Job Mobility, Wage Dispersion, and Asymmetric Information

Jin Li*
Kellogg School of Management, Northwestern University

September 26, 2011

Abstract

This paper develops a model of job mobility and wage dispersion with asymmetric information. Contrary to the existing models in which the superior information of current employers lead to market collapse, this model generates a unique equilibrium outcome in which a) positive turnover exists and b) identical workers can be paid differently. The model implies that, in the presence of technological change that is skill-biased and also favors general skills over firm-specific skills, the wage distribution will become more spread out (corresponding to greater inequality) and job mobility will increase.

JEL Classifications: D82, J31, J63
Keywords: Asymmetric Information, Inequality, Turnover

*I thank Daron Acemoglu, Bengt Holmstrom, and particularly Robert Gibbons for invaluable discussions, encouragements, and supports. I also thank Josh Angrist, Ben Jones, Jeanne Lafortune, Kevin Lang, Jim Rebitzer, Peter Schnabl, Eric Van den Steen, Jose Tessada, as well as the seminar participants at Boston University, Case Western Reserve, Cornell, Kellogg, L.S.E., and M.I.T. for helpful comments and discussions. All remaining errors are mine.
1 Introduction

There are two well-documented empirical regularities about the labor market for young workers. First, young workers frequently change jobs. A typical U.S. worker holds seven jobs in the first ten years of his labor market experience; see for example Topel and Ward (1992). Second, the wage distribution of young workers widens over time, both unconditionally and conditionally (on the observables); see for example Farber and Gibbons (1996). Moreover, the wage dispersion has increased significantly over the past thirty years; see for example Autor, Katz, Kearney (2005).

Job mobility and wage dispersion have both been studied extensively and important insights have been gained on both questions.\(^1\) The existing models of job mobility, however, rarely discuss their implications for wage dispersion, nor do models of wage dispersion typically consider job mobility. This is surprising because models of either mobility or dispersion often share two common sets of assumptions. First, information about the abilities of workers is imperfect and is revealed over time. Second, workers can receive multiple offers from different firms at the same time. These assumptions reflect a set of underlying economic forces, namely imperfect information, learning, and multiplicity of wage offers, common to both job mobility and wage dispersion. Therefore, one might expect that patterns in job mobility and in wage dispersion are connected.

In this paper, we develop a single model that offers a comprehensive framework to study the joint determination of wage and job changes as a result of asymmetric learning. The model applies to asymmetric-information settings where the current employer learns more about a worker’s ability than prospective employers do.\(^2\) The superior information of the current employer creates a standard lemons problem. In this model, however, the presence of the lemons problem does not lead to market collapse. Instead, there exists a unique equilibrium allocation in which the current employer offers a wage equal to the average outside output of all types below the worker’s ability, and outside firms compete for the workers by using mixed strategies. The unique equilibrium allocation determines both the allocation of workers with heterogeneous abilities to different firms and how wages change when workers change jobs.

In the unique equilibrium allocation of this model, the current employer’s wage offer is strictly increasing in the worker’s ability. The outside firms randomize their wage offers between the minimal and the maximal wage offer by the current employer. This randomization implies that some workers will receive outside offers higher than the current employer’s offer. Consequently, endogenous turnover can arise in this model without requiring exogenous movers or differences in match qualities. Moreover, in this model the equilibrium turnover probability of a worker decreases continuously in worker’s ability, which stands in

---

\(^1\)Important models on job mobility include, for example, Jovanovic (1979), Greenwald (1986), Gibbons and Katz (1991), and Acemoglu and Pischke (1998,1999a,b). Important models on wage dispersion include, for example, Burdett and Judd (1983) and Burdett and Mortensen (1998). Mortensen (2001) offers an excellent survey of models on wage dispersion.

\(^2\)Several recent papers test the empirical implications of the asymmetric information in the labor market; see for example, Kahn (2008), Pinkston (2009), and Schoenberg (2007).
contrast to earlier asymmetric information models of turnover; see for example Greenwald (1986) and Gibbons and Katz (1991).

This model also generates wage dispersion for workers. The wage dispersion occurs at two levels. The first level says that different workers are in general paid differently because the current employer’s wage is increasing in worker’s ability. This prediction differs from earlier models of adverse selection; see for example Greenwald (1986), where all workers receive the same wage regardless of their abilities. The second level says that workers of identical abilities can also be paid differently because outside firms randomize in their wage offers so, for two identical workers, one may receive a high outside offer and the other may not. This second level of wage dispersion explain why similar workers are paid differently, which is an extensively studied question, see for example Burdett and Judd (1983) and Burdett and Mortensen (1998).

The setup of this model is similar to Greenwald’s (1986) classic asymmetric-information model, but with one important difference (described below). There are two periods, a single worker, and finitely many firms. At the beginning of period 1, firms compete for the worker using wage competition a la Bertrand. The worker picks a period 1 employer. At the end of period 1, the incumbent (period 1 employer) learns about the worker’s ability while outside firms do not. Moreover, the worker may accumulate firm-specific human capital, which makes him more productive at the incumbent in period 2. Fully aware of the informational and production advantage of the incumbent, all firms offer wage contracts simultaneously to the worker at the beginning of period 2.

The key difference between our model and Greenwald’s involves the timing of offers at the beginning of period 2. Greenwald assumes (as do many subsequent models) that the incumbent knows all the offers made by the outside firms before making a counteroffer. This offer timing makes the lemons problem in the second-hand labor market so severe that the market collapses when there are no exogenous movers or differences in match qualities. In our model, in contrast, we assume that the incumbent and outside firms make their offers simultaneously. This difference in timing alleviates the lemons problem and produces endogenous turnover and a non-degenerate wage dispersion.

This assumption about the simultaneity of offers has been used in the literature; see for example Waldman (1984). We believe this assumption is reasonable because employees typically cannot credibly communicate the total value of outside offers to the incumbent: even if the exact monetary values of the outside offers are known to the incumbent, it is unlikely that the worker’s non-monetary preferences can be known exactly. In other words, this difference, while stated formally in terms of the model’s timing, should actually be interpreted in terms of the information structure. For example, we obtain identical results

---

3In earlier asymmetric information models, the turnover probability of a worker typically takes a zero-one form: a worker leaves the firm with probability 1 if his ability is below certain threshold and otherwise leaves with the firm with probability 0.

4Most of the papers in this literature assume that the incumbent can make the counteroffer; see for example Greenwald (1986), Lazear (1986), Gibbons and Katz (1991), and Golan (2005). In an interesting paper that provides direct evidence on the prevalence of counteroffer, Barron, Berger, and Black (2006) report that 344 employers they survey will consider counteroffers, 437 report that they will not, 46 are unsure, and 5 refuse to answer the question regarding counteroffer.
even if the incumbent makes its offer after the outside firms do, as long as the incumbent does not know the outside offers.

The unique equilibrium allocation of this model leads to an explicit formula of turnover probabilities for workers of all ability levels. The turnover probabilities are determined by a ratio of the firm-specific output to the difference between the marginal and the average general output of worker. The firm-specific human capital of the worker reflects the incumbent’s production advantage; the difference between the marginal and the average general output of the worker reflects the incumbent’s information advantage because the incumbent’s wage offer equals the worker’s average general output. Therefore, we can interpret this ratio as a comparison of the incumbent’s profit from the worker’s firm-specific output to its profit from the worker’s general output. The turnover probability of workers becomes uniformly larger if this ratio becomes smaller for all ability levels. One case in which this can happen is when there is a technological change that is both log-skill-biased and general-skill-biased.

We also derive explicit formulas for the wage distributions of the movers, of the stayers, and of the two types combined. These formulas enable us to compare the wage distribution of the movers with that of the stayers. It is unclear ex ante which group has a higher average wage because the stayers are of higher average ability, while the movers are luckier in receiving higher outside wage offers. It turns out that the profit ratio plays a key role in the comparison. In particular, if the ratio is increasing (decreasing) in the worker’s ability, the wage distribution of the stayers first order stochastic dominates (dominated by) that of the movers, and thus the average wage of the stayers is higher (lower). If the firm-specific output is more sensitive to ability in larger firms, this suggests that the tenure effect is more likely to be observed in larger firms. Although we couldn’t compare the wage dispersion between the stayers and movers, we show that, when the firm-specific human capital is absent, the wage distributions of the movers and stayers are identical.

The explicit formulas for turnover probabilities and the wage distributions enable us to apply the model to study the joint evolution of within-group inequality and job mobility in the U.S. in the past 30 years. Recent increases in wage inequality in the U.S. have attracted attention from both the popular press and academic researchers. The popular press also suggests that job mobility has increased, although evidence from economic research is less clear cut. Nevertheless, it appears that job-to-job mobility has increased significantly; see for example Stewart (2002).

Many hypotheses for the cause of increases in wage inequality involve changes in technology. In Section 5, we take two such hypotheses seriously. First, we suppose that the technological change has been log-skill-biased, so it favors workers of higher abilities over those of lower abilities. Second, we suppose that the technological change has been general-skill-biased, so it favors general skills over specific skills in production. In the presence of such technological changes, our model predicts that: 1) the wage distribution becomes more spread out in the sense of Bickel and Lehmann (1979), which corresponds to greater

---

5 While job-to-job mobility is the only type of mobility considered in this paper, it is possible to extend the paper to allow for job-to-unemployment mobility. See the discussion on layoff decisions of the firm in the Section 6 for more details.
inequality; 2) job mobility of all types of workers increases; and 3) the proportionate increase in job mobility is larger for workers with higher levels of firm-specific human capital. These patterns are consistent with recent empirical evidence on changes in job mobility in the United States, which will be reviewed in Section 5.

The structure of this model is closely related to two classes of models. First, in an important class of macro-labor models, including Burdett and Judd (1983) and Burdett and Mortensen (1998), there is also a mixed strategy equilibrium where firms randomize in their wage offers and thus generate wage dispersion. In the Burdett-Judd-Mortensen models, workers are identical and the number of wage offers they receive follows an exogenous Poisson process. Firms randomize to trade off the probability of hiring a workers against the profit made from the worker. In equilibrium, all firms make the same positive expected profit. As the number of firms increases to infinity, the equilibrium allocation converges to the efficient competitive outcome.

In our model, in contrast, workers are heterogeneous and receive offers from all firms. Firms randomize to trade off the average quality of the workers hired against the wage paid to the workers, taking into account the productivity and selection effects that arise from worker heterogeneity and asymmetric information. In equilibrium, all outside firms make zero expected profit. As the number of firms increases to infinity, the equilibrium allocation remains inefficient.

Second, the analytical structure of this paper’s basic model is equivalent to that of a first-price auction with privately-informed bidders; see for example Engelbrecht-Wiggans, Milgrom, and Weber (EMW) (1983). While this type of auction has been studied in the auction literature, the implications of EMW have never been explored in the labor market context before. Moreover, the focus of the auction literature has been on common-value auctions. Here, we mainly examine the case where firm-specific human capital is present (corresponding to an auction in which the better informed bidder also has a higher value). We focus on how the comparison between the production advantage and the informational advantage of the incumbent can affect job mobility and the wage distribution. In Section 6, we illustrate how to apply the insights of EMW to various labor market issues, including training, layoff rules, career choice, minimum wages, inter-industry wage differentials, and wage changes from job changes.

In the rest of the paper, we proceed as follows. We set up the model in Sector 2. Section 3 solves the mixed equilibrium and proves its uniqueness (in outcomes). We derive the equilibrium turnover probabilities and wage distributions in Section 4. Section 5 explores the model’s predictions about job mobility and wage distribution in the face of technological changes. Section 6 concludes and discusses further applications of the model.

---

6Gabor (2007) considers a repeated common value first price auction with one-sided private information and considers its application in labor market. He shows that the bidders bid more aggressively in a repeated auction than in a static auction.
2 Model Setup

We set up the model formally in this section. Subsection 2.1 describes the model basics, including the types of players and their respective objective functions. Subsection 2.2 specifies the timing and information structure of the model and introduces notations for the strategies of the players. The solution concept of the model is given in Subsection 2.3.

2.1 Worker and Firms

There is a single worker who lives for two periods. The worker has ability \( a \), unknown at the beginning of period 1, that is drawn uniformly from \([0, 1]\). The worker is risk neutral, has no disutility of effort, and does not discount the future. His utility is given by

\[
u = w_1 + w_2,
\]

the sum of his wage incomes in the two periods. Note that in this setup, the same analysis can be carried out if there is a continuum of workers as long as the firm has constant return to scale in production.\(^7\)

There are \( N \) ex ante identical firms, \( 2 < N < \infty \). The payoff of each firm is the sum of its payoffs in the two periods. A firm’s period 1 payoff is 0 if it does not hire the worker. If it hires the worker, its period 1 payoff is

\[
\pi_1 = y(a, t) - w_1
\]

where \( y(a, t) \), the period 1 output of the worker, depends on the worker’s ability \( a \) and an index \( t \) that reflects the state of technology common to all firms. We assume that

\[
y(a, t) \geq 0, \quad \frac{\partial y(a, t)}{\partial a} > 0, \quad \frac{\partial y(a, t)}{\partial t} > 0 \quad \text{for all } a \text{ and } t,
\]

so the output is (uniformly) strictly higher if the worker is more able or the technology index is larger. The technology index \( t \) plays no role in the basic model analyzed in Section 3, but is central to the comparative static results presented in Section 4.

A firm’s period 2 payoff is 0 if it does not hire the worker. If it hires the worker, its period 2 payoff is

\[
\pi_2 = y(a, t) + 1_{\text{incumbent}} s(a, t) - w_2
\]

where \( 1_{\text{incumbent}} \) is an indicator function that takes the value of 1 if the firm is an incumbent (the worker’s period 1 employer) and 0 otherwise. In other words, a firm’s output equals \( y(a, t) + s(a, t) \) if it is an incumbent; its output is \( y(a, t) \) if it is an outside firm (i.e. a firm that does not hire the worker in period 1). We assume that

\[
s(a, t) \geq 0, \quad \frac{\partial s(a, t)}{\partial a} \geq 0, \quad \frac{\partial s(a, t)}{\partial t} \geq 0 \quad \text{for all } a \text{ and } t,
\]

\(^7\)The uniform assumption is without loss of generality because we can interpret the worker’s ability as his relative rank in the distribution.

\(^8\)When the production function displays decreasing return to scale and the workers are risk averse, Laing (1994) shows that involuntary layoffs can arise and may exceed efficient levels.

\(^9\)When \( N = 2 \), the existence of the equilibrium remains, but we lose the uniqueness.
so the $s(a, t)$ is weakly higher if the worker is more able or the technology index is larger. In this paper, we interpret $y(a, t)$ as the general output (from the worker’s general-purpose human capital) and $s(a, t)$ as the firm-specific output (from the worker’s firm-specific human capital).\(^{10}\)

We also assume that

$$y(0, t) + s(0, t) < \int_0^1 y(a, t) da \equiv E[y(a, t)],$$

so the lowest inside output is smaller than the average outside output. This is a standard assumption in the literature to rule out trivial cases; see for example Gibbons and Katz (1991).

2.2 Timing and Information Structure

At the beginning of period 1, all firms simultaneously offer contracts to the worker. The contracts are restricted to be nonnegative, non-contingent, single-period wage offers. We allow for the firms to play mixed strategies, so each firm can draw its wage offer according to its own choice of offer distribution.\(^{11}\) After all wage offers are made, the worker picks one firm from the $N$ wage offers it receives. Contrary to the existing literature, this setup does not restrict the worker’s behavior when a tie (multiple highest offers) occurs. Once the worker picks a firm, period 1 production takes place and the wage is paid. Through production, the incumbent observes the exact ability level of the worker. On the other hand, outside firms receives no information about the worker’s ability. These information assumptions are extreme and are made for simplicity. The analysis can be adapted to more a general setting, which assumes that the information set of the incumbent is finer than that of outside firms.

At the beginning of period 2, all firms simultaneously offer contracts to the worker. Since the incumbent observes the worker’s ability, its wage offer may depend on the worker’s ability. On the other hand, such dependence is not possible for outside firms. We again allow both the incumbent and the outside firms to play mixed strategies. After all offers are made in period 2, the worker picks one offer from the $N$ period 2 wage offers and works for the firm whose offer is chosen. The worker can randomize in his choice. Finally, after the worker chooses a period 2 employer, period 2 production takes place, the wage is paid, and the game ends.

Our assumption of the simultaneity of offers in period 2 is the key difference from earlier models (Greenwald (1986), Gibbons and Katz (1991)), which assumes that the incumbent

\(^{10}\)Other interpretations are also possible. For example, we can also interprete $s(a, t)$ as the match quality between workers and firms if we allow $s$ to be negative.

\(^{11}\)The restrictions on the available contracts are standard in the literature; see for example Greenwald (1986) and Gibbons and Katz (1991). In particular, the non-negativity assumption is made to rule out equilibria in which the wages accepted by the worker are not bounded below. The non-contingent assumption is made to fit with the assumption that outside firms cannot observe the output of the worker (so contracts based on outputs cannot be verified by courts). The single-period assumption is made to reflect the lack of commitment power of the firms (and the worker) and the associated lack of enforceability of long-term contracts.
observes all outside offers to the worker and can make counteroffers. This simultaneity assumption is a standard one in the literature; see for example Waldman (1984). In the current context, it captures that employees typically cannot credibly communicate the total value of outside offers to the incumbent: even if the exact monetary values of the outside offers are known to the incumbent, it is unlikely that the worker’s non-monetary preferences can be known exactly.\textsuperscript{12} In other words, this difference, while formally about the model’s timing, should actually be interpreted in terms of difference about information structure. For example, we obtain identical results even if the incumbent makes its offer after the outside firms, as long as the incumbent does not know the outside offers.

More formally, we summarize the sequence of events as follows:

1. At the beginning of period 1, firms simultaneously offer contracts to the worker. Each firm \( j (j \in \{1, \ldots, N\}) \) draws its offer from \([0, \infty]\) according to its choice of offer distribution \( G_j \in \Delta R^+ \), where \( \Delta R^+ \) is the set of probability distributions on non-negative real numbers.

2. The worker makes a decision \( D_1 : (R^+)^N \rightarrow \Delta\{1, \ldots, N\} \) (the set of probability distributions on \( \{1, \ldots, N\} \)) to choose a firm from the \( N \) wage offers. Production takes place and wage is paid. The incumbent observes the ability of the worker while outside firms don’t.

3. At the beginning of period 2, all firms simultaneously offer contracts to the worker. For each firm \( j \in \{1, \ldots, N\} \), if firm \( j \) hires the worker in period 1, it chooses a wage offer \( w_j : [0, 1] \rightarrow \Delta R^+ \) (the set probability distribution on nonnegative real numbers) based on the worker’s ability. If firm \( j \) does not hire the worker in period 1, it draws its offer randomly from \([0, \infty]\) according to its choice of offer distribution \( F_j \).

4. The worker makes a decision \( D_2 : (R^+)^N \rightarrow \Delta\{1, \ldots, N\} \) to choose a firm from the wage offers. Production takes place and the wage is paid.

\textbf{2.3 Perfect Bayesian Equilibrium}

According to the timing and information structure, the strategy of the worker is a 2-tuple \((D_1, D_2)\), and the strategy of firm \( j \in \{1, \ldots, N\} \) is a 3-tuple \((G_j, F_j, w_j)\). Given the strategies, we solve the Perfect Bayesian Equilibrium (PBE) of the model. The PBE requires that the strategies of the worker and the firms to be sequentially optimal given their beliefs and that their beliefs are determined from Bayes Rule wherever possible. In particular, we require the solution to satisfy

1. The worker’s equilibrium period 2 contract choice \( D_2^* \) are optimal given any period 1 strategy \( D_1 \) of the worker and any strategies of the firms \( \prod_{j=1}^{N} (G_j, F_j, w_j) \).

\textsuperscript{12}Note that we do not model the wage bargaining process between the firm and the worker. Golan (2009) shows that wage dispersion can occur for workers with the same observables when there is asymmetric information. Her insight is that firms are willing to offer higher wages to workers of higher productivities since the failing to reach an agreements with such workers are more costly.
2. For each firm $j$, its period 2 strategy $(F_j^*, w_j^*)$ is optimal given any period 1 strategy $D_1$ of the worker, any period 1 strategy of the firms $\prod_{j=1}^N (G_j)$, period 2 equilibrium strategy of the worker $D_2^*$, and period 2 strategies of all other firms $(F_{-j}^*, w_{-j}^*)$.

3. The worker’s choices $(D_1^*, D_2^*)$ are optimal given the strategies of the firms: $\prod_{j=1}^N (G_j, F_j^*, w_j^*)$.

4. For each firm $j$, its strategy $(G_j^*, F_j^*, w_j^*)$ is optimal given the worker’s equilibrium strategy $(D_1^*, D_2^*)$ and the equilibrium strategies of all other firms: $(G_{-j}^*, F_{-j}^*, w_{-j}^*)$.

5. For each firm $j$, at the beginning of period 1, its belief about the worker’s ability is the prior distribution of the worker’s ability. In period 2, if $j$ is the incumbent, it knows the exact ability of the worker. Otherwise, $j$’s belief is the prior distribution of the worker’s ability. The worker does not know his ability in period 1. The worker knows his ability in period 2.

3 **Equilibrium of the Model**

We solve the equilibrium of the model in this section. First, we show that the model does not have a pure strategy PBE. Next, we show in Theorem 1 that there exists a mixed strategy PBE. Finally, we show in Theorem 2 that the equilibrium is essentially unique: every equilibrium of the model leads to the same outcome in job mobility and the wage distribution. All of the proofs are in the appendix.

For ease of exposition, we drop the technology index $t$ in this section to write $y(a)$ and $s(a)$ instead of $y(a, t)$ and $s(a, t)$ because $t$ plays no role in establishing the equilibrium. In addition, since all firms are ex ante identical, we let firm 1 be the incumbent in period 2 and write its equilibrium wage offer as $w_{1n}(a)$.

Before describing the mixed strategy PBE, our first observation, reported in Lemma 1, is that the model does not have a pure strategy PBE. The absence of pure strategy equilibrium here stands in contrast with earlier results; see for example Greenwald (1986) and Gibbons and Katz (1991). The difference arises because the incumbent in this model cannot make counteroffers. When counteroffers are allowed, the incumbent has the opportunity to match outside offer when the worker is worth more than the outside offer. Such matching exacerbates the adverse selection problem and helps sustain a low wage equilibrium by discouraging the outside firms from making offers. When the incumbent cannot use counteroffers to respond to outside offers, such low wage equilibrium is no longer sustainable because outside firms can deviate by raising wage.

**Lemma 1** There is no pure strategy PBE.
Although no pure strategy PBE exists in this model, Theorem 1 constructs a mixed strategy PBE. Theorem 1 determines a) the wage offer made by the incumbent and b) the distribution of wage offers made by outside firms, which are the two building blocks to our analysis of the turnovers and the wage distribution in Section 4. We will start by describing these two building blocks and giving intuitions for them. Since the equilibrium involves mixed strategy, it implies that 1): the offer by the incumbent in period 2 makes the outside firms willing to randomize in their offers, and 2): the randomization of the outside firms in period 2 makes the incumbent’s offer optimal. The description of the incumbent’s offer can be simplified by the following definition.

**Definition 1:** The average outside output\(^{13}\) of a worker of ability \(a\) is defined as

\[
B(a) = \frac{\int_0^a y(x)dx}{a} = E[y(x)|x \leq a].
\]  

(7)

The first building block from Theorem 1 states that in period 2 the incumbent offers a wage that equals the average outside output of the worker. More formally,

\[
w_{In}(a) = B(a), \text{ for all } a > 0.
\]  

(8)

It is easy to see that the incumbent’s offer \(B(a)\) is strictly increasing in worker’s ability. More importantly, by offering \(w_{In}(a) = B(a)\), the incumbent makes the outside firms indifferent in their wage offers because the expected payoff of an outside firm by offering any \(w \in [0, E[y(a)]]\) is always 0. To see this, suppose \(y(a) = a\), then \(B(a) = \frac{1}{2}a\). Now suppose an outside firm offers a wage of \(w \in [0, \frac{1}{2}]\). If the outside firm does not hire the worker, its profit is zero. If it manages to hire the worker, the worker’s ability must satisfy \(a \leq 2w\) (because \(w_{In}(a) = B(a) = \frac{1}{2}a\)). Recall that the ability is uniformly distributed, so the expected output of the worker is \(w\). Therefore, the expected profit of the firm is also zero when it hires the worker. Theorem 1 generalizes the logic up to the case when the output of a worker does not equal to his ability.

The second building block of Theorem 1 describes the pattern of randomization of outside firms. To simplify the description, we introduce the following definition.

**Definition 2:** The maximum outside offer distribution \(F\) is defined as

\[
F(w) = \prod_{j=2}^{N} F_j^*(w),
\]  

(9)

where \(F_j^*(w)\) is the probability that firm \(j\) choosing an offer less than or equal to \(w\) in equilibrium. For the maximum outside offer to be less than \(w\), it must be that the offer from each outside firm is less than \(w\), so the maximum outside offer distribution is the multiplication of the equilibrium offer distribution of all outside firms. Also define the "boundary" of support of \(F\) as

\[
\frac{w}{w} = \inf\{w : F(w) > 0, w \geq w_{In}(a) \text{ for some } a\};
\]  

\[
\frac{w}{w} = \sup\{w : F(w) < 1\}.
\]  

\(^{13}\)We omit the modifier "below his ability" in the definition for simplicity. This also draws parallel with the average cost term in price theory.
The concept of the maximum outside offer plays an important role in our analysis below because its distribution summarizes all the relevant information of the outside offers from the incumbent’s point of view: the incumbent keeps a worker if and only if the maximum outside offer is less than its offer.

The second building block of the equilibrium states that in period 2 outside firms randomize their offers in \([0, E[y(a)])\) so that the maximum outside offer distribution \(F\) satisfies

\[
\frac{1}{y(a) + s(a) - B(a)} = \frac{f(B(a))}{F(B(a))}, \quad \text{for all } a > 0. \tag{11}
\]

This pattern of randomization implies that when the distribution of maximum outside offer satisfies (11), the incumbent finds it optimal to offer a wage that equal to the average outside output.

To see this, the incumbent’s expected payoff in period 2 by offering \(w\) to a worker of ability \(a\) is given by

\[
F(w)(y(a) + s(a) - w), \tag{12}
\]

where \(y(a) + s(a) - w\) is its profit for keeping the worker and \(F(w)\) is the probability that it keeps him. Profit maximization implies that the incumbent’s optimal wage choice \(w^*\) must satisfy

\[
\frac{1}{y(a) + s(a) - w^*} = \frac{f(w^*)}{F(w^*)}. \tag{13}
\]

Now note that equation (11) is identical to the equation above, except \(w^*\) is replaced with \(B(a)\). This means that when (11) holds, it is optimal for the incumbent to offer \(w^* = B(a)\), the wage policy described in the first feature.

In summary, when the incumbent offers \(w_{In}(a) = B(a)\), any wage offer in \((0, E[y(a)])\) is an optimal response for each outside firm. When the maximum offer distribution satisfies the differential equation (11), the incumbent finds it optimal to offer \(w_{In}(a) = B(a)\). Therefore, this is an equilibrium strategy profile in period 2. To fully describe a PBE, we specify in Theorem 1 below the period 1 strategies and beliefs.

**Theorem 1:** The following strategies and beliefs form a PBE:

(i) In period 2, the worker chooses the maximum wage offer. If there are multiple maximum offers, the worker a) stays with the incumbent if its offer is one of the maximum offers; b) randomize otherwise.

(ii) At the beginning of period 2, the incumbent firm offers

\[
w_{In}(a) = B(a) \quad \text{for all } a, \tag{14}
\]

and each outside firm \(j (j \in \{2, \ldots N\})\) offers a wage drawn independently from the distribution

\[
F^*_j(w) = F(w)^{\frac{1}{N-1}}, \tag{15}
\]

where

\[
\begin{align*}
F(w) &= 0 \quad \text{for } w < \underline{w}; \\
F(w) &= C \exp \left( \int_{\underline{w}}^{w} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} \, dx \right) \quad \text{for } w \in [\underline{w}, \overline{w}]; \\
F(w) &= 1 \quad \text{for } w > \overline{w},
\end{align*}
\]
\[ [w, \bar{w}] = [y(0), E[y(a)]] \] and \( C = \exp \left( -\int_{w}^{\bar{w}} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} dx \right). \]

(iii) In period 1, the worker chooses the maximum wage offer. If there are multiple maximum offers, the worker randomizes among them.

(iv) At the beginning of period 1, all firms offer
\[ E[y(a)] + \int_{0}^{1} F(B(a))(y(a) + s(a) - B(a)) da. \] (17)

(v) Each firm’s belief about the worker’s ability equals the prior if it has not hired the worker. The incumbent knows the worker’s ability at the end of period 1. The worker does not know his ability in period 1. The worker knows his ability in period 2.

One prominent feature of the equilibrium is the randomization of wage offers. The randomization of wage offers is fundamental to our results because job mobility and wage dispersion follow directly from it. However, it may appear that the randomization is unappealing because in real life firms do not randomize their wage offers. Nevertheless, it can be shown that the wage distribution from the randomization is equivalent to a modified game in which outside firms independently form estimates of the ability of the worker (say through job interviews) and make nonrandomized wage offers contingent on the estimates.

In Theorem 1, the offer distribution of each outside firm is identical with \( F_j^* = F \frac{\pi^j}{N} \) for all \( j \in \{2, ...N\} \). In general, any set of outside offer distributions that satisfy \( \prod_{j=2}^{N} F_j^*(w) = F \) can be sustained as an equilibrium and there are infinitely many of them. Therefore, the PBE in Theorem 1 is just one of the infinitely many PBEs of the game. However, all PBEs of the model, as shown by Theorem 2 below, share the same two properties. First, the incumbent’s wage offer must equal the average outside output. Second, the maximum outside offer distribution is unique and must satisfy the differential equation in (11). Note that the worker’s wage and his mobility decision is completely determined by the incumbent’s wage offer and the highest outside wage offer. Therefore, the identical incumbent offer and maximum outside wage offer distributions across all equilibria implies that the worker’s mobility and wage distribution is also identical across all equilibria. In other words, this model has a unique equilibrium allocation in terms of job mobility and the wage distribution.

**Theorem 2:** In each PBE, all outside firms have zero expected profits in period 2. The incumbent’s offer must satisfy \( w_{1n}(a) = B(a) \) for all \( a \in (0, 1] \), and the distribution of the maximum outside offer must satisfy
\[
\begin{align*}
F(w) &= 0 \quad \text{for } w < \frac{w}{\bar{w}}, \\
F(w) &= C \exp \left( -\frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} \int_{w}^{\bar{w}} dx \right) \quad \text{for } w \in [w, \bar{w}], \\
F(w) &= 1 \quad \text{for } w > \bar{w},
\end{align*}
\]
where \( [w, \bar{w}] = [y(0), E[y(a)]] \) and \( C = \exp \left( -\int_{w}^{\bar{w}} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} dx \right) \).
4 Job Mobility and Wage Dispersion

In this section, we apply the unique equilibrium allocation to study job mobility and wage dispersion. We emphasize on results that help us study how technological changes affect job mobility and wage inequality. In Subsection 4.1, we derive explicit formulas for turnover probability of workers of all ability levels (Proposition 1). We show that the turnover probability is completely determined by a profit ratio that compares the incumbent’s production advantage through the worker’s firm-specific human capital with its informational advantage over the worker’s general human capital. In Subsection 4.2, we obtain explicit formulas for the wage distributions of the stayers, the movers, and the two types combined. We compare the wage distribution of the stayers with that of the movers and give sufficient conditions on when the average wage of stayers is higher or lower.

Before proceeding to the next subsection, we make two remarks here. First, we use the word “workers” in our discussion as if the model had multiple workers. We interpret the turnover probability of the worker as the proportion of workers who move to outside firms and \( F(w) \) as the proportion of workers whose maximum outside wage offer is less than or equal to \( w \). Second, we re-include in this section the technology index \( t \) into the expressions and write \( y(a, t) \), \( s(a, t) \), \( B(a, t) \) to study the role of technology on turnover probability and wage distribution.

4.1 Turnover Probability

We study the turnover probability of the worker in this subsection. The formulas we derive on turnover not only describes the overall turnover pattern, but also allows us to carry out comparative statics on the likelihood of turnovers. Because the formulas are easier for the probability that a worker stays with the firm, we state the results in terms of staying probability.

Definition 3: Let \( P(a, t) \) be the equilibrium probability that a worker of ability \( a \) stays with the incumbent in period 2 when the technology level is \( t \).

Proposition 1: For all \( a \) and \( t \), let

\[
    r(a, t) = \frac{s(a, t)}{y(a, t) - B(a, t)},
\]

then

\[
    P(a, t) = \exp(-\int_a^1 \frac{1}{x(1 + r(x, t))} dx).
\]

The formula of turnover probability in this model differs qualitatively with results in the literature; see for example Greenwald (1986) and Gibbons and Katz (1991). In earlier asymmetric information models, the turnover probability of a worker typically takes a zero-one form: a worker leaves the firm with probability 1 if his ability is below certain threshold and otherwise leaves with the firm with probability 0. Here, the turnover probability decreases continuously with the worker’s ability and only the highest ability worker has a zero turnover probability.
Moreover, the turnover probability depends completely on \( r(a,t) \), which has a natural economic interpretation as profit ratios. The numerator of \( r(a,t) \) is equal to the incumbent’s "profit" from keeping the worker had the ability of the worker been known. In other words, the numerator reflects the incumbent’s "profit" from its production advantage because of the firm-specific human capital. The denominator of \( r(a,t) \) is equal to the incumbent’s "profit" from keeping the worker had there been only general human capital. In other words, the denominator reflects the incumbent’s profit source from its informational advantage on the worker’s general output. This profit ratio \( r(a,t) \) of production over information advantage is key to the comparative statics results on technological changes.

For example, when there is a technological change that favors high ability worker over low ability worker in terms of their general outputs but leaves their firm-specific outputs unchanged, the profit ratio can be shown to become smaller and thus turnover increases. Corollary 1 states this formally.

**Corollary 1:** If \( \frac{\partial y(a,t)}{\partial a} > 0 \) and \( s(a,t_1) = s(a,t_2) \) for all \( a \) and \( t \), then for \( t_1 < t_2 \), \( P(a,t_1) > P(a,t_2) \) for all \( a > 0 \).

### 4.2 Wage Distribution

In period 2, the worker either leaves the incumbent and becomes a mover, or continues to work for the firm and becomes a stayer. We characterize the wage distributions of the movers, of the stayers, and of the two types combined in this subsection. We also compare the wage distribution of the movers with that of the stayers and provide sufficient conditions on when one dominates the other in the sense of First Order Stochastic Dominance (FOSD). In the special case where the firm-specific output is missing, we show that the two distributions are identical. We first introduce the notations for wage distributions of different types.

**Definition 4:** Let \( G(w) \) be the wage distributions of the stayers and movers combined; let \( G_S(w) \) be the wage distribution of the stayers; let \( G_M(w) \) be the distribution of the movers.

Proposition 2 derives the wage distributions of the three types. To avoid complicated expressions, the formulas are written in terms of the average outside output \( (B(a,t)) \) and the staying probability \( (P(a,t)) \).

**Proposition 2:**

\[
\begin{align*}
G(B(a,t)) &= aP(a,t); \\
G_S(B(a,t)) &= \frac{\int_0^a P(x,t)dx}{\int_0^1 P(x,t)dx}; \\
G_M(B(a,t)) &= \frac{\int_0^a (1 - P(x,t))dx}{\int_0^1 (1 - P(x,t))dx}.
\end{align*}
\]
The expressions of wage distributions in Proposition 2 enables us to compare the wage distribution of the movers with that of the stayers. This comparison reflects two conflicting forces. On the one hand, stayers on average have higher abilities because the incumbent’s offer is increasing in ability. On the other hand, movers on average are luckier in receiving outside offers because for two workers of the same ability (so they receive the same offer from the incumbent), the mover must have received a better outside offer. Therefore, the comparison of the mover-stayer wage distribution sheds light on the source of wage growth from learning (by the incumbent) and luck (from receiving good outside offers).

The key to this comparison is again the profit ratio $r(a,t)$ in (18). Corollary 2 shows that when this ratio is increasing in ability, the wage distribution of the stayers FOSD that of the movers, so the stayers have a higher average wage. And if this ratio is decreasing, the wage distribution of the movers FOSD that of the stayers, so the movers have a higher average wage.

**Corollary 2:** The wage distribution of the stayers FOSD that of the movers if $r(a,t) = \frac{s(a,t)}{y(a,t)-B(a,t)}$ is increasing in $a$; the wage distribution of the movers FOSD that of the stayers if $r(a,t) = \frac{s(a,t)}{y(a,t)-B(a,t)}$ is decreasing in $a$.

In this model, a positive mover-stayer wage gap corresponds to a positive tenure effect. Corollary 2 suggests that the tenure effect is positive if the ratio of firm-specific output to general output is increasing in ability. If more able worker can better leverage their abilities in larger firms possibly because a larger size allows better matching (of ability to specific positions), then Corollary 2 suggests that we are more likely to observe a positive tenure effect in larger firms, a well documented empirical regularity.

It will also be interesting to know how the dispersion of the wage distribution of the movers compares with that of the stayers. Unfortunately, we do not have general results for such comparison. However, when there is no firm-specific output ($s(a,t) \equiv 0$), Corollary 3 shows that the wage distributions of the stayers and the movers are identical and the worker moves to an outside firm with ex ante probability $\frac{1}{2}$.

**Corollary 3:** If $s(a,t) \equiv 0$, then $P(a,t) = a$ for all $a$, and $G_s(w) = G_m(w)$ for all $w$. Therefore,

$$\int_0^1 P(a,t)da = \frac{1}{2}.$$  \hspace{1cm} (22)

5 Inequality and Job Mobility

This section applies the predictions of the model on turnovers and wage distributions in Section 4 to shed light on the joint evolution of wage inequality and job mobility in the U.S. in the past 30 years. In Subsection 5.1, we review the facts and theories on wage inequality and technological changes. In addition, we discuss why a technological change that favors general skills may lead to a decrease in job mobility in symmetric learning.
models. In Subsection 5.2, we show that if technological changes are skill-biased and also favor general skills, this model predicts that the wage distribution will become more spread out (corresponding to greater inequality) and job mobility will rise (Theorem 3 and Theorem 4). Moreover, the proportional increase in job mobility will be larger for workers with more firm-specific human capital (Theorem 5). These predictions are broadly consistent with the empirical findings of Stewart (2002).

5.1 Background Facts and Theories

Wage inequality in the U.S. has increased substantially in the past 30 years (Bound and Johnson (1992), Katz and Murphy (1992), Murphy and Welch (1993)). At least half of the increase in wage inequality results from the rise in residual inequality, the dispersion of wages in observationally equivalent groups (Juhn, Murphy, and Pierce (1993)). Moreover, much of the change in within-group inequality appears to concentrate on the top end of the distribution. For example, Lemieux (2006) shows that the within-group inequality has increased for the college-educated and but has changed little for other groups since the 1990s. Autor, Katz, and Kearney (2005) also report that the within-group inequality has continued to rise after the 1990s for the 90/50 wage ratio while it has declined for the 50/10 ratio.

One hypothesis attributes the increase in within-group inequality to skill-biased technology change (SBTC). SBTC states that, if we order workers by their abilities, the recent technological change has favored high-ability workers and raised the relative demand for them (Berman, Bound, and Griliches (1994), Autor, Katz, and Krueger (1998)). The computer and internet revolution, together with associated technological and organizational changes, lend direct support to this hypothesis.

It has been recognized, however, that ability is multidimensional (Gardner (1983)). Therefore, ordering workers in a one dimensional ability is an oversimplification that sometimes does not capture the complexity and the full impacts of technological changes (Acemoglu (2002)). For example, different industries and occupations require different mixes of ability types, so the effect of technological change may differ across sectors (where a sector here can either be an industry or an occupation).

One approach to capture the effect of technological change on different sectors is to decompose a worker’s ability into a general component and a specific component and to use the specific component to reflect the worker’s ability specific to that sector. Casual empiricism suggests that the technological innovation has become more frequent over the years and changes in production has favored workers who are quick to learn and are more flexible, i.e. those with higher general ability. In the same vein, sociologists have argued that aptitude, the general capacity for learning, has become increasingly valuable; see for example Sennett (2003).

We take seriously the hypothesis that technological change has made general ability more important in production and study its implication on the labor market in the next subsection. The validity and consequence of this hypothesis has been studied by Gould
Gould (2002) examines a two-sector model of comparative advantage a la Roy (1951). In his model, this general-skill-biased technological change (GSBTC) is modeled as an increase in the correlation of abilities in the two sectors. When the abilities in the two sectors become more correlated, wage inequality increases because the insurance role of having multiple sectors is now diminished. Gould tests this increasing correlation assumption using CPS data by decomposing the whole economy into three sectors (professional, service, and blue-collars). He shows that the ratio of the variance of log-wages within each sector to the variance of log-wage of the whole economy has decreased and the implied abilities in these sectors have become more correlated over the years.

The increased correlation of abilities across sectors naturally affects the job mobility of workers. When abilities required in different sectors become more similar, one might think that mobility between jobs should become larger because skills are now more substitutable. This view focuses on the labor demand side of job changes. If the abilities required by different firms are more substitutable, then a firm with a positive labor demand shock can more easily find workers from other firms. Taking this view, Kambourov and Manovskii (2008) argue that the increases in variability of productivity shocks to occupations can explain the increases in between-occupation mobility in the U.S.

It has been argued, however, that an important source of job mobility comes from labor supply reasons that involve firm-worker matching; see for example Jovanovic and Moffitt (1990). By embedding Gould’s idea into a simplified turnover model of Jovanovic (1979), we find that GSBTC decreases, rather than increases turnovers. This is because the increased correlation of abilities between sectors implies that the productivity of the worker is more likely to be similar across sectors, thus the return of switching into a new sector becomes smaller. For example, suppose there are two jobs: farming and hunting. If abilities used in farming and hunting are independent, then a bad farmer is as likely to be a good hunter as a good farmer. Now if abilities in these two sectors become correlated, then a bad farmer is more likely to be a bad hunter. This reduces the incentive for the bad farmer to change jobs so over all job mobility becomes smaller.

It is worthwhile to note that in Jovanovic (1979) and subsequent symmetric information models on turnover, the absolute level of general ability does not affect job mobility. In symmetric information models, a worker’s turnover decision is determined by differences in his expected outputs (which are equal to the wage offers) at different firms. Because general ability is a component of output common to all firms, changes general ability do not affect the relative productivity of the worker in different firms and thus do not affect the worker’s turnover decision. In contrast, the absolute level of general ability matters in job mobility in this model. As is discussed in Subsection 4.1, the turnover probability in this model depends on the ratio of the firm-specific output to the difference between marginal and average general output, which relates to the absolute level of general ability. In Subsection 5.2, we show that GSBTC decreases this ratio and hence increases turnover and discuss the relevant empirical evidence on turnover.
5.2 Implications for Inequality and Job-to-Job Mobility

In this subsection, we show formally how the predictions of our model relate with the empirical findings on wage inequality and job mobility. Theorem 3 below shows that if a technological change is log-skill-biased and general-skill-biased, then job mobility increases and the wage distribution becomes more spread out in the sense of Bickel and Lehmann (1979). Theorem 4 shows that if general output is sufficiently important in production, and if a technological change is skill-biased, then the increase in job mobility is larger for workers with higher levels of firm-specific human capital. Before stating these theorems, we first define the different types of technological changes.

**Definition 5:** A technological change is skill-biased if the increase in technology raises the output of the higher ability workers more than that of the lower ability workers:
\[
\frac{\partial^2 y(a,t)}{\partial a \partial t} > 0 \quad \text{for all } a \text{ and } t. \tag{23}
\]

**Definition 6:** A technological change is log-skill-biased if the increase in technology raises the output of the higher ability workers proportionately more than that of the lower ability workers:
\[
\frac{\partial^2 \log y(a,t)}{\partial a \partial t} > 0 \quad \text{for all } a \text{ and } t. \tag{24}
\]

One can show that a log-skill-biased technological change is also skill-biased, so log-skill-biased is a stronger notion. Both of the definitions above compare how technological changes affect the general output at different ability levels. The next definition focuses on how technological changes affect the general component and firm-specific component of the output differently.

**Definition 7:** A technological change is general-skill-biased if an increase in technology raises the general component of output proportionately more than the firm-specific part:
\[
\frac{\partial \log y(a,t)}{\partial t} > \frac{\partial \log s(a,t)}{\partial t} \quad \text{for all } a \text{ and } t. \tag{25}
\]

The definitions above introduce different types of technological changes. To state the theorems of how technological changes affect the wage distributions, we next introduce the definition by Bickel and Lehmann (1979) of "spread out", which compares the dispersion of distributions.

**Definition 8:** A distribution $H_1$ to be more spread out than distribution $H_2$ if
\[
H_1^{-1}(q) - H_1^{-1}(q') \geq H_2^{-1}(q) - H_2^{-1}(q') \quad \text{for all } 0 \leq q' < q \leq 1. \tag{26}
\]

In other words, the distance between the values corresponding to the two quantiles is larger for any two quantiles in the more "spread out" distribution. Alternatively, suppose there
is a class of distributions $H(x,t)$, let $x(q,t)$ be the value of the $q$th quantile under index $t$. The distributions are more spread out as $t$ increases if

$$\frac{\partial^2 x(q,t)}{\partial q \partial t} > 0 \quad \text{for all } q \text{ and } t. \quad (27)$$

A unique feature of "spread out" is that its order is preserved under translation. In other words, if $H_1(x)$ is more spread out than $H_2(x)$, then $H_1(x+t)$ is also more spread out than $H_2(x)$ for all $t$. This stands in contrast with Second Order Stochastic Dominance (SOSD), a popular measure of dispersion, which only compares distributions with the same mean. When two symmetric distributions have the same mean, then "spread out" implies SOSD. For our purpose, technological changes typically affect both the mean and the dispersion of the wage distribution, so we use spread out as the measure of wage inequality.

The major result in this section is that when a technological change is both log-skill-biased and general-skill-biased, turnover probability increases for workers of all ability levels and the wage distribution becomes more spread out. The intuition of the increased turnover is that the technological change makes the firm specific output relatively less important than the general output, so turnover increases. In particular, recall in Proposition 1 that the turnover probability decreases with $r(a,t) = \frac{s(a,t)}{y(a,t) - B(a,t)}$. A technological change that is both log-skill-biased and general-skill-biased decreases this ratio and thus increases turnover. On the other hand, the intuition of the increased wage dispersion is that the technological change makes the output distribution more spread out. It is natural to guess that shape of the wage distribution changes in the same direction as the output distribution, and we confirm this conjecture in Theorem 3 using explicit calculation made possible by Proposition 2.

**Theorem 3:** Let the period 2 outside output be $y(a,t) > 0$ and the inside output be $y(a,t) + s(a,t)$, where $s(a,t) > 0$. If a technological change is log-skill-biased ($\frac{\partial^2 \log y(a,t)}{\partial a \partial t} > 0$) and general-skill-biased ($\frac{\partial \log y(a,t)}{\partial t} > \frac{\partial \log s(a,t)}{\partial t}$), then the turnover probability increases and the wage distribution (of the movers and stayers combined) becomes more spread out as $t$ increases.

The media has suggested that job mobility in the U.S. has increased recently, although evidence from economic research has been less definitive. A special symposium on job stabilities and securities is reported in the Journal of Labor Economics (1999). Evidence from different authors using different data sources suggests that the “short-term” job stability, defined as shares of workers with less than 18 months of tenure (Jaeger and Stevens (1999)), four year retention rate (Neumark, Polsky, and Hansen (1999)), or one year retention rate (Gottschalk and Moffit (1999)), has not changed much since 1983. On the other hand, the “long-term” job stability, defined as shares of workers with less than ten years of tenure (Jaeger and Stevens (1999)) or as eight year retention rate (Neumark, Polsky, and Hansen (1999)), has decreased somewhat since the 1990s.

Perhaps the more relevant empirical measure of turnover probability for this model is job-to-job mobility, where workers directly work for new firms when they change jobs.
Stewart (2002) examines job changes in this direction and classifies a job change into changes that occur from a): employment to employment (EE), b): employment to unemployment (EU), and c): employment to not-in-labor-force-participation (EN). He shows that the EU and EN transition rates in the U.S. have decreased over the years. But most strikingly, the EE transition rate in the U.S. has increased dramatically. According to Stewart (2002), between 1975-2000 the EE transition rate in the U.S. has increased 45% for men and 58% for women.

In Stewart (2002), in addition to the finding that job to job mobility rate has increased, he finds that the proportionate increase in employment to employment (EE) transition rate has been strongly increasing with experience even if the absolute EE transition rate decreases with experience. From 1975 to 2000, the proportionate increase in EE rate for men is 23% for the age group 25-35, 79% for the age group of 35-45, and 144% for the age group of 45-55.

If older workers also have more firm-specific human capital, then this pattern of increasing job mobility for older workers can also be explained by the technological changes using this model. In particular, the next theorem shows that if general ability is sufficiently important in production, when there is a SBTC, the increase in turnover is larger for workers with higher levels of firm-specific human capital.

**Theorem 4**: Let the output be \( y(a,t) > 0 \) for the outside firms and \( y(a,t) + ks(a) \) for the incumbent, where \( ks(a) > 0 \). If the technological change is skill-biased (\( \frac{\partial^2 y(a,t)}{\partial a \partial t} > 0 \)) and \( y(a,t) - B(a,t) > ks(a) \) for all \( a \), then the proportionate increase in turnover increases with \( k \), i.e.,

\[
\frac{\partial^2 \log(1 - P(a,t,k))}{\partial t \partial k} > 0 \quad \text{for all} \quad a.
\]

The intuition of Theorem 4 may be best gained through looking at the benchmark case in Corollary 3 where there is no firm-specific human capital. In this case, the aggregate turnover probability is always \( \frac{1}{2} \). Consequently, for workers with little firm-specific human capital, their average turnover probability is always close to \( \frac{1}{2} \), and thus any technological change cannot have a large effect on the aggregate turnover probability of such workers. When workers have more firm-specific human capital, their aggregate turnover probability is lower, and technological changes can have a larger impact on it.

### 6 Conclusion and Discussion

This paper develops a framework to study job mobility and wage dispersion under asymmetric information. Contrary to existing work, the lemons problem from the superior information of the current employer does not lead to market collapse. Instead, there exists a unique equilibrium allocation in which the current employer offers a wage equal to the average output of all types below the ability of the worker and outside firms compete for the worker by using mixed strategies. These mixed strategies lead to a non-degenerate wage distribution for all types of workers. This unique equilibrium allocation determines both
the allocation of workers with heterogeneous abilities to different firms and also how wages change when workers change jobs (due to both selection and productivity effects).

We apply the framework to study how technological changes affect the joint evolution of wage inequality and job mobility in the United States over the past 30 years. The model implies that, in the presence of technological change that is both skill-biased and general-skill-biased, the wage distribution will become more spread out (corresponding to greater inequality) and job mobility will increase. The model also suggests that mobility should increase more for older workers. These patterns are consistent with recent empirical evidence on changes in job mobility in the United States.

The framework developed in this paper can also help us study many other questions about the labor market when information asymmetry matters. We point below to several areas where this framework is useful and discuss some relevant results.

First, this framework can be applied to study training decisions of firms. Earlier analysis shows that, when training is general, firms will not pay for it. Acemoglu and Pischke (1998, 1999ab) show that when asymmetric information is present, then firms will pay for some of the general training but the general training is still underprovided (in the sense that the marginal benefit of training in productivity exceeds the marginal cost). On the other hand, casual empiricism, see for example Baron and Kreps (1999) on firm’s paying for an executive MBA education, suggests that general training may actually be overprovided. Excessive general training is possible in an extended version of this model because the profitability of a firm depends on the entire productivity distribution of its workers and the firm can make more profit if this distribution becomes more dispersed. Therefore, firms may overprovide training if training also makes the productivity more dispersed.

Second, this framework enables us to study the layoff decisions of firms. For example, if we enrich the basic model by allowing the firms to offer contracts that specify a wage floor in the second period, then one can ask about what the optimal level of wage floor is. One can show that the optimal wage floor is zero when the general output is convex in ability. When output is concave in ability, the optimal wage floor may be positive. These results can be tested empirically by comparing the frequency of layoffs to the distribution of outputs in different sectors.

Third, this framework sheds new light on the career choice of workers. Conventional wisdom in economic theory suggests that young workers should start out with risky sectors because risky sectors have higher option value. This view does not take into account the possibility that asymmetric learning limits the subsequent career choice of a worker because outside firms no longer learns about his ability and understand that they are subject to adverse selection. It may well be then that a worker who chooses a risky sector performs poorly and is then “trapped” in that sector, and this disadvantage may prompt the worker to choose a less risky sector in the first place.

Fourth, we can use this framework to study the spillover effect in minimum wage level increases. It is well-known that when minimum wage level increases, workers with initial

---

14I thank Kevin Lang for raising this question and helpful discussions.
wage levels above the minimum wage often experience wage gains as well. This spillover effect is possible in this framework because as the minimum wage increases, firms no longer hire workers whose output is below the minimum wage and thus increases output distributions of the workers. Since the wage of a worker is determined by the complete output distribution, this change in output distribution can increase a worker’s wage even if his productivity remains unchanged.

Fifth, we can extend the model to multiple sectors and study the implications of asymmetric information on how wages differ across sectors and how wages change when workers switch jobs. In a multi-sector setting in which sectors are ordered by their returns to ability, we show that inter-sector wage differential occurs, where a sector can be interpreted either as an industry or an occupation. Firms in a sector only offer wages in the range where the “average output” in that sector of the worker is higher than his “average output” in all other sectors.

In addition, we show that, for workers who are initially in the highest sector, the average wage of within-sector movers is higher than both stayers and between-sector movers when production depends only on general human capital. Moreover, the between-sector movers from the highest sector may have higher wage than the stayers. These two predictions are difficult to obtain from other theories of wage determination, including human capital theory, symmetric learning with matching, and search theory.

Finally, the multi-sector model can also provide more detailed predictions on how technological changes affect both job mobility and inequality. In particular, we can show that if a technological change is both skill-biased and general skill-biased, the wage distribution becomes more dispersed, overall job mobility increases, and within-sector mobility increases. Moreover, between-sector mobility increases more for workers with more sector-specific human capital, and the overall between-sector mobility increases if the sector specific human capital is sufficiently large. These predictions are consistent recent evidence on changes in job mobility in the U.S.; see for example Kambourov and Manovskii (2009).

References


Lemma 1: There is no pure strategy PBE.

Proof. We prove by contradiction. Suppose instead there is a pure strategy PBE. Let \( w \) be the highest outside offer in period 2 in this equilibrium, so the incumbent can keep the worker if it offers any wage above \( w \). Because the incumbent earns zero (in period 2) if it does not keep the worker, the incumbent will keep the worker in equilibrium if his inside output \( (y(a) + s(a)) \) is greater than \( w \). This implies that outside firms never get the worker when \( y(a) + s(a) > w \), so the expected profit of any outside firm, conditional on hiring the worker, is at most

\[
E[y(a)] = y(0) - s(0) - \varepsilon.
\]

Therefore, if an outside firm hires the worker with positive probability, it must have a negative payoff. Since outside firms can always guarantee themselves nonnegative payoffs (by offering zero wages), this implies that the outside firms must hire the worker with zero probability, or equivalently, the incumbent must keep the worker with probability 1 in this pure strategy PBE.

This implies that the incumbent must offer a wage greater or equal to \( w \) with probability 1. But any wage offer strictly greater than \( w \) cannot be optimal for the incumbent (because it can be replaced by a smaller wage, say the average of \( w \) and itself, that also keeps the worker but is smaller in amount), the incumbent must offer \( w \) with probability 1 in this equilibrium. Now consider an outside firm that deviates by offering \( w' = w + \varepsilon \) for some \( \varepsilon > 0 \). This deviation hires the worker with probability 1 and gives to the deviating firm an expected profit of

\[
E[y(a)] - w - \varepsilon \geq E[y(a)] - y(0) - s(0) - \varepsilon.
\]
The deviation is profitable for small enough $\varepsilon$ because $E[y(a)] - y(0) - s(0) > 0$ by the production assumption (6). This leads to a contradiction. 

**Theorem 1:** The following strategies and beliefs form a PBE:

(i) In period 2, the worker chooses the maximum wage offer. If there are multiple maximum offers, the worker $a)$ stays with the incumbent if its offer is one of the maximum offers; $b)$ randomize otherwise.

(ii) At the beginning of period 2, the incumbent firm offers

$$w_{In}(a) = B(a) \quad \text{for all } a,$$

and each outside firm $j$ ($j \in \{2, \ldots, N\}$) offers a wage drawn independently from the distribution

$$F_j^*(w) = F(w) \frac{1}{\bar{w} - w},$$

where

$$F(w) = C \exp \left( \int_{w}^{\bar{w}} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} dx \right) \quad \text{for } w < \bar{w};$$

$$F(w) = 1$$

for $w > \bar{w}$,

$$[w, \bar{w}] = [y(0), E[y(a)]] \quad \text{and } C = \exp \left( - \int_{w}^{\bar{w}} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} dx \right).$$

(iii) In period 1, the worker chooses the maximum wage offer. If there are multiple maximum offers, the worker randomizes among them.

(iv) At the beginning of period 1, all firms offer

$$E[y(a)] + \int_{0}^{1} F(B(a)) (y(a) + s(a) - B(a)) da.$$  

(v) Each firm’s belief about the worker’s ability equals the prior if it has not hired the worker. The incumbent knows the worker’s ability at the end of period 1. The worker does not know his ability in period 1. The worker knows his ability in period 2.

**Proof.** We start with the beliefs. Since the worker does not know his own ability in period 1, the beliefs described by (v) is the only beliefs consistent with the (degenerated) Bayes rule.

Next, we examine the strategies in period 2. In period 2, the worker’s utility is maximized by choosing the maximum wage offer. Therefore, the strategy described in (i) is optimal for the worker.

Now given the equilibrium strategy of the worker and the maximum outside offer distribution, the incumbent’s payoff by offering $w$ to a worker of ability $a$ is

$$(y(a) + s(a) - w) F(w).$$

Maximization of the incumbent’s profit gives the following first order condition:

$$\frac{1}{y(a) + s(a) - w_{In}(a)} = f(w_{In}(a)) \frac{f(w_{In}(a))}{F(w_{In}(a))},$$

(32)
The first order condition is a necessary condition for optimality, and it is easy to check that if a solution satisfies it and is also increasing in $a$ (which is the case for $B(a)$), then the second order condition is also satisfied and the solution is optimal. Let $w_{1n}(a) = B(a)$, then by the definition of $F$ in (ii) we can check that

$$\frac{1}{y(a) + s(a) - B(a)} = \frac{f(B(a))}{F(B(a))}.$$ 

Therefore, $w_{1n}(a) = B(a)$ satisfies (32) and thus maximizes the incumbent’s expected pay-off.

Given the incumbent’s wage offer and the worker’s strategy, the expected profit of an outside firm by offering $w$ is

$$P(\text{Not Hiring}) \cdot 0 + P(\text{Hiring})(E[y(a)|B(a) < w] - w). \tag{33}$$

This expression is zero for all $w \leq E[y(a)]$ because $E[y(a)|B(a) < w] = B(B^{-1}(w)) = w$. When $w > E[y(a)]$, this expression is negative because $E[y(a)|B(a) < w] - w = E[y(a)] - w < 0$. Therefore, the optimal response of each outside firm is to randomize over $(0, E[y(a)])$. This completes proving the optimality of the strategies in period 2.

Moving back to period 1, it is clear that the worker maximizes his utility by choosing the maximum wage offer because all the firms are ex ante identical. Therefore, the worker’s strategy described in (iii) is optimal.

For the firms, wage competition implies that the equilibrium wage offer will be bid up to the expected output of the worker in period 1 plus the expected profit the firm makes if it is the incumbent in period 2. Therefore, it is an equilibrium that each firm offers in period 1

$$E[y(a)] + \int_0^1 F(B(a))(y(a) + s(a) - B(a))da.$$

This finishes the proof. ■

To prove Theorem 2, we first show that the incumbent’s wage offer is strongly increasing in the worker’s ability. Strongly increasing is an order on sets$^{15}$. We use this order here because we have not shown that the incumbent’s wage offer is single-valued yet.

**Lemma 2** If $w_{1n}(a_2) > w$, we have $w_{1n}(a_1) \geq w_{1n}(a_2)$ for all $a_1 > a_2$.

**Proof.** We prove by contradiction. Take two arbitrary ability levels $a_1 > a_2$ such that $w_{1n}(a_2) > w$. Let $w_1$ and $w_2$ be two equilibrium wage offers of the incumbent when the worker is of ability $a_1$ and $a_2$ respectively; i.e. $w_1 < w_{1n}(a_1)$ and $w_2 < w_{1n}(a_2)$. Suppose instead we have $w_1 < w_2$.

Define $\bar{F}(w)$ as the probability that the incumbent keeps the worker if it offers $w$. Clearly we have $\bar{F}(w_1) \leq \bar{F}(w_2)$, where recall $\bar{F}(w)$ is the probability that the incumbent keeps the worker if it offers $w$. Now if $\bar{F}(w_1) = \bar{F}(w_2)$, the incumbent’s payoff by offering $w_1$ is

$$\bar{F}(w_1)(y(a_2) + s(a_2) - w_1) > \bar{F}(w_2)(y(a_2) + s(a_2) - w_2) \tag{34}$$

$^{15}$Let $X$ and $Y$ be two sets. $X \geq Y$ if $x \geq y$ for all $x \in X$ and $y \in Y$. 

27
since $\bar{F}(w_1) = \bar{F}(w_2) > 0$. In other words, the incumbent strictly prefers offering $w_1$ to $w_2$ at $a_2$ if $\bar{F}(w_1) = \bar{F}(w_2)$, violating the assumption that $w_2 \in w_{In}(a_2)$. Therefore, we must have $\bar{F}(w_1) < \bar{F}(w_2)$.

Now let $y_{In}(a) = y(a) + s(a)$. The expected profit of the incumbent by offering $w_2$ when the worker has ability $a_1$ is

$$
(y_{In}(a_1) - w_2)\bar{F}(w_2) = (y_{In}(a_1) - y_{In}(a_2))\bar{F}(w_2) + (y_{In}(a_2) - w_2 - y_{In}(a_2))\bar{F}(w_2) \\
\geq (y_{In}(a_1) - y_{In}(a_2))\bar{F}(w_2) + (y_{In}(a_2) - w_1 - y_{In}(a_2))\bar{F}(w_1) \\
> (y_{In}(a_1) - y_{In}(a_2))\bar{F}(w_1) + (y_{In}(a_2) - w_1 - y_{In}(a_2))\bar{F}(w_1) \\
= (y_{In}(a_1) - w_1)\bar{F}(w_1).
$$

The first inequality follows from $w_2 \in w_{In}(a_2)$, and the strict inequality follows because $y_{In}(a_1) - y_{In}(a_2) > 0$ and $\bar{F}(w_2) > \bar{F}(w_1)$.

This chain of inequality implies that the incumbent would strictly prefer offering $w_2$ to $w_1$ when the worker has ability $a_1$, thus contradicting the assumption that $w_1 \in w_{In}(a_1)$. Therefore, we must have $w_1 \geq w_2$. □

Using the monotonicity property in Lemma 2, we show that each outside firm must have zero expected profits. This is the content of Lemma 3.

**Lemma 3** The expected profit of each outside firm is zero.

**Proof.** It suffices to show that the expected profit of one outside firm is zero. Because if one outside firm earns zero expected profit and another one earns positive profit, the one with zero expected profit can always mimic the one with positive profit to earn positive profit as well. We consider two cases.

Case 1: $F(w) = 0$, where recall that $F$ is the CDF of the maximal outside offers.

Since $F(w) = 0$, the definition of $w$ together with right-continuity of a CDF implies that for each $n > 0$, there exists $w_n > w$ such that a): $w_n$ is offered by an outside firm and b): $F(w_n) < \frac{1}{n}$.

By offering $w_n$, the expected profit of the outside firm is less than

$$
F(w_n)(E[y(a)|w_{In}(a) \leq w_n] - w_n)
$$

because the probability that the firm hires the worker is no greater than $F(w_n)$ and the expected profit from hiring the worker is no greater than $E[y(a)|w_{In}(a) \leq w_n] - w_n$ because the incumbent’s offer is monotone in the worker’s ability (Lemma 2).

Equation (36) goes to zero as $n$ goes to infinity because $F(w_n) < \frac{1}{n}$ and $E[y(a)|w_{In}(a) \leq w_n] - w_n$ is bounded. Because we have only finite many outside firms, this implies that there must be one outside firm whose profit is less than $\frac{1}{n}$ for all $n$ and its profit must thus be zero. This finishes the proof for the case with $F(w) = 0$.

Case 2: $F(w) > 0$.

First, if an outside firm offers a wage less than $w$, this wage offer hires the worker with probability zero (by the definition of $w$ ) and gives the firm a profit of zero. Since all offers in the support of its equilibrium offers must yield the same profit for the firm, this implies
the outside firm must have an expected profit zero (and we are done). Therefore, we may assume all outside firms offer wages greater than or equal to \( w \).

When all outside firms offer wages greater than or equal to \( w \), our assumption \( F(w) > 0 \) implies that each outside firm must offer \( w \) with positive probability. In other words, \( F_j^*(w) > 0 \) for all \( j \in \{2, \ldots, N\} \) (recall the firm 1 is the incumbent). When all outside firms offer \( w \) (and the incumbent offers no greater than \( w \)), without loss of generality, we may assume that the worker chooses firm 2 with probability \( q \leq \frac{1}{2} \).

Now consider the payoff of firm 2 when it offers \( w \). Its expected output is less than or equal to

\[
q \prod_{j=3}^{N} F_j^*(w) \Pr(w_{I_n}(a) \leq w)(E[y(a)|w_{I_n}(a) \leq w] - w).
\]

(37)

This expression follows because firm 2 can hire the worker only if all other outside firms offer \( w \) (with probability \( \prod_{j=3}^{N} F_j^*(w) \)), the incumbent offers less than or equal to \( w \) (with probability \( \Pr(w_{I_n}(a) \leq w) \)), and the worker chooses firm 2 in face of this tie (with probability \( q \)). Hence \( q \prod_{j=3}^{N} F_j^*(w) \Pr(w_{I_n}(a) \leq w) \) is the probability that firm 2 hires the worker with wage \( w \). And the firm 2’s expected profit conditional on hiring the worker is at most \( E[y(a)|w_{I_n}(a) \leq w] - w \).

Suppose firm 2 offers \( w + \varepsilon \) instead. Firm 2’s expected profit with this offer is at least

\[
\prod_{j=3}^{N} F_j^*(w) \Pr(w_{I_n}(a) \leq w)(E[y(a)|w_{I_n}(a) \leq w] - w - \varepsilon).
\]

If \( E[y(a)|w_{I_n}(a) \leq w] - w > 0 \), this is a profitable deviation for \( \varepsilon \) small enough because there is a discrete increase in the probability of hiring the worker. Therefore, we must have \( E[y(a)|w_{I_n}(a) \leq w] - w = 0 \), so firm 2 must have zero expected profit and we are done.

The zero expected profits result in Lemma 3 plays two important roles in our proof of Theorem 2. First, it is the key property used in Lemma 7 below that determines the equilibrium wage offer of the incumbent. Second, it helps prove Lemma 5, which states that the maximum outside offer distribution does not have an atom except possibly at the bottom.

Before proving Lemma 5, we first prove a technical lemma that states that the incumbent’s offer does not have an atom at any \( w > w \). This is the content of Lemma 4.

**Lemma 4** If \( w > w \), then \( \Pr(w_{I_n}(a) = w) = 0 \).

**Proof.** We prove by contradiction. If instead there exists a wage \( w > w \) such that \( \Pr(w_{I_n}(a) = w) > 0 \), then Lemma 2 implies that there exists \( a < a_2 \), such that \( w_{I_n}(a_1) = w_{I_n}(a_2) = w \).

Now take any \( \varepsilon > 0 \) and compare the expected profit of an outside firm between offering \( w + \varepsilon \) and \( w - \varepsilon \). Since by offering \( w + \varepsilon \) \((w - \varepsilon)\) the incumbent keeps the worker with a probability of at least(most) \( F(w) \), the extra expected profit in offering \( w + \varepsilon \) is at least

\[
F(w)(B(a_2) - w - \varepsilon) - F(w)(B(a_1) - w + \varepsilon)
\]

\[
= F(w)(B(a_2) - B(a_1) - 2\varepsilon).
\]

29
Since $F(w) > 0$ and $B(a_2) - B(a_1) > 0$, for small enough $\varepsilon$ the term above is positive. In other words, for small enough $\varepsilon$, offering $w - \varepsilon$ is strictly dominated by offering $w + \varepsilon$. Therefore, no outside firm would offer a wage in $(w - \varepsilon, w)$ for small enough $\varepsilon$.

Since no outside firms offer wage in $(w - \varepsilon, w)$, the incumbent that offers $w$ has a profitable deviation by offering $w - \frac{\varepsilon}{2}$ for small enough $\varepsilon$, as long as the maximum outside offer distribution does not have a mass in $w$ (so deviating to $w - \frac{\varepsilon}{2}$ does not affect the incumbent’s probability of retaining the worker).

When the maximum outside offer distribution has a mass at $w$, let $q$ be the probability that the worker leaves the incumbent if both the incumbent offer and the maximum outside offer is $w$. If $q = 1$, the incumbent never keeps the worker when the worker randomizes over $w$. In this case, the incumbent strictly prefers offering $w - \frac{\varepsilon}{2}$ to $w$, because it is cheaper and keeps the worker with the same probability.

When $q < 1$, the outside firm that offers $w$ (there is such a firm because the maximum outside offer distribution has a mass at $w$) can profitably deviate by offering $w + \delta$, which increase the probability of hiring the worker and discretely increase the expected ability of the workers hired.

These cases combined show that there is always a profitable deviation when $\text{Pr}(a | \text{w}_{I_n}(a) = w) > 0$ for $w > w$. Therefore, we must have $\text{Pr}(\text{w}_{I_n}(a) = w) = 0$ if $w > w$. ■

**Lemma 5** $F(w) = F(w_-)$ for all $w > w$, where $F(w_-)$ is the left limit of $F$ at $w$.

**Proof.** We prove by contradiction. Suppose instead $F$ has an atom at $w > w$. To get a contradiction, it suffices to show that the incumbent will not offer a wage in $(w - \varepsilon, w)$. Because in this case, an outside firm that offers $w - \frac{\varepsilon}{2}$ has an expected profit (conditional on hiring the worker) of at least

$$E[B(a) | \text{w}_{I_n}(a) \leq w - \varepsilon] - w + \frac{\varepsilon}{2}$$

$$= E[B(a) | \text{w}_{I_n}(a) \leq w] - w + \frac{\varepsilon}{2}$$

$$= \frac{\varepsilon}{2},$$

where the first equality follows from that the incumbent will not offer a wage in $(w - \varepsilon, w)$ and that incumbent’s offer does not have an atom (Lemma 3), and the second equality follows from that the expected payoff of all outside firms must be zero (so the conditional profit of an outside firm offering $w$ must be zero). This leads to a contradiction because it implies that an outside firm can have positive expected profit by offering $w - \frac{\varepsilon}{2}$, violating Lemma 3.

Now we show that the incumbent will not offer a wage in $(w - \varepsilon, w)$ for small enough $\varepsilon$ and thus complete the proof. Since the incumbent’s expected profit is increasing in the worker’s ability, there exists an $M$ such that if $\text{w}_{I_n}(a) \in (w - \frac{1}{M}, w)$, we must have $y_{I_n}(a) > w + \frac{1}{M}$. Now take $\varepsilon = \frac{1}{2M}(F(w) - F(w_-))$. Note that $\varepsilon < \frac{1}{M}$, so if the incumbent offers a worker of ability $a$ a wage in $(w - \varepsilon, w)$, we must have $y_{I_n}(a) > w + \frac{1}{M}$.

Suppose the incumbent offers a worker of ability $a$ a wage in $(w - \varepsilon, w)$, its expected profit is at most $F(w_-)(y_{I_n}(a) - w + \varepsilon)$. If the incumbent deviates and offers $w + \varepsilon$ instead, its expected profit is at least $F(w)(y_{I_n}(a) - w - \varepsilon)$. Therefore, the difference in the expected profit after the deviation is at least
We prove by contradiction. Suppose instead we have $F(w_1) = F(w_2)$ for some $w_1 < w_2 \in [\underline{w}, \overline{w}]$. Without loss of generality, we may assume that $w_2$ is the largest wage such that no outside firms makes offer between $(w_1, w_2)$ (with positive probability), i.e. $w_2 = \sup\{w : F(w) = F(w_1)\}$. This implies that, for any $\epsilon > 0$, we can find a wage $w \in [w_2, w_2 + \epsilon)$ that is offered in equilibrium by an outside firm.

Now take an outside firm that offers wage $w \in [w_2, w_2 + \epsilon)$, its expected profit is at most

$$F(w_2 + \epsilon) \Pr(w_{In}(a) \leq w_2 + \epsilon)(E[B(a)|w_{In}(a) \leq w_2 + \epsilon] - w_2),$$

(41)

where $F(w_2 + \epsilon) \Pr(w_{In}(a) \leq w_2 + \epsilon)$ is an upper bound of the probability of hiring the worker and $E[B(a)|w_{In}(a) \leq w_2 + \epsilon]$ is an upper bound of the expected output of the worker hired.

Suppose the firm deviates and offers $\frac{w_1 + w_2}{2}$ instead. Its expected profit is at least

$$F(w_1) \Pr(w_{In}(a) \leq w_1)(E[B(a)|w_{In}(a) \leq w_1] - \frac{w_1 + w_2}{2}),$$

(42)

where $F(w_1) \Pr(w_{In}(a) \leq w_1)$ is a lower bound of hiring the worker and $E[B(a)|w_{In}(a) \leq w_1]$ is a lower bound of the expected output of the worker hired.

Since $F(w_1) = F(w_2)$, it is clear that the incumbent would not offer a wage in $(w_1, w_2)$ in equilibrium. Furthermore, Lemma 4 implies that the incumbent offers $w_2$ with probability zero. Therefore, we have

$$\Pr(w_{In}(a) \leq w_1) = \Pr(w_{In}(a) \leq w_2);$$

(43)

$$E[B(a)|w_{In}(a) \leq w_1] = E[B(a)|w_{In}(a) \leq w_2].$$

(44)

Substituting these expressions in (41) and (42), we see that the expected profit from offering $\frac{w_1 + w_2}{2}$ is

$$F(w_2) \Pr(w_{In}(a) \leq w_2)(E[B(a)|w_{In}(a) \leq w_2] - \frac{w_1 + w_2}{2})$$

$$> F(w_2 + \epsilon) \Pr(w_{In}(a) \leq w_2 + \epsilon)(E[B(a)|w_{In}(a) \leq w_2 + \epsilon] - w_2)$$

for small enough $\epsilon$. This implies that the outside firm has a profitable deviation, and we have a contradiction. ■
Lemma 6, together with the zero expected profit lemma (Lemma 3), implies that the incumbent’s wage offer must equal the average output for workers for all ability levels \( a \) such that \( B(a) \in (\underline{w}, \overline{w}] \). This is stated formally in Lemma 7.

**Lemma 7** If \( B(a) \in (\underline{w}, \overline{w}] \), then \( w_{I_n}(a) = B(a) \).

**Proof.** Suppose an outside firm offers a wage \( w \in (\underline{w}, \overline{w}] \). It hires the worker with positive probability. This implies that the firm’s conditional expected profit of hiring the worker is zero (Lemma 3). Since the incumbent’s wage offer is strictly increasing by Lemma 2 and Lemma 4, the conditional expected profit of the firm is

\[
E[y(a) | w_{I_n}(a) \leq w] - w = E[a | a \leq w_{I_n}^{-1}(w)] - w = B(w_{I_n}^{-1}(w)) - w = 0,
\]

where the last equality follows from Lemma 3.

By Lemma 6, the support of maximum outside offer is dense in \( (\underline{w}, \overline{w}] \). Therefore, \( B(w_{I_n}^{-1}(w)) = w \) for all \( w \in (\underline{w}, \overline{w}] \). Since \( B \) is strictly increasing, this implies that \( w_{I_n}(a) = B(a) \) if \( B(a) \in (\underline{w}, \overline{w}] \). □

**Lemma 8** The equilibrium maximum outside offer distribution satisfies

\[
F(w) = C \exp \left( \int_{\underline{w}}^{w} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} dx \right) \quad \text{for } w \in [\underline{w}, \overline{w}],
\]

where \( C = \exp \left( - \int_{\underline{w}}^{\overline{w}} \frac{1}{y(B^{-1}(x)) + s(B^{-1}(x)) - x} dx \right) \).

**Proof.** Since \( F \) does not have an atom above \( \overline{w} \) (Lemma 5), the incumbent’s expected profit by offering \( w \) to a worker of ability \( a \) is

\[
(y(a) + s(a) - w)F(w).
\]

Since \( F \) is strictly increasing, it is differentiable almost everywhere so the incumbent’s profit maximization condition leads to the following first order condition:

\[
\frac{1}{y(a) + s(a) - w_{I_n}(a)} = \frac{f(w_{I_n}(a))}{F(w_{I_n}(a))}.
\]

Since we have \( w_{I_n}(a) = B(a) \) for all \( B(a) \in [\underline{w}, \overline{w}] \), the expression above can be written as

\[
\frac{1}{y(B^{-1}(w)) + s(B^{-1}(w)) - w} = \frac{f(w)}{F(w)}.
\]

Integrating this equation, we obtain (47) by using that \( F(\overline{w}) = 1 \) and \( F \) is right-continuous at \( \overline{w} \). □

Lemma 8 pins down the distribution of the maximum outside offer in its support. To completely determine the maximum outside distribution, we need to specify the end points of \( F \). This is done in Lemma 9 and 10.
Lemma 9 \( \bar{w} = E[y(a)] \).

**Proof.** Outside firms that offer \( w > E[y(a)] \) clearly have negative profits, so we must have \( \bar{w} \leq E[y(a)] \).

Suppose instead \( \bar{w} < E[y(a)] \). Let \( a^* = B^{-1}(\bar{w}) \). Since the incumbent’s offer is strongly increasing in worker’s ability (Lemma 2), all workers with ability greater than \( a^* \) must be offered at least \( \bar{w} \). Moreover, identical reasoning as in Lemma 1 shows that the incumbent will choose \( w_{In}(a) = \bar{w} \) for all \( a > a^* \).

Now if an outside firm offers \( w = \bar{w} + \varepsilon \), it hires the worker with probability 1 and its expected profit is
\[
E[y(a)] - \bar{w} - \varepsilon > 0.
\]
for small enough \( \varepsilon \). This is a contradiction to the zero expected profit condition (Lemma 3) of outside firms. □

Lemma 10 \( \bar{w} = y(0) \).

**Proof.** We prove by contradiction. If \( w < y(0) \), then an outside firm can deviate by offering \( w + \varepsilon < y(0) \) for some small \( \varepsilon \). This outside firm hires the worker with positive probability (by the definition of \( \bar{w} \)) and earns positive conditional profit because the wage is lower than the lowest possible output. This violates the zero expected profit of outside firms (Lemma 3).

If \( w > y(0) \), let \( a_1 = B^{-1}(w) > 0 \), so \( w_{In}(a_1) = w \). Also define \( a_2 \) as the unique ability level such that: \( y(a_2) + s(a_2) = w \). Because \( B(a) < y(a) \) for all \( a > 0 \), it follows that \( a_1 > a_2 \).

Now we must have \( w_{In}(a) \geq w \) for \( a > a_2 \) because otherwise the incumbent keeps such worker with probability zero. On the other hand, we also have \( w_{In}(a_1) = w \). Therefore, by the monotonicity lemma (Lemma 2), we must have \( w_{In}(a) = w \) for all \( a \in (a_2, a_1] \).

Furthermore, we must have \( F(w) > 0 \) because otherwise when it offer \( \bar{w} \) its profit is zero. This implies that when an outside firm offers \( w \) (which happens with positive probability because \( F(w) > 0 \)), it cannot hire any worker with ability greater than \( a_2 \). Therefore, its conditional expected profit is
\[
E[y(a)|a \leq a_2] - w < E[y(a)|a \leq a_1] - w = B(a_1) - w = 0.
\]

Furthermore, the outside firm hires the worker with positive probability because \( \bar{w} > y(0) \). This leads to a contradiction because all outside firms must have expected profit of zero by Lemma 3. □

Lemma 8-10 completely determine the distribution of the maximum outside offer by specifying the differential equation that governs it (Lemma 8) and its two end points (Lemma 9 and 10). Therefore, the distribution of the maximum outside offer is unique. This finishes the proof of Theorem 2.
Proposition 1: For all \(a\) and \(t\), let

\[
 r(a, t) = \frac{s(a, t)}{y(a, t) - B(a, t)},
\]

then

\[
 P(a, t) = \exp\left(- \int_a^1 \frac{1}{x(1 + r(x, t))} dx\right).
\]

Proof. For a worker with ability \(a > 0\), his offer from the incumbent is \(B(a, t)\) (Lemma 7). This worker stays with the incumbent if and only if the maximum outside offer he receives is less than \(B(a, t)\), which occurs with probability \(F(B(a, t))\). In other words, we have

\[
 P(a, t) = F(B(a, t)) \quad \text{(53)}
\]

Let \(p(a, t) = \frac{\partial P(a, t)}{\partial a}\). Then (42) implies that \(p(a, t) = f(B(a, t))B'(a, t)\).

By Lemma 8,

\[
 \frac{p(a, t)}{P(a, t)} = \frac{f(B(a, t))B'(a, t)}{F(B(a, t))B'(a, t)} = \frac{B'(a, t)}{y(a, t) + s(a, t) - B(a, t)} = \frac{B'(a, t)}{s(a, t) + aB'(a, t)},
\]

where the last equality uses \(B'(a, t) = \frac{y(a, t) - B(a, t)}{a}\).

Integrating equation (54), we obtain that

\[
 \ln P(1, t) - \ln P(a, t) = \int_a^1 \frac{B'(x, t)}{s(x, t) + xB'(x, t)} dx \quad \text{for all } a > 0.
\]

Since \(P(1, t) = F(B(1, t)) = 1\), the equation above gives

\[
 P(a, t) = \exp\left(- \int_a^1 \frac{B'(x, t)}{s(x, t) + xB'(x, t)} dx\right) \quad \text{for all } a > 0.
\]

Finally, for a worker with ability \(a = 0\), the comment following Lemma 10 implies that \(P(0, t) = P(0+, t) = F(B(0+, t))\), so the formula remains correct (but the integral maybe improper and should be interpreted as a limit).

Corollary 1: If \(\frac{\partial^2 y(a, t)}{\partial a \partial t} > 0\) and \(s(a, t_1) = s(a, t_2)\) for all \(a\) and \(t\), then for \(t_1 < t_2\), \(P(a, t_1) > P(a, t_2)\) for all \(a > 0\).

Proof. Note that \(y(a, t) - B(a, t) = aB'(a, t)\). Moreover

\[
 \frac{\partial (aB'(a, t))}{\partial t} = \frac{\partial (ay(a, t) - \int_0^a y(x, t) dx)}{\partial t} = \int_0^a \left( \frac{\partial y(a, t)}{\partial t} - \frac{\partial y(x, t)}{\partial t} \right) dx
\]

\[
 = \int_0^a \int_x^a \frac{\partial^2 y(z, t)}{\partial z \partial t} dz \, dx
\]

\[
 > 0.
\]
Therefore, when \( s(a, t_1) = s(a, t_2) \), we have \( \frac{B'(a, t_1)}{s'(a) + aB'(a, t_1)} < \frac{B'(a, t_2)}{s'(a) + aB'(a, t_2)} \) for all \( a \). Proposition 1 immediately implies that \( P(a, t_1) > P(a, t_2) \) for all \( a > 0 \).

**Proposition 2:**

\[
\begin{align*}
G(B(a, t)) &= aP(a, t); \\
G_S(B(a, t)) &= \int_0^a P(x, t)dx; \\
G_M(B(a, t)) &= \frac{\int_0^a (1 - P(x, t))dx}{\int_0^1 (1 - P(x, t))dx}.
\end{align*}
\]

**Proof.** We first calculate \( G(B(a, t)) \). If a worker’s wage is less than \( B(y(a, t)) \), it must be the case that both the incumbent offer and the maximum outside offer he has received are less than \( B(a, t) \). The probability that the incumbent offer is less than \( B(a, t) \) is \( a \). The probability that the maximum outside offer is less than \( B(a, t) \) is \( F(B(a, t)) = P(a, t) \). Since the outside offers are independent of the incumbent’s offer, this gives that \( G(B(a, t)) = aP(a, t) \).

Next, we calculate the wage distribution of the stayers. Since a worker of ability \( a \) (who is offered \( B(a, t) \) by the incumbent) stays with the incumbent with probability \( P(a, t) \), the total number of stayers who receive less than \( B(a, t) \) is \( \int_0^a P(x, t)dx \). The wage distribution of the stayers in (20) follows immediately.

Finally, \( G \) is a weighted average of \( G_S \) and \( G_M \), so that we have

\[
G_s(B(a, t)) \int_0^1 P(x, t)dx + G_m(B(a, t)) \int_0^1 (1 - P(x, t))dx = G(B(a, t)).
\]

This gives the wage distribution of the movers.

**Corollary 2:** The wage distribution of the stayers FOSD that of the movers if \( r(a, t) = \frac{s(a, t)}{y(a, t) - B(a, t)} \) is increasing in \( a \); the wage distribution of the movers FOSD that of the stayers if \( r(a, t) = \frac{s(a, t)}{y(a, t) - B(a, t)} \) is decreasing in \( a \).

**Proof.** By Proposition 2, the average wage of the stayers is

\[
\int_0^1 B(a, t) \left( \frac{P(a, t)}{\int_0^1 P(x, t)dx} \right) da.
\]

The average wage of the movers is

\[
\int_0^1 B(a, t) \left( \frac{ap(a, t)}{\int_0^1 (1 - P(x, t))dx} \right) da.
\]

The likelihood ratio of movers over stayers is

\[
\left( \frac{\int_0^1 P(x, t)dx}{\int_0^1 (1 - P(x, t))dx} \right) \left( \frac{ap(a, t)}{P(a, t)} \right) = \left( \frac{\int_0^1 P(x, t)dx}{\int_0^1 (1 - P(x, t))dx} \right) \frac{1}{\frac{s(a, t)}{y(a, t) - B(a, t)} + 1}
\]
The term in the big parenthesis is a constant. Therefore, if \( \frac{s(a,t)}{B'(a,t)} \) is increasing, this implies that the likelihood ratio is decreasing in \( a \), so the order of monotone likelihood ratio implies the order of FOSD. 

**Corollary 3:** If \( s(a,t) \equiv 0 \), then \( P(a,t) = a \) for all \( a \), and \( G_s(w) = G_m(w) \) for all \( w \). Therefore,
\[
\int_0^1 P(a,t)da = \frac{1}{2}.
\]

**Proof.** By Proposition 1, \( P(a,t) = \exp(-\int_a^1 \frac{B'(x,t)}{s(x,t)+xB'(x,t)}dx) \). When \( s(a,t) \equiv 0 \),
\[
\frac{B'(x,t)}{s(x,t)+xB'(x,t)} \equiv \frac{1}{x}.
\]
This implies that \( P(a,t) = a \) for all \( a \). Now by Proposition 2, the wage distribution of the stayers and movers combined is
\[
G(B(a,t)) = aP(a,t) = a^2 = \frac{\int_0^a P(x,t)dx}{\int_0^1 P(x,t)dx} = G_s(B(a,t)).
\]
This implies that the movers and stayers have the same wage distribution. Equation (22) follows immediately. 

**Theorem 3:** Let the period 2 outside output be \( y(a,t) > 0 \) and the inside output be \( y(a,t)+s(a,t), \) where \( s(a,t) > 0 \). If a technological change is log-skilled-biased (\( \frac{\partial^2 \log y(a,t)}{\partial a \partial t} > 0 \)) and general-skill-biased (\( \frac{\partial \log y(a,t)}{\partial t} > \frac{\partial \log s(a,t)}{\partial t} \)), then the turnover probability increases and the wage distribution (of the movers and stayers combined) becomes more spread out as \( t \) increases.

**Proof.** We first show that the turnover probability increases with \( t \). First note that \( a^2B'(a,t) = ay(a,t) - \int_0^a y(x,t)dx \). Therefore,
\[
\frac{\partial \log(yB'(a,t))}{\partial t} = \frac{\partial \log y(a,t)}{\partial t} = \frac{a(\partial y(a,t)/\partial t) - \int_0^a (\partial y(x,t)/\partial t)dx}{ay(a,t) - \int_0^a y(x,t)dx} - \frac{\partial y(a,t)/\partial t}{y(a,t)}
\]
\[
= \left( \frac{y(a,t)\int_0^a y(x,t)dx}{ay(a,t) - \int_0^a y(x,t)dx} \right) \frac{\partial y(a,t)/\partial t}{y(a,t)} - \frac{\int_0^a (\partial y(x,t)/\partial t)dx}{\int_0^a y(x,t)dx} > 0,
\]
where the last inequality follows from the log-skilled biased assumption, which implies that \( \frac{\partial y(a,t)/\partial t}{y(a,t)} > \frac{\partial \log y(a,t)}{\partial t} \) is increasing in \( a \), so \( \frac{\partial y(a,t)/\partial t}{y(a,t)} \) is greater than an average of smaller ratios. The inequality above together with the general-skill biased assumption (\( \frac{\partial \log y(a,t)}{\partial t} > \frac{\partial \log s(a,t)}{\partial t} \)) implies that
\[
\frac{\partial \log aB'(a,t)}{\partial t} > \frac{\partial \log s(a,t)}{\partial t}.
\]
With the inequality above, we see that
\[
\frac{\partial\left(\frac{aB'(a,t)}{s(a,t) + aB'(a,t)}\right)}{\partial t} = \frac{1}{(s(a,t) + aB'(a,t))^2}\left(\frac{\partial aB'(a,t)}{\partial t}s(a,t) - \frac{\partial s(a,t)}{\partial t}aB'(a,t)\right)
\]
\[
= \frac{aB'(a,t)s(a,t)}{(s(a,t) + aB'(a,t))^2}\left(\frac{\partial \log aB'(a,t)}{\partial t} - \frac{\partial \log s(a,t)}{\partial t}\right)
\]
\[
> 0.
\]
This immediately implies that turnover probability increases with \( t \) (by Proposition 1).

Next, we show that the wage distributions become more spread out as \( t \) increases. Let \( w(q,t) \) be the value of the \( q \)th quantile in the wage distribution with technology index \( t \). We need to show that \( \frac{\partial^2 w(q,t)}{\partial q \partial t} > 0 \). Take \( w = B(a,t) \), then Proposition 2 implies that the proportion of workers receiving wage less than \( w \) is \( q(a,t) = a \exp(-\int_a^1 \frac{r(z,t)}{z} dz) \), where \( r(a,t) = \frac{aB'(a,t)}{s(a,t) + aB'(a,t)} \). In other words, we have
\[
w(q(a,t),t) = B(a,t),
\]
where \( q(a,t) = a \exp(-\int_a^1 \frac{r(z,t)}{z} dz) \). This implies that
\[
\frac{\partial w}{\partial q} = \frac{\partial B}{\partial a} = \frac{aB'(a,t)}{a \exp(-\int_a^1 \frac{r(z,t)}{z} dz)(1 + r(a,t))},
\]
and the term in the bracket is
\[
\frac{aB'(a,t)}{a \exp(-\int_a^1 \frac{r(z,t)}{z} dz)(1 + r(a,t))}.
\]
Take derivative with respect to \( t \) to the above expression, we have
\[
\frac{\partial^2 w}{\partial q \partial t} = \exp(-\int_a^1 \frac{r(z,t)}{z} dz)\left\{\frac{\partial aB'(a,t)}{\partial t}(1 + r(a,t)) - aB'(a,t)[\frac{\partial r(a,t)}{\partial t} - \int_a^1 \frac{\partial r(z,t)/\partial t}{z} dz]\right\}
\]
\[
\frac{a \exp(-\int_a^1 \frac{r(z,t)}{z} dz)(1 + r(a,t))^2}{a \exp(-2\int_a^1 \frac{r(z,t)}{z} dz)(1 + r(a,t))^2}.
\]
The denominator of is clearly positive. In the numerator, \( \exp(-\int_a^1 \frac{r(z,t)}{z} dz) > 0 \), and the term in the bracket is
\[
\frac{\partial aB'(a,t)}{\partial t}(1 + r(a,t)) - aB'(a,t)[\frac{\partial r(a,t)}{\partial t} - \int_a^1 \frac{\partial r(z,t)/\partial t}{z} dz]
\]
\[
= \frac{\partial aB'(a,t)}{\partial t} + aB'(a,t) \int_a^1 \frac{\partial r(z,t)/\partial t}{z} dz + \left(\frac{\partial aB'(a,t)}{\partial t} r(a,t) - aB'(a,t) \frac{\partial r(a,t)}{\partial t}\right)
\]
\[
> 0,
\]
where the inequality holds because \( \frac{\partial aB'(a,t)}{\partial t} > \frac{\partial \log s(a,t)}{\partial t} \geq 0, aB'(a,t) \int_a^1 \frac{\partial r(z,t)/\partial t}{z} dz > 0 \) (\( \partial r(z,t)/\partial t > 0 \) follows from the mobility part of the proof), and
\[
\left(\frac{\partial aB'(a,t)}{\partial t} r(a,t) - aB'(a,t) \frac{\partial r(a,t)}{\partial t}\right)
\]
\[
= \left(\frac{\partial aB'(a,t)}{\partial t}\right)^2 r(a,t)^2
\]
\[
= \left(\frac{\partial aB'(a,t)}{\partial t} + \frac{\partial s(a,t)}{\partial t}\right) r(a,t)^2
\]
\[
> 0
\]
This shows that the numerator is also positive, so we have $\frac{\partial^2 w}{\partial q \partial t} > 0$. Thus the wage distribution becomes more spread out when $t$ increases. 

**Theorem 4:** Let the output be $y(a, t) > 0$ for the outside firms and $y(a, t) + ks(a)$ for the incumbent, where $ks(a) > 0$. If the technological change is skill-biased ($\frac{\partial^2 y(a, t)}{\partial a \partial t} > 0$) and $y(a, t) - B(a, t) > ks(a)$ for all $a$, then the proportionate increase in turnover increases with $k$, i.e,

$$\frac{\partial^2 \log(1 - P(a, t, k))}{\partial t \partial k} > 0 \text{ for all } a.$$ 

**Proof.** The proof is based entirely on computation. It is easy to check that

$$\frac{\partial^2 \log(1 - P(a, t, k))}{\partial t \partial k} = -\frac{\partial^2 P(a, t, k)}{\partial t \partial k} (1 - P(a, t, k)) - \frac{\partial P(a, t, k)}{\partial t} \frac{\partial P(a, t, k)}{\partial k}.$$

By Corollary 2 and Theorem 4, we have

$$\frac{\partial P(a, t, k)}{\partial k} > 0 \text{ and } \frac{\partial P(a, t, k)}{\partial t} < 0,$$

so $\frac{\partial P(a, t, k)}{\partial t} \frac{\partial P(a, t, k)}{\partial k} < 0$. Therefore, it suffices to show that $\frac{\partial^2 P(a, t, k)}{\partial t \partial k} < 0$ to prove the theorem.

Now since $P(a, t) = \exp(-\int_a^1 B'(x, t) \frac{dx}{s(x, t) + xB'(x, t)})$, it can be checked that

$$\frac{\partial^2 P(a, t, k)}{\partial t \partial k} = \frac{\partial^2 P(a, t, k)}{\partial k} \frac{\partial P(a, t, k)}{\partial t} + P(a, t, k) \frac{\partial^2 \log(P(a, t, k))}{\partial t \partial k},$$

where $\frac{\partial P(a, t, k)}{\partial k} \frac{\partial P(a, t, k)}{\partial t} < 0$, so it suffices to show that $\frac{\partial^2 \log(P(a, t, k))}{\partial t \partial k} < 0$ to prove the theorem. To do so, we check that

$$\frac{\partial^2 \log P(a, t, k)}{\partial t \partial k} = -\int_a^1 ks(x) \frac{\partial^2 B(x, t)}{\partial x \partial t} \frac{xB'(x, t) - ks(x)}{(ks(x) + xB'(x, t))^3} dx < 0,$$

where the inequality follows because $\frac{\partial^2 B(x, t)}{\partial x \partial t} > 0$ (by Theorem 4) and $xB'(x, t) = y(x, t) - B(x, t) > ks(x)$ (by assumption). This finishes the proof. 

38