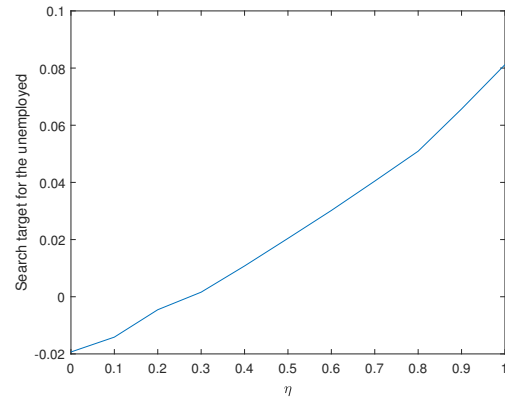


(b) Job-finding rate of the employed (given V)

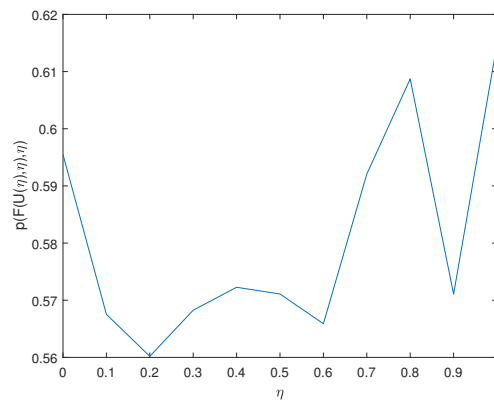
Figure 3: The employed worker's search decisions



(a) Value of the unemployed



(b) Search target of the unemployed



(c) Job-finding rate of the unemployed

Figure 4: The unemployed worker's search decision

Figure 3(a) and 3(b) illustrate the employed worker's optimal search decision as a function of η . The figures show that better workers search for a better job, and they move to the next job at a faster rate when the promised value is the same. Note that these two figures are drawn given V , therefore, it does not mean that better workers leave a job faster in equilibrium.

Figure 4(a) illustrates the value of the unemployed $U(\eta)$. The value of unemployed is increasing and slightly convex in η , meaning that the unemployed workers are better off by uncertainty resolution. It is an additional incentive to search for a job. Because the value of unemployed is increasing in η , the job-finding probability of the unemployed is not monotone. This is because a higher current value pushes up the search target, which reduces the job-finding probability. It means that the uncertain worker quality is not a sufficient condition for the negative duration effect (Doppelt, 2016).

5.1.3 Firm value function

Figure 5 illustrates the firm value function $J(V, \mu, \eta)$. The firm value is increasing in both μ and η . Intuitively, the firm value is increasing in μ because it directly increases the expected output (Figure 5(c)). While its direction of change is always the same, the convexity differs by the level of V . When V is low so that the worker can search for a job, the firm value is convex in μ . This is because the firm can adjust the effective discount rate through the optimal wage contract, so that the benefit exceeds the linear portion that comes from the increased output. When the worker cannot search because V is too high, the firm value is linear in μ because the only effect is through expected output. In this region of V , the optimal wage is also flat with respect to μ .

Figure 5(d) shows that the firm value is also increasing in η . As mentioned earlier, there are two opposite effects of η on the firm value: the higher job separation rate and lower total compensation. Figure 5(d) shows that the latter exceeds the former effect so that the firm value is increasing in η . Similarly to μ case, the shape differs by the promised value V . When V is high so that the worker cannot search, the marginal benefit of an increase in η is proportional to $\frac{1}{r+\delta} \frac{dU(\eta)}{d\eta}$. Thus, the firm value function and $U(\eta)$ have a similar shape. When the worker can search for a job, the firm

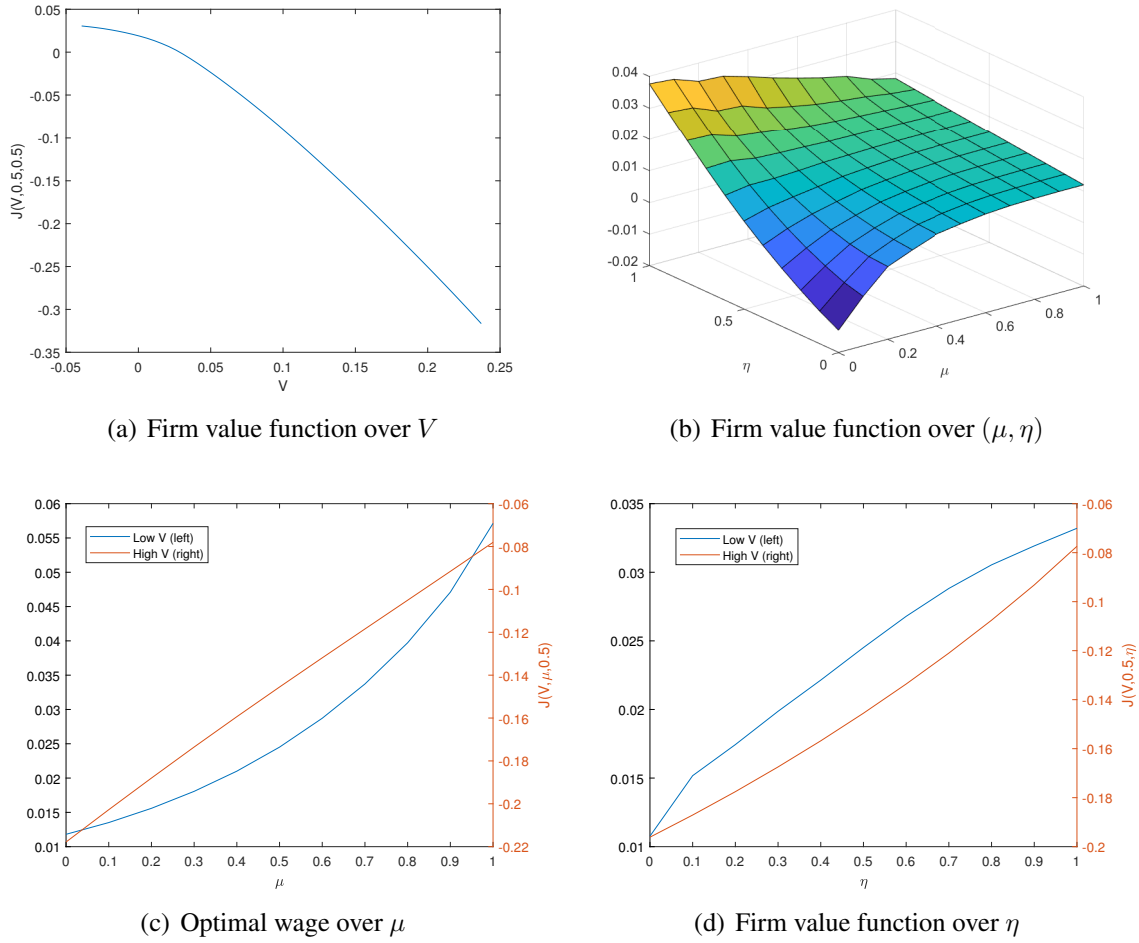
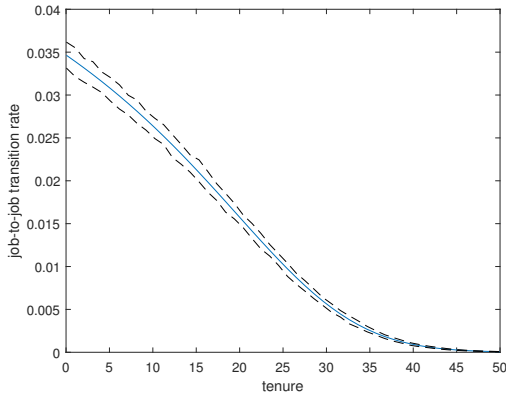


Figure 5: Firm value function $J(V, \mu, \eta)$

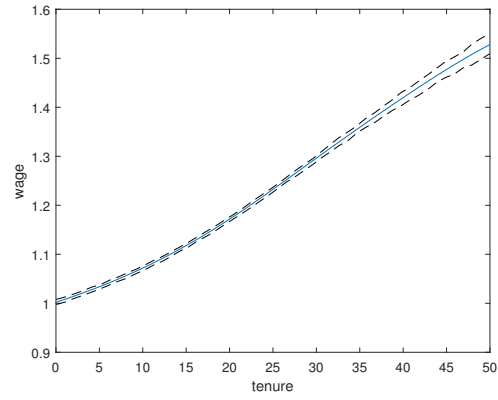
value change is a composite of two effects, search rate and compensation. It leads to a concave relationship between the firm value and η .

5.2 Simulated Economy

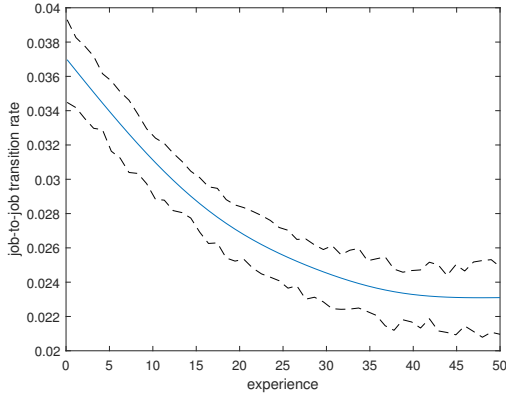
Table 2 summarizes important moments of the simulated economy. The unemployment rate and E–E transition rate are within a reasonable range. The median tenure of the employed is also close to the U.S data counterpart, which is close to 4 years. The model generates relative high wage dispersion in terms of mean-min ratio. Both uncertainty and on-the-job search contribute to the higher wage dispersion.



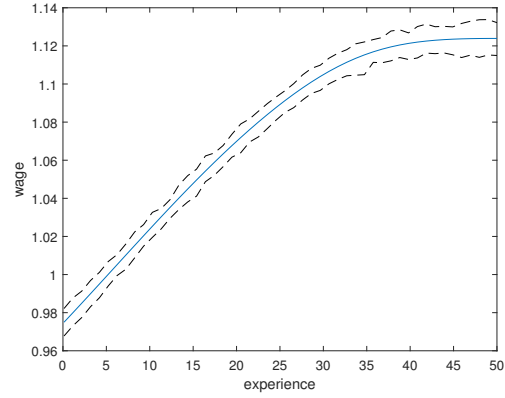
(a) Job-to-job transition rate over tenure



(b) Wage over tenure



(c) Job-to-job transition rate over experience



(d) Wage over experience

Figure 6: Effects of tenure and experience

The solid line is the kernel estimation, and dotted lines are 95% confidence intervals calculated by bootstrapping.

The model successfully generates well-known empirical regularities regarding tenure effects, such as declining job-to-job transition rate (Figure 6(a)) and increasing wage (Figure 6(b)). The model also generates life-cycle features although we do not introduce finitely lived agents, such as declining job-to-job transition rate over experience (Figure 6(c)), and concave wage profile (Figure 6(d)). It suggests that the resolution of uncertainty about ability plays an important role in shaping life-cycle patterns of wage and job transitions.

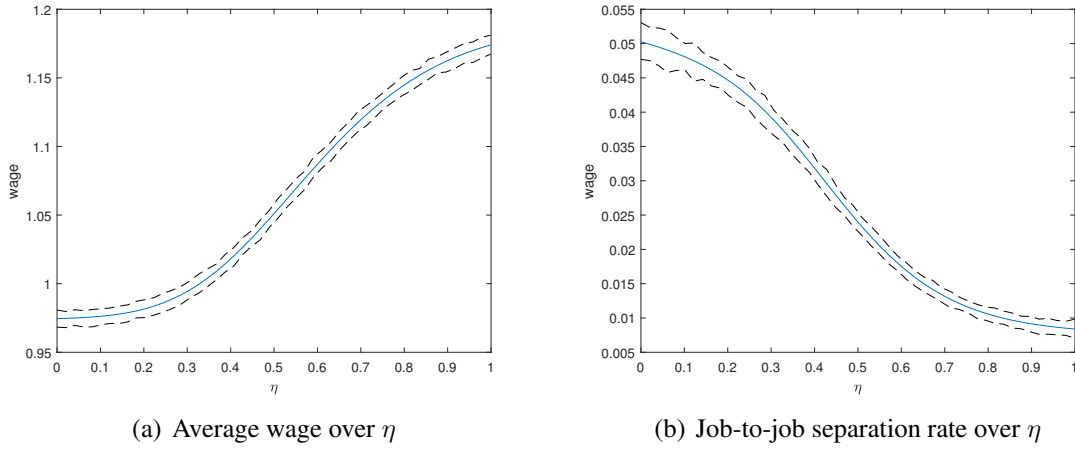


Figure 7: Worker ability and average wage, job-to-job transition

The solid line is the kernel estimation, and dotted lines are 95% confidence intervals calculated by bootstrapping.

5.2.1 Do better workers stay longer?

Figure 7 illustrates the kernel estimation of the average wage and job-to-job transition rate as a function of η at a specific point in time of the simulated economy. Figure 7(a) illustrates that better workers are on average paid well in equilibrium. Recall that the optimal wage policy itself is decreasing in η given V . Thus, Figure 7(a) implies that the positive association between η and V dominates the negative policy effect.

Figure 7(b) shows that better workers stay longer at a job. Two opposite forces indicate that the direction between job-to-job transition rate and worker ability can go either way. On the one hand, better workers have better outside opportunities (Figure 3(b)). On the other hand, better workers are paid better on average (Figure 7(a)). Figure 7(b) shows that the latter force dominates the former force. For this result, the optimal wage contract plays a crucial role. For instance, the relationship between worker ability and job-to-job transition becomes negative if wage contracts are restricted to a fixed-wage contract. It highlights the importance of endogenous payment schemes in shaping workers' transition dynamics.

Figure 7(b) explains why the previous number of job-to-job transitions is negatively correlated with the current wage (Light and McGarry, 1998; Carrillo-Tudela and Kaas, 2015). Figure 8

illustrates the average V and η over the number of job-to-job transitions. We count the previous job-to-job transitions uninterrupted by an unemployment spell. Standard on-the-job search models without uncertainty predict a positive relationship between V and job-to-job transitions regardless of contractual environments and search methods (Burdett and Mortensen, 1998; Burdett and Coles, 2003; Delacroix and Shi, 2006; Shi, 2009). When uncertainty presents and wages endogenously respond, the past job-to-job transitions is associated with lower value (Figure 8(a)) and ability (Figure 8(b)).

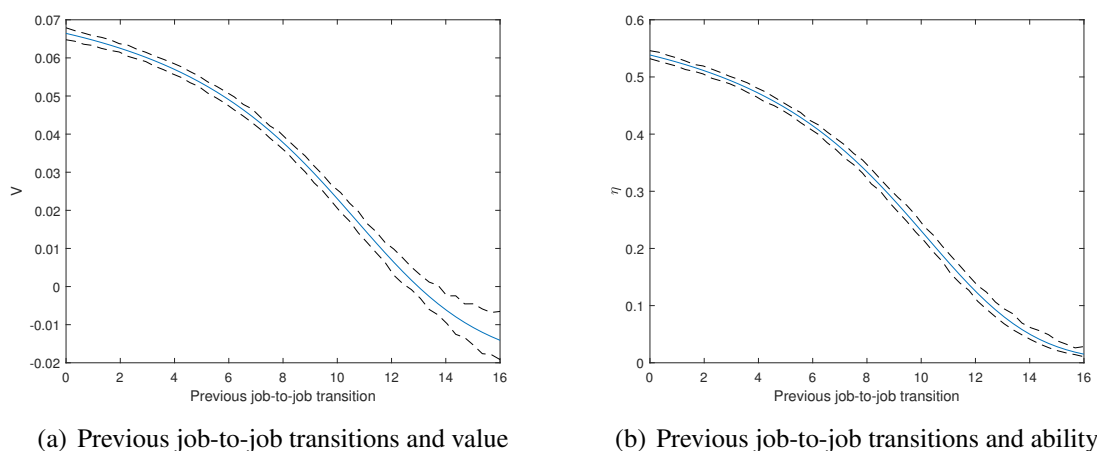


Figure 8: Previous job-to-job transitions, value, and ability

The solid line is the kernel estimation, and dotted lines are 95% confidence intervals calculated by bootstrapping.

The lower value and ability do not imply the lower wage because the optimal wage policy is decreasing in η . In the simulation, the number of job-to-job transitions and wage are negatively associated while the relationship is weaker than that of V or η (Figure ??).

To further examine the relationship between the number of job-to-job transitions and wage, I run a wage regression using the simulated data. The first column shows that the number of job-to-job transitions is negatively correlated with the current wage after controlling for tenure and labor market experience. This negative correlation comes from the worker's unobserved ability. Once η is controlled in the second column, the relationship becomes a positive. (Carrillo-Tudela and Kaas, 2015).

	wage	wage
N. of Job-to-job	−0.0025*(0.0012)	0.0125*** (0.0012)
η		0.1582*** (0.0060)
Tenure	0.0429*** (0.0006)	0.0395*** (0.0006)
Tenure*Tenure	−0.0007*** (0.0000)	−0.0007*** (0.0000)
Exp	0.0030*** (0.0005)	0.0034*** (0.0004)
Exp*Exp	−0.0000*** (0.0000)	0.0001*** (0.0000)
Constant	0.8282*** (0.0047)	0.7463*** (0.0053)
N	4743	4743
R^2	0.7151	0.7520

Table 3: Wage regression using simulated data

Note: Numbers in parentheses are standard deviations. * < 0.05, ** < 0.01, *** < 0.001.

5.2.2 Incidence of wage cut

Contrary to the model without uncertainty, this model allows a possibility of wage cut of job stayers. In the simulated data, the fraction of stayers whose wage falls at a point in time is 19.76%. We view that this magnitude is reasonable considering the data because [Elsby et al. \(2016\)](#) finds that the fraction is around 10% – 20% using CPS data. It is a bit smaller than [Park and Shin \(2017\)](#) reports from Korean data which is in range of 20% – 30%.

Note that the incidence of wage cut varies much by how one defines wage cut. For instance, if it is defined as a reduction of baseline wages, then the fraction of wage cut becomes much smaller, which is about 2% ([Grigsby et al., 2021](#)). Still, they report that 35% of workers experience the baseline wage unchanged, implying the real wage reduction.

	1(wage cut)
μ	−8.3193*** (0.3377)
η	3.9784*** (0.2846)
wage	3.3892*** (0.5106)
Tenure	−0.1878** (0.0647)
Tenure*Tenure	−0.0078 (0.0049)
Exp	0.0234 (0.0132)
Exp*Exp	−0.0005*** (0.0003)
Constant	−1.4682*** (0.3587)
N	4743

Table 4: Regression on the incidence of wage cut

Note: Numbers in parentheses are standard deviations. * < 0.05, ** < 0.01, *** < 0.001.

To examine the characteristics of workers who experience wage cut, we run a logit regression on the incidence of wage cut using the simulated data. The dependent variable is the dummy variable which takes value one if the worker's wage growth rate is negative, and explanatory variables are wage, tenure, experience, and beliefs. The results (Table 4) show that better workers (higher η) who are highly paid, but perform poorly in the current position (lower μ) are likely to experience wage cut. This is intuitive because the wage cut occurs when the worker has an outside opportunity, but the firm does not want to keep the worker.

How is this model result related to the empirical finding? While there are not many studies that systematically analyze the characteristics of workers who experience wage cut, [Park and Shin \(2017\)](#) documents that many establishments cut wages of a subset of workers, instead of a few establishment cut wages of all workers. It highlights an importance of match quality in the incidence of wage cut, which is consistent with this model. Also, [Park and Shin \(2017\)](#) documents that within an establishment, highly paid workers are more likely to experience wage cut, which is also consistent with the model prediction.

Note that the necessary condition behind the wage reduction is the presence of match quality uncertainty. The effect of worker ability uncertainty is a bit subtle. Even if workers are homogeneous, the wage cut can occur because the inside and outside values can differ, although the incidence of wage cut falls to 11% in this case. It does not mean the incidence of wage cut is increasing in ρ . In the extreme case when $\rho = 1$, there is no difference between the inside and outside values because $\mu = \eta$ always, meaning that no wage reduction occurs in equilibrium. It suggests that the incidence of wage cut is a hump-shaped with respect to ρ .

6 Conclusion

To be added

References

- Balke, Neele and Thibaut Lamadon**, “Productivity shocks, long-term contracts and earnings dynamics,” Technical Report, National Bureau of Economic Research 2020.
- Bernard, Benjamin**, “Continuous-Time Stochastic Games with Imperfect Public Monitoring,” *Theoretical Economics*, 2016, 11, 411–453.
- Blumen, Isadore, Marvin Kogan, and Philip J. Maccarthy**, *The Industrial Mobility of Labor as a Probability Process*, Cornell University, 1955.
- Burdett, Ken and Melvyn Coles**, “Equilibrium wage-tenure contracts,” *Econometrica*, 2003, 71 (5), 1377–1404.
- Burdett, Kenneth and Dale T Mortensen**, “Wage differentials, employer size, and unemployment,” *International Economic Review*, 1998, pp. 257–273.
- Carrillo-Tudela, Carlos and Leo Kaas**, “Worker mobility in a search model with adverse selection,” *Journal of Economic Theory*, 2015, 160, 340–386.
- Delacroix, Alain and Shouyong Shi**, “Directed search on the job and the wage ladder,” *International Economic Review*, 2006, 47 (2), 651–699.
- Doppelt, Ross**, “The hazards of unemployment: A macroeconomic model of job search and résumé dynamics,” *Unpublished working paper*, 2016.
- Elsby, Michael WL, Donggyun Shin, and Gary Solon**, “Wage adjustment in the Great Recession and other downturns: Evidence from the United States and Great Britain,” *Journal of Labor Economics*, 2016, 34 (S1), S249–S291.
- Fawcett, Kevin**, “Match quality, on-the-job search and wage-tenure contracts,” *Unpublished manuscript*, 2021.
- **and Shouyong Shi**, “Learning, On-the-Job Search and Wage-Tenure Contracts,” *Unpublished Working Paper*, 2018.
- Franceschelli, Ignacio, Sebastian Galiani, and Eduardo Gulmez**, “Performance pay and productivity of low-and high-ability workers,” *Labour Economics*, 2010, 17 (2), 317–322.
- Gonzalez, Francisco M and Shouyong Shi**, “An equilibrium theory of learning, search, and

- wages,” *Econometrica*, 2010, 78 (2), 509–537.
- Grigsby, John, Erik Hurst, and Ahu Yildirmaz**, “Aggregate nominal wage adjustments: New evidence from administrative payroll data,” *American Economic Review*, 2021, 111 (2), 428–71.
- Hom, Peter W, Loriann Roberson, and Aimee D Ellis**, “Challenging conventional wisdom about who quits: Revelations from corporate America.,” *Journal of Applied Psychology*, 2008, 93 (1), 1.
- Jardim, Ekaterina S, Gary Solon, and Jacob L Vigdor**, “How prevalent is downward rigidity in nominal wages? Evidence from payroll records in Washington State,” Technical Report, National Bureau of Economic Research 2019.
- Jovanovic, Boyan**, “Job matching and the theory of turnover,” *Journal of political economy*, 1979, 87 (5, Part 1), 972–990.
- , “Matching, turnover, and unemployment,” *Journal of political Economy*, 1984, 92 (1), 108–122.
- Lazear, Edward P**, “Performance pay and productivity,” *American Economic Review*, 2000, 90 (5), 1346–1361.
- Li, Fei and Xi Weng**, “Efficient learning and job turnover in the labor market,” *International Economic Review*, 2017, 58 (3), 727–750.
- Light, Audrey and Kathleen McGarry**, “Job change patterns and the wages of young men,” *Review of Economics and Statistics*, 1998, 80 (2), 276–286.
- Menzio, Guido and Shouyong Shi**, “Block recursive equilibria for stochastic models of search on the job,” *Journal of Economic Theory*, 2010, 145 (4), 1453–1494.
- **and** —, “Efficient search on the job and the business cycle,” *Journal of Political Economy*, 2011, 119 (3), 468–510.
- Moscarini, Giuseppe**, “Job matching and the wage distribution,” *Econometrica*, 2005, 73 (2), 481–516.
- Papageorgiou, Theodore**, “Learning Your Comparative Advantages.,” *Review of Economic Studies*, 2014, 81 (3), 1263–1295.

- Park, Seonyoung and Donggyun Shin**, “The extent and nature of downward nominal wage flexibility: An analysis of longitudinal worker/establishment data from Korea,” *Labour Economics*, 2017, 48, 67–86.
- Postel-Vinay, Fabien and Jean-Marc Robin**, “Equilibrium wage dispersion with worker and employer heterogeneity,” *Econometrica*, 2002, 70 (6), 2295–2350.
- Salamin, Alain and Peter W Hom**, “In search of the elusive U-shaped performance-turnover relationship: Are high performing Swiss bankers more liable to quit?,” *Journal of applied psychology*, 2005, 90 (6), 1204.
- Sannikov, Yuliy**, “Games with imperfectly observable actions in continuous time,” *Econometrica*, 2007, 75 (5), 1285–1329.
- Shi, Shouyong**, “Directed search for equilibrium wage–tenure contracts,” *Econometrica*, 2009, 77 (2), 561–584.
- Staudigl, Mathias and Jan-Henrik Steg**, “On repeated games with imperfect public monitoring: From discrete to continuous time,” *Journal of Dynamics and Games*, 2017, 4 (1), 1–23.

A Deriving the continuous-time limit equation

In general, the continuous-time HJB equation can be derived by the limit of the discrete-time model when $\Delta \rightarrow 0$. Following the same logic, recall the finite-time version of optimization problem [with a constant wage choice \$w\$ over the time period of \$\Delta\$](#) :

$$J(V, \mu, \eta) = \max_{\pi} E_{\pi} \left[(\lambda \alpha(\mu) y - w) \Delta + \frac{1}{1 + r \Delta} (\lambda \Delta \bar{J}(V^S, V^F, \mu, \eta) + Q(V^C, \eta) J(V^C, \mu, \eta)) \right] \quad (32)$$

where $\bar{J}(V^S, V^F, \mu, \eta) = \alpha(\mu) J(V^S, \mu^S, \eta^S) + (1 - \alpha(\mu)) J(V^F, \mu^F, \eta^F)$

$$Q(V^C, \eta) = 1 - (\lambda + \delta + s_e p_F(V^C, \eta)) \Delta$$

subject to

$$V = E_{\pi} \left[u(w) \Delta + \frac{1}{1 + r \Delta} (\lambda \Delta \bar{V} + \delta \Delta U(\eta) + s_e p_F(V^C, \eta) \Delta F(V^C, \eta) + Q(V^C, \eta) V^C) \right] \quad (33)$$

where $\bar{V} = \alpha(\mu) V^S + (1 - \alpha(\mu)) V^F$

Dividing both sides of (32) by Δ , we get

$$r J(V) = \max_{\pi} E_{\pi} \left[(\lambda \alpha(\mu) y - w) + \lambda (J(V^C) - \bar{J}) - (\delta + s_e p_F(V^C)) J(V^C) + \frac{J(V^C) - J(V)}{\Delta} \right] \quad (34)$$

where I omit the dependence of J on (μ, η) .

Let's denote the solutions to the above problem as $\pi := \pi(\Delta)$, $w := w(\Delta)$, $V^S := V^S(\Delta)$, and $V^F := V^F(\Delta)$ for each given $\Delta > 0$. Since the set of admissible (w, V^S, V^F, V^C) is compact in the Euclidean space, we know that the space of admissible π over the compact set of admissible (w, V^S, V^F, V^C) is pre-compact. Therefore, we could pick a convergent sub-sequence from the tight family of solutions $\{\pi(\Delta)\}_{\Delta > 0}$ with corresponding quadruples $\{(w(\Delta), V^S(\Delta), V^F(\Delta), V^C(\Delta))\}_{\Delta > 0}$. From now on, we will say that the limit of the convergent sub-sequence when $\Delta \rightarrow 0$ as the limits of solutions. For the simplicity, let's denote the limits of these solutions as π , w , V^S , V^F , and V^C , respectively, as $\Delta \rightarrow 0$. We know that this is same as

the solution π^* to the (13), that is, $w = w^*$.

The main difficulty of getting $\Delta \rightarrow 0$ limit is that $\frac{J(V^C(\Delta)) - J(V)}{\Delta}$ may not have a well-defined limit. As $\Delta \rightarrow 0$, $E_{\pi(\Delta)}[V^C(\Delta)] \rightarrow V$ from (33), but it does not guarantee $V^C \approx V$ for every realization of V^C . Indeed, $V^C(\Delta)$ does not converge to V point-wise as $\Delta \rightarrow 0$ when the firm randomizes over V^C . In this appendix, we will prove that the continuous-time limit equation is well-defined even though the firm randomizes over V^C .

Define $\epsilon(\Delta)$ as the difference between J evaluated at $E_{\pi(\Delta)}[V^C(\Delta)]$ and expected value of J over $\pi(\Delta)$.

$$\epsilon(\Delta) = J(E_{\pi(\Delta)}[V^C(\Delta)]) - E_{\pi(\Delta)}[J(V^C(\Delta))] \quad (35)$$

We will show that $\epsilon(\Delta)/\Delta \rightarrow 0$ when $\pi(\Delta)$ converges to a degenerate π , thus one can exchange the order of J and E when calculating the limit of $E_{\pi(\Delta)} \left[\frac{J(V^C(\Delta)) - J(V)}{\Delta} \right]$. By definition of $\epsilon(\Delta)$,

$$\left(\frac{1}{\Delta} - \lambda - \delta \right) \epsilon(\Delta) \leq s_e p_F(E_{\pi(\Delta)}[V^C(\Delta)]) J(E_{\pi(\Delta)}[V^C(\Delta)]) - E_{\pi(\Delta)}[s_e p_F(V^C(\Delta)) J(V^C(\Delta))] \quad (36)$$

Suppose the marginal distribution of V^C of π is degenerate at V . Then, the limit of RHS of Equation (36) is 0 because p_F and J are continuous. Moreover, $\epsilon(\Delta) \geq 0$ because J is weakly concave in V , therefore $\epsilon(\Delta)/\Delta \rightarrow 0$ as $\Delta \rightarrow 0$. It means that

$$\lim_{\Delta \rightarrow 0} \frac{E_{\pi(\Delta)}[J(V^C(\Delta))] - J(V)}{\Delta} = \lim_{\Delta \rightarrow 0} \frac{J(E_{\pi(\Delta)}[V^C(\Delta)]) - J(V)}{\Delta} = J_V(V) \mathcal{A}V \quad (37)$$

Consider the other case when π^* puts a positive measure on more than one element. We will first show that it is enough to consider that $\pi(\Delta)$ converges to a two-point lottery. To show that it

is the case, let's define an interim value function $\tilde{J}(V, \mu, \eta)$ by the following.

$$\begin{aligned} \tilde{J}(V, \mu, \eta) &= \max_{w, V^C, V^F, V^S} \left[(\lambda \alpha(\mu) y - w) \Delta + \frac{1}{1 + r \Delta} (\lambda \Delta \bar{J}(V^S, V^F, \mu, \eta) + Q(V^C, \eta) J(V^C, \mu, \eta)) \right] \\ \text{subject to} \\ V &= \left[u(w) \Delta + \frac{1}{1 + r \Delta} (\lambda \Delta \bar{V} + \delta \Delta U(\eta) + s_e p_F(V^C, \eta) \Delta F(V^C, \eta) + Q(V^C, \eta) V^C) \right] \end{aligned} \quad (38)$$

The difference from J and $\tilde{J}$ is that J allows to choose a general distribution over (w, V^C, V^F, V^S) , while $\tilde{J}$ only allow a degenerate distribution. By definition,

$$J(V, \mu, \eta) = \max_{\tilde{\pi}} \tilde{J}(V, \mu, \eta) \quad (39)$$

where $\tilde{\pi}$ is a distribution over V . Intuitively, π can be decomposed into two-step: first, randomize over V , and then choose (w, V^C, V^F, V^S) accordingly. Given this definition, $\tilde{J}$ is weakly decreasing in V . Therefore, the maximum of Equation (39) can be achieved by either a degenerate or two-point lottery $\tilde{\pi}$. Because each $\tilde{\pi}$ implies π in the original problem, it is loss of generality to assume that π^* puts a positive measure on at most two V^C .

Denote these two points by V_i for $i = 1, 2$. Consider the optimal contract problem [represented by \(18\) and \(19\) corresponding to determine \$J\(V_i\)\$ for given \$\Delta\$ and \$i = 1, 2\$](#) . Denote the optimal solution by $\pi^i(\Delta)$. Then, the limit of $\pi^i(\Delta)$ must have a degenerate distribution over $V^C(\Delta)$. If not, $J(V_i) > \tilde{J}(V_i)$ holds at the limit of $\Delta \rightarrow 0$, meaning that $J(V)$ can be increased by mixing V_{-i} and $\pi^i(\Delta)$, instead of using $\tilde{\pi}(\Delta)$. Because $\pi^i(\Delta) \circ \tilde{\pi}(\Delta)$ attains the maximum,

$$\lim_{\Delta \rightarrow 0} \frac{E_{\pi(\Delta)}[J(V_i^C(\Delta))] - J(V_i)}{\Delta} = \lim_{\Delta \rightarrow 0} E_{\tilde{\pi}(\Delta)} \left(\frac{E_{\pi^i(\Delta)}[J(V_i^C(\Delta))] - J(V_i)}{\Delta} \right) \quad (40)$$

$$= \lim_{\Delta \rightarrow 0} E_{\tilde{\pi}(\Delta)} \frac{J(E_{\pi^i(\Delta)}[V_i^C(\Delta)]) - J(V_i)}{\Delta} \quad (41)$$

$$= E_{\pi^*} (J_V(V^C) \mathcal{A} V^C) \quad (42)$$

Note that $\frac{J(E_{\pi^i(\Delta)}[V_i^C(\Delta)]) - J(V_i)}{\Delta}$ converges to $J_V(V_i) \mathcal{A} V_i$ as $\pi^i(\Delta)$ converges to a degenerate distri-

bution. Also, $\tilde{\pi}(\Delta)$ converges to the marginal distribution of V^C of π^* because the promise-keeping constraint converges to $V = V^C$. Therefore, the third equality holds.

It shows that the limit of Equation (34) is the following.

$$rJ(V) = E_{\pi^*} \left[(\lambda\alpha(\mu)y - w) + \lambda(J(V^C) - \bar{J}) - (\delta + s_e p_F(V^C))J(V^C) + J_V(V^C)\mathcal{A}V^C \right] \quad (43)$$

Note that π^* satisfies $E_{\pi^*}[V^C] = V$ and $E_{\pi^*}[J(V^C)] = J(V)$. Therefore,

$$\begin{aligned} rJ(V) &\leq \max_{\pi} E_{\pi} \left[(\lambda\alpha(\mu)y - w) + \lambda(J(V^C) - \bar{J}) - (\delta + s_e p_F(V^C))J(V^C) + J_V(V^C)\mathcal{A}V^C \right] \\ &\text{subject to } J(V) = E_{\pi}[J(V^C)], \quad V = E_{\pi}[V^C] \end{aligned}$$

On the other hand,

$$rJ(V) \geq E_{\pi} \left[(\lambda\alpha(\mu)y - w) + \lambda(J(V^C(\Delta)) - \bar{J}) - (\delta + s_e p_F(V^C(\Delta)))J(V^C(\Delta)) + \frac{J(V^C(\Delta)) - J(V)}{\Delta} \right] \quad (44)$$

for any π, Δ that satisfies the promise-keeping constraint. Given that π satisfies the promise-keeping constraint, the upper bound of the limit of RHS as $\Delta \rightarrow 0$ is $E_{\pi} \left[(\lambda\alpha(\mu)y - w) + \lambda(J(V^C) - \bar{J}) - (\delta + s_e p_F(V^C))J(V^C) + J_V(V^C)\mathcal{A}V^C \right]$ because $\epsilon(\Delta) \geq 0$. It means that

$$rJ(V) \geq E_{\pi} \left[(\lambda\alpha(\mu)y - w) + \lambda(J(V^C) - \bar{J}) - (\delta + s_e p_F(V^C))J(V^C) + J_V(V^C)\mathcal{A}V^C \right] \quad (45)$$

for any π , thus the inequality holds when the maximum is taken over π . It proves Equation (13).

$$\begin{aligned} rJ(V, \mu, \eta) &= \max_{\pi} E_{\pi} \left[(\lambda\alpha(\mu)y - w) + \lambda(\bar{J}(V^S, V^F, \mu, \eta) - J(V^C, \mu, \eta)) \right. \\ &\quad \left. - (\delta + s_e p_F(V^C, \eta))J(V^C, \mu, \eta) + \mathcal{A}J(V^C, \mu, \eta) \right] \\ &\text{subject to } J(V, \mu, \eta) = E_{\pi}[J(V^C, \mu, \eta)], \quad V = E_{\pi}[V^C] \end{aligned} \quad (46)$$