PERUVIAN ECONOMIC ASSOCIATION

Worker Mobility in a Search Model with Adverse Selection

Carlos Carrillo-Tudela

Leo Kaas

Worker Mobility in a Search Model
with Adverse Selection *

Carlos Carrillo-Tudela [†]   Leo Kaas [‡]

February 12, 2015

Abstract

We analyze the effects of adverse selection on worker turnover and wage dynamics in a frictional labor market. We consider a model of on-the-job search where firms offer promotion wage contracts to workers of different ability, which is unknown to firms at the hiring stage. With sufficiently strong information frictions, low-wage firms offer separating contracts and hire all types of workers in equilibrium, whereas high-wage firms offer pooling contracts, promoting high-ability workers only. Low-ability workers have higher turnover rates and are more often employed in low-wage firms. The model replicates the negative relationship between job-to-job transitions and wages observed in the U.S. labor market.

Keywords: Adverse Selection, On-the-job search, Worker mobility, Wage dynamics.

JEL: D82; J63; J64

1 Introduction

1.1 Motivation and Summary

The ability of the labor market to allocate resources hinges upon the type and severity of the frictions that prevent workers and firms in forming the most efficient matches. On the one hand, theories of search frictions emphasize the costs associated with finding the right worker or the right job. Theories of adverse selection, on the other hand, stress the importance of asymmetric information at the hiring stage as an impediment for labor turnover.[1] Taken together these frictions can present formidable barriers for worker turnover and efficient resource allocation. Lockwood (1991), for example, suggests that adverse selection exacerbates the negative effects of search frictions by reducing the re-employment chances of unemployed workers. With almost no exceptions, however, current contributions on labor search with adverse selection abstract from job-to-job flows,[2] although these transitions account for a sizeable part of worker flows. Furthermore, the rate at which workers change jobs is an important determinant of wage dynamics (see, e.g., Topel and Ward (1992)). Thus one would expect that asymmetric information not only has non-trivial implications for workers' job turnover, but also for how their wages evolve over time.

In this paper we present a theoretical analysis of the interaction between search frictions, on-the-job search and asymmetric information. Our objective is to study how asymmetric information about workers' abilities affects the mobility of workers within and across firms in a frictional labor market. A key implication of our model is that high-wage firms offer more attractive employment conditions to high-ability workers than to low-ability workers. This implies that low-ability workers have higher turnover rates even though all workers face the same degree of search frictions.[3] We show that our model is quantitatively consistent with the observed negative relationship between wages and the number of job-to-job transitions we uncover using the National Longitudinal Survey of Youth (NLSY). In particular, workers that undertake substantial job changes have on average lower earnings than workers who change relatively fewer times. This is in contrast with standard theories of on-the-job search such as Burdett and Mortensen (1998), which we show predict a positive relationship between wages and the number of job-to-job transitions. We further show that the negative relationship between those variables observed in the data is generated by worker unobserved heterogeneity, as implied by our model.

We consider a frictional labor market similar to Burdett and Mortensen (1998), where workers search randomly for job opportunities and firms commit to long-term wage contracts. In deviation from this benchmark, information is asymmetrically distributed in our model: while workers are perfectly informed about their ability, firms learn workers' ability slowly over time.

---

[1]Search models of the labor market are surveyed in Rogerson et al. (2005). For labor market implications of adverse selection, see e.g. Salop and Salop (1976), Greenwald (1986), Gibbons and Katz (1991).

[2]We review some of this literature in Section 1.2 below.

[3]The empirical findings of Kahn (2013) indeed suggest that workers with higher job mobility patterns are on average of lower ability than those who move relatively less.

Further, firms post a menu of promotion contracts, one for each worker type, to which they are committed. Any contract offers an initial wage based on the worker's reported ability. Upon learning the worker's type, the firm promotes the worker if the worker reported his type truthfully; otherwise the worker is demoted.

When a meeting takes place, the worker chooses whether to accept the job and the terms of employment based on the reported ability. By misreporting his type, a low-ability worker earns a higher initial wage but faces the possibility of demotion accompanied by a wage cut. By reporting truthfully, the worker earns a lower starting wage but faces the prospects of a promotion with a wage rise. This trade-off determines the incentive-compatibility constraint that firms must satisfy if they want to separate workers at the hiring stage. A key result is that the firms' willingness to separate their applicants depends on the degree of information frictions relative to search frictions. Indeed, adverse selection adds a novel trade-off in the firm's problem. By offering high initial wages to attract and retain more high-ability workers, it becomes more costly to satisfy incentive compatibility so as to separate workers. As information frictions increase relative to search frictions, these separation costs become larger. Stronger information frictions (or lower search frictions) imply that workers have a higher chance of moving to another job before the firm learns their type, making it even more attractive for workers to misreport and more difficult for firms to separate at the hiring stage.

In equilibrium, firms follow one of two strategies. Either they decide to offer separating contracts or they offer pooling contracts. Firms offering pooling contracts hire all workers at the same initial wage, promoting high-ability workers and demoting low-ability workers after their types are revealed. Firms offering separating contracts hire workers at different initial wages and promote all workers who self-select into the right contracts. The offer distribution of initial wages for each ability type is always dispersed for similar reasons as in Burdett and Mortensen (1998). We find a cutoff value of the firms' learning rate such that when firms learn sufficiently fast, a *separating equilibrium* emerges where all firms offer separating contracts. Otherwise, there exists a *segmented equilibrium* in which low-wage firms offer separating contracts while high-wage firms offer pooling contracts. The segmented equilibrium has particularly interesting qualitative properties: because high-wage firms demote workers of low-ability, these workers have higher turnover rates than high-ability workers. Hence, they are more frequently employed in low-wage firms who then end up with a less productive workforce.[4] The segmented equilibrium also features rich individual wage dynamics, including wage cuts and wage gains, both within and between firms.

To analyze the quantitative implications, we calibrate our model to match the monthly rates and the wage gains/losses associated with job-to-job transitions, job-to-nonemployment transi-

---

[4]Idson and Oi (1999) suggest that smaller firms are less productive because they attract and retain less productive workers than larger firms. They argue that this is due to complementarities between capital and labor, while they are silent on the way recruitment and retention policies firms use to achieve a more productive workforce. Our theory provides a possible explanation. See Lentz (2010) for a model in which high-ability workers search harder and hence have a higher chance of being employed in high-wage firms.

tions and within-firm promotions, as observed in the U.S. for workers with ten years of potential work experience. Our estimate for the firm's learning rate follows Lange (2007). The segmented equilibrium that arises in the calibration has novel features relative to standard labor search models without information frictions, such as Burdett and Mortensen (1998). In the calibrated model, for example, larger firms have higher internal mobility, pay higher wages, employ a more productive workforce and exhibit less separations (see Idson (1989) and Papageorgiou (2014), among others).

The main implication of our calibrated model pertains to the relationship between the cumulative count of workers' job-to-job transitions and (log) wages. We show using NSLY data that an OLS regression generates a negative relationship between these two variables, after controlling for a large set of observable characteristics including the cumulative count of non-employment spells.[5] Our calibrated model generates this negative relationship, while a Burdett and Mortensen (1998) model predicts a positive relationship.

We further analyze whether worker unobserved heterogeneity might be driving the negative relationship between the number of job-to-job transitions and wages observed in the data. This exercise is important as our model predicts that this should be the case. To test this prediction we control for possible correlations between a time-invariant worker effect in the error term and the cumulative count of job-to-job transitions. Once we control for such a correlation, we find a positive relationship between the cumulative count of job-to-job transitions and wages. Hence, standard job ladder models seem to be fully consistent with the data once we account for worker unobserved heterogeneity. However, those models miss the fact that some workers churn a lot in the labor market and yet remain largely unsuccessful.

The rest of the paper is organized as follows. After a brief review of the related literature, we set out the basic framework in Section 2. In Section 3 we characterize separating and segmented equilibria. Particularly, we show that all firms separate their applicants when the firms' learning rate is high enough; but when information frictions are sufficiently severe, a fraction of high-wage firms offer pooling contracts and end up retaining more high-ability workers. In Section 4 we calibrate our model and explore its implications for worker turnover and wage dynamics. Section 5 discusses alternative learning processes and analyzes an extension where firms can condition their contracts on workers' employment status. In the latter case, we show that employment status gives firms further monopsony power, which in turn makes it easier to separate workers at the hiring stage. Section 6 concludes. All proofs, tedious derivations and data analysis are relegated to the Appendix.

1.2 Related Literature

In their recent survey paper about hiring and incentives, Oyer and Schaefer (2011) call for further theoretical exploration about the effects of asymmetric information at the hiring stage

---

[5]This is inline with previous empirical studies that find that more job turnover is associated with lower average wages (see, e.g., Mincer and Jovanovic (1981) and Light and McGarry (1998)).

in a frictional labor market. We contribute to this literature by developing an equilibrium model of random search in which firms commit to self-selecting contracts to study the effects on worker turnover and wage dynamics. We allow firms to face the risk of losing workers to other firms through competition brought about by workers' on-the-job search. In general, firms may also use other screening devices like aptitude tests, but those can be costly to implement or might not reveal the desired information. Salop and Salop (1976), for example, show how firms can use deferred compensation contracts to separate workers with high and low quit rates. Lazear (2000) and Oyer and Schaefer (2005) emphasize the importance of self-selection contracts in explaining the use of performance-based pay in organizations. Here we explore the role of promotion/demotion contracts as a self-selection device that helps firms to separate applicants at the hiring stage.

Besides a few earlier contributions (e.g. Lockwood (1991), Albrecht and Vroman (1992), Montgomery (1999)), a number of more recent papers study the interrelation between search frictions and adverse selection. Guerrieri et al. (2010) analyze existence and efficiency properties of competitive search models with adverse selection, characterizing separating equilibria where different worker types are employed in different contracts. As they consider a static environment, they cannot discuss worker turnover or wage dynamics. Inderst (2005) analyzes existence of separating equilibria in a model of random search with adverse selection. In his model, the composition of the pool of searching individuals evolves over time. However, once a productive match is formed and a contract agreed, the pair leaves the market. To the best of our knowledge, there are only two papers with on-the-job search under adverse selection. Kugler and Saint-Paul (2004) analyze the effects of firing cost on different types of workers in a model with search on-the-job, assuming however an ad-hoc wage schedule. This is very different from this paper which is interested in optimal wage policies under adverse selection. Visschers (2007) considers a model with random search based on Stevens (2004) and assumes that both the worker and the employer do not observe the worker's (match-specific) ability at the start of the relation. Although the employer learns faster than the worker, it offers the same wage contract to all its new hires.

A few papers consider the interaction of search frictions and adverse selection to study firms' decisions to offer a take-it-or-leave-it wage offer or to engage in bilateral bargaining with their job applicants. Camera and Delacroix (2004), for example, consider a random search model, while Michelacci and Suarez (2006) consider a directed search model to address this issue. As in our paper, firms choose between different contracts which impacts the type of workers they employ. Michelacci and Suarez (2006) show that the labor market segments and firms posting wages attract workers with low productivity, while firms that bargain attract high productivity workers. In our paper we restrict attention to wage posting and we let firms choose between offering separating contracts to hire both types of workers at different wages or posting a pooling contract that provides a higher retention rate for high-ability workers.

This paper also relate to the literature that studies employer learning to analyze worker reallocation. Jovanovic (1979), Moscarini (2005) and Papageorgiou (2014) provide insightful

examples. This literature typically focuses on bilateral asymmetric information where the firm and the worker jointly observe signals about the match quality over time. An important empirical literature has developed from such insights, aiming to test the degree of asymmetric information and the speed of employer learning (e.g. Altonji and Pierret (1996), Lange (2007) and Kahn (2013)). These papers imply that low-ability or poorly matched workers are negatively selected, facing a higher turnover rate. Our paper constructs a tractable equilibrium wage-posting model with adverse selection, on-the-job search and firm learning. In this setting we analyze how the degree of frictions in the labor market and firms' speed of learning affect their hiring and retention strategies. We show that the resulting workers' job mobility and wage patterns are quantitatively consistent with evidence for the U.S. labor market.

2 Basic Framework

2.1 Workers and Firms

Consider the steady state of a continuous time economy which comprises a continuum of workers and a continuum of firms. The life of any worker has uncertain duration and follows an exponential distribution with parameter $\phi > 0$. To keep the population of workers constant, ϕ also describes the rate at which new workers enter the labor market. Firms are infinitely lived. All agents have a zero rate of time preference and are risk neutral. The objective of any worker is to maximize total expected lifetime income, and the objective of any firm is to maximize the expected steady-state profit flow.

There are two types of workers who differ in their innate ability. A mass $\alpha_H = 1$ of workers has high ability p_H and a mass $\alpha_L = \alpha$ has low ability p_L. All firms operate the same constant returns to scale technology, producing flow output p_i with every employed worker of type $i = H, L$. While workers are perfectly informed about their type upon entering the labor market, the firm does not know a worker's ability at the hiring stage. We assume that firms monitor the output of a particular worker at exogenous rate ρ, which describes the firm's learning rate.[6] Once the firm has learned the worker's ability, the latter can be verified in a court of law.

In other respects the information structure mirrors that of the Burdett and Mortensen (1998) model. In particular, firms cannot make offers contingent on the applicants' employment histories. As it will become clear later, this assumption is important since a firm could use information on its applicants' current wages or their wage histories to update its beliefs.[7] However, allowing offers to be contingent on these characteristics also involves dealing with the worker's decision

---

[6]The implicit assumption here is that the firm observes total output, but since it employs a mass of workers it is too costly to observe the output of each individual worker immediately. In Section 5.2 we further discuss the firms' learning process.

[7]In Section 5.1 we discuss a variation of the model in which firms condition their contracts on workers' employment status and analyze whether this information makes it easier for firms to separate workers.

of whether to reveal this information truthfully to the firm. In principle workers would be unwilling to do so since the firm would then be able to condition its contract on workers' current reservation wages, individualizing their offers, and extracting further rents. We do not pursue this possibility here as it would complicate our analysis even further, raising difficult signalling issues at the recruitment stage. Instead we assume that firms must treat equally all applicants that report a given type.

Unemployed and employed workers meet firms according to a Poisson process with parameter $\lambda > 0$. There are also job destruction shocks in that each employed worker is displaced into unemployment according to a Poisson process with parameter $\delta > 0$. Once unemployed, any worker receives a payoff of $b < p_L$ per unit of time. For simplicity we do not allow that workers of different abilities obtain different payoffs when unemployed. The flow payoff b can be interpreted as flow income from unemployment benefits (imposing equal treatment across workers) or as flow utility from leisure (imposing identical leisure preferences).

2.2   Contracts

We consider flat–wage contracts that allow for promotions or demotions when the firm verifies the worker's type. All other transfers between workers and firms (at the hiring stage or thereafter) are ruled out. That is, a contract contains a commitment to three wages: the initial wage (paid prior to learning), a promotion wage and a demotion wage. When a firm meets a worker, the firm offers a menu of two contracts, indexed by the worker's reported ability $i = H, L$. We denote these contracts as $\omega_i = (w_i, w_i^+, w_i^-)$ where w_i denotes the initial wage that is paid in the probation period, i.e. before learning the worker's type. When worker i reports his ability truthfully, he is promoted to receive the promotion wage w_i^+ as long as he stays employed in the firm. If worker $j \neq i$ misreported his type in contract ω_i, the worker is demoted in which case the worker is paid the demotion wage w_i^- for as long as he stays employed in the firm. The firm commits to pay the initial wage w_i until the worker's type is revealed and to follow the promotion/demotion commitment thereafter.[8] While firms commit to wage profiles, they cannot commit to retain workers that yield negative expected profit value.[9]

Upon meeting the firm, the worker observes the posted contracts and can choose one of them, but nothing restricts the worker from choosing the contract the firm designs for workers of a different ability level. If both contracts are rejected, however, the worker remains in his current state with no option to recall previously met firms. We make the following tie-breaking assumptions: an unemployed worker accepts a wage offer if indifferent to accepting it or remaining unemployed, while an employed worker quits only if the offer is strictly preferred.

---

[8]In a previous working paper version (Carrillo-Tudela and Kaas (2011)), we consider a contractual environment in which firms are restricted to offer flat wage contracts without promotions, but can threaten to fire workers who misreport their type upon learning. We also consider contracts allowing wage cuts for misreporting workers.

[9]The underlying assumption here is that both the worker and the firm are free to initiate a separation at any time. But even if firms were able to commit to employ unprofitable workers, our main segmentation results would survive; see also footnote 18 below.

To summarize, the main restrictions we impose on the contract space are: (i) equal treatment at the hiring stage to all workers that report a given type; (ii) flat-wage contracts where promotions and demotions occur together with learning events; and (iii) no side payments.

2.3 Equilibrium

When a worker of type $i = H, L$ encounters a new employer offering contract menu $[\omega_H, \omega_L] = [(w_H, w_H^+, w_H^-), (w_L, w_L^+, w_L^-)]$, the worker may report his type truthfully, which leads to continuation utility $V_{ii}(\omega_i)$, or the worker may misreport his type with continuation utility $V_{ij}(\omega_j)$ for $i \neq j$. The flow gain from misreporting the type is the difference in initial wages, $w_j - w_i$. After the firm learns the worker's type, the worker either receives continuation utility $V_i(w_i^+)$, if he reported his true type, or $V_i(w_j^-)$, if the worker misreported his type. Since learning occurs at Poisson rate ρ, worker i's incentive constraint $V_{ii}(\omega_i) \geq V_{ij}(\omega_j)$ is equivalent to

$$w_j - w_i \leq \rho[V_i(w_i^+) - V_i(w_j^-)] . \tag{1}$$

This incentive constraint describes the main trade-off faced by a worker when meeting a firm. By misreporting his type, worker i earns potentially the higher initial wage w_j but faces the possibility of demotion with continuation utility $V_i(w_j^-)$. By reporting truthfully, the worker possibly earns a lower starting wage but faces the prospects of a promotion, yielding continuation utility $V_i(w_i^+)$. The worker will report his type truthfully and self-select into the right contract when the flow gain from misreporting does not exceed the expected gain from a promotion relative to a demotion.[10] In Appendix A we present the recursive equations for V_{ii}, V_{ij} and V_i and derive this incentive constraint formally. We write U_i for the utility value of unemployment for a type-i worker.

Firms choose contract pairs (ω_H, ω_L) to maximize the steady-state profit value $\Omega_H(\omega_H) + \Omega_L(\omega_L)$, where $\Omega_i(.)$ denotes the firm's profit from hires in contract ω_i. Firms take into account that workers may misreport their type whenever the incentive constraint (1) fails in which case the firm pools all workers in the same contract. In such cases, the profit value $\Omega_i(\omega_i)$ also includes the gains (or losses) from workers of type j who are later demoted to wage w_i^-.

A *market equilibrium* is a joint distribution Ψ of contract offers, and value functions U_i, V_i, V_{ii}, V_{ij}, Ω_i, $i, j = H, L$, for workers and firms such that (i) every (ω_H, ω_L) in the support of Ψ maximizes firms' profits subject to optimal turnover and truth-telling behavior of workers; and (ii) workers' turnover and reporting strategies are optimal given that new contract offers are drawn randomly from distribution Ψ at Poisson rate λ.

In the following subsection, we consider those market equilibria where the promotion and demotion wages are maximally differentiated so that the incentives to report truthfully are as large as possible, given the constraint that workers and firms may voluntarily quit any employment

---

[10]Note that this trade-off is similar to the one found in efficiency wage models, in which shirking (misreporting) yields a higher utility (initial wage) but implies a higher probability of being caught and fired (demoted).

relationship. These constraints imply in particular that $w_i^+ \leq p_i$ (otherwise a firm would fire the worker) and $w_i^- \geq b$ (otherwise the worker would quit to unemployment). In spite of strong incentives, however, a market equilibrium does not always give rise to a perfect separation of workers. In fact, as we show in Section 3, some firms may decide to pool all workers in the same contract, so that low-ability workers are demoted in equilibrium. Those workers then experience higher turnover rates as well as upward and downward wage mobility, both within and between firms.

2.4   Equilibrium Restriction

In the following we restrict the analysis to those market equilibria where all firms offer contracts of the form $\omega_i = (w_i, p_i, b)$, which we label *simple contracts*. That is, firms demote misreporting workers by cutting their pay to the reservation wage,[11] while they promote truth-telling workers to their marginal product. This maximally feasible differential between promotion and demotion wages provides the greatest possible incentives for workers to report their type truthfully upon hiring. Further, backloading of wages for truth-telling workers serves the purpose of minimizing turnover while earning positive profits on lower initial wages prior to learning.

In Section 3, we characterize all *candidate equilibria* where firms are restricted to choose simple contracts. Before we do that, however, we show in the following lemma that those candidate equilibria are indeed market equilibria (where firms are allowed to choose arbitrary contracts (w_i, w_i^+, w_i^-)). This result holds generally when all firms offer separating contracts. In those cases where some firms pool workers, a specific parameter restriction is required to guarantee that the candidate equilibrium in simple contracts is indeed a market equilibrium. For the plausibly calibrated parameter combinations that we consider in Section 4, we verify that this condition is satisfied so that the considered promotion/demotion strategies are in fact equilibrium outcomes.[12] For a given candidate equilibrium in simple contracts, we write F_i for the marginal distribution of starting wage offers to workers of type i. Further, let R_i denote the reservation starting wage for an unemployed worker of type i who reports truthfully when meeting a firm. This reservation starting wage is smaller than unemployment income b, since truth-telling workers can expect a promotion to wage $p_i > b$ with positive probability. In Appendix A we show that all workers indeed follow a standard reservation-wage strategy.

Lemma 1: *Consider a candidate equilibrium in simple contracts where F_i is the marginal distribution of wage offers to workers of type i and R_i is worker i's reservation starting wage, i.e. $V_{ii}(R_i, p_i, b) = U_i$. Then the candidate equilibrium is a market equilibrium if all firms offer separating contracts. If some firms pool all workers in the same contract, the candidate*

---

[11]Since the offer arrival rate is independent of a worker's employment status, b is the common reservation wage for both worker types.

[12]If the parameter condition fails, pooling firms would offer dispersed demotion wages which considerably complicates the equilibrium analysis without providing additional economic insights. We therefore only consider equilibria in simple contracts which already give rise to rich wage dynamics between and across firms.

equilibrium is a market equilibrium if the condition $\Gamma(w) \leq \Gamma(R_L)$ holds for all $w \in [R_L, p_L]$, where

$$\Gamma(w) \equiv \frac{p_L - b - \int_{R_L}^{w} \frac{\phi + \delta + \lambda(1 - F_L(w'))}{\phi + \delta + \rho + \lambda(1 - F_L(w'))} dw'}{\phi + \delta + \lambda(1 - F_L(w))} \ .$$

The condition $\Gamma(w) \leq \Gamma(R_L)$ is needed to ensure that the demotion wage $w_L^- = b$ is indeed optimal for all pooling firms. For equilibria in simple contracts, we further obtain a convenient formulation of the incentive constraint, stated as follows.

Lemma 2: *In any equilibrium in simple contracts, the incentive constraint for type i workers is*

$$w_j - w_i \leq b - R_i \ . \tag{2}$$

3 Equilibrium Analysis

3.1 Preliminary Considerations

Before we analyze the different equilibria in simple contracts, we discuss a few general properties. A first insight is that firms offer contracts with dispersed starting wages which gives rise to worker turnover before the promotion/demotion events. That is, the distributions F_i of starting wages are non-degenerate. Workers employed at starting wages quit to other contracts if and only if they provide strictly higher continuation value. The reason for contracts with dispersed starting wages is that promotions and demotions are linked one-to-one to employer learning which is a stochastic event. Since promotion dates are uncertain, workers quit to contracts offering higher starting wages during the probation period. Firms, in turn, respond to these incentives by offering dispersed contracts, trading off higher flow profits against higher attraction and retention rates, similar to Burdett and Mortensen (1998).[13] Promoted workers of type H (earning $w_H^+ = p_H$) never quit since no firm can offer a greater continuation value. However, there can still be turnover of low-ability workers after they have been promoted or demoted, as we see below.

We restrict the analysis to rank-preserving wage policies: firms that offer higher starting wages to high-ability workers also offer higher starting wages to low-ability workers; that is, there is an increasing function $\hat{w}(.)$ such that $F_L(\hat{w}(w_H)) = F_H(w_H)$ for all wages w_H in the support of the offer distribution F_H. If any of the incentive constraints (2) binds (on some interval), rank preservation follows directly because w_L increases one-for-one with w_H (no matter which constraint binds). But if incentive constraints are slack in some range of the wage offer

---

[13]See Stevens (2004) for the case in which firms pre-commit to a promotion date without facing the adverse selection problem we introduce in our paper.

distributions, any wage offer w_L in this range satisfying incentive compatibility yields the same profit, so that wage offers across firms can be ordered in a rank-preserving way.[14] Before we proceed, note that offering initial wage $w_i = R_i$ strictly dominates offering any $w_i < R_i$ as it generates strictly positive profit while the latter leads to zero profit. And because contract (R_i, p_i, b) attracts unemployed workers, R_i is the lower bound of the support of F_i.

It is useful to note that in any equilibrium the incentive constraint does not bind at low wages and hence firms offering these wages will always separate workers at the hiring stage. To see this consider the lowest contract offered in equilibrium, (R_H, R_L). Condition (2) implies that when meeting a firm offering this contract, a high-ability worker's incentive constraint does not bind whenever $R_L < b$ and a low-ability worker's incentive constraint does not bind whenever $R_H < b$. Since $R_i < b$ for $i = L, H$, the previous conditions are always satisfied. Rank preservation then implies slack incentive constraints in the neighborhood of (R_H, R_L).

A similar argument shows that incentive constraints do not bind for high-ability workers. For these workers, (2) is given by $w_L - w_H = \hat{w}(w_H) - w_H \leq b - R_H$. As we show in the proofs of Propositions 1 and 2, the slope of function $\hat{w}$ does not exceed one, hence the left-hand side is weakly decreasing in w_H, so that the incentive constraint holds for all wages in the support of the wage offer distribution if it is satisfied at $w_H = R_H$. The latter condition amounts to $\hat{w}(R_H) = R_L \leq b$, which is always satisfied.

Lastly, note that the arguments of Burdett and Mortensen (1998) imply that the offer distribution F_H is continuous and has connected support when firms are able to separate workers. When some firms offer pooling contracts, however, we will see that F_H exhibits a mass point in the range of wages in which firms offer pooling contracts.

We now proceed to fully characterize the equilibria in simple contracts that can arise in this model. First we consider equilibria where the contracts offered by all firms are incentive compatible. In those situations, which requires the learning rate ρ to be sufficiently large, all firms promote workers who stay long enough with the firm, and they never exercise a demotion option. In contrast, if the learning rate is sufficiently low, the market equilibrium is segmented, featuring some pooling contracts at the top of the wage offer distribution, with promotions of high-ability workers and demotions of low-ability workers. To save notation, we write instead of the contract pair (ω_H, ω_L) the pair of initial wages (w_H, w_L), and we also write all value function expressions in those starting wages.

---

[14]When incentive constraints are slack, our restriction to rank-preserving wage policies has implications for the relationship between firm size, wages and productivity we discuss in Section 4.2. Without it, firms that offer contracts with slack incentive constraints might not exhibit a positive relationship between wages, size and workforce productivity.

3.2 Separating Equilibrium

Workers

In a candidate separating equilibrium the value functions of unemployed workers $i = L, H$ are given by

$$(\phi + \lambda)U_i = b + \lambda \int_{R_i}^{\overline{w}_i} V_{ii}(w) \, dF_i(w) \ . \tag{3}$$

For employed workers $i = L, H$ in probation contracts with initial wage w_i, we have

$$[\phi + \delta + \rho + \lambda(1 - F_i(w_i))]V_{ii}(w_i) = w_i + \delta U_i + \rho V_i(p_i) + \lambda \int_{w_i}^{\overline{w}_i} V_{ii}(w) \, dF_i(w) \ . \tag{4}$$

For promoted workers we have

$$[\phi + \delta]V_i(p_i) = p_i + \delta U_i \ ,$$

since these workers never quit in a separating equilibrium.[15] Note again that in equation (4) the reservation wage of an employed workers is given by the current (initial) wage since all firms offer the same promotion wage (for a given type). The reservation (initial) wages of unemployed workers are given by

$$R_i = b - \rho \frac{p_i - \phi U_i}{\phi + \delta} < b \ . \tag{5}$$

The incentive constraint for low-ability workers is

$$w_H - w_L \leq b - R_L \ . \tag{6}$$

As discussed before, high-ability workers do not misreport their type so that their incentive constraint never binds. Also incentive constraints for low-ability workers are always slack at the lowest wage of the offer distribution. Possibly, however, incentive constraints for these workers are binding at higher wages. We denote the critical threshold wages by $\tilde{w}_i$ such that the incentive constraint (6) is binding for $w_i \geq \tilde{w}_i$ and slack otherwise.

Steady-State Measures

Simple steady-state considerations yield the unemployment rate for both types,

$$u = \frac{(\phi + \delta)}{\phi + \delta + \lambda} \ ,$$

as well as the earnings distribution of workers employed at initial wages below or equal to w_i:

$$G_i(w_i) = \frac{(\phi + \delta)F_i(w_i)}{\phi + \delta + \rho + \lambda(1 - F_i(w_i))} \ .$$

---

[15]Since all firms offer incentive-compatible contracts, low-ability workers earning p_L cannot gain utility by accepting an outside offer. This will be different when firms at the top of the offer distribution post pooling contracts.

Profit Maximization

Firms' steady-state profit flows are given by $\Omega_H(w_H) + \Omega_L(w_L)$ where

$$\Omega_L(w_L) = \ell_L(w_L)(p_L - w_L) \qquad , \qquad \Omega_H(w_H) = \ell_H(w_H)(p_H - w_H) \ ,$$

and $\ell_i(w_i)$ denotes the steady-state workforce of workers of ability i who are employed at initial wage w_i. Steady state and optimal turnover then imply that in this candidate equilibrium

$$\ell_i(w_i) = \frac{\lambda \alpha_i [u + G_i(w_i)(1 - u)]}{\phi + \delta + \rho + \lambda(1 - F_i(w_i))} \ .$$

The firms' profit is therefore given by

$$\Omega_i(w_i) = \frac{A_0 \alpha_i}{[\phi + \delta + \rho + \lambda(1 - F_i(w_i))]^2}(p_i - w_i) \ , \text{ where } A_0 \equiv \frac{\lambda(\phi + \delta + \rho + \lambda)(\phi + \delta)}{\phi + \delta + \lambda} \ . \tag{7}$$

Profit-maximizing firms are indifferent between all contracts in the support of the wage-offer distribution. When incentive constraints do not bind (i.e., $w_i < \tilde{w}_i$), the constant profit condition holds for each ability type independently. Equating $\Omega_i(w_i) = \Omega_i(R_i)$ yields offer distributions with a similar functional form as in Burdett and Mortensen (1998). Namely,

$$F_i(w_i) = \frac{(\phi + \delta + \rho + \lambda)}{\lambda}\left[1 - \left(\frac{p_i - w_i}{p_i - R_i}\right)^{1/2}\right], \ w_i \le \tilde{w}_i \ . \tag{8}$$

For wages above $\tilde{w}_i$, the incentive constraint (6) binds. Substitution into the firm's profit function gives, for $w_H \ge \tilde{w}_H$,

$$\Omega_H(w_H) + \Omega_L(w_H - b + R_L) = \frac{A_0[p_H - w_H + \alpha(p_L - w_H + b - R_L)]}{[\phi + \delta + \rho + \lambda(1 - F_H(w_H))]^2} \ .$$

Now the constant-profit condition $\Omega_H(w_H) + \Omega_L(w_H - b + R_L) = \Omega_H(R_H) + \Omega_L(R_L)$ yields the wage-offer distribution

$$F_H(w_H) = \frac{(\phi + \delta + \rho + \lambda)}{\lambda}\left[1 - \left(\frac{\overline{p} - w_H(1 + \alpha) + \alpha(b - R_L)}{\overline{p} - \overline{R}}\right)^{1/2}\right], \ w_H \ge \tilde{w}_H \ , \tag{9}$$

where we define $\overline{p} \equiv p_H + \alpha p_L$ and $\overline{R} \equiv R_H + \alpha R_L$.

Rank Preservation

Using (8) and $F_L(\hat{w}(w_H)) = F_H(w_H)$, we obtain that

$$w_L = \hat{w}(w_H) = p_L - \left(\frac{p_L - R_L}{p_H - R_H}\right)[p_H - w_H] \ , \ w_H \le \tilde{w}_H \ . \tag{10}$$

This equation then implies that the incentive constraint (6) is fulfilled for all wages

$$w_H \le \tilde{w}_H \equiv \frac{b(p_H - R_H) - R_H(p_L - R_L)}{p_H - R_H - p_L + R_L} \ . \tag{11}$$

For wages above $\tilde{w}_H$, the binding incentive constraint defines $w_L = \hat{w}(w_H) = w_H - b + R_L$.

Slack Incentive Constraints

If the firms' learning rate is sufficiently large (above threshold level ρ_1 defined below), the incentive constraint is slack for all wages in the wage distribution. In this situation, the threshold wage $\tilde{w}_H$ exceeds the upper bound $\overline{w}_H$ of the distribution of initial wage offers. Using (8) and $F_i(\overline{w}_i) = 1$, we can solve for the upper bounds

$$\overline{w}_i = p_i - C^2(p_i - R_i) , \ i = L, H, \text{ where } C \equiv \frac{\phi + \delta + \rho}{\phi + \delta + \rho + \lambda} . \tag{12}$$

To find reservation wages, rewrite equations (3) and (5) to obtain

$$\frac{\phi + \delta}{\rho}(R_i - b) + p_i - b = \lambda \int_{R_i}^{\overline{w}_i} [V_{ii}(w) - V_{ii}(R_i)] \, dF_i(w) = \int_{R_i}^{\overline{w}_i} \frac{\lambda(1 - F_i(w))}{\phi + \delta + \rho + \lambda(1 - F_i(w))} dw , \tag{13}$$

where the last equation uses partial integration and the derivative of (4). Solving the integral using (8) yields

$$R_i = \frac{(\phi + \delta + \rho)(\phi + \delta + \rho + \lambda)^2 b - \rho[(\phi + \delta + \rho + \lambda)^2 - \lambda^2]p_i}{(\phi + \delta)(\phi + \delta + \rho + \lambda)^2 + \lambda^2 \rho} . \tag{14}$$

Since promotion wages for high-ability workers are higher, we have $R_H < R_L$.

From (12) and (14) we obtain the top initial wages:

$$\overline{w}_i = p_i - \frac{(\phi + \delta + \rho)^3}{(\phi + \delta)(\phi + \delta + \rho + \lambda)^2 + \lambda^2 \rho}[p_i - b] . \tag{15}$$

Since $w_H - w_L = w_H - \hat{w}(w_H)$ is increasing in w_H, the incentive constraint (6) is slack for all wages if it is slack at $\overline{w}_H$. Using (14) and (15), this is true if and only if

$$\frac{p_L - b}{p_H - p_L} > \frac{\lambda^2 + 2\lambda(\phi + \delta) - \rho(\phi + \delta + \rho)}{\rho(\rho + \phi + \delta + 2\lambda)} . \tag{16}$$

This inequality defines a unique threshold level $\rho_1 > 0$ for the firms' learning rate such that incentive constraints are slack for all $\rho > \rho_1$.[16]

Binding Incentive Constraints

If $\rho \leq \rho_1$, incentive constraints must be binding for some wages at the top of the wage offer distribution so that $\tilde{w}_H \leq \overline{w}_H$. In this case we can use $F_H(\overline{w}_H) = 1$ with (9) to find

$$\overline{w}_H = \frac{1}{1 + \alpha}\left[\overline{p} + \alpha(b - R_L) - C^2(\overline{p} - \overline{R})\right] . \tag{17}$$

---

[16]To verify this claim rewrite (16) as $\frac{p_L - b}{p_H - p_L}[\rho(\rho + \phi + \delta + 2\lambda)] = \lambda^2 + 2\lambda(\phi + \delta) - \rho(\phi + \delta + \rho)$. Note that the LHS of this equation equals zero at $\rho = 0$ and is strictly increasing in ρ, while the RHS is positive at $\rho = 0$ and is strictly decreasing in ρ. Continuity implies that there exists a unique $\rho_1 > 0$ that solves this equation, such that incentive constraints are slack for all $\rho > \rho_1$.

Reservation wages for the two worker types are again obtained from equations (13). In Appendix B (proof of Proposition 1) we show how these equations can be solved for R_L and R_H. Given these solutions, we can back out $\tilde{w}_i$ from (11), $\overline{w}_H$ from (17), $\overline{w}_L = \overline{w}_H - b + R_L$, and the wage-offer distributions from (8) and (9). All these equilibrium objects are uniquely defined.

The candidate equilibrium characterized by these equations only constitutes an equilibrium if the highest initial wage offered to low-ability workers $\overline{w}_L$ does not exceed their marginal product p_L. If $\overline{w}_L > p_L$, firms could not credibly offer such contracts to low-ability workers since they would prefer to layoff the worker after he signed up. In fact, if the learning rate is sufficiently low, and given a sufficiently high job-arrival rate λ, it is possible that the highest incentive compatible wage offer to low-ability workers exceeds their marginal product. We denote $\rho_2 < \rho_1$ as the threshold value for the learning rate such that $\overline{w}_L \leq p_L$ if $\rho \geq \rho_2$. This threshold value is strictly positive if the parameter condition

$$p_H - p_L > \left(\frac{\phi + \delta}{\lambda + \phi + \delta} \right)^2 [\overline{p} - (1 + \alpha)b] \tag{18}$$

is satisfied, which requires the job arrival rate λ to be sufficiently large.[17] Intuitively, for larger values of λ, firms compete more fiercely for workers which drives up the highest wage offers. At the same time, it becomes harder to separate workers because misreporting workers quit faster to another job before they can be demoted. Furthermore, the threat of a demotion is weaker if workers quickly find other jobs. For these reasons, it may not be feasible for firms to offer separating contracts at the top of the wage-offer distribution when λ is relatively large.

To verify that there is no profitable deviation for firms from the characterized candidate equilibrium, we still need to make sure that firms do not find profitable deviations to a simple pooling contract. In Appendix B, we prove that this claim is fulfilled provided that $\rho \geq \rho_2$, and hence $\overline{w}_L \leq p_L$. By virtue of Lemma 1, the candidate equilibrium with separating firms is a market equilibrium. Thus, we have

Proposition 1: *There are threshold values $\rho_1 > \rho_2 \geq 0$, such that for any $\rho \geq \rho_2$, there exists a unique market equilibrium in simple contracts with dispersed offers in initial wages w_i drawn from distributions F_i and support $[R_i, \overline{w}_i]$, with $R_i < b$ and $\overline{w}_i \leq p_i$, and separation of workers such that*

(a) if $\rho > \rho_1$, incentive constraints are slack for all firms;

(b) if $\rho \leq \rho_1$, incentive constraints are binding for firms offering $w_H \geq \tilde{w}_H$ with $\tilde{w}_H \leq \overline{w}_H$, while they are slack for all other firms.

Moreover, ρ_2 is strictly positive if condition (18) is satisfied.

---

[17]In the proof of Proposition 1, we establish the existence of the threshold value ρ_2. While cumbersome analytical expressions complicate a uniqueness proof, all our numerical examples suggest a unique value ρ_2 since the defining equation is monotonic in ρ.

3.3 Segmented Equilibrium with Pooling at Top Wages

If $\rho < \rho_2$, there is no equilibrium in which all firms are able to separate workers by offering an incentive-compatible contract menu. In such situations, firms at the top of the wage-offer distribution decide to offer pooling contracts designed to attract and retain high-ability workers. Low-ability workers accept these contracts, anticipating that firms demote them to their reservation wage after learning their type.

Firms at the top of the wage distribution pool workers of high and low ability in the same contract at initial wages $w_H > w_H^* \equiv \hat{w}^{-1}(p_L) = p_L + b - R_L$. Such situations occur if the job-arrival rate is sufficiently large (inducing firms to compete more fiercely for workers) and the learning rate is sufficiently small (so that incentive-compatible initial wages are sufficiently close).[18]

Whenever firms offer pooling contracts $w_H > w_H^*$, they attract all workers of low-ability that have been promoted to p_L at another (lower wage) firm since these workers strictly prefer the job with a higher initial wage even though they expect to be demoted later on. Since there is a positive mass of such workers which yield negative expected profit value when employed in a pooling contract $w_H > w_H^*$, the profit value of firms offering $w_H^* + \varepsilon$ would jump down discontinuously, *unless* the wage offer distribution has a mass point at w_H^*. With a mass point of the wage offer distribution at wage w_H^*, a positive mass of high-ability workers employed at w_H^* quit their job to outside offers $w_H^* + \varepsilon$, such that in equilibrium this profitable inflow of high-ability workers exactly offsets the unprofitable inflow of low-ability workers.[19] Furthermore, at the mass of firms offering the incentive-compatible contract pair (w_H^*, p_L), equilibrium requires that a positive fraction of low-ability workers do not self-select into contract p_L but instead misreport high ability by choosing w_H^*.[20] In fact, we prove the following:

Lemma 3: *If $\rho < \rho_2$, a positive mass of firms offer contract menu (w_H^*, p_L). Low-ability workers contacted by these firms report high ability with positive probability.*

In what follows we characterize an equilibrium where a positive mass of firms offer contract

---

[18]It is important to note that pooling at top wages will still occur even if we were to allow firms to offer low-ability workers wages above their marginal product. This arises because firms offering the highest wages to high-ability workers will also need to offer high wages to low-ability workers in order to satisfy their incentive constraint. Given a small enough ρ relative to λ, such a strategy becomes too costly for high-wage firms which then prefer to hire all workers at the same initial wage, promoting high-ability workers and demoting low-ability ones after learning their types. See Carrillo-Tudela and Kaas (2011) for the formal argument using constant wage contracts.

[19]In the Burdett and Mortensen (1998) model, mass points in the offer distribution are ruled out because higher wage offers would lead to a profitable inflow of workers employed at the mass point. Here this inflow is needed precisely to compensate for the losses on low-ability hires.

[20]Even though the low-wage contract is incentive-compatible, low-ability workers are indifferent between the two contracts and they may equally well accept the contract with the higher initial wage. In principle, such deviations could also occur at lower wages with binding incentive constraints, but firms would easily prevent those by paying $\varepsilon > 0$ more to workers of low ability. At the contract pair (w_H^*, p_L), however, such counter deviations are not possible, since firms cannot credibly offer wages above p_L to workers of low ability.

pair (w_H^*, p_L) such that a fraction of low-ability workers misreport their type, so that some pooling occurs at these firms. Possibly there is also a mass of firms offering pooling contracts $w_H > w_H^*$. We can equivalently interpret such a pooling contract as a menu of contract pairs (w_H, p_L) where p_L is so unattractive that low-ability workers do not accept this offer but instead report high ability with starting wage w_H. Without loss of generality and to keep the notation consistent throughout, we specify the analysis in terms of such contract pairs where initial wages are linked according to $w_L = \hat{w}(w_H)$, such that $p_L = \hat{w}(w_H)$ for $w_H > w_H^*$, thus violating incentive compatibility (6).

In a segmented equilibrium, the different types of firms can be ranked according to their wage offers $w_H \in [R_H, \overline{w}_H]$ as follows.

1. For $R_H \leq w_H < \tilde{w}_H$, firms offer separating contracts with slack incentive constraints.

2. For $\tilde{w}_H \leq w_H < w_H^* = p_L + b - R_L$, firms offer separating contracts with binding incentive constraints.

3. Mass $\eta > 0$ of firms offer the contract menu (w_H^*, p_L). Low-ability workers misreport their type with probability $\xi > 0$.

4. Firms offering $w_H > w_H^*$ pool all workers in the same contract, promoting high-ability workers and demoting low-ability workers.

The last group of firms only exists if the learning rate is sufficiently low. In fact, in our numerical examples, we determine a threshold value of the learning rate, denoted ρ_3 ($< \rho_2$), such that a positive mass offer pooling contracts at $w_H > w_H^*$ if $\rho < \rho_3$. For $\rho \in [\rho_3, \rho_2)$, in contrast, the highest pooling wage is at w_H^*. Although we do not have an existence proof for ρ_3, in Proposition 2 we prove the existence of a pooling equilibrium for both these cases together.

To describe the equilibrium, suppose that mass $\eta > 0$ of firms offer the contract menu (w_H^*, p_L), and that fraction ξ of low-ability workers who are offered these contracts opt for w_H^*, thus pooling with high-ability workers. Denote by $\varphi = F_{L-}(p_L) = F_{H-}(w_H^*)$ the fraction of firms offering separating contracts strictly below (w_H^*, p_L). When the mass point is the highest offered wage ($\rho \geq \rho_3$), we have that $\varphi + \eta = 1$; otherwise $\varphi + \eta < 1$.

In Appendix B (proof of Proposition 2) we characterize a candidate equilibrium in simple contracts by a set of equations determining the vector of equilibrium objects $\mathcal{E} \equiv (\varphi, \xi, \eta, R_L, R_H)$ and we also prove that such a candidate equilibrium exists. Under the conditions of Lemma 1, the candidate equilibrium is a market equilibrium:

Proposition 2: *For any $\rho < \rho_2$, there exists a candidate equilibrium in simple contracts with dispersed offers in initial wages w_i drawn from distributions F_i and support $[R_i, \overline{w}_i]$, with $R_i < b$. The candidate equilibrium is a market equilibrium if the condition stated in Lemma 1 holds. Firms with $w_H < w_H^* = p_L + b - R_L$ offer separating contracts to all workers. There is*

17

also a positive mass of firms offering $w_H \geq w_H^*$ who pool low-ability workers in the same contract as high-ability workers.

We briefly discuss some of the implications for labor turnover in an equilibrium with pooling at top wages. First, low-ability workers have a higher degree of turnover. While high-ability workers are promoted at rate ρ and subsequently leave the job at rate $\phi + \delta$, low-ability workers expect a demotion to wage b at flow rate ρ, with subsequent quits at rate λ (in addition to separation and labor market exit risks). Second, the segmented equilibrium features wage gains and wage cuts, both within firms (promotions and demotions) and between firms. Particularly, demoted low-ability workers are willing to quit to new employers offering promotion contracts with starting wages below b. Third, low-ability workers are underrepresented in high-wage (pooling) firms, and therefore high-wage firms are more productive. Conversely, firms offering separating contracts have a higher proportion of low-ability workers in their workforces. Furthermore, since quit rates are falling in wages, the proportion of high-ability workers is increasing in $w_H > w_H^*$ among pooling firms. The intuition is that high-wage firms are able to attract and retain more workers of both types, while they demote misreporting low-ability workers at the same rate ρ (who subsequently quit at rate λ), independent of the offered initial wage.[21] We summarize these observations as follows.

Corollary 1: *If $\rho \geq \rho_2$, both worker types have the same turnover patterns, no worker experiences wage cuts without intervening unemployment spells, and all firms have the same ability composition of the workforce. If $\rho < \rho_2$, low-ability workers have higher turnover rates, and they experience wage cuts within and between firms. Firms offering pooling contracts (high-wage firms) have a more productive workforce than firms offering separating contracts (low-wage firms). Among high-wage firms with $w_H > w_H^*$, the workforce productivity is increasing in w_H.*

4   Quantitative Implications

In this section we show that our model delivers predictions about the relationships between job mobility and wages and between internal mobility and firm size which are in line with empirical evidence. At the same time, we show that those predictions are not easily picked up by standard job-ladder models.

We calibrate our model to match the following statistics for the U.S. labor market. Set the time period to a month and let $\phi = 0.002$ reflect an average working life of 40 years. We set $\rho = 0.027$ to reflect an average learning period of three years, which we take from Lange (2007).[22]

---

[21]Due to the mass point at w_H^*, the relationship between firm size and workforce productivity is generally non-monotone.

[22]Lange (2007) estimates a model in which employers learn gradually the ability of their employees. His estimate implies that employers' initial expectation errors decline by 51% during the first three years of employment and decline to 36% of their initial value after five years. This suggests that a large part of employer learning occurs in the first few years of an employment relationship. Although we consider a different learning process, we take

We set $\delta = 0.02$ to reflect the adjusted layoff rate of Davis et al. (2008) and normalize $b = 1$. We set the remaining four parameters, λ (meeting rate), p_L and p_H (worker productivities), and α (mass of low-ability workers) to match the following four calibration targets: (1) A monthly job-to-job transition rate of 2.8%, which corresponds to the average monthly EE rate reported by Nagypal (2005) based on Current Population Survey data for all workers and which is also close to her estimates for workers aged 25–34. (2) An average promotion gain of 8%, which corresponds to the estimate of Pergamit and Veum (1999) based on workers with approximately 10 years in the labor market in the NLSY (see Table 5, specification 2 in their paper).[23] (3) An average quit gain of 3%. (4) An average layoff loss of 4%. These last two calibration targets correspond to our own estimation results using the NLSY for workers with approximately 10 years in the labor market.[24]

To calculate the corresponding model statistics, we solve our model for the stationary equilibrium characterized above. We then simulate the labor market transitions of 100,000 workers during the first ten years of their work life. The parameters are estimated by minimizing the sum of squared distances between the simulated moments and the data moments. Table 1 shows the calibrated parameter values and how our model matches the calibration targets stated above.

Table 1: Parameter choices and calibration targets.

Parameter	Value	Model statistics	Value
λ	0.118	Monthly EE flow	2.7%
p_L	1.004	Quit gain	3.0%
p_H	1.131	Promotion gain	8.0%
α	0.324	Layoff loss	4.2%

The calibrated model implies that in equilibrium 29% of firms offer separating contracts,

---

the Lange (2007) estimate as the best available for our purposes and assume an average learning period of three years. Using an average learning period of four or five years does not change our main conclusions. In Section 5.2 we briefly discuss the possible implications of gradual learning in our setting.

[23]Consistent with the interpretation of our model, Pergamit and Veum (1999) find that in the NLSY most worker-reported promotions do not involve a change in job title and nearly all promotions involve a wage increase. In particular, they find that around 56% of those who received a promotion did not change job title and essentially performed the same duties as before. Only 14% were promoted to a higher-level job in a different section, while the rest obtained new jobs because of reorganization, a lateral move or because they took their supervisor's job. Of all promotions, 89% led to a wage increase. The reported promotion gain of 8% is the average wage change for all types of promotions. But even promotions that involved performing the same duties as before raised wages by about 7%.

[24]These estimates are obtained by using first differences on the regressions used in Section 4.1 and described in Appendix C for the period 1987-1994, which is chosen so that the average potential labour market experience is 10 years. The wage gain estimate after a job-to-job transition is consistent with that of Pergamit and Veum (1999), Table 5, specification 2.

while 71% of firms offer pooling contracts.[25] The equilibrium unemployment rate is 16.8%, and reservation wages are $R_H = 0.935 < R_L = 0.999$. Among employed high-ability workers, 47.7% are in the initial period of a contract, while the rest are promoted to $w = p_H$. Among employed low-ability workers, 83.7% are in probation (where 89.4% of these workers are found in pooling firms), 2.2% are promoted to $w = p_L$, while the remaining 14.1% are demoted to $w = b$. Further, 4.6% of low-ability workers experience wage cuts when making a job-to-job transition.

Our calibrated model also implies that the monthly promotion rate is 1.2%. This statistics is a bit smaller than the average monthly promotion rate of 2% obtained by Pergamit and Veum (1999) for workers with around 10 years in the labor market in the NLSY.[26] We also obtain a monthly demotion rate of 0.5%, which implies a promotion-to-demotion ratio of 2.4. Given that one tends to consider demotions as infrequent events, this number might seem low. However, Belzil and Bognanno (2008), analyzing a survey of executives in U.S. firms, find that within-firm promotions are only slightly more frequent than demotions.[27]

4.1 Worker Mobility and Wage Dynamics

We now analyze the implications for workers' job mobility and wage dynamics. Is it the case that workers move to better paying jobs over time, as predicted, for example, by the Burdett and Mortensen (1998) model? An important finding of Light and McGarry (1998) is that more separations are negatively related to wages. But as they do not distinguish between separations into non-employment and job-to-job transitions, those results cannot be directly used to evaluate our model predictions against other theories of job mobility.

To take different labor market transitions into account, we use a similar sample of the NLSY as Light and McGarry (1998) and regress the log of real hourly wages on cumulative counts of job-

---

[25]Among the firms offering separating contracts, $\eta = 0.05\%$ are at the mass point (p_L, w_H^*), where $\xi = 53\%$ of low-ability workers misreport their type. Note that our wage offer distributions has an increasing an convex density as in the Burdett and Mortensen (1998) model. The threshold values for ρ described in the previous sections are $\rho_1 = 0.123$, $\rho_2 = 0.119$ and $\rho_3 = 0.116$. Further, the condition in Lemma 1 holds in this calibration.

[26]Note that in the calibration presented above promoted high-ability workers will not experience a second promotion without first losing their jobs. The latter arises because high-ability workers will not quit their jobs once they are promoted. This creates a tension with the data in which many workers have more than one promotion, and these are likely workers of "high ability". As an alternative calibration, we use the 2% promotion rate of Pergamit and Veum (1999) as a calibration target which delivers $\rho = 0.05$. This calibration still produces a segmented equilibria with very similar properties as the ones discussed in Sections 4.1 and 4.2. In this case we find that 61% of all employed high-ability workers are promoted, aggravating the tension with the data. Although not pursued here, one way to address this issue is to introduce a reallocation shock as in Jolivet et al. (2006) or add firm heterogeneity as suggested in Carrillo-Tudela and Kaas (2011).

[27]When defining a promotion (demotion) as an upward (downward) change in reporting levels, the promotion-to-demotion ratio is 1.08. When considering only level changes that were accompanied by job title changes, the promotion/demotion ratio increases to 1.6. If one only considers job title changes, disregarding changes in reporting level, the promotion/demotion ratio is 5.05. Further, Kramarz et al. (2014), using French administrative data, find that the promotion/demotion ratio within firms is 3.6 based on occupational changes. See also Lazear (1992) and Seltzer and Merrett (2000) for evidence on the extent of demotions based on firm case studies.

to-job transitions and non-employment spells, together with the same number of further controls that Light and McGarry (1998) used. As Table 2 shows, both coefficients are *negative* in the OLS wage regression.[28] In light of the literature on earnings losses after displacements (see Jacobson et al. (1993) for a seminal study), the negative coefficient on the count of non-employment spells is not surprising. But we also observe a negative correlation between the cumulative count of job-to-job transitions and current earnings. This goes against the intuition from standard theories of on-the-job search where workers generally climb up the wage distribution as they move between employers. Taking intervening non-employment spells into account, as we do in the wage regressions, one should therefore expect a positive relationship between the number of past job-to-job transitions and current wages. Indeed, when we run the same wage regression on simulated data from a Burdett and Mortensen (1998) model, we confirm this insight: more job-to-job transitions correlate positively with wages.[29] We further prove in the Appendix (Lemma A.2) that more job-to-job transitions (counted from the last non-employment spell) in the Burdett and Mortensen (1998) model indeed lead to higher expected wages.[30]

The last column in Table 2 shows that the same regression applied to our model can account for the empirical observations: more job-to-job transitions are negatively correlated with wages, while again non-employment spells go together with lower earnings. Note that this finding does not contradict the average wage gain of 3% of a job-to-job transition in our calibration (Table 1). We argue that the negative link between the *count* of job-to-job transitions and wages is driven by worker (unobserved) heterogeneity, both in our model and in the data. To test this, we follow the procedure described in Light and McGarry (1998) and run IV regressions on our data sample to account for unobserved worker heterogeneity in the error term. We find that in this case the coefficient on job-to-job transitions turns positive, while the coefficient on non-employment spells stays negative (see Appendix C for further details). Hence, standard job ladder models appear to be fully consistent with the data once worker heterogeneity is accounted for. On the other hand, those models miss the fact that some workers churn a lot in the labor market and yet remain largely unsuccessful.[31]

---

[28]The coefficient estimates shown in Table 2 are obtained using all available years in the NLSY (1979-2010). Light and McGarry (1998) use the first 8 years of workers' labor market history. Using the first 8, 10 or 15 years of workers labor market history does not change our general conclusion. The job-to-job transition coefficients at 8, 10 and 15 years are -0.007, -0.0075 and -0.0081, respectively, all significant at a 1% level. The non-employment spell coefficients at 8, 10 and 15 years are -0.0137, -0.0140 and -0.0159, respectively, again all significant at a 1% level. See Appendix C for the data description and a discussion of the regression specifications used.

[29]The reported regression results are based on a homogeneous worker version of the model, but it remains true when we use the two-worker version as in our calibration. It is also not sensitive to the parametrization. We use the same calibrated parameters as for our model (with $\alpha = 0$ so that all workers have productivity p_H). In the simulated data both for our model and for the Burdett and Mortensen (1998) model, workers have very similar average experience, job duration, number of job-to-job transitions, number of job-to-nonemployment transitions and non-employment durations; see Appendix C for further details.

[30]This theoretical finding is not trivial since workers with few job transitions are predominantly workers who initially find a good job and hence have less reasons to quit.

[31]Light and McGarry (1998) also find considerable variation in mobility patterns. Particularly, the mobility of some workers does not decline over time, while other workers undergo no or only little job mobility.

Table 2: Wage regressions

	Data	Models	
		B-M	CT-K
JTJ	-0.0073	0.0009	-0.0051
NESP	-0.0185	-0.0034	-0.0040
EXP	0.0431	0.0021	0.0055
EXP^2/10	-0.0075	-0.0005	-0.0010
R^2	0.331	0.064	0.161

Notes: Data regressions are based on the NLSY, regressions for the Burdett-Mortensen model (B-M) and for our model (CT-K) are based on simulations of 100,000 workers; for further details about the sample, control variables and robustness, see Appendix C. JTJ and NESP stand for the cumulative counts of job-to-job transitions and non-employment spells, EXP is actual labor market experience. All reported coefficients are statistically significant at the 1% level.

In our model, workers of high-ability climb up the distribution of initial wages with every quit until they are promoted to their marginal product in which case no further job-to-job transitions occur unless the worker is displaced. In contrast, low-ability workers misreport their ability level sometimes, which can lead to demotions and to further job transitions to low-wage employers.[32] Hence, low-ability workers undergo much higher job mobility.[33]

4.2 Internal Mobility and Firm Size

Another feature present in our model that does not come out of a standard job ladder model like Burdett and Mortensen (1998), is the presence of within firm worker mobility and its relationship with firm size. Our model is able to generate the positive relationship between firm size, internal mobility, wages and job stability observed in the data. In the calibration larger (pooling) firms have a 2% higher internal mobility rate (defined as promotions and demotions relative to employment) than smaller (separating) firms, which is consistent with the evidence obtained by Idson (1989).[34] Alongside this result, we also find that workers in smaller (separating) firms have

---

[32]Consistent with Gibbs et al. (2002) and Belzil and Bognanno (2008), demotions in our model are accompanied by pay cuts and by subsequent increases in job-transition probabilities.

[33]Kahn (2013) also suggests that such composition effects are present in the NLSY data, finding that movers are on average of lower ability than stayers. In particular, she finds that movers have lower AFQT scores, years of school, and tenure in the year before they moved, relative to stayers. See Table 2 of her paper for further details.

[34]In the calibration the relative difference in internal mobility between larger and smaller firms depends quite heavily on the value of ρ. For example, with $\rho = 0.05$ (as obtained when targeting an average monthly promotion rate of 2%) implies that larger (pooling) firms have around 17% higher internal mobility than smaller (separating firms).

18% lower tenure than workers in larger (pooling) firms, which is consistent with the empirically well documented fact that larger employers have lower job separation rates.[35]

Furthermore, as we establish in Corollary 1, pooling firms have a more productive workforce than separating firms. We confirm this finding in the calibrated example but note that productivity differences are tiny which is explained by the feature that most workers in the calibration have high ability and that productivity differences between workers are also not too large. We do reproduce, however, the usual positive relationships between firm size and wages, consistent with the empirical evidence (e.g. Brown and Medoff (1989), Idson and Oi (1999)). Although firms in our model are ex-ante identical, we expect that a similar conclusion obtains if firms differ in their exogenous productivity level, in which case positive sorting between workers and firms should obtain in equilibrium.[36]

Recently, Papageorgiou (2014) presents a theory that is also able to explain the positive relationship between firm size, internal mobility, wages and job stability. A key difference is that Papageorgiou (2014) uses occupational mobility within the firm to measure and model the internal mobility of workers (see also Kramarz et al. (2014)). Further, his theory is based on workers' learning about their match quality in a given occupation within the firm and hence is closely related to that of Jovanovic (1979) and Moscarini (2005). Instead, here we build on search models in the Burdett and Mortensen (1998) tradition and focus on firms' learning about the productivity of a worker. Also Papageorgiou (2014) does not consider long-term wage contracts, but instead uses spot wages that are determined by Nash bargaining. We see our contributions as complementary. Using different approaches, both our paper and Papageorgiou (2014) relate within and between firm mobility in an equilibrium framework in the context of labor market frictions.

5   Further Discussion

Here we discuss some variations to our assumptions. In Section 5.1 we consider the case in which firms can observe their applicants' employment status to analyze whether using this information makes it easier to separate workers. In this context we analyze two types of contracts: (i) up-or-out contracts, where workers that reported their type truthfully get promoted and misreporting workers are laid off; and (ii) up-or-down contracts as analyzed in the previous sections.[37] An important conclusion from these extensions is that information on employment status strengthens firms' monopsony power which makes it easier to separate workers. In Section 5.2 we discuss alternative assumptions about the firms' learning process and their implications.

---

[35]Note that in our model the relationship between firm size and job stability is generally ambiguous. This is because high-wage firms have higher quit rates of demoted low-ability workers.

[36]In a previous version of this paper with flat-wage contracts (Carrillo-Tudela and Kaas (2011)), we prove this assertion.

[37]Up-or-out contracts are common practice in some labor markets, such as those for academics, consultants and lawyers.

5.1 Information on Employment Status

Consider the same setup as before, but now firms offer a menu of four contracts $\omega_i^s = (w_i^s, w_i^{+s}, w_i^{-s})$ indexed by the worker's reported ability level $i = L, H$ and employment status $s = u, e$. The first component, w_i^s, denotes the initial wage offered to a worker of employment status s that reports type i. When a firm learns a worker's type, the worker receives $w_i^{+s} = p_i$ if he reported his type truthfully. Otherwise, he receives w_i^{-s}. In the case of up-or-out contracts, it is notationally convenient to assume that firms set w_i^{-s} below the worker's reservation wage, which is equivalent to a layoff. In the case of up-or-down contracts, w_i^{-s} equals the worker's reservation wage as analyzed in previous sections. Let F_i^s denote the corresponding offer distribution of initial wages, where $\underline{w}_i^s$ and $\overline{w}_i^s$ denote the infimum and supremum of its support, for $i = L, H$ and $s = u, e$.

The workers' value functions are similar to the ones described in equations (19)-(22) and are described in Appendix D. The main difference is that the offer distributions and their supports are now indexed by s. The incentive-compatibility constraint is now $w_j^s - w_i^s \leq \rho[V_i(p_i) - U_i]$. The problem of the firm is to maximize $\sum_{s=u,e} \Omega_H^s(w_H^s) + \Omega_L^s(w_L^s)$ by choosing w_H^s, w_L^s for $s = u, e$. Since firms can perfectly differentiate workers by employment status, they choose (w_L^u, w_H^u) and (w_L^e, w_H^e) independently. Hence the problem of hiring workers in the unemployment market can be treated independently from that of hiring workers in the employment market. Equilibrium requires that the optimal choices of w_i^s must be consistent with the offer distributions F_i^s and the associated function $\hat{w}^s$. Once again we analyze a rank preserving equilibrium such that $\hat{w}^s$ is increasing and $F_H^s(w) = F_L^s(\hat{w}^s(w))$ for each $s = u, e$ and for all w in the support of F_H^s.

Despite this change, we show that for a sufficiently low learning rate, a segmented equilibrium emerges with higher turnover of low-ability workers and higher workforce productivity of high-wage firms. The key difference between the up-or-out contracts and the up-or-down contracts is that when a segmented equilibrium arises, up-or-out contracts imply that the unemployment pool is biased towards low-ability workers.

5.1.1 Up-or-out contracts

Lemma 4: *Consider the model where firms offer up-or-out contracts and condition their offers on employment status. Then, for any initial wages earned by workers of type $i = H, L$ who were hired from employment (w_i^e) or from unemployment (w_i^u) it holds that $w_i^e > w_i^u$.*

The above result shows that the supports of the offer distributions F_H^u and F_H^e do not overlap. Importantly, this result is independent from whether we have a separating or a segmented equilibrium. Non-overlapping supports imply that workers hired from unemployment quit as soon as they get an outside offer during the probation period. Firms then maximize profits by offering unemployed workers their reservation wage. This leads F_i^u to degenerate to a mass point at $w_i^u = R_i$ for $i = L, H$. Further, the incentive-compatibility constraint for low-ability workers hired from unemployment becomes $R_H - R_L \leq \rho[V_L(p_L) - U_L]$, which never binds since $p_H > p_L$ implies that $R_L > R_H$ holds for any ρ. This result shows that information on employment status

enables firms to exert their full monopsony power when hiring these workers, leading to an outcome similar to that of Diamond (1971), which in turn makes it easier to separate unemployed workers.

Since workers earning their reservation wage face offer distributions F_i^e, the infimum of the support is $\underline{w}_i^e = R_i$ which is not offered in equilibrium, so that all firms recruit workers hired from unemployment upon contact. In Appendix D we fully characterize the equilibrium where firms observe workers' employment status and use up-or-out contracts. There we show that the problem faced by firms hiring employed workers turns out to be isomorphic to the one faced by firms offering promotion/demotion contracts without conditioning on employment status considered in the previous sections. In particular, the incentive-compatibility constraint for low-ability workers hired from employment is now given by $w_H^e - w_L^e \leq \rho[p_L - b]/(\phi + \delta)$. As we decrease ρ, this constraint starts to bind at high wages. Decreasing ρ further then implies that firms start offering pooling contracts when the incentive compatible w_L^e exceeds p_L. Therefore F_H^e and F_L^e can be derived in the same way as before and the main insights still apply here. The difference, however, is that with up-or-out contracts firms do not extract rents from workers caught misreporting their type, whereas conditioning contracts on employment status allows firms to strictly increase profits by segmenting its hiring markets.

To illustrate the properties of this version, we provide a quantitative example using the same calibration targets as in Section 4. When firms can condition their offers on employment status, reservation wages are lower. This reduces all initial wages in the offer distributions which makes it easier for firms to separate workers. As a result, the threshold values for parameter ρ where incentive constraints start and where pooling is the preferred choice for some firms are lower than in the benchmark model. Indeed, we have that $\rho_1 = 0.0404 > \rho_2 = 0.0395 > \rho_3 = 0.0393$. Given our a value of $\rho = 0.027$, a segmented equilibrium with pooling at top wages is the outcome. In this equilibrium low-ability workers have an unemployment rate of $u_L = 32.8\%$, while high-ability workers have an unemployment rate of $u_H = 28.7\%$. Reservation wages are $R_H = 0.911$, $R_L = 0.993$. Different from the benchmark model, however, this model predicts a positive effect of the count of job-to-job transitions on wages.

In the environment analyzed in this paper, up-or-out contracts could arise, for example, because firms face downward wage rigidities that do not allow them to cut wages. However, without such imposed constraints, firms will always prefer to use up-or-down contracts. To understand this, first note that both contracts offer worker the same incentives to report truthfully their type. Since with up-or-out contracts misreporting workers are laid off, firms make strictly less profits ex-post relative to up-or-down contracts where firms continue to make profits on demoted workers. From an ex-ante perspective, firms will also prefer to use up-or-down contracts as they provide workers with the same incentives as up-or-out contracts. Given these arguments, we now turn to the case in which firms observe their applicants' employment status and use up-or-down contracts.

5.1.2 Up-or-down contracts

The analysis of the case of up-or-down contracts is very similar to the one of up-or-out contracts described in Appendix D.[38] With non-overlapping supports of the offer distributions F_H^u and F_H^e, we have that information on employment status once again enables firms to exert their full monopsony power when hiring unemployed workers, which in turn makes it easier to separate these workers. Further, the threshold values ρ_1 and ρ_2 as well as all separating equilibria are the same as in the previous subsection. The difference arises in the case of segmented equilibria where demotions (instead of layoffs) occur in equilibrium. A segmented equilibrium with up-or-down contracts therefore has very similar properties to those of the benchmark model considered in the previous sections.

To illustrate these properties, we calibrate this version of the model using the same targets as we used for our benchmark model. The model matches the targets reasonably well, exhibiting a job-to-job transition rate of 2.5%, a quit gain, a layoff loss and a promotion gain of 3.0%, 4.5% and 6.7%, respectively. The cutoff values for the learning rate are $\rho_1 = 0.0394 > \rho_2 = 0.0383 > \rho_3 = 0.0380$, which implies the calibrated model is characterized by a segmented equilibrium. In this equilibrium 80.2% of firms offer separating contracts, while 19.8% of firms offer pooling contracts, reflecting that firms find it easier to separate when they can exert a higher degree of monopsony power at the hiring stage. The latter is illustrated in the lower values of unemployed workers' reservations wages relative to the benchmark model, where R_L and R_H are now 0.994 and 0.916, respectively. Further, this calibration exhibits a promotion rate of 1.4% and a demotion rate of 0.2%, yielding a higher promotion-to-demotion ratio relative to the benchmark model.

Using this calibration we revisit the implications for wage dynamics and job mobility and the relationship between firm size and internal mobility. In both cases we find that the predictions of the benchmark model survive when firms condition their contracts on workers' employment status. In particular, we perform the same exercise as described in Table 2 and obtain that the coefficient for the cumulative count of job-to-job transitions remains negative, although it is now one order of magnitude lower relative to the benchmark case (-0.0002 vs -0.005). The coefficient for the cumulative count of non-employment spells also remains negative and of similar magnitude as before (-0.007 vs -0.004). Regarding internal mobility and firm size, we find that the internal mobility rate is 14% higher in larger (pooling) firms relative to smaller (separating) firms. We also find that jobs in smaller (separating) firms have 2.6% lower tenure than those in larger (pooling) firms, as well as a positive relationship between wages and firm size.

---

[38]Given the similarity between the two analyses, we do not present the case of up-or-down contracts. The only substantial difference is that the profits of pooling firms include an additional term which reflects the profits earned on demoted workers, which is similar to the proof of Proposition 2.

5.2 Firms' Learning Process

In the framework of our model, firms learn worker ability with certainty at Poisson rate ρ. This is clearly a strong abstraction which masks many important features of real-world employer learning. In a more general model, the firm would draw signals which are imperfectly correlated with the worker's ability, so that the firm has to update its beliefs about worker ability with every new signal. While such learning processes are often applied in simpler environments, they tremendously complicate the analysis in our model, given that we allow wage contracts to be contingent on all verifiable information. A wage contract in such a general model should specify sequences of flow wages $(w(s^n))_{n \geq 0}$ which are contingent on the signal histories $s^n = (s_0, s_1, \ldots, s_n)$ where signals arrive at an exogenous Poisson rate. Even with a finite (e.g. two state) signal space, the firm's contracting problem and the corresponding equilibrium analysis would complicate greatly. Alternatively, contracts could be restricted to a one time promotion/demotion only. In such a world, making signals reveal partial information restrains the firm's ability to promote or demote the worker; thus it should be harder for the firm to satisfy incentive compatibility. This, we conjecture, would lead to more pooling in equilibrium.

One can also think about further elaborating the learning process by letting the firm decide ρ. A natural way to endogenize ρ is by letting firms choose costly monitoring intensity m, such that an increasing function $\hat{\rho}(m)$ determines the learning rate for any given $m \geq 0$. An obvious question in such a setting is which firms in the wage-offer distribution use more intensive monitoring. Given that incentive constraints are tighter in the upper part of the wage-offer distribution, we conjecture that high-wage firms spend more on monitoring and therefore are able to promote (or demote) workers faster; this mechanism would strengthen our previous result that larger firms have higher internal mobility rates. By varying parameters of the monitoring cost function, the optimal choices of ρ will change and we could perform the same exercises as done so far to understand firms' ability to separate workers when information frictions change relative to search frictions.

6 Conclusions

In this paper we consider a model of the labor market in which search frictions coexist with information frictions. The latter arise as firms do not observe worker ability upon hiring but gradually learn it over time. Given this adverse selection problem, we show that when the learning rate is sufficiently low, an equilibrium emerges in which low-wage firms attempt to hire both low- and high-ability workers by offering incentive-compatible separating contracts. These contracts offer initially a low wage and then promote the worker by increasing his wage to marginal productivity. High-wage firms, however, offer contracts that pay the same initial wage to all workers, but after learning their employees' types they promote high-ability workers and demote low-ability workers. This implies that low-ability workers can experience a wage cut or a wage rise when undertaking a job-to-job transition, while high-ability workers can only

experience wage rises when changing jobs without an intervening spell of unemployment.

In such a segmented equilibrium, low-ability workers have a higher degree of job turnover and lower average earnings than high-ability workers. To gain further insights on this relationship, we calibrate our model and show that it generates indeed a negative overall relationship between the number of job-to-job transitions and wages, which is consistent with the evidence that we obtain in NLSY data, extending the empirical findings of Light and McGarry (1998). Such a negative relationship is not easily picked up by standard models of job-to-job mobility, such as Burdett and Mortensen (1998), where workers generally climb up the wage ladder as they move between employers.

It should be an interesting extension to include different occupations and assignment problems within the firm. For example, suppose that firms aim to assign workers into one of two different occupations, such that only high-ability workers can efficiently perform the "high occupation", yielding output p_H, whereas low-ability workers should be assigned to the "low occupation", producing output $p_L < p_H$. Any mismatch between workers' abilities and occupations yields (weakly) lower output levels. If a worker self-selects to the right occupation/contract, the firm promotes the worker after learning his type, but if he is caught misreporting, the firm re-assigns the worker to the correct occupation and pays the demotion wage. In such a setting, this model could speak to occupational mobility within and across firms.

References

Albrecht, J. and S. Vroman (1992), "Non-existence of single-wage equilibria in search models with adverse selection." *Review of Economic Studies*, 59, 617–624.

Altonji, J. and C. Pierret (1996), "Employer learning and the signaling value of education." NBER Working Paper No. 5438.

Belzil, C. and M. Bognanno (2008), "Promotions, demotions, halo effects, and the earnings dynamics of American executives." *Journal of Labor Economics*, 26, 287–310.

Brown, C. and J. Medoff (1989), "The employer size-wage effect." *Journal of Political Economy*, 97, 1027–1059.

Burdett, K. and D. Mortensen (1998), "Wage differentials, employer size, and unemployment." *International Economic Review*, 39, 257–273.

Camera, G. and A. Delacroix (2004), "Trade mechanism selection in markets with frictions." *Review of Economic Dynamics*, 4, 851–868.

Carrillo-Tudela, C. and L. Kaas (2011), "Wage dispersion and labor turnover with adverse selection." IZA Discussion Paper No. 5936.

Davis, S., J. Faberman, and J. Haltiwanger (2008), "Adjusted estimates of worker flows and job openings in JOLTS." NBER Working Paper No. 14137.

Diamond, P. (1971), "A model of price adjustment." *Journal of Economic Theory*, 3, 156–168.

Gibbons, R. and L. Katz (1991), "Layoffs and lemons." *Journal of Labor Economics*, 9, 351–380.

Gibbs, M., K. Ierulli, and E. Milgrom (2002), "Occupational labor markets." Unpublished manuscript, University of Chicago.

Greenwald, B. (1986), "Adverse selection in the labour market." *Review of Economic Studies*, 53, 325–347.

Guerrieri, V., R. Shimer, and R. Wright (2010), "Adverse selection in competitive search equilibrium." *Econometrica*, 78, 1823–1862.

Idson, T. (1989), "Establishment size differentials in internal mobility." *The Review of Economics and Statistics*, 721–724.

Idson, T. and W. Oi (1999), "Workers are more productive in large firms." *American Economic Review*, 89, 104–108.

Inderst, R. (2005), "Matching markets with adverse selection." *Journal of Economic Theory*, 121, 145–166.

Jacobson, L., R. LaLonde, and D. Sullivan (1993), "Earnings losses of displaced workers." *American Economic Review*, 83, 685–709.

Jolivet, G., F. Postel-Vinay, and J-M. Robin (2006), "The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US." *European Economic Review*, 50, 877–907.

Jovanovic, B. (1979), "Job matching and the theory of turnover." *Journal of Political Economy*, 87, 972–990.

Kahn, L. (2013), "Asymmetric information between employers." *American Economic Journal: Applied Economics*, 5, 165–205.

Kramarz, F., F. Postel-Vinay, and J.-M. Robin (2014), "Occupational mobility and wage dynamics within and between firms." Unpublished Manuscript, University College London.

Krishna, V. (2009), *Auction Theory*, 2nd edition. Academic Press.

Kugler, A. and G. Saint-Paul (2004), "How do firing costs affect worker flows in a world with adverse selection?" *Journal of Labor Economics*, 22, 553–583.

Lange, F. (2007), "The speed of employer learning." *Journal of Labor Economics*, 25, 1–35.

Lazear, E. (1992), "The job as a concept." In *Performance Measurement, Evaluation, and Incentives* (William J. Bruns, ed.), 183–215, Harvard Business School Press.

Lazear, E. (2000), "Performance pay and productivity." *American Economic Review*, 90, 1346–1361.

Lentz, R. (2010), "Sorting by search intensity." *Journal of Economic Theory*, 145, 1436–1452.

Light, A. and K. McGarry (1998), "Job change patterns and the wages of young men." *Review of Economics and Statistics*, 80, 276–286.

Lockwood, B. (1991), "Information externalities in the labour market and the duration of unemployment." *Review of Economic Studies*, 58, 733–753.

Manning, A. (2003), *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Michelacci, C. and J. Suarez (2006), "Incomplete wage posting." *Journal of Political Economy*, 114, 1098–1123.

Mincer, J. and B. Jovanovic (1981), "Labor mobility and wages." In *Studies in Labor Markets* (S. Rosen, ed.), 21–63, University of Chicago Press.

Montgomery, J. (1999), "Adverse selection and employment cycles." *Journal of Labor Economics*, 17, 281–297.

Moscarini, G. (2005), "Job matching and the wage distribution." *Econometrica*, 73, 481–516.

Nagypal, Eva (2005), "Worker reallocation over the business cycle: The importance of job-to-job transitions." Unpublished manuscript, Northwestern University.

Oyer, P. and S. Schaefer (2005), "Why do some firms give stock options to all employees? An empirical examination of alternative theories." *Journal of Financial Economics*, 76, 99–133.

Oyer, P. and S. Schaefer (2011), "Personnel economics: Hiring and incentives." In *Handbook of Labor Economics Vol. 4b* (O. Ashenfelter and D. Card, eds.), 1769–1823, Elsevier, Amsterdam.

Papageorgiou, T. (2014), "Large firms and internal labor markets." Unpublished manuscript, Pennsylvania State University.

Pergamit, M. and J. Veum (1999), "What is a promotion?" *Industrial and Labor Relations Review*, 581–601.

Rogerson, R., R. Shimer, and R. Wright (2005), "Search-theoretic models of the labor market: A survey." *Journal of Economic Literature*, 43, 959–988.

Salop, J. and S. Salop (1976), "Self-selection and turnover in the labor market." *Quarterly Journal of Economics*, 90, 619–627.

Seltzer, A. and D. Merrett (2000), "Personnel policies at the Union Bank of Australia: Evidence from the 1888–1900 entry cohorts." *Journal of Labor Economics*, 18, 573–613.

Stevens, M. (2004), "Wage-tenure contracts in a frictional labour market: Firms' strategies for recruitment and retention." *Review of Economic Studies*, 71, 535–551.

Topel, R. and M. Ward (1992), "Job mobility and the careers of young men." *Quarterly Journal of Economics*, 107, 439–479.

Visschers, L. (2007), "Employment uncertainty and wage contracts in frictional labor markets." Mimeo, University of Pennsylvania.

Appendix A: Worker strategies

Let U_i denote the value of unemployment of a worker with ability $i = L, H$. Note that once this worker encounters a potential employer, the firm does not observe his ability, so that the worker can claim to be of different ability. For any contract $\omega = (w, w^+, w^-)$, let $V_{ij}(\omega)$ denote the maximum expected value of employment for a worker with ability i employed in probation at initial wage w after reporting type j. We can think of these workers as randomly meeting firms who draw contract offers (ω_H, ω_L) from joint distribution Ψ. The worker then decides which contract to choose (if any) to maximize expected lifetime income. Using this insight and standard dynamic programming arguments, the Bellman equation that describes U_i is given by

$$\phi U_i = b + \lambda \int \max[V_{iL}(\omega_L) - U_i, V_{iH}(\omega_H) - U_i, 0] d\Psi(\omega_H, \omega_L) . \tag{19}$$

The value of employment for a worker of ability i employed in contract $\omega_i = (w_i, w_i^+, w_i^-)$ at the initial wage is given by

$$\phi V_{ii}(\omega_i) = w_i + \lambda \int \max[V_{iL}(\omega_L') - V_{ii}(\omega_i), V_{iH}(\omega_H') - V_{ii}(\omega_i), 0] d\Psi(\omega_H', \omega_L') \\ + \delta(U_i - V_{ii}(\omega_i)) + \rho(V_i(w_i^+) - V_{ii}(\omega_i)) , \tag{20}$$

where the continuation value after promotion is given by

$$\phi V_i(w_i^+) = w_i^+ + \lambda \int \max[V_{iL}(\omega_L') - V_i(w_i^+), V_{iH}(\omega_H') - V_i(w_i^+), 0] d\Psi(\omega_H', \omega_L') + \delta(U_i - V_i(w_i^+)) . \tag{21}$$

Now consider a worker of type i that misreported his type and is currently earning the initial wage w_j in contract $\omega_j = (w_j, w_j^+, w_j^-)$. This worker's value function is described by

$$\phi V_{ij}(\omega_j) = w_j + \lambda \int \max[V_{iL}(\omega_L') - V_{ij}(\omega_j), V_{iH}(\omega_H') - V_{ij}(\omega_j), 0] d\Psi(\omega_H', \omega_L') \\ + \delta(U_i - V_{ij}(\omega_j)) + \rho(V_i(w_j^-) - V_{ij}(\omega_j)) . \tag{22}$$

Finally, the continuation value after a demotion is given by

$$\phi V_i(w_j^-) = w_j^- + \lambda \int \max[V_{iL}(\omega_L') - V_i(w_j^-), V_{iH}(\omega_H') - V_i(w_j^-), 0] d\Psi(\omega_H', \omega_L') + \delta(U_i - V_i(w_j^-)) . \tag{23}$$

From (20) and (22), it is easy to see that worker i's incentive constraint $V_{ii}(\omega_i) \geq V_{ij}(\omega_j)$ can be expressed as $w_j - w_i \leq \rho[V_i(w_i^+) - V_i(w_j^-)]$, which is (1) in the main text.

When we consider simple contracts $\omega_i = (w, p_i, b)$, we simplify notation by expressing workers' value functions as $V_{ii}(w)$ etc. rather than $V_{ii}(\omega_i)$. In those situations it is straightforward to verify that any worker's optimal search strategy is described by a reservation wage. Let $R_{ijk}(x)$ denote the reservation (initial) wage of a worker in a probation contract who (i) currently receives flow payoff x, (ii) is of type $i = L, H$, (iii) has reported type $j = L, H$, and (iv) when meeting a firm decides to report type $k = L, H$. Thus, $R_{ijk}(x)$ is defined by $V_{ij}(x) = V_{ik}(R)$. Similarly, $R_{ik}(x)$ is the reservation (initial) wage of a promoted/demoted worker with flow income x, which is defined by $V_i(x) = V_{ik}(R)$. The above value functions imply that an unemployed worker of type i accepts a wage offer w' if and only if $w' \geq R_{ik}(b)$ for $i, k = L, H$.

We write $R_i \equiv R_{ii}(b)$ for the reservation (initial) wage of an unemployed worker of type i who reports his true type. Those reservation wages are smaller than b since truth-telling workers can expect a promotion to $w_i = p_i > b$ at rate ρ.

Consider an employed worker of type i that reported his true type and is earning initial wage w_i. Given contact with a firm and revealing his true type once again (i.e. $k = i$), (20) implies that this worker accepts employment if and only if the firm offers a wage $w_i' > R_{iii}(w_i) = w_i$. If the worker decides to misreport his type (i.e. $k \neq i$), however, (20) and (22) imply that the worker accepts employment if and only if the firm offers a wage $w_k' > R_{iik}(w_i) = w_i + \rho[V_i(p_i) - U_i]$. In this case, the worker must be compensated by the expected loss of misreporting his type. Next consider the same worker after his employer has learned his true type and is earning wage p_i. Given contact with a firm and revealing his true type (i.e. $k = i$), (20) and (21) imply that this worker accepts employment if and only if the firm offers an initial wage $w_i' > R_{ii}(p_i) = p_i$. If this worker decided to misreport his type, (21) and (22) imply that the worker accepts employment if and only if the firm offers a wage $w_k' > R_{ik}(p_i) = p_i + \rho[V_i(p_i) - U_i]$.

Now suppose that an employed worker of type i misreported type $j \neq i$ and is earning the initial wage w_j. Given contact with a firm and reporting his true type ($k = i$), (20) and (22) imply that the worker accepts employment if and only if the firm offers a wage $w_i' > R_{iji}(w_j) = w_j - \rho[V_i(p_i) - U_i]$. In this case, the worker voluntarily accepts a wage cut since the truth-telling worker expects a promotion gain relative to the demotion loss. On the other hand, if the worker misreports his type once again ($k = j$), the worker accepts employment if and only if the firm offers a wage $w_j' > R_{ijj}(w_j) = w_j$. Finally consider the same worker after his employer has learned his true type, earning demotion wage b. This worker then behaves like an unemployed worker and accepts any contract that provides utility greater than the continuation value from unemployment.

Appendix B: Proofs

Proof of Lemma 1

Consider a candidate equilibrium in simple contracts. That is, for any contract pair $(\omega_i)_{i=H,L} = (w_i, p_i, b)_{i=H,L}$ in the support of Ψ, it holds that $\Omega^* \equiv \sum_{i=H,L} \Omega_i(\omega_i) \geq \sum_{i=H,L} \Omega_i(\omega_i')$ for all pairs of simple contracts $(\omega_i')_{i=H,L} = (w_i', p_i, b)_{i=H,L}$. To prove the Lemma, we need to show that no other feasible contract pair of arbitrary form $(\hat{\omega}_i)_{i=H,L} = (\hat{w}_i, \hat{w}_i^+, \hat{w}_i^-)_{i=H,L}$ can lead to higher profit value than Ω^*. Feasibility requires that $w_i, w_i^+, w_i^- \leq p_i$ and $w_i^+, w_i^- \geq b$ (otherwise the firm or the worker would quit the contract).

Consider first the situation where the deviating contract pair $(\hat{\omega}_i)_{i=H,L}$ is incentive compatible, i.e. for $i \neq j$,

$$\hat{w}_j - \hat{w}_i \leq \rho[V_i(\hat{w}_i^+) - V_i(\hat{w}_j^-)] . \tag{24}$$

Note that $V_{ii}(\hat{\omega}_i) \in [U_i, V_i(p_i)]$: $V_{ii}(\hat{\omega}_i)$ cannot be smaller than U_i (otherwise the contract would not attract any workers), and it cannot exceed $V_i(p_i)$ because of the feasibility restrictions $\hat{w}_i, \hat{w}_i^+ \leq p_i$. Hence, for both $i = H, L$, there exist simple contracts $\omega_i' = (w_i', p_i, b)$ such that $V_{ii}(\omega_i') = V_{ii}(\hat{\omega}_i)$. The starting wage w_i' is then pinned down by

$$w_i' + \rho V_i(p_i) = \hat{w}_i + \rho V_i(\hat{w}_i^+) . \tag{25}$$

Now it follows that the pair of simple contracts $(\omega_i')_{i=H,L}$ is also incentive-compatible: for any $j \neq i$

$$
\begin{aligned}
w_j' &= \hat{w}_j + \rho[V_j(\hat{w}_j^+) - V_j(p_j)] \\
&\leq \hat{w}_i + \rho[V_i(\hat{w}_i^+) - V_i(\hat{w}_j^-)] \\
&= w_i' + \rho[V_i(p_i) - V_i(\hat{w}_j^-)] \\
&\leq w_i' + \rho[V_i(p_i) - V_i(b)] .
\end{aligned}
$$

Here the two equalities make use of (25), the first inequality makes use of (24) and $\hat{w}_j^+ \leq p_j$ (which implies $V_j(\hat{w}_j^+) \leq V_j(p_j)$), and the second inequality makes use of $\hat{w}_j^- \geq b$ (which implies $V_i(\hat{w}_j^-) \geq V_i(b)$). Because of $\sum_{i=H,L} \Omega_i(\omega_i') \leq \Omega^*$, contract pair $(\hat{\omega}_i)_{i=H,L}$ does not lead to higher profit than Ω^* if we can prove that $\Omega_i(\hat{\omega}_i) \leq \Omega_i(\omega_i')$ for $i = H, L$. To show this assertion, note that

$$\Omega_i(\omega_i') = \ell_i'[p_i - w_i'] \qquad \text{and} \qquad \Omega_i(\hat{\omega}_i) = \hat{\ell}_i\left[p_i - \hat{w}_i + \rho\frac{p_i - \hat{w}_i^+}{\phi + \delta + \lambda \hat{q}_i}\right] ,$$

where ℓ_i' and $\hat{\ell}_i$ are the masses of workers of type i who are employed in the starting phases of the two contracts. Because of $V_{ii}(\omega_i') = V_{ii}(\hat{\omega}_i)$, both contracts attract the same number of type i workers who have the same quit rates in the starting phase of the contracts; hence $\ell_i' = \hat{\ell}_i$. The terms in [.] are the flow profits for each of these workers; these contain the flow profit $p_i - w_i'$ ($p_i - \hat{w}_i$) and a continuation profit value which is realized at flow rate ρ when the worker is promoted (which is zero for the simple contract ω_i'). $\hat{q}_i$ is the probability that worker i who is promoted to wage $\hat{w}_i^+$ quits this contracts when meeting another firm. Because of these expressions and (25), the requirement $\Omega_i(\hat{\omega}_i) \leq \Omega_i(\omega_i')$ is equivalent to showing that

$$V_i(p_i) - V_i(\hat{w}_i^+) \geq \frac{p_i - \hat{w}_i^+}{\phi + \delta + \lambda \hat{q}_i} . \tag{26}$$

From workers' Bellman equations follows

$$V_i(p_i) = \frac{p_i + \delta U_i + \lambda E\{\max[V_i(p_i), \tilde{V}_i]\}}{\phi + \delta + \lambda} \qquad \text{and} \qquad V_i(\hat{w}_i^+) = \frac{\hat{w}_i^+ + \delta U_i + \lambda E\{\max[V_i(\hat{w}_i^+), \tilde{V}_i]\}}{\phi + \delta + \lambda} ,$$

where $\tilde{V}_i$ is worker i's continuation utility from a random outside offer. It follows

$$
\begin{aligned}
V_i(p_i) - V_i(\hat{w}_i^+) &= \frac{p_i - \hat{w}_i^+ + \lambda \mathrm{E}\{\max[V_i(p_i), \tilde{V}_i] - \max[V_i(\hat{w}_i^+), \tilde{V}_i]\}}{\phi + \delta + \lambda} \\
&\geq \frac{p_i - \hat{w}_i^+ + \lambda[V_i(p_i) - V_i(\hat{w}_i^+)](1 - \hat{q}_i)}{\phi + \delta + \lambda} \; .
\end{aligned}
$$

Rearranging proves (26).

Second, suppose that the deviating contract pair $(\hat{\omega}_i)_{i=H,L}$ is not incentive compatible so that all workers pool in contract $\hat{\omega}_H$. The alternative scenario where both workers pool in contract $\hat{\omega}_L$ is then captured as well: if both workers would pool in contract $\hat{\omega}_L = (\hat{w}_L, \hat{w}_L^+, \hat{w}_L^-)$, we can define contract $\tilde{\omega}_H = (\tilde{w}_H, \tilde{w}_H^+, \tilde{w}_H^-)$ by $\tilde{w}_H = \hat{w}_L$, $\tilde{w}_H^+ = \hat{w}_L^-$, $\tilde{w}_H^- = \hat{w}_L^+$, and pick an arbitrary unattractive contract $\tilde{\omega}_L$, so that all workers pool in contract $\tilde{\omega}_H$. We can therefore consider the case where all workers pool in some contract $\hat{\omega}_H$ and we need to prove that $\Omega_H(\hat{\omega}_H) \leq \Omega^*$. We can write

$$
\Omega_H(\hat{\omega}_H) = \hat{\ell}_H \underbrace{\left[p_H - \hat{w}_H + \rho \frac{p_H - \hat{w}_H^+}{\phi + \delta + \lambda \hat{q}_H} \right]}_{\equiv \hat{\pi}_H} + \hat{\ell}_L \underbrace{\left[p_L - \hat{w}_H + \rho \frac{p_L - \hat{w}_H^-}{\phi + \delta + \lambda \hat{q}_L} \right]}_{\equiv \hat{\pi}_L} \, ,
$$

where $\hat{\ell}_i$ is steady-state employment of type i workers in the starting phase of this contract and $\hat{q}_i$, $i = H, L$, are the probabilities that promoted/demoted workers of type i quit contract $\hat{\omega}_H$ when outside offers arrive. As in the first part of the proof, because of $V_{HH}(\hat{\omega}_H) \in [U_H, V_H(p_H)]$, there exists a simple contract $\omega'_H = (w'_H, p_H, b)$ such that $V_{HH}(\omega'_H) = V_{HH}(\hat{\omega}_H)$. This implies that

$$
w'_H + \rho V_H(p_H) = \hat{w}_H + \rho V_H(\hat{w}_H^+) \; . \tag{27}
$$

Because contract ω'_H attracts the same mass of H-workers as $\hat{\omega}_H$, and quit rates in the starting phase are also the same, it follows that $\ell'_H = \hat{\ell}_H$. The simple contract ω'_H yields higher profit per worker of type H if $\pi'_H \equiv p_H - w'_H \geq \hat{\pi}_H$ which is equivalent to $\hat{w}_H - w'_H \geq \rho \frac{p_H - \hat{w}_H^+}{\phi + \delta + \lambda \hat{q}_H}$, which because of (27) is equivalent to

$$
V_H(p_H) - V_H(\hat{w}_H^+) \geq \frac{p_H - \hat{w}_H^+}{\phi + \delta + \lambda \hat{q}_H} \; . \tag{28}
$$

Note that $V_H(p_H) = \frac{p_H + \delta U_H}{\phi + \delta}$ and $V_H(\hat{w}_H^+) = \frac{\hat{w}_H^+ + \delta U_H + \lambda \hat{q}_H \hat{V}_H}{\phi + \delta + \lambda \hat{q}_H}$ where $\hat{V}_H$ is the expected continuation utility for a H-worker quitting w_H^+. Substitution into (28) and rearranging yields the equivalent inequality

$$
\hat{V}_H \leq \frac{p_H + \delta U_H}{\phi + \delta} = V_H(p_H) \; ,
$$

which is fulfilled since no firm can offer a greater continuation utility than $V_H(p_H)$. This proves that the simple contract yields (weakly) greater profit on H workers:

$$
\ell'_H \pi'_H \geq \hat{\ell}_H \hat{\pi}_H \; . \tag{29}
$$

To prove that $\Omega_H(\hat{\omega}_H) \leq \Omega^*$, we distinguish between two cases: $\hat{\pi}_L \geq 0$ and $\hat{\pi}_L < 0$. In the first one, we can construct a simple contract ω'_L that is incentive-compatible to ω'_H and such that contract pair (ω'_H, ω'_L) dominates pooling contract $\hat{\omega}_H$. In the second case other arguments are required.

Suppose *first* that $\hat{\pi}_L \geq 0$. By defining $w'_L = \hat{w}_H - \rho[V_L(p_L) - V_L(\hat{w}_H^-)]$, the simple contract $\omega'_L = (w'_L, p_L, b)$ yields the same utility to L-workers as the pooling contract: $V_{LL}(\omega'_L) = V_{LH}(\hat{\omega}_H)$. Moreover, the simple contract is feasible, i.e. $w'_L \leq p_L$. This follows because of

$$
V_L(p_L) - V_L(\hat{w}_H^-) \geq \frac{p_L - \hat{w}_H^-}{\phi + \delta + \lambda \hat{q}_L} \, ,
$$

which follows from a similar argument as (26), and because of $\hat{\pi}_L \geq 0$:

$$w'_L = \hat{w}_H - \rho[V_L(p_L) - V_L(\hat{w}_H^-)] \leq p_L - \hat{\pi}_L + \rho \frac{p_L - \hat{w}_H^-}{\phi + \delta + \lambda \hat{q}_L} - \rho \frac{p_L - \hat{w}_H^-}{\phi + \delta + \lambda \hat{q}_L} \leq p_L .$$

This also proves that $\pi'_L = p_L - w'_L \geq \hat{\pi}_L$. The pair of simple contracts (ω'_H, ω'_L) is incentive-compatible since $V_{LL}(\omega'_L) = V_{LH}(\hat{\omega}_H) \geq V_{LH}(\omega'_H)$. Here the inequality follows because contract ω'_H promises a lower starting wage and a lower continuation wage to L-workers than contract $\hat{\omega}_H$ (i.e. $\hat{w}_H \geq w'_H$ and $\hat{w}_H^- \geq b$). Therefore, and because of $V_{LL}(\omega'_L) = V_{LH}(\hat{\omega}_H)$, simple contract ω'_L when offered jointly with ω'_H attracts the same number of L-workers as pooling contract $\hat{\omega}_H$. Hence $\ell'_L = \hat{\ell}_L$ which implies

$$\ell'_L \pi'_L \geq \hat{\ell}_L \hat{\pi}_L . \tag{30}$$

Because the pair of simple contracts (ω'_H, ω'_L) yields profit not greater than Ω^*, it follows from (29) and (30) that $\Omega_H(\hat{\omega}_H) \leq \Omega^*$.

Suppose *second* that $\hat{\pi}_L < 0$. Here we again need to distinguish between two situations. The first is the case where only separating contracts are offered in the candidate equilibrium. In this situation, $\ell'_H \pi'_H$ is smaller than Ω^* and it provides an upper bound for the deviating pooling contract $\hat{\omega}_H$, as we show in the next paragraph. In the other situation where some pooling occurs in the candidate equilibrium, further arguments are required.

Consider first the case in which all firms offer separating contracts in the candidate equilibrium (cf. Proposition 1). Because of (29) and $\hat{\pi}_L < 0$, $\ell'_H \pi'_H$ provides an upper bound for $\Omega_H(\hat{\omega}_H)$. Hence, the deviating contract $(\hat{\omega}_H, \hat{\omega}_L)$ does not lead to higher profit than Ω^* if we can show that $\ell'_H \pi'_H \leq \Omega^*$. Write $\overline{w}_H$ for the highest equilibrium starting wage offered to H-workers, and let $\overline{\ell}_H$ denote the employment of H-workers in the starting phase of this contract. $\overline{\pi}_H = p_H - \overline{w}_H$ is the profit flow for each of these workers. If $w'_H \geq \overline{w}_H$, it follows that $\pi'_H \leq \overline{\pi}_H$. Moreover, since contract (w'_H, p_H, b) does not attract or retain any more H-workers than contract $(\overline{w}_H, p_H, b)$ (because the equilibrium distribution F_H has no mass point at $\overline{w}_H$ as shown in Proposition 1), it follows that $\ell'_H = \overline{\ell}_H$. Hence,

$$\ell'_H \pi'_H \leq \overline{\ell}_H \overline{\pi}_H < \overline{\ell}_H \overline{\pi}_H + \overline{\ell}_L \overline{\pi}_L = \Omega^* .$$

Conversely, if $w'_H < \overline{w}_H$, w'_H is in the support of the equilibrium wage-offer distribution (cf. Proposition 1). Hence, a feasible incentive-compatible contract (w'_L, p_L, b) exists such that $\ell'_H \pi'_H < \ell'_H \pi'_H + \ell'_L \pi'_L = \Omega^*$.

Consider now the case in which $\hat{\pi}_L < 0$ and some firms offer pooling contracts in the candidate equilibrium so that the arguments of the previous paragraphs do not apply. In this case there may not exist a feasible simple contract ω'_L which is incentive compatible to contract ω'_H and yields the same utility to L-workers as pooling contract $\hat{\omega}_H$. Because of $w'_H \leq \hat{w}_H$ and $b \leq \hat{w}_H^-$, the simple contract $\omega'_H = (w'_H, p_H, b)$, when offered as a pooling contract, is less attractive to L workers than pooling contract $\hat{\omega}_H$ and it also retains fewer of those workers; hence $\ell'_L \leq \hat{\ell}_L$. We further claim that the profit from any L worker in the simple contract is no smaller than the one with contract $\hat{\omega}_H$, i.e.

$$\pi'_L = p_L - w'_H + \frac{p_L - b}{\phi + \delta + \lambda} \geq \hat{\pi}_L = p_L - \hat{w}_H + \rho \frac{p_L - \hat{w}_H^-}{\phi + \delta + \lambda \hat{q}_L} .$$

Because of $w'_H \leq \hat{w}_H$, this follows if

$$\frac{p_L - b}{\phi + \delta + \lambda} \geq \frac{p_L - \hat{w}_H^-}{\phi + \delta + \lambda \hat{q}_L} . \tag{31}$$

Consider any arbitrary demotion wage w^- and define function

$$H(w^-) \equiv \frac{p_L - w^-}{\phi + \delta + \lambda[1 - F_L(\check{w}(w^-))]} \; ,$$

where $\check{w}(w^-)$ is the initial wage of a simple contract that yields the same continuation payoff to worker L as the flat demotion wage w^-, so that $[1 - F_L(\check{w}(w^-))]$ is the probability that L-workers quit demotion wage w^-. Therefore, $H(\hat{w}_H^-)$ equals the right-hand side in condition (31). The wage $\check{w}(w^-)$ satisfies $V_{LL}(\check{w}(w^-), p_L, b) = V_L(w^-)$ (and hence $\check{w}(b) = R_L$). With this notation, we can also define $\Gamma(w) \equiv H(\check{w}^{-1}(w))$ for initial wages $w \geq R_L$. Condition (31) for any feasible value of $\hat{w}_H^-$ is then equivalent to $\Gamma(w) \leq \Gamma(R_L)$ for all $w \in [R_L, p_L]$. The relation $V_{LL}(w, p_L, b) = V_L(\check{w}^{-1}(w))$ can also be expressed as

$$\check{w}^{-1}(w) = w + \rho[V_L(p_L) - V_{LL}(w, p_L, b)] \; ,$$

so that we have

$$\check{w}^{-1\prime}(w) = 1 - \rho \frac{d}{dw} V'_{LL}(w, p_L, b) = \frac{\phi + \delta + \lambda(1 - F_L(w))}{\phi + \delta + \rho + \lambda(1 - F_L(w))} \; ,$$

which makes use of the derivative of (4) which applies to all wages $w \leq p_L$. Therefore, integration implies

$$\check{w}^{-1}(w) = b + \int_{R_L}^{w} \frac{\phi + \delta + \lambda(1 - F_L(w'))}{\phi + \delta + \rho + \lambda(1 - F_L(w'))} dw' \; ,$$

so that

$$\Gamma(w) = \frac{p_L - b - \int_{R_L}^{w} \frac{\phi + \delta + \lambda(1 - F_L(w'))}{\phi + \delta + \rho + \lambda(1 - F_L(w'))} dw'}{\phi + \delta + \lambda(1 - F_L(w))} \; .$$

With the condition of Lemma 1, the requirement $\Gamma(w) \leq \Gamma(R_L)$ is fulfilled so that $\hat{\pi}_L \leq \pi'_L$. This, together with $\ell'_L \leq \hat{\ell}_L$ and $\hat{\pi}_L < 0$ proves that $\ell'_L \pi'_L \geq \hat{\ell}_L \hat{\pi}_L$. Therefore, the simple pooling contract ω'_H yields total profit $\Omega_H(\omega'_H) = \ell'_H \pi'_H + \ell'_L \pi'_L$ which is at least as large as $\Omega_H(\hat{\omega}_H) = \hat{\ell}_H \hat{\pi}_H + \hat{\ell}_L \hat{\pi}_L$. This proves that $\Omega_H(\hat{\omega}_H) \leq \Omega_H(\omega'_H) \leq \Omega^*$. $\qquad \square$

Proof of Lemma 2

Note that a worker does not misreport his type whenever the incentive-compatibility constraint $V_{ii}(w_i) \geq V_{ij}(w_j)$ holds for any offered pair (w_i, w_j). Using (20) and (22), it follows that this is equivalent to

$$w_j - w_i \leq \rho[V_i(p_i) - U_i] \; ,$$

which is identical to the incentive constraint (1). We can simplify this equation by making use of the reservation wage equation

$$R_i = b - \rho[V_i(p_i) - U_i] \; , \tag{32}$$

where $R_i = R_{ii}(b) < b$ is the reservation initial wage of an unemployed worker of type i who reports truthfully. Using this equation, we can rewrite the incentive constraints as (2). $\qquad \square$

Proof of Proposition 1

To express the dependance on ρ, we write $R_i(\rho)$, $i = H, L$, for the reservation wages, and $\overline{w}_i(\rho)$ for the highest wages in the support of the offer distributions. The separating equilibrium characterized in Section 3.2 is feasible provided that $\overline{w}_L(\rho) \leq p_L$ which is equivalent to

$$\overline{w}_H(\rho) \leq p_L + b - R_L(\rho) \; . \tag{33}$$

37

This inequality is strictly fulfilled as long as incentive constraints do not bind (that is, if $\rho \geq \rho_1$), because of $\overline{w}_L < p_L$. On the other hand, when ρ tends to zero, $R_L(\rho) \to b$, while

$$\overline{w}_H(\rho) \to \overline{w}_H(0) = \frac{1}{1+\alpha}\left[\overline{p} - \frac{(\phi+\delta)^2(\overline{p} - (1+\alpha)b)}{(\lambda+\phi+\delta)^2}\right] .$$

Because (18) is equivalent to $\overline{w}_H(0) > p_L$, inequality (33) fails if $\rho < \rho_1$ is sufficiently small. Hence, under condition (18), there exists $\rho_2 \in (0, \rho_1)$ such that the separating equilibrium is feasible for any $\rho \geq \rho_2$. On the other hand, if (18) fails, either (33) holds for all $\rho \in [0, \rho_1]$, in which case the separating equilibrium is feasible for all $\rho \geq \rho_2 = 0$, or condition (33) fails for some values $\rho \in [0, \rho_1]$ in which case $\rho_2 > 0$ is defined as the supremum of those values of ρ where (33) holds with equality.

Conditional on $\rho \geq \rho_2$, we first solve for equilibrium reservation wages when incentive constraints bind. Define

$$h_i(w) \equiv \frac{\lambda(1 - F_i(w))}{\phi + \delta + \rho + \lambda(1 - F_i(w))} , \quad i = L, H ,$$

and split the integral expressions in the reservation wage equations (13) as follows:

$$\int_{R_H}^{\tilde{w}_H} h_H(w)dw = \tilde{w}_H - R_H + 2C(Y^{1/2} - 1)(p_H - R_H) ,$$

$$\int_{\tilde{w}_H}^{\overline{w}_H} h_H(w)dw = \overline{w}_H - \tilde{w}_H + \frac{2C}{1+\alpha}\left[C(\overline{p} - \overline{R}) - (\overline{p} - \overline{R})^{1/2}(\overline{p} + \alpha(b - R_L) - (1+\alpha)\tilde{w}_H)^{1/2}\right] ,$$

$$\int_{R_L}^{\tilde{w}_L} h_L(w)dw = \tilde{w}_L - R_L + 2C(Y^{1/2} - 1)(p_L - R_L) ,$$

$$\int_{\tilde{w}_L}^{\overline{w}_L} h_L(w)dw = \overline{w}_L - \tilde{w}_L + \frac{2C}{1+\alpha}\left[C(\overline{p} - \overline{R}) - (\overline{p} - \overline{R})^{1/2}(\overline{p} + \alpha(b - R_L) - (1+\alpha)\tilde{w}_H)^{1/2}\right] ,$$

where we define $Y \equiv \frac{p_H - b - p_L + R_L}{p_H - R_H - p_L + R_L}$. Adding up the integral expressions gives

$$\int_{R_H}^{\overline{w}_H} h_H(w)dw = \frac{1}{1+\alpha}\left[(\overline{p} - \overline{R})(1 + C^2) + \alpha(b - R_H)\right]$$

$$+2C\frac{\alpha}{1+\alpha}[p_H - b - p_L + R_L]^{1/2}[p_H - R_H - p_L + R_L]^{1/2} - 2C(p_H - R_H)$$

and

$$\int_{R_L}^{\overline{w}_L} h_L(w)dw = \frac{1}{1+\alpha}\left[(\overline{p} - \overline{R})(1 + C^2) + R_H - b\right]$$

$$-2C\frac{1}{1+\alpha}[p_H - b - p_L + R_L]^{1/2}[p_H - R_H - p_L + R_L]^{1/2} - 2C(p_L - R_L) .$$

Substitution of these terms into the reservation wage equations (13) yields two nonlinear equations that determine R_L and R_H simultaneously. Adding the equation for R_H to the one for R_L multiplied by α, we see that the nonlinear term disappears so that we can solve for $\overline{R} = R_H + \alpha R_L$:

$$\overline{R} = \frac{\overline{p}\rho C(C - 2) + b(1 + \alpha)(\rho + \phi + \delta)}{\phi + \delta + \rho(1 - C)^2} .$$

We can now substitute $R_H = \overline{R} - \alpha R_L$ into the reservation wage equation for R_L, which is quadratic in R_L. The relevant root is obtained as follows:

$$R_L = \frac{4C^2(F(1+\alpha) + G) - 2DE + \sqrt{[4C^2(F(1+\alpha) + G) - 2DE]^2 - 4(4C^2(1+\alpha) - E^2)(4C^2FG - D^2)}}{2E^2 - 8C^2(1+\alpha)} ,$$

with

$$D \equiv b[(1+\alpha)\frac{\phi+\delta}{\rho} + \alpha] + \overline{p}(1+C^2) - p_L(1+\alpha)(1+2C) - \overline{R}C^2 \ ,$$

$$E \equiv (1+\alpha)[2C - \frac{\phi+\delta}{\rho}] - \alpha \ , \quad F \equiv p_H - b - p_L \ , \quad G \equiv p_H - \overline{R} - p_L \ .$$

These reservation wages, together with $\tilde{w}_i$ from (11), $\overline{w}_H$ from (17), $\overline{w}_L = \overline{w}_H - b + R_L$, and the wage offer distributions from (8) and (9), characterizes the candidate equilibrium with binding incentive constraints. To prove that the candidate equilibrium indeed is an equilibrium, we still need to show that no separating firm finds it optimal to deviate to a simple pooling contract. We formulate this assertion as Lemma A.1 below. Because of this Lemma and Lemma 1, the candidate equilibrium is a market equilibrium. □

Lemma A.1: *For any $\rho \geq \rho_2$, firms do not find profitable deviations to a simple pooling contract.*
    Proof: Consider a firm offering a simple pooling contract $\tilde{\omega} = (w_H, p_H, b)$ such that $w_H \leq \overline{w}_H$, to retain high-ability workers. Note that offering a pooling contract (w_H, p_H, b) in which $w_H > \overline{w}_H$ is not profitable since this contract has the same hiring and retention rate of workers as the pooling contract $(\overline{w}_H, p_H, b)$, but leads to lower profit per worker.
    Instead of offering the pooling contract $\tilde{\omega} = (w_H, p_H, b)$, the firm can also offer a menu of separating contracts (ω_H, ω_L) with $\omega_H = \tilde{\omega}$ and $\omega_L = (\hat{w}(w_H), p_L, b)$, $\hat{w}(w_H) = w_H - b + R_L \leq p_L$. The pooling (separating) contracts, conditional on hiring a worker, yield expected profit values $\tilde{\Omega}$ (Ω) satisfying

$$\begin{aligned}
\tilde{\Omega} &= \tilde{\ell}_H[p_H - w_H] + \tilde{\ell}_L\left[p_L - w_H + \rho\frac{p_L - b}{\phi + \delta + \lambda}\right] \ , \\
\Omega &= \ell_H[p_H - w_H] + \ell_L[p_L - \hat{w}(w_H)] \ ,
\end{aligned}$$

where $\tilde{\ell}_i$ and ℓ_i are the numbers of workers of ability i in the starting phase of the two contracts. Note that both contracts have the same hiring and quit rates for both worker types in the pre-promotion stage since $F_L(\hat{w}(w_H)) = F_H(w_H)$ and low-ability workers are indifferent between reporting type L or type H at the offered contract wages. Therefore, $\ell_i = \tilde{\ell}_i$ for $i = H, L$, and the pooling contract strictly dominates the separating contract if and only if

$$p_L - \hat{w}(w_H) < p_L - w_H + \rho\frac{p_L - b}{\phi + \delta + \lambda} \ ,$$

which is equivalent to (see (5) and (6))

$$w_H - \hat{w}(w_H) = \rho\frac{p_L - \phi U_L}{\phi + \delta} < \rho\frac{p_L - b}{\phi + \delta + \lambda} \ .$$

This in turn is the same as

$$\phi U_L > \frac{\lambda p_L + (\phi + \delta)b}{\phi + \delta + \lambda} \ . \tag{34}$$

On the other hand, we have that

$$(\phi + \lambda)U_L = b + \lambda \int_{R_L}^{\overline{w}_L} V_{LL}(w)dF_L(w) < b + \lambda\frac{p_L + \delta U_L}{\phi + \delta} \ ,$$

since no contract offered to low-ability workers yields utility value greater than $(p_L + \delta U_L)/(\phi + \delta)$. But this inequality is equivalent to

$$\phi U_L < \frac{\lambda p_L + (\phi + \delta)b}{\phi + \delta + \lambda} \ , \tag{35}$$

which contradicts (34) and thus proves that the separating contract dominates the pooling contract when $\rho \geq \rho_2$. This completes the proof of Lemma A.1. □

Proof of Lemma 3

We first show that firms make negative expected profit on any low-ability worker hired in a pooling contract with initial wage $w_H \geq w_H^* = \hat{w}^{-1}(p_L)$, i.e.

$$p_L - w_H + \rho \frac{p_L - b}{\phi + \delta + \lambda} < 0 \ ,$$

for all $w_H \geq p_L + b - R_L$. This is true if and only if

$$\rho \frac{p_L - \phi U_L}{\phi + \delta} = b - R_L > \rho \frac{p_L - b}{\phi + \delta + \lambda} \ .$$

In the proof of Proposition 1 we show that this inequality is fulfilled (see equation (35)).

To prove the first claim, suppose there is no mass point. Then $\rho < \rho_2$ implies that $\overline{w}_H > w_H^* = \hat{w}^{-1}(p_L)$ so that some firms offer pooling contracts at $w_H^* + \varepsilon$. In the limit $\varepsilon \to 0$, the inflow (quit) rates of high-ability workers at these firms are identical to the inflow (quit) rates of high-ability workers at firms with the highest separating contract at w_H^*. However, for $\varepsilon \to 0$, the inflow rate of low-ability workers is strictly larger at $w_H^* + \varepsilon$ than at w_H^* since the former contract attracts the mass of promoted low-ability workers (earning p_L) whereas the latter contract does not. Hence, profit would jump down discontinuously since firms make negative profits on low-ability workers in a pooling contract $w_H \geq w_H^*$. This contradicts profit maximization.

To prove the second claim, suppose that all low-ability workers who are offered separating contracts (w_H^*, p_L) report truthfully. Then firms offering this contract earn zero profits on low-ability workers and positive profits on high-ability workers. A firm offering a separating contract at $w_H^* - \varepsilon$ and $w_L = \hat{w}(w_H^* - \varepsilon)$, however, also earns (nearly) zero profit on low-ability workers and it has the same inflow rate of high-ability workers in the limit $\varepsilon \to 0$; however, the quit rate of high-ability workers jumps down discontinuously from $w_H^* - \varepsilon$ to w_H^* because workers earning $w_H^* - \varepsilon$ (w_H^*) quit (do not quit) to another firm offering w_H^*. Since there is a mass of firms offering w_H^*, profits jump up discontinuously from $w_H^* - \varepsilon$ to w_H^*, which again contradicts profit maximization.

This completes the proof of Lemma 3. □

Proof of Proposition 2

The proof proceeds in two steps. We first characterize the vector of endogenous variables $\mathcal{E} \equiv (\varphi, \xi, \eta, R_L, R_H)$ by a set of equilibrium conditions. Second, we prove that the candidate equilibrium exists using Brouwer's fixed-point theorem.

Steady State Measures

We write $G_i(w)$ for the earnings distribution of *initial* wages and $G_i^*(w)$ for the earnings distribution of wages *after promotions/demotions*. Since the latter has mass points at p_i and at b and zero density elsewhere, we write $g_i^*(p_i)$, $i = H, L$, and $g_L^*(b)$ for the measures of employed workers after promotion/demotion. Since the distribution of initial wages has a mass point at (w_H^*, p_L), we write $g_L(p_L)$, $g_L(w_H^*)$ and $g_H(w_H^*)$ for the measures of workers earning p_L (low ability) or w_H^* (high and low ability) before promotion/demotion decisions. We calculate these earnings distribution measures as functions of equilibrium variables $\mathcal{E}$ as follows.

1. Low-ability workers

Write $g_1 = G_{L-}(p_L)$, $g_2 = G_L(p_L) - G_{L-}(p_L) = g_L(p_L)$, $g_3 = G_L(w_H^*) - G_{L-}(w_H^*) = g_L(w_H^*)$, $g_4 = G_L(\overline{w}_H) - G_L(w_H^*)$, $g_5 = g^*(b)$, $g_6 = g^*(p_L)$. Thus, fraction $G_0 \equiv g_1 + g_2 + g_3 + g_4$ of employed low-ability workers receive initial wages and fractions g_5 (g_6) have been demoted (promoted). Hence,

$G_0 + g_5 + g_6 = 1$. Note that no low-ability worker earns a wage in the interval (p_L, w_H^*), so that $G_{L-}(w_H^*) = G_L(p_L)$. Remember that fraction $\varphi > 0$ of firms offer wages strictly below p_L and fraction $1 - \eta - \varphi \geq 0$ offer wages strictly above w_H^*. Also remember that fraction ξ of low-ability workers accept w_H^* when offered (w_H^*, p_L).

G_0, g_1, g_2, g_3, g_5 and g_6 satisfy the set of linear steady-state equations

$$g_1[\phi + \delta + \rho + \lambda(1 - \varphi)] = [\phi + \delta + \lambda g_5]\varphi , \tag{36}$$

$$g_2[\phi + \delta + \rho + \lambda(1 - \varphi - \eta)] = [\phi + \delta + \lambda(g_1 + g_5)]\eta(1 - \xi) , \tag{37}$$

$$g_3[\phi + \delta + \rho + \lambda(1 - \varphi - \eta)] = [\phi + \delta + \lambda(g_1 + g_5)]\eta\xi , \tag{38}$$

$$G_0[\phi + \delta + \rho] = [\phi + \delta + \lambda g_5] + \lambda g_6(1 - \varphi - \eta) , \tag{39}$$

$$g_5[\phi + \delta + \lambda] = \rho(G_0 - g_1 - g_2) , \tag{40}$$

$$g_6[\phi + \delta + \lambda(1 - \varphi - \eta)] = \rho(g_1 + g_2) . \tag{41}$$

These can be solved for

$$g_1 + g_2 = \frac{C(E + \rho) + DE}{B(E + \rho) + DF} , \quad G_0 = \frac{E - F(g_1 + g_2)}{E + \rho} ,$$

with

$$
\begin{aligned}
B &= (\phi + \delta + \rho + \lambda(1 - \varphi - \eta))(\phi + \delta + \rho + \lambda(1 - \varphi))(\phi + \delta + \lambda) \\
&\quad + \lambda\varphi\rho(\phi + \delta + \rho + \lambda(1 - \varphi - \eta\xi)) + \lambda\rho\eta(1 - \xi)(\phi + \delta + \rho + \lambda(1 - \varphi)) , \\
C &= (\phi + \delta)(\phi + \delta + \lambda)\Big[(\varphi + \eta(1 - \xi))(\phi + \delta + \rho + \lambda(1 - \varphi)) - \lambda\varphi\eta\xi\Big] , \\
D &= \lambda\rho\Big[(\varphi + \eta(1 - \xi))(\phi + \delta + \rho + \lambda(1 - \varphi)) - \lambda\varphi\eta\xi\Big] , \\
E &= \phi + \delta + \lambda , \\
F &= \frac{\rho\lambda(\varphi + \eta)}{\phi + \delta + \lambda(1 - \varphi - \eta)} .
\end{aligned}
$$

The other variables g_n, $n = 1, \ldots, 6$ follow immediately from (36)–(41). The unemployment rate is $u = (\phi + \delta)/(\phi + \delta + \lambda)$. For the earnings distributions of initial wages, the steady state relations are

$$G_L(w) = \frac{(\phi + \delta + \lambda g_5)F_L(w)}{\phi + \delta + \rho + \lambda(1 - F_L(w))} \quad \text{if} \quad w < p_L , \tag{42}$$

$$G_L(w) = \frac{(\phi + \delta + \lambda g_5)F_H(w) + \lambda g_6(F_H(w) - \varphi - \eta)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} \quad \text{if} \quad w > w_H^* . \tag{43}$$

2. High-ability workers

For the distribution of initial wages, steady state relations yield

$$G_H(w) = \frac{(\phi + \delta)F_H(w)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} , \quad w \neq w_H^* ,$$

and

$$g_H(w_H^*) = G_H(w_H^*) - G_{H-}(w_H^*) = \frac{(\phi + \delta + \lambda G_{H-}(w_H^*))\eta}{\phi + \delta + \rho + \lambda(1 - \varphi - \eta)} .$$

Profit Maximization

Firms are indifferent between all contracts in the offer distribution. In the following, we derive this indifference relation for different wage offers to high-ability workers.

For convenience, we define $\hat{\alpha} \equiv \alpha \frac{\phi+\delta+\lambda g_L^*(b)}{\phi+\delta}$, which is the ratio of low-ability worker who are either unemployed or demoted at wage b, relative to the mass of unemployed high-ability workers.[39] We also define parameter A_0 as in Section 3, equation (7).

1. $w_H \in [R_H, \tilde{w}_H)$: Firms offer separating contracts with slack incentive constraints.
When the firm offers w_i to workers of ability $i = H, L$, profit is

$$\Omega(w_H, w_L) = \frac{A_0[p_H - w_H]}{[\phi+\delta+\rho+\lambda(1-F_H(w_H))]^2} + \hat{\alpha}\frac{A_0[p_L - w_L]}{[\phi+\delta+\rho+\lambda(1-F_L(w_L))]^2} \ .$$

With slack incentive constraints, profit is constant for both worker types independently. This gives rise to the same wage offer distributions (8) as in the previous section. Rank preservation implies that the two initial wages are linked according to (10), and the incentive constraint (6) implies that the threshold wage $\tilde{w}_H$ satisfies (11).

2. $w_H \in [\tilde{w}_H, w_H^*)$: Firms offer separating contracts with binding incentive constraints.
Incentive constraints are binding so that $w_L = \hat{w}(w_H) = w_H - b + R_L$. The firm's profit is

$$\Omega(w_H, \hat{w}(w_H)) = \frac{A_0}{(\phi+\delta+\rho+\lambda(1-F_H(w_H)))^2}\left[p_H - w_H + \hat{\alpha}(p_L - w_H + b - R_L)\right] \ .$$

The constant profit condition $\Omega(w_H, \hat{w}(w_H)) = \Omega(R_H, R_L)$ yields the wage offer distribution

$$F_H(w_H) = \frac{\phi+\delta+\rho+\lambda}{\lambda}\left[1 - \left(\frac{\hat{p} - w_H(1+\hat{\alpha}) + \hat{\alpha}(b - R_L)}{\hat{p} - \hat{R}}\right)^{1/2}\right], \quad w_H \in [\tilde{w}_H, w_H^*) , \qquad (44)$$

where we define $\hat{p} \equiv p_H + \hat{\alpha}p_L$ and $\hat{R} \equiv R_H + \hat{\alpha}R_L$. We define

$$C(x) \equiv \frac{\phi+\delta+\rho+\lambda(1-x)}{\phi+\delta+\rho+\lambda} \ ,$$

and obtain, because of $F_{H-}(w_H^*) = F_{L-}(p_L) = \varphi$, an equilibrium condition for φ

$$C(\varphi)^2 = \frac{p_H - p_L - b + R_L}{\hat{p} - \hat{R}} \ . \qquad (45)$$

3. $w_H = w_H^*$: Mass $\eta > 0$ of firms offer the contract menu (w_H^*, p_L). Fraction $\xi > 0$ of low-ability workers misreport their type.
These firm offer p_L to low-ability workers of whom fraction $1 - \xi$ accept this contract so that firms make zero profits on these workers. Fraction ξ of low-ability workers, being indifferent between the two contracts, report the wrong type, earn initial wage $w_H^* = p_L + b - R_L$ and are demoted to wage b at the firm's learning rate ρ. On each worker of low ability hired into such a contract, the firm makes expected profit

$$J_L(w_H^*) = \frac{(\phi+\delta+\lambda)(R_L - b) + \rho(p_L - b)}{(\phi+\delta+\lambda)(\phi+\delta+\rho+\lambda(1-\varphi-\eta))} \ .$$

Note that low-ability workers before demotion quit at rate $\lambda(1-\varphi-\eta)$ since fraction $1-\varphi-\eta$ of firms offer initial wages above w_H^*. After demotion, these workers quit at rate λ. Note also that $J_L(w_H^*)$ is negative (see the proof of Lemma 3).

---

[39]The mass $\alpha\frac{\phi+\delta+\lambda g_L^*(b)}{\phi+\delta+\lambda}$ of low-ability workers are unemployed or employed at demotion wage b, whereas there are $\frac{\phi+\delta}{\phi+\delta+\lambda}$ unemployed high-ability workers.

The rate at which low-ability workers are hired into this contract is

$$h_L(w_H^*) = \frac{\alpha\lambda\xi}{\phi+\delta+\lambda}\Big[\phi+\delta+\lambda(g_L^*(b)+G_{L-}(p_L))\Big] = \frac{\lambda\hat{\alpha}\xi(\phi+\delta)(\phi+\delta+\rho+\lambda)}{(\phi+\delta+\lambda)[\phi+\delta+\rho+\lambda(1-\varphi)]} \;.$$

The first expression is the flow of low-ability workers who are either unemployed or who are employed at wages below p_L who report the wrong type when meeting this firm (flow rate $\lambda\xi$). The second expression makes use of the steady-state earnings distribution of low-ability workers (42). Therefore the firm's profit on low-ability workers at wage w_H^* is

$$\Omega_L(w_H^*) = h_L(w_H^*)J_L(w_H^*) = \frac{A_0\hat{\alpha}\xi[(\phi+\delta+\lambda)(R_L-b)+\rho(p_L-b)]}{[\phi+\delta+\rho+\lambda(1-\varphi)][\phi+\delta+\rho+\lambda(1-\varphi-\eta)](\phi+\delta+\lambda)} \;.$$

Workers of high ability yield a profit flow before promotion $p_H - w_H^*$, quit at rate $\lambda(1-\varphi-\eta)$, and they are hired at rate $\frac{\lambda(\phi+\delta)(\phi+\delta+\rho+\lambda)}{(\phi+\delta+\lambda)(\phi+\delta+\rho+\lambda(1-\varphi))}$. Thus, for the firm's profit, we have a similar expression:

$$\Omega_H(w_H^*) = \frac{A_0[p_H-w_H^*]}{[\phi+\delta+\rho+\lambda(1-\varphi)][\phi+\delta+\rho+\lambda(1-\varphi-\eta)]} \;.$$

The constant-profit condition $\Omega_L(w_H^*) + \Omega_H(w_H^*) = \Omega(R_H, R_L)$ then implies

$$C(\varphi)C(\varphi+\eta) = \frac{p_H - p_L - b + R_L + \xi\hat{\alpha}[R_L - b + \frac{\rho}{\phi+\delta+\lambda}(p_L-b)]}{\hat{p}-\hat{R}} \;, \tag{46}$$

which is an equilibrium condition for the endogenous variable ξ.

4. $w_H > w_H^*$: **Firms pool all workers in the same contract, promoting high-ability workers and demoting low-ability workers.**

If the equality

$$1 - \varphi - \eta = 0 \tag{47}$$

holds, there are no firms offering wages above w_H^*. Otherwise, positive mass $1-\varphi-\eta$ of firms offer wages $w_H > w_H^*$. These firms hire all low-ability workers who currently earn w_H^* or less, including those low-ability workers who have been promoted to wage p_L. Similar to the previous case, each worker of low ability hired at w_H yields negative profit

$$J_L(w_H) = \frac{(\phi+\delta+\lambda)(p_L-w_H)+\rho(p_L-b)}{(\phi+\delta+\lambda)(\phi+\delta+\rho+\lambda(1-F_H(w_H)))} \;,$$

since these workers quit at rate $\lambda(1-F_H(w_H))$ (rate λ) before (after) demotion.

The rate at which low-ability workers are hired into this contract is

$$h_L(w_H) = \frac{\lambda\alpha}{\phi+\delta+\lambda}\Big[\phi+\delta+\lambda(g_L^*(b)+g_L^*(p_L)+G_L(w_H))\Big]$$

$$= \frac{\lambda\alpha\Big[(\phi+\delta+\rho+\lambda)(\phi+\delta+\lambda(g_L^*(b)+g_L^*(p_L)))-\lambda^2 g_L^*(p_L)(\varphi+\eta)\Big]}{(\phi+\delta+\lambda)[\phi+\delta+\rho+\lambda(1-F_H(w_H))]} \;.$$

The first expression is the flow of low-ability workers who are either unemployed or who are employed at wages below w_H meeting this firm. The second expression makes use of the steady-state earnings distribution of low-ability workers (43). We define

$$\hat{\hat{\alpha}} \equiv \alpha\frac{\phi+\delta+\lambda(g_L^*(b)+g_L^*(p_L))-\frac{\lambda^2 g_L^*(p_L)(\varphi+\eta)}{\phi+\delta+\rho+\lambda}}{\phi+\delta} \;,$$

43

and find that the firm's expected (negative) profit on low-ability workers at wages $w_H > w_H^*$ is

$$\Omega_L(w_H) = h_L(w_H) J_L(w_H) = \frac{A_0 \hat{\hat{\alpha}}[(\phi + \delta + \lambda)(p_L - w_H) + \rho(p_L - b)]}{[\phi + \delta + \rho + \lambda(1 - F_H(w_H))]^2 (\phi + \delta + \lambda)} .$$

For workers of high ability, the firm's profit is

$$\Omega_H(w_H) = \frac{A_0[p_H - w_H]}{[\phi + \delta + \rho + \lambda(1 - F_H(w_H))]^2} .$$

Now the constant-profit condition $\Omega_L(w_H) + \Omega_H(w_H) = \Omega(R_H, R_L)$ yields the wage offer distribution

$$F_H(w_H) = \frac{\phi + \delta + \rho + \lambda}{\lambda} \left[1 - \left(\frac{\hat{\hat{p}} - w_H(1 + \hat{\hat{\alpha}}) + \hat{\hat{\alpha}} \frac{\rho}{\phi + \delta + \lambda}(p_L - b)}{\hat{p} - \hat{R}} \right)^{1/2} \right], \quad w_H \in (w_H^*, \overline{w}_H], \quad (48)$$

where we define $\hat{\hat{p}} \equiv p_H + \hat{\hat{\alpha}} p_L$. The top wage follows from $F_H(\overline{w}_H) = 1$:

$$\overline{w}_H = \frac{1}{1 + \hat{\hat{\alpha}}} \left[\hat{\hat{p}} + \frac{\hat{\hat{\alpha}} \rho}{\phi + \delta + \lambda}(p_L - b) - \left(\frac{\phi + \delta + \rho}{\phi + \delta + \rho + \lambda} \right)^2 (\hat{p} - \hat{R}) \right] .$$

Since the distribution of wage offers must have connected support, $F_{H+}(w_H^*) = \lim_{w_H \searrow w_H^*} F_H(w_H) = \varphi + \eta$, we obtain an implicit condition for variable η:

$$C(\varphi + \eta)^2 = \frac{p_H - p_L - b + R_L + \hat{\hat{\alpha}}[R_L - b + \frac{\rho}{\phi + \delta + \lambda}(p_L - b)]}{\hat{p} - \hat{R}} . \quad (49)$$

Reservation Wages

It remains to derive reservation wages for the two workers types. For both types, we make use of (32) and calculate the difference $V_i(p_i) - U_i$.

For workers of low ability we have from (19) and (21)

$$(\phi + \delta) V_L(p_L) = p_L + \delta U_L + \lambda \int_{w_H^*}^{\overline{w}_H} [V_{LH}(w) - V_L(p_L)] \, dF_H(w) , \quad (50)$$

$$\phi U_L = b + \lambda \int_{R_H}^{\overline{w}_H} [\max[V_{LL}(\hat{w}(w)), V_{LH}(w)] - U_L] \, dF_H(w) \quad (51)$$

$$= b + \lambda \int_{R_H}^{\to w_H^*} [V_{LL}(\hat{w}(w)) - U_L] \, dF_H(w)$$

$$+ \lambda \int_{w_H^*}^{\overline{w}_H} [V_{LH}(w) - V_L(p_L) + V_L(p_L) - U_L] \, dF_H(w)$$

$$= b + \lambda \int_{R_H}^{\to w_H^*} [V_{LL}(\hat{w}(w)) - U_L] \, dF_H(w)$$

$$+ \lambda \int_{w_H^*}^{\overline{w}_H} [V_{LH}(w) - V_L(p_L)] \, dF_H(w) + \lambda(1 - \varphi)[V_L(p_L) - U_L] .$$

Subtracting (51) from (50) and substitution into (32) gives

$$\frac{\phi + \delta + \lambda(1 - \varphi)}{\rho}(b - R_L) = p_L - b - \lambda \int_{R_H}^{\to w_H^*} [V_{LL}(\hat{w}(w)) - U_L] \, dF_H(w) . \quad (52)$$

44

The integral expression in this equation can be calculated as follows:

$$\lambda \int_{R_H}^{\to w_H^*} V_{LL}(\hat{w}(w)) - U_L \, dF_H(w) = \lambda \int_{R_L}^{\to p_L} [V_{LL}(w) - V_{LL}(R_L)] \, dF_L(w)$$

$$= \int_{R_L}^{p_L} \frac{\lambda[\varphi - F_L(w)]}{\phi + \delta + \rho + \lambda(1 - F_L(w))} dw$$

$$= (p_L - R_L)(1 - 2C(\varphi)) + \frac{2C(\varphi)^2}{1 + \hat{\alpha}}(\hat{p} - \hat{R}) - \frac{2C(\varphi)}{1 + \hat{\alpha}}[p_H - R_H - p_L + R_L]^{1/2}[p_H - p_L - b + R_L]^{1/2} .$$

The last equation makes use of (8), (44), and $F_L(w) = F_H(w + b - R_L)$ for $w_L \geq \tilde{w}_L = \tilde{w}_H - b + R_L$, similar to the calculations in the proof of Proposition 1. Substitution back into (52) yields an equation which defines R_L implicitly.

For workers of high ability, we obtain a similar reservation wage equation:

$$\frac{\phi + \delta}{\rho}(b - R_H) = p_H - b - \lambda \int_{R_H}^{\overline{w}_H} [V_{HH}(w) - U_H] \, dF_H(w) . \tag{53}$$

For the integral expression, we obtain

$$\int_{R_H}^{\overline{w}_H} [V_{HH}(w) - V_{HH}(R_H)] dF_H(w) = \int_{R_H}^{\to w_H^*} [V_{HH}(w) - V_{HH}(R_H)] \, dF_H(w) + \eta[V_{HH}(w_H^*) - V_{HH}(R_H)]$$

$$+ \int_{\to w_H^*}^{\overline{w}_H} [V_{HH}(w) - V_{HH}(w_H^*)] dF_H(w) + (1 - \varphi - \eta)[V_{HH}(w_H^*) - V_{HH}(R_H)]$$

$$= \int_{R_H}^{w_H^*} \frac{\varphi - F_H(w)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} dw + \int_{w_H^*}^{\overline{w}_H} \frac{1 - F_H(w)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} dw$$

$$+ \int_{R_H}^{w_H^*} \frac{1 - \varphi}{\phi + \delta + \rho + \lambda(1 - F_H(w))} dw$$

$$= \int_{R_H}^{w_H^*} \frac{1 - F_H(w)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} dw + \int_{w_H^*}^{\overline{w}_H} \frac{1 - F_H(w)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} dw .$$

Using (8), (44) and (48), the sum of the last two integrals can be further calculated to obtain

$$\lambda \int_{R_H}^{\overline{w}_H} [V_{HH}(w) - V_{HH}(R_H)] dF_H(w) = \overline{w}_H - R_H - 2C(1)(p_H - R_H) + \frac{2C(1)C(\varphi)}{1 + \hat{\alpha}}(\hat{p} - \hat{R})$$

$$+ \frac{2C(1)\hat{\alpha}}{1 + \hat{\alpha}}[p_H - p_L - R_H + R_L]^{1/2}[p_H - b - p_L + R_L]^{1/2} + \frac{2C(1)(C(1) - C(\varphi + \eta))}{1 + \hat{\hat{\alpha}}}(\hat{p} - \hat{R}) .$$

Substitution back into (53) yields the implicit equation for R_H.

This completes the first part of the proof, namely the derivation of all equilibrium conditions.

Equilibrium Existence

There are two types of candidate equilibria in simple contracts. The first case is that the mass point at $w_H^* = \overline{w}_H$ is the highest offered initial wage. In this case the equilibrium variables $(\varphi, \xi, \eta, R_L, R_H)$ satisfy equations (45), (46), (47), (52) and (53). Second, some firms offer initial wages above w_H^* in which case $(\varphi, \xi, \eta, R_L, R_H)$ satisfies (45), (46), (49), (52) and (53) (so that $\eta < 1 - \varphi$). Typically the second case describes an equilibrium if the learning rate falls below some threshold ρ_3. If this threshold value exists, both (47) and (49) are jointly satisfied at $\rho = \rho_3$.

To prove equilibrium existence, we apply Brouwer's fixed-point theorem by defining a suitable continuous map $H : X \to X$ on a compact, convex subset $X \subset \mathbb{R}^5$, such that every fixed point of H satisfies the equilibrium conditions. First, we define upper and lower bounds on reservation wages. For workers of high ability, note that the integral on the right-hand side of (53) is bounded below by zero and it is bounded above by

$$\int_{R_H} [V_{HH}(w) - U_H] \, dF_H(w) < \frac{p_H - \phi U_H}{\phi + \delta} = \frac{p_H - b - \lambda \int_{R_H} [V_{HH}(w) - U_H] \, dF_H(w)}{\phi + \delta} \; ,$$

which implies

$$\int_{R_H} [V_{HH}(w) - U_H] \, dF_H(w) < \frac{p_H - b}{\phi + \delta + \lambda} \; .$$

Therefore, (53) defines upper and lower bounds on R_H:

$$\underline{R}_H \equiv b - \frac{\rho}{\phi + \delta}(p_H - b) < R_H < \overline{R}_H \equiv b - \frac{\rho}{\phi + \delta + \lambda}(p_H - b) \; .$$

For workers of low ability, the integral on the right-hand side of (52) is bounded below by zero which defines the lower bound

$$R_L > b - \frac{\rho}{\phi + \delta + \lambda(1 - \varphi)}(p_L - b) \geq \underline{R}_L \equiv b - \frac{\rho}{\phi + \delta}(p_L - b) \; .$$

The integral in (52) is bounded above by

$$\int_{R_L}^{\to w_H^*} [V_{LL}(\hat{w}(w)) - U_L] \, dF_H(w) < \varphi[V_L(p_L) - U_L] = \frac{\varphi(b - R_L)}{\rho} \; ,$$

which defines the upper bound

$$R_L < \overline{R}_L \equiv b - \frac{\rho}{\phi + \delta + \lambda}(p_L - b) \; .$$

We can now define

$$X \equiv \left\{ (\varphi, \xi, \eta, R_L, R_H) \in [0,1]^3 \times [\underline{R}_L, \overline{R}_L] \times [\underline{R}_H, \overline{R}_H] \; , \; \eta + \varphi \leq 1 \right\} \; ,$$

which is a compact, convex subset of $\mathbb{R}^5$.

Define RHS(N) (LHS(N)) for the right-hand side (left-hand side) of generic equation (N), and define functions $G_i : X \to \mathbb{R}$ by

$$
\begin{aligned}
G_1(\varphi, \xi, \eta, R_L, R_H) &= \text{LHS}(45) - \text{RHS}(45) \; , \\
G_2(\varphi, \xi, \eta, R_L, R_H) &= \text{LHS}(46) - \text{RHS}(46) \; , \\
G_3(\varphi, \xi, \eta, R_L, R_H) &= \text{LHS}(49) - \text{RHS}(49) \; , \\
G_4(\varphi, \xi, \eta, R_L, R_H) &= \text{LHS}(52) - \text{RHS}(52) \; , \\
G_5(\varphi, \xi, \eta, R_L, R_H) &= \text{LHS}(53) - \text{RHS}(53) \; .
\end{aligned}
$$

Since all steady-state measures and the other expressions defined above depend continuously on $(\varphi, \xi, \eta, R_L, R_H)$, functions G_i are continuous. Then define the map $H : X \to X$ by

$$
\begin{aligned}
H_1(x) &= \max\{0, \min\{1, \varphi + G_1(x)\}\} \; , \\
H_2(x) &= \max\{0, \min\{1, \xi - G_2(x)\}\} \; , \\
H_3(x) &= \max\{0, \min\{1 - \varphi, \eta + G_3(x)\}\} \; , \\
H_4(x) &= \max\{\underline{R}_L, \min\{\overline{R}_L, R_L + G_4(x)\}\} \; , \\
H_5(x) &= \max\{\underline{R}_H, \min\{\overline{R}_H, R_H + G_5(x)\}\} \; ,
\end{aligned}
$$

for any $x = (\varphi, \xi, \eta, R_L, R_H) \in X$. It is obvious that H maps X into itself and that it is continuous. Hence it has a fixed-point, denoted $x^* = (\varphi^*, \xi^*, \eta^*, R_L^*, R_H^*) \in X$. It still needs to be shown that this fixed-point is a candidate equilibrium. For this we need to prove that $G_i(x^*) = 0$ for $i = 1, 2, 4, 5$, and that $G_3(x^*) \geq 0$ if $\eta^* + \varphi^* = 1$.

To show $G_1(x^*) = 0$, we have to prove that $\varphi^* \in (0, 1)$. Observe that, due to Proposition 1, $\rho < \rho_2$ is equivalent to $\overline{w}_H > p_L + b - R_L$, with $\overline{w}_H$ defined in (17). This is equivalent to

$$C(1)^2 < \frac{p_H - p_L + R_L - b}{\overline{p} - \overline{R}} \ . \tag{54}$$

Hence, at $\varphi = 1$, we have that $g_L^*(b) = 0$, $\hat{\alpha} = \alpha$, so that $\hat{p} = \overline{p}$, $\hat{R} = \overline{R}$. Then (54) implies that LHS(45) < RHS(45) at $\varphi = 1$ provided that $\rho < \rho_2$. Now suppose that $\varphi^* = 1$, so that $G_1(x^*) < 0$. But this contradicts $H_1(x^*) = \varphi^* = 1$. Second, suppose that $\varphi^* = 0$, so that LHS(45) = 1 while RHS(45) < 1, and hence $G_1(x^*) > 0$, which contradicts $H_1(x^*) = \varphi^* = 0$.

Next we show that $\eta^* > 0$ and that $G_3(x^*) \geq 0$, $\varphi^* + \eta^* \leq 1$ hold with complementary slackness. Suppose first that $\eta^* = 0$. Then, since $G_1(x^*) = 0$ and because of $[R_L - b + \frac{\rho}{\phi + \delta + \lambda}(p_L - b)] < 0$ (proof of Lemma 1),

$$C(\varphi^*)^2 = \frac{p_H - p_L - b + R_L}{\hat{p} - \hat{R}} > \text{RHS(49)} \ .$$

This implies $G_3(x^*) > 0$, which contradicts $H_3(x^*) = \eta^* = 0$. From $H_3(x^*) = \eta^*$ it is also easy to show that $G_3(x^*) \geq 0$ and that $\eta^* + \varphi^* \leq 1$. Moreover, one of these inequalities must bind if $H_3(x^*) = \eta^*$.

To prove $G_2(x^*) = 0$, we have to show that $\xi^* \in (0, 1)$. Suppose first that $\xi^* = 0$; then, because of $G_1(x^*) = 0$ and $\eta^* > 0$,

$$C(\varphi^*)C(\varphi^* + \eta^*) < C(\varphi^*)^2 = \frac{p_H - p_L - b + R_L}{\hat{p} - \hat{R}} = \text{RHS(46)} \ ,$$

so that $G_2(x^*) < 0$. This contradicts $H_2(x^*) = \xi^* = 0$. Now suppose $\xi^* = 1$, which implies $g_L^*(p_L) = 0$ and therefore $\hat{\alpha} = \hat{\hat{\alpha}}$. This implies that

$$\text{RHS(46)} = \text{RHS(49)} \leq C(\varphi^* + \eta^*)^2 < C(\varphi^*)C(\varphi^* + \eta^*) \ ,$$

because of $\eta^* > 0$ and $G_3(x^*) \geq 0$. Hence $G_2(x^*) > 0$, which contradicts $H_2(x^*) = \xi^* = 1$.

Finally, it is straightforward to show that $G_4(x) < 0$ ($G_5(x) < 0$) if $R_L = \overline{R}_L$ ($R_H = \overline{R}_H$) and that $G_4(x) > 0$ ($G_5(x) > 0$) if $R_L = \underline{R}_L$ ($R_H = \underline{R}_H$). This proves that $G_i(x^*) = 0$ for $i = 4, 5$.

This completes the proof of Proposition 2. $\qquad\square$

Lemma A.2: *In a Burdett-Mortensen (1998) model with matching rates λ_0 and λ_1 for unemployed and employed workers, separation rate δ, birth/death rate ϕ, let G_i denote the earnings distribution of workers who had $i \geq 0$ job-to-job transitions since the last unemployment spell. Then G_i first-order stochastically dominates G_j for any $i > j$. Hence a worker's expected wage is increasing in the number of job-to-job transitions since the last unemployment spell.*

Proof: Write F for the wage-offer distribution with support $[R, \overline{w}]$ where R is the reservation wage. Let e_i be the mass of employed workers with exactly i job-to-job transitions since the last unemployment spell, and let u be the mass of unemployed workers. Let $G_i(w)$ be the earnings distribution of workers in group e_i, and write g_i for the density.

For group $e_i G_i(w)$, $i \geq 1$, the outflow=inflow condition is

$$e_i \left[(\Phi + \delta)G_i(w) + \lambda_1 \int_R^w g_i(w')[1 - F(w')]dw'\right] = e_{i-1}\lambda_1 \int_R^w g_{i-1}(w')[F(w) - F(w')]dw' \ . \tag{55}$$

For group $e_0 G_0(w)$, the corresponding condition is

$$e_0 \left[(\Phi + \delta)G_0(w) + \lambda_i \int_R^w g_0(w')[1 - F(w')]dw' \right] = u\lambda_0 F(w) . \tag{56}$$

Partial integration and differentiation yields the following formulas for the densities:

$$g_i(w) = \frac{\lambda_1 e_{i-1} G_{i-1}(w)}{e_i} f(w)\psi(w) , \ i \geq 1 ,$$
$$g_0(w) = \frac{\lambda_0 u}{e_0} f(w)\psi(w) ,$$

with $\psi(w) \equiv [\Phi + \delta + \lambda(1 - F(w))]^{-1}$. Because $\psi(w)$ is strictly increasing, it first follows that $g_0(w)/f(w) \approx \psi(w)$ is strictly increasing. Hence G_0 likelihood-ratio dominates F. Second, since $g_0(w)$ is proportional to $f(w)\psi(w)$, it follows that $g_1(w)/g_0(w) \approx G_0(w)$ is strictly increasing. Hence, G_1 likelihood-ratio dominates G_0, which in turn implies reverse-hazard-rate dominance: $g_1(w)/G_1(w) \geq g_0(w)/G_0(w)$ (see Krishna (2009)). Third, it follows recursively that any G_i likelihood-ratio dominates G_{i-1} for $i \geq 2$. To see this, consider

$$\frac{g_i(w)}{g_{i-1}(w)} = \frac{e_{i-1}^2 G_{i-1}(w)}{e_i e_{i-2} G_{i-2}(w)} .$$

Here the right-hand side is increasing iff $g_{i-1}(w)G_{i-2}(w) \geq g_{i-2}(w)G_{i-1}(w)$ for all w which is true iff G_{i-1} reverse-hazard-rate dominates G_{i-2} which is true if G_{i-1} likelihood-ratio dominates G_{i-2}. As this is true for $i = 2$, it holds for all $i > 2$ by induction. It follows that any G_i first-order stochastically dominates G_j for $i > j$ (Krishna (2009)). $\qquad\square$

Appendix C: Data and Simulations

Our discussion in Section 4.1 extends the work of Light and McGarry (1998) by considering a worker's cumulative count of the number of job-to-job transitions and cumulative count of the number of non-employment spells. To do so, we use the National Longitudinal Survey of Youth 1979 (NLSY), focusing on white male and female workers. Following Light and McGarry (1998), we drop from this group those workers with indeterminate entry dates, those with entry dates that preceded their 16th birthday or 01/01/78 (the earliest date included in our regressions timeframe), those who stayed in school throughout 1979-1993, those who were observed for less than 8 years after their entry date and those without employment data during the 1979-1993 period. Also following Light and McGarry (1998), to determine the entry date we use as a guide the start of the first non-enrolment (in education) spell lasting more than 12 months. Because of the lack of detailed enrolment information for the survey years 1979-1980, we use a different method in those years relative to the 1981-2010 survey years. In the 1979-1980 survey years, the NLSY only provides information on the date last enrolled. For these two years, we use the number of months between the date last enrolled and the interview date, and identify the entry date for those who had 12 months in between. From 1981 onwards, the NLSY provides a dummy variable for each month since the last interview which is equal to 1 if the respondent was enrolled in that month and zero otherwise. In this case we use the first 12-months non-enrolled streak and took the first month of that streak as the entry month.

An important difference with Light and McGarry (1998) is that we create a "main job" dummy variable for a particular month to compute transitions, instead of using all of the overlapping jobs the worker could have held in the same month. In particular, for months where more than one job was held, we follow these tie-breaking rules: (i) The job that had the most hours worked per week is taken to be the main job. (ii) If there were two or more jobs that month with the same maximum hours, the job that began earlier (earlier "jobstart") is chosen as the main job. (iii) If two or more jobs that month had the same maximum hours and "jobstart", then the one with higher wages is chosen as the main job. (iv) If two or more jobs had the same hours, start and wages, then the one which lasted longer (later "jobstop") is considered the main job. (v) If there were still two jobs that had the same hours, start, stop and wages, we assume these to be exactly the same jobs, in which case we choose arbitrarily the one with a smaller job id. Job-to-job transitions are then computed by identifying the months for which the main job changed such that the time gap between these jobs is less than a month. The non-employment spell variable is created when the main job variable is either missing or zero.

Using this data set, we take the same econometric specification proposed by Light and McGarry (1998), except that instead of including the total number of job separations at 2 years and 8 years as regressors, we replace them by the cumulative counts of the numbers of job-to-job transitions and non-employment spells. Other regressors include a quadratic on actual experience and a quadratic on tenure, years of schooling, marital status dummy, health status dummy, a part-time/full-time employment dummy, a government sector dummy, a union job dummy, one-digit industry dummies, unemployment rate and wage index as measures of aggregate conditions, a dummy if the individual lived in a city, a dummy if the individual lived in the South and a gender dummy. The dependent variable is the real log hourly wage. All these variables are computed following Light and McGarry (1998).

We then run the OLS and IV regressions considered by these authors, using their same error structure and procedure for the IV regressions. The only difference is that we use the cumulative numbers of job-to-job transitions and non-employment spells as part of the IV procedure. We perform these regressions considering all available years (1979-2010), all years up to 1994 and all years up to 1989 in the NLSY. We also perform these regressions considering the first 15 years of workers' labor market histories, the first 10 years of workers' labor market histories and the first 8 years of workers' labor market histories. Further, following Light and McGarry (1998) we consider different specifications for the error

structure in the IV regressions. We first consider the case in which the error structure only captures the effects of time-invariant individual characteristics. We then consider the cases in which the error structure considers the individual and match-specific effects together in the error term. Table 3 presents the results. We obtain a positive coefficient on the cumulative number of job-to-job transitions when considering only individual characteristics and when considering individual characteristics in conjunction with match-specific characteristics. Finally, to make sure we indeed follow the same procedure as Light and McGarry (1998), we replicate their original exercise, using separately all jobs and the main jobs, and obtain very similar results.

In the main text, Table 2 reports the estimates using the OLS regressions without the tenure variable. See Table 4 for the full regression result under this specification. The reason we present the regressions results without tenure in Table 2 is because in all our simulations we find that the Burdett and Mortensen (1998) model has a very strong positive relationship between wages and labor market experience, but not necessarily between wages and tenure. In particular, returns to tenure turn negative when the displacement shock, δ, becomes sufficiently high. With small enough δ and ϕ shocks, the Burdett and Mortensen (1998) model generates an aggregate positive relationship between wages and tenure because workers in high-paying jobs stay in those jobs for a long time. However, when δ shocks become more frequent, including our benchmark value of $\delta = 0.02$, workers that climb the wage ladder and end up in high-paying jobs are not able to stay in those jobs for a long time. In these cases we observe a negative relationship between tenure and wages (see Manning (2003) -Appendix 6A- for a formalization of this argument using a partial equilibrium set up). Nevertheless, even if we include a quadratic on tenure in the Burdett and Mortensen (1998) simulated regressions, the same general conclusion drawn from Table 2 holds.

Finally the regressions for the Burdett-Mortensen (B-M) model and for our (CT-K) model reported in Table 2 are based on simulations of 100,000 workers who we follow for up to 30 years, where all Poisson shocks (matching, separation, learning, exit) arrive at monthly frequency. We also run the same regressions following workers for the first 15, 10 and 8 years of workers' labor market histories (as also done in the data) and find that our results do not qualitatively change. Note that for these simulations we make sure that we generate the same data format as in the NLSY. Further, Table 5 reports the average cumulative counts of job-to-job transitions (JTJ) and non-employment spells (NESP) as well as workers' actual experience (EXP), job duration (J-DUR) and non-employment duration (NE-DUR) implied by the NLSY data and the (B-M) and (CT-K) simulated data sets we use for Table 2. Both our model and the calibrated B-M model fit these data averages rather well.

Table 3: OLS and IV regressions with cumulative job-to-job transitions and non-employment spells

| | All years | | | First 15 years | | | First 10 years | | | First 8 years | | |
VARIABLES	OLS	IV/GLS WE	IV/GLS WME	OLS	IV/GLS WE	IV/GLS WME	OLS	IV/GLS WE	IV/GLS WME	OLS	IV/GLS WE	IV/GLS WME
Job-to-job transitions	-0.0040***	0.0258***	0.0251***	-0.0036***	0.0183***	0.0161***	-0.0029**	0.0261***	0.0123***	-0.0022*	0.0168	0.0079
Non-employment spells	-0.0154***	-0.0578***	-0.0287***	-0.0114***	-0.0514***	-0.0137***	-0.010***	-0.0484***	-0.0065	-0.0098***	-0.0461***	-0.0020
Years of work experience X	0.0338***	0.0300***	0.0268***	0.0471***	0.0370***	0.0372***	0.0566***	0.0284**	0.0588***	0.0581***	0.0263**	0.0731***
$X^2/10$	-0.0057***	-0.0051***	-0.0043***	-0.0119***	-0.0056***	-0.0066***	-0.0193***	0.0029	-0.0200***	-0.0224***	0.0130	-0.0308**
Years of job tenure T	0.0211***	0.0234***	0.0169***	0.0390***	0.0390***	0.0318***	0.0454***	0.0599***	0.0428***	0.0490***	0.0672***	0.0606***
$T^2/10$	-0.0058***	-0.0066***	-0.0050***	-0.0203***	-0.0236***	-0.0197***	-0.0296***	-0.0523***	-0.0407***	-0.0341***	-0.0757***	-0.0759***
Transitions * T	0.0005	-0.0020***	-0.0014**	0.0002	-0.0043***	-0.0028***	-0.0001	-0.0092***	-0.0034	0.001	-0.0085	-0.0030
Transitions * T^2	-2.23×10^{-5}	0.0009***	0.0007***	0.0002	0.0030***	0.0020***	0.001	0.0097***	0.0041	-0.0019	0.0100	0.0041
1 if <12 years of schooling	-0.2942***	-0.2380***	-0.2302***	-0.2663***	-0.1935***	-0.1783***	-0.2667***	-0.1711***	-0.1672***	-0.2596***	-0.1726***	-0.1541***
1 if 12 years of schooling	-0.1951***	-0.1774***	-0.1684***	-0.1769***	-0.1163***	-0.1144***	-0.1722***	-0.1081***	-0.1144***	-0.1655***	-0.1159***	-0.1141***
1 if 16 years of schooling	0.2528***	0.2311***	0.2443***	0.2410***	0.2460***	0.2454***	0.2298***	0.2303***	0.2314***	0.2248***	0.2329***	0.2226***
1 if ≥ 17 years of schooling	0.3516***	0.2326***	0.2733***	0.3700***	0.3268***	0.3026***	0.3467***	0.3082***	0.2805***	0.3450***	0.3309***	0.2787***
1 if in school	-0.0746***	-0.1181***	-0.0810***	-0.0593***	-0.1033***	-0.0656***	-0.0431***	-0.0885***	-0.0471***	-0.0371***	-0.0842***	-0.0392***
1 if married	0.0625***	0.0548***	0.0597***	0.0310***	0.0326***	0.0340***	0.0218***	0.0309***	0.0291***	0.0178***	0.0223***	0.0210***
1 if divorced	0.0194**	0.0236**	0.0110	0.0192*	0.0152	0.0004	0.0271**	-0.0165	-0.0043	0.0187	-0.0018	-0.0080
1 if has health limitations	-0.1204***	-0.0687***	-0.0928***	-0.0967***	-0.0300***	-0.0553***	-0.1071***	-0.0279*	-0.0598***	-0.1130***	-0.0411***	-0.0606***
1 if works < 35 hrs/wk	-0.0196***	0.0181***	0.0481***	0.0023	0.0386***	0.0576***	-0.0111	0.0262***	0.0423***	-0.0170**	0.0206**	0.0445***
1 if government job	-0.2821***	-0.1650***	-0.2070***	-0.1458***	-0.0566***	-0.0911***	-0.1227***	-0.0339**	-0.0625***	-0.1257***	-0.0430***	-0.0690***
1 if union job	0.1194***	0.1344***	0.1514***	0.1557***	0.1642***	0.1703***	0.1590***	0.1695***	0.1687***	0.1616***	0.1643***	0.1674***
1 if lives in city	0.0518***	0.0092*	0.0153***	0.0818***	0.0214***	0.0364***	0.0745***	0.0230***	0.0419***	0.0672***	0.0251***	0.0419***
1 if lives in the south	-0.0372***	-0.0297***	-0.0251***	-0.0268***	-0.0246**	-0.0200***	-0.0254***	-0.0232**	-0.0207***	-0.0181***	-0.0145	-0.0127
Unemp rate	-0.0009	-0.0032***	-0.0006	0.0017	-0.0130***	-0.0101***	-0.0057***	-0.0105***	-0.0059***	-0.0056***	-0.0100***	-0.0049***
Log wage index	0.0312***	0.0439***	0.0327***	0.0068	0.0221***	0.0071	0.0068	0.0233***	0.0075	0.0200***	0.0369***	0.0174***
1 if agriculture	-0.6141***	-0.3367***	-0.4472***	-0.4483***	-0.2169***	-0.3471***	-0.4327***	-0.2145***	-0.3396***	-0.4383***	-0.2514***	-0.3431***
1 if mining	-0.1782***	-0.0186	-0.0209	0.0175	-0.1059***	-0.1059***	0.0732***	0.1140***	0.1383***	0.1065***	0.1504***	0.1664***
1 if construction	-0.1449***	-0.0232	-0.0597***	0.0089	0.0670***	0.0394	0.0563***	0.0941***	0.0713***	0.0585***	0.0899***	0.0792***
1 if manufacturing	-0.2269***	-0.0751***	-0.1175***	-0.0853***	0.0039	-0.0354	-0.0580***	0.0213	-0.0172	-0.0481**	0.0217	-0.0080
1 if transportation	-0.1372***	-0.0671***	-0.0943***	-0.0063	0.0058	0.0023	0.0347	0.0336	0.0427	0.0375*	0.0397	0.0508
1 if wholesale & retail	-0.3961***	-0.2368***	-0.3220***	-0.2536***	-0.1519***	-0.2279***	-0.2219***	-0.1286***	-0.1928***	-0.2136***	-0.1290***	-0.1781***
1 if info, finance & insurance	0.1325***	-0.0439**	-0.0257	-0.0251	0.0210	0.0394	-0.0127	0.0340	0.0448	-0.0351	0.0082	0.0268
1 if services	-0.3257***	-0.2085***	-0.2587***	-0.1847***	-0.1247***	-0.1636***	-0.1733***	-0.1229***	-0.1558***	-0.1769***	-0.1328***	-0.1580***
1 if male	0.1704***	0.1694***	0.1753***	0.1456***	0.1601***	0.1594***	0.1366***	0.1521***	0.1519***	0.1263***	0.1406***	0.1410***
Constant	1.6531***	1.5094***	1.4840***	1.6905***	1.5935***	1.5677***	1.5773***	1.5034***	1.4332***	1.4833***	1.4265***	1.3117***
Obs	57,047	57,047	57,047	42,604	42,604	42,604	33,833	33,833	33,833	28,212	28,212	28,212
R^2	0.3071	0.2687	0.2810	0.3445	0.3093	0.3269	0.3320	0.2929	0.3183	0.3447	0.3134	0.3287
Adj R^2	0.3067			0.3440			0.3314			0.3439		
Num of caseid		4,538			4,519			4,498			4,473	
Num of indivjob			20,325			16,744			14,629			13,198

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Sample: Male and female white individuals. Dependent variable: log real hourly wage. WE: Error term including only time-invariant worker heterogeneity. WME: Error term including time-invariant worker and match-specific heterogeneity.

Table 4: OLS regressions without tenure

VARIABLES	All years	First 15 years	First 10 years	First 8 years
Job-to-job transitions	-0.0073***	-0.0081***	-0.0075***	-0.0070***
Non-employment spells	-0.0185***	-0.0159***	-0.0140***	-0.0137***
Years of work experience X	0.0431***	0.0657***	0.0792***	0.0835***
$X^2/10$	-0.0075***	-0.0190***	-0.0300***	-0.0360***
1 if <12 years of schooling	-0.3036***	-0.2721***	-0.2726***	-0.2658***
1 if 12 years of schooling	-0.1931***	-0.1807***	-0.1748***	-0.1683***
1 if 16 years of schooling	0.2653***	0.2520***	0.2422***	0.2372***
1 if ≥ 17 years of schooling	0.3688***	0.3868***	0.3659***	0.3646***
1 if in school	-0.0872***	-0.0596***	-0.0426***	-0.0359***
1 if married	0.0698***	0.0349***	0.0242***	0.0197***
1 if divorced	0.0184*	0.0146	0.0210	0.0113
1 if has health limitations	-0.1166***	-0.0949***	-0.1040***	-0.1096***
1 if works < 35 hrs/wk	-0.0396***	-0.0162***	-0.0281***	-0.0336***
1 if government job	-0.2487***	-0.1433***	-0.1183***	-0.1208***
1 if union job	0.1341***	0.1641***	0.1636***	0.1656***
1 if lives in city	0.0500***	0.0820***	0.0751***	0.0674***
1 if lives in the south	-0.0362***	-0.0310***	-0.0299***	-0.0224***
Unemp rate	-0.001	-0.0089***	-0.0055***	-0.0051***
Log wage index	0.0331***	0.0057	0.0087	0.0230***
1 if agriculture	-0.5755***	-0.4434***	-0.4283***	-0.4301***
1 if mining	-0.1473***	0.009	0.0620*	0.0980***
1 if construction	-0.1228***	0.0011	0.0509**	0.0558**
1 if manufacturing	-0.2008***	-0.0855***	-0.0570***	-0.0452**
1 if transportation	-0.1151***	-0.0102	0.0327	0.0376*
1 if wholesale & retail	-0.3607***	-0.2573***	-0.2234***	-0.2123***
1 if inf, finance & insurance	-0.1164***	-0.0299*	-0.0166	-0.0370*
1 if services	-0.3198***	-0.1902***	-0.1767***	-0.1777***
1 if male	0.1707***	0.1516***	0.1423***	0.1311***
Constant	1.6476***	1.7072***	1.5934***	1.4874***
Obs	62,269	42,604	33,833	28,212
R^2	0.3309	0.339	0.3275	0.3415

*** p<0.01, ** p<0.05, * p<0.1. Sample: Male and female white individuals. Dependent variable: log real hourly wage.

Table 5: Data and model averages

Means	Data	Models	
		B-M	CT-K
JTJ	4.2	3.7	3.3
NESP	3.9	3.6	3.6
EXP	11.6	10.8	10.9
J-DUR	2.3	3.4	2.5
NE-DUR	0.87	0.69	0.68

Appendix D: Offers Contingent on Employment Status

In this Appendix we fully characterize the equilibrium when firms condition their up-or-out contracts on employment status. We first present workers value functions. Then we prove that the offer distributions faced by unemployed and employed workers do not have overlapping supports. Finally, we characterize the equilibrium. The arguments used here can then be used to analyze the equilibrium when firms condition their up-or-down contracts on employment status.

Workers Value Functions: It is convenient to think of workers randomly meeting firms who draw the high-ability wage offer w_H^s from distribution F_H^s and who offer $w_L^s = \hat{w}^s(w_H^s)$ to low-ability workers. This implies that the value of unemployment is given by

$$\phi U_i = b + \lambda \int_{\underline{w}_H^u}^{\overline{w}_H^u} \max[V_{iL}(\hat{w}^u(x)) - U_i, V_{iH}(x) - U_i, 0] dF_H^u(x) \ . \tag{57}$$

The value of employment at initial wage w for a worker of ability i that reported truthfully his type is given by

$$\phi V_{ii}(w_i) = w_i + \lambda \int_{\underline{w}_H^e}^{\overline{w}_H^e} \max[V_{iL}(\hat{w}^e(x)) - V_{ii}(w_i), V_{iH}(x) - V_{ii}(w_i), 0] dF_H^e(x) \\ + \delta(U_i - V_{ii}(w_i)) + \rho(V_i(p_i) - V_{ii}(w_i)) \ , \tag{58}$$

where the continuation value after promotion is given by (21) with F_H^e. The value of employment of a worker of type i that misreported his type and is currently earning the initial wage w_j is given by

$$\phi V_{ij}(w_j) = w_j + \lambda \int_{\underline{w}_H^e}^{\overline{w}_H^e} \max[V_{iL}(\hat{w}^e(x)) - V_{ij}(w_j), V_{iH}(x) - V_{ij}(w_j), 0] dF_H^e(x) \\ + (\delta + \rho)(U_i - V_{ij}(w_j)) \ .$$

The incentive-compatibility constraint is given by $w_j^s - w_i^s \le \rho[V_i(p_i) - U_i]$.

Proof of Lemma 4: To show this, it is useful to note that in any equilibrium $\underline{w}_i^e > \underline{w}_i^u \ge R_i$. This follows since, for example, $\underline{w}_i^e < \underline{w}_i^u$ implies firms offering $\underline{w}_i^e$ will not hire any worker and will make positive profits by offering a wage slightly above $\underline{w}_i^u$. We divide the proof in two steps. The first step constructs the profits of firms when hiring unemployed workers and derives some preliminary results. The second step compares the firms' profits when hiring employed workers under the assumption of overlapping supports.

Step 1:

Consider a firm offering w_H^u and $w_L^u = \hat{w}^u(w_H^u)$ to unemployed workers. Note that this firm could be separating or pooling workers. When this firm pools we have that $w_H^u = w_L^u$, while when it separates workers $w_H^u \ne w_L^u$. In either case its steady-state profit is $\Omega^u(w_H^u, w_L^u) = \Omega_H^u(w_H^u) + \Omega_L^u(w_L^u)$, where

$$\Omega_i^u(w_i^u) = \frac{\lambda u_i \alpha_i (p_i - w_i^u)}{\phi + \delta + \rho + \lambda(1 - F_H^e(w_H^u))} \ , \tag{59}$$

for $i = L, H$. Let $N_i^u(w_i^u)$ denote the mass of workers hired from unemployment of ability $i = L, H$ earning an initial wage no greater than w_i^u. Steady state turnover implies that

$$N_i^u(w_i^u) = \frac{\lambda u_i \alpha_i F_i^u(w_i^u)}{\phi + \delta + \rho + \lambda(1 - F_H^e(w_H^u))} \ .$$

Now we show under what conditions firms would prefer to offer separating contracts when hiring unemployed workers. First, if satisfying the incentive-compatibility constraint of low-ability workers implies $w_H^u > w_L^u$ and $w_i^u \leq p_i$ for $i = L, H$, this firm will strictly prefer to separate workers. To verify this claim, suppose this firm offered a wage $w_L'^u$ to low-ability workers that did not satisfy their incentive constraint, while keeping w_H^u constant. (59) then implies this firm will strictly reduce its profits by doing so, as it will hire low-ability workers at a higher wage while keeping the hiring and retention rates constant. Second, if satisfying incentive compatibility implies $w_H^u \leq w_L^u$, but $w_H^u < R_L$, this firm will also prefer to separate workers, as offering w_H^u to all unemployed workers will only hire high-ability workers. However, if satisfying incentive compatibility implies $w_H^u \leq w_L^u$, but $w_H^u \geq R_L$, then this firm prefers to pool workers and offer both types w_H^u.

Also note that $\underline{w}_i^e > \underline{w}_i^u \geq R_i$ and (59) implies that in equilibrium $\underline{w}_i^u = R_i$ and firms separate workers at low wages. Further, (59) also implies that any firm offering wages $w_i^u \in [\underline{w}_i^u, \underline{w}_i^e)$ for each $i = L, H$, must offer $w_i^u = R_i$, as this firm faces a constant hiring rate and all workers employed in this firm quit as soon as they get an outside offer.

Finally note that in equilibrium $\overline{w}_i^e \geq \overline{w}_i^u$ for $i = L, H$. To verify this argument fix $\overline{w}_i^e$ for $i = L, H$ and suppose $\overline{w}_i^e < \overline{w}_i^u$ for $i = L, H$. Note that (59) then implies that the firm offering $(\overline{w}_L^u, \overline{w}_H^u)$ will optimally set $\overline{w}_i^u = \overline{w}_i^e$ for $i = L, H$, since by doing so it increases flow profits while keeping constant its hiring and retention rates. Reducing the $\overline{w}_i^u$ below $\overline{w}_i^e$ will be optimal, for example, when in equilibrium $\underline{w}_i^e > \overline{w}_i^u$ for $i = L, H$.

Step 2:

In this step we prove the non-overlapping support result by contradiction. Suppose there exists an equilibrium in which $\overline{w}_i^u \in (\underline{w}_i^e, \overline{w}_i^e]$ for $i = L, H$. We want to show that $\overline{w}_i^u > \underline{w}_i^e$ violates the constant-profit condition in the employed workers' market and hence cannot occur in any equilibrium. To do so consider a firm offering wages $w_i^e = \underline{w}_i^e$ for $i = L, H$. This firm's steady-state profit is then given by $\Omega^e(\underline{w}_H^e, \underline{w}_L^e) = \sum_{i=L,H} \Omega_i^e(\underline{w}_i^e)$, where

$$\Omega_i^e(\underline{w}_i^e) = \frac{\lambda N_i^u(\underline{w}_i^e)[p_i - \underline{w}_i^e]}{\phi + \delta + \rho + \lambda} = \frac{\lambda F_i^u(R_i)}{\phi + \delta + \rho + \lambda} \Omega_i^u(R_i) \ . \tag{60}$$

The second equality follows from noting that this firm only hires workers that are currently earning R_i, and it maximises profits by setting $\underline{w}_i^e = R_i + \varepsilon$, where $\varepsilon > 0$ is sufficiently small, for $i = L, H$. Continuity of (59) then implies $\Omega_i^u(R_i) = \Omega_i^u(\underline{w}_i^e)$. These results also imply that a firm offering $\underline{w}_i^e$ for $i = L, H$ will separate workers in equilibrium.

Now consider a firm offering a wage $w_H^e = \overline{w}_H^u + \varepsilon \leq p_H$ to employed workers. Here we have several possible cases depending on whether the firms decide to pool or separate when hiring unemployed and employed workers. Note that firms could follow different strategies when hiring workers of different employment status.

First assume that the incentive constraint of low-ability workers is binding at $\overline{w}_H^u$, such that $\overline{w}_L^u = \overline{w}_H^u - \rho(p_L - \phi U_L)/(\phi + \delta)$, and suppose that $\overline{w}_L^u < p_L$. The firm offering $w_H^e = \overline{w}_H^u + \varepsilon \leq p_H$ to employed workers will then offer $w_L^e = \hat{w}^e(w_H^e) = \overline{w}_L^u + \varepsilon$ to low-ability workers. Since $\overline{w}_H^u > \overline{w}_L^u$, firms offering these wages separate workers hired from unemployment and employment. Let $N_i^e(w)$ denote the mass of workers hired from employment with ability i earning a wage no greater than w. The steady-state profit from hiring employed workers at these wages is given by $\Omega^e(\overline{w}_H^u, \overline{w}_L^u) = \sum_{i=L,H} \Omega_i^e(\overline{w}_i^u)$, where

$$\begin{aligned}
\Omega_i^e(\overline{w}_i^u) &= \frac{\lambda[N_i^u(\overline{w}_i^u) + N_i^e(\overline{w}_i^u)](p_i - \overline{w}_i^u)}{\phi + \delta + \rho + \lambda(1 - F_i^e(\overline{w}_i^u))} \\
&= \frac{\lambda F_i^u(\overline{w}_i^u)\Omega_i^u(\overline{w}_i^u)}{\phi + \delta + \rho + \lambda(1 - F_i^e(\overline{w}_i^u))} + \frac{\lambda N_i^e(\overline{w}_i^u)(p_i - \overline{w}_i^u)}{\phi + \delta + \rho + \lambda(1 - F_i^e(\overline{w}_i^u))} \ .
\end{aligned} \tag{61}$$

55

Since in equilibrium $\sum_{i=L,H} \Omega_i^u(R_i) = \sum_{i=L,H} \Omega_i^u(\overline{w}_i^u)$, comparing the profits from offering $(\underline{w}_H^e, \underline{w}_L^e)$ and $(\overline{w}_H^u, \overline{w}_L^u)$ to employed workers implies $\Omega^e(\overline{w}_H^u, \overline{w}_L^u) > \Omega^e(\underline{w}_H^e, \underline{w}_L^e)$, contradicting the constant profit condition on the employed workers' market.

Now suppose that offering a wage $w_H^e = \overline{w}_H^u + \varepsilon \leq p_H$ to unemployed workers of high ability implies that satisfying the incentive-compatibility constraint of low-ability workers yields a wage $w_L^u > p_L$. In this case the firms will offer pooling contracts with initial wage $\overline{w}_H^u$ to both unemployed and employed workers. Noting that we can still use (61), with $\overline{w}_i^u = \overline{w}_H^u$ for $i = L, H$ and that in equilibrium $\sum_{i=L,H} \Omega_i^u(R_i) = \sum_{i=L,H} \Omega_i^u(\overline{w}_H^u)$, one can verify that $\Omega^e(\overline{w}_H^u, \overline{w}_L^u) > \Omega^e(\underline{w}_H^e, \underline{w}_L^e)$, once again contradicting the constant-profit condition on the employed workers' market.

Finally, assume that the incentive constraint of low-ability workers is slack at wages $\overline{w}_i^u$ for $i = L, H$. In this case, we could have that incentive compatibility implies $\overline{w}_H^u < \overline{w}_L^u$ and firms offering $\overline{w}_H^u$ to unemployed workers of high ability can decide to pool workers, while the firm offering $\overline{w}_H^u$ to employed workers of high ability decides to separate. Note, however, that when the incentive constraint is slack, the constant-profit condition can be applied to high-ability workers independently in both the employed and unemployed workers' market. Hence using $\Omega_H^u(R_H) = \Omega_H^u(\overline{w}_H^u)$ and (60) and (61), we obtain that $\Omega_H^e(R_H) < \Omega_H^e(\overline{w}_H^u)$, which contradicts the constant-profit condition on the market for employed high-ability workers.

Taking the above arguments together, they show that in any equilibrium $\underline{w}_i^e > \overline{w}_i^u$ for all $i = L, H$.
$\square$

An implication of Lemma 4 is that all firms offer unemployed workers the separating pair of reservation initial wages (R_H, R_L), while employed workers are offered initial wages $w_i > R_i$, drawn from cdf F_i, such that the infimum of the support of F_i is equal to R_i. We now proceed to characterize separating and pooling equilibria for this economy.

General considerations

Since firms offer unemployed workers their reservation wage, unemployment utility for both types is identical, satisfying $U_L = U_H = U = b/\phi$. Bellman equations for (truthful and misreporting) low-ability workers and for (truth-telling) high-ability workers are

$$[\phi + \delta + \rho + \lambda]V_{LL}(w) = w + \delta U + \rho V_L(p_L) + \lambda \int \max[V_{LL}(\hat{w}(w')), V_{LH}(w'), V_{LL}(w)]dF_H(w') \ ,$$

$$[\phi + \delta + \rho + \lambda]V_{LH}(w) = w + [\delta + \rho]U + \lambda \int \max[V_{LL}(\hat{w}(w')), V_{LH}(w'), V_{LH}(w)]dF_H(w') \ ,$$

$$[\phi + \delta + \rho + \lambda]V_{HH}(w) = w + \delta U + \rho V_H(p_H) + \lambda \int \max[V_{HL}(\hat{w}(w')), V_{HH}(w'), V_{HH}(w)]dF_H(w') \ .$$

From $V_{ii}(R_i) = U$ follow the reservation wage equations

$$R_i = b - \rho[V_i(p_i) - U] - \lambda \int \max[V_{iL}(\hat{w}(w)), V_{iH}(w)] - U \ dF_H(w) \ . \tag{62}$$

Since promoted high-ability workers do not quit, it follows that $\rho[V_H(p_H) - U] = \frac{\rho(p_H - b)}{\phi + \delta}$. Further define $x \equiv \rho[V_L(p_L) - U]$. Then the incentive constraint for low-ability workers is

$$w_H - w_L \leq x \ . \tag{63}$$

Separating equilibrium

Since all firms offer separating contracts with initial wages $w_L \leq p_L$, promoted low-ability workers do not quit. Therefore,

$$x = \frac{\rho(p_L - b)}{\phi + \delta} . \tag{64}$$

Since no low-ability worker misreports the type, unemployment rates for both types are identical,

$$u_L = u_H = u = \frac{(\phi + \delta)}{\phi + \delta + \lambda} .$$

Let h_i denote the fraction of employed workers of type i hired from unemployment and earning initial reservation wage R_i. Further, let $G_i(w)$ denote the cumulative distribution of initial wages for workers hired from another employer. In steady state this distribution satisfies

$$G_i(w) = \frac{\lambda h_i F_i(w)}{\phi + \delta + \rho + \lambda(1 - F_i(w))} .$$

The steady-state workforce of workers of ability i employed at initial wage w_i is

$$\ell_i(w_i) = \frac{\lambda \alpha_i (1 - u)(h_i + G_i(w_i))}{\phi + \delta + \rho + \lambda(1 - F_i(w_i))} .$$

It follows that firms' profits are again given by (7), with a redefinition of parameter $A_0 \equiv \lambda^2(\phi + \delta)/(\phi + \delta + \lambda)$. It follows that the wage-offer distribution for $w_i \leq \tilde{w}_i$ is again given by (8). For wages above $\tilde{w}_i$, the binding incentive constraint (63) and $\Omega_H(w_H) + \Omega_L(w_H - x) = \Omega_H(R_H) + \Omega_L(R_L)$ yield the wage-offer distribution

$$F_H(w_H) = \frac{(\phi + \delta + \rho + \lambda)}{\lambda} \left[1 - \left(\frac{\overline{p} - w_H(1 + \alpha) + \alpha x}{\overline{p} - \overline{R}} \right)^{1/2} \right] , \quad w_H \geq \tilde{w}_H ,$$

where $\overline{p}$ and $\overline{R}$ are defined as in the main text.

We again consider a rank-preserving equilibrium. From (10) follows that the incentive constraint (63) is slack for wages

$$w_H \leq \tilde{w}_H \equiv \frac{(p_H - R_H)x + p_H R_L - p_L R_H}{p_H - R_H - p_L + R_L} , \tag{65}$$

while for wages above $\tilde{w}_H$, $w_L = \hat{w}(w_H) = w_H - x$.

To characterize an equilibrium with slack incentive constraints and to derive threshold parameter ρ_1, note that

$$\lambda \int_{R_i}^{\overline{w}_i} [V_{ii}(w) - V_{ii}(R_i)]dF_i(w) = \int_{R_i}^{\overline{w}_i} \frac{\lambda(1 - F_i(w))}{\phi + \delta + \rho + \lambda(1 - F_i(w))}dw = (p_i - R_i)(1 + C^2 - 2C) ,$$

with parameter C defined as in (12). Reservation wages can then be obtained from (62) and (64):

$$R_i = p_i - (p_i - b)\frac{\phi + \delta + \rho}{(\phi + \delta)C(2 - C)} . \tag{66}$$

Since promotion wages for high-ability workers are higher, we have $R_H < R_L$. Incentive constraints are slack if $\overline{w}_H - \overline{w}_L \leq x$, where top initial wages $\overline{w}_i$ are related to reservation wages via (12). Hence, threshold parameter ρ_1 is the implicit solution of

$$p_H - p_L - C^2(p_H - p_L - R_H + R_L) = x ,$$

with reservation wages (66).

If $\rho \leq \rho_1$, incentive constraints bind at the top of the wage offer distribution. A separating equilibrium with this feature is similarly characterized as in the main text. Instead of (17) we have

$$\overline{w}_H = \frac{1}{1+\alpha}\left[\overline{p} + \alpha x - C^2(\overline{p} - \overline{R})\right] \, , \, \overline{w}_L = \overline{w}_H - x \, .$$

This outcome only describes a separating equilibrium if $\overline{w}_L$ does not exceed marginal productivity of low-ability workers. Hence, the threshold value ρ_2 is defined as the highest value for which $\overline{w}_H - x = p_L$ and hence ensures that a separating equilibrium exists for all $\rho \geq \rho_2$.

In a separating equilibrium with binding incentive constraints, reservation wages for the two worker types can be similarly calculated as in the proof of Proposition 1. Define again $h_i(w) \equiv \frac{\lambda(1-F_i(w))}{\phi+\delta+\rho+\lambda(1-F_i(w))}$, and write (62) as

$$R_i - b + \frac{\rho}{\phi+\delta}(p_i - b) = \int_{R_i}^{\overline{w}_i} h_i(w)dw \, .$$

As in the proof of Proposition 1, these integrals can be calculated as follows:

$$\int_{R_H}^{\overline{w}_H} h_H(w)dw = \frac{1}{1+\alpha}\left[(\overline{p} - \overline{R})(1 + C^2) + \alpha(x + R_L - R_H)\right]$$

$$+2C\frac{1}{1+\alpha}[p_H - p_L - x]^{1/2}[p_H - R_H - p_L + R_L]^{1/2} - 2C(p_H - R_H) \, ,$$

$$\int_{R_L}^{\overline{w}_L} h_L(w)dw = \frac{1}{1+\alpha}\left[(\overline{p} - \overline{R})(1 + C^2) + R_H - R_L - x\right]$$

$$-2C\frac{1}{1+\alpha}[p_H - p_L - x]^{1/2}[p_H - R_H - p_L + R_L]^{1/2} - 2C(p_L - R_L) \, .$$

Adding the equation for R_H to the one for R_L multiplied by α, the resulting equation can be solved for $\overline{R} = R_H + \alpha R_L$:

$$\overline{R} = \overline{p} - [\overline{p} - (1+\alpha)b]\frac{\phi+\delta+\rho}{(\phi+\delta)C(2-C)} \, .$$

Substitution of $R_H = \overline{R} - \alpha R_L$ into the reservation wage equation for R_L yields a quadratic equation in R_L whose relevant root is

$$R_L = \frac{-D + \sqrt{D^2 - 4E}}{2}$$

with

$$N \equiv x - b - 2Cp_L + \frac{1}{1+\alpha}\left[(\overline{p} - \overline{R})(1 + C^2) + \overline{R} - x\right] \, ,$$

$$D \equiv \frac{N}{C} - \frac{1}{1+\alpha}(p_H - p_L - x) \, ,$$

$$E \equiv \frac{N^2}{4C^2} - \frac{1}{(1+\alpha)^2}(p_H - p_L - \overline{R})(p_H - p_L - x) \, .$$

These reservation wages together with the previous equations for $\tilde{w}_i$, $\overline{w}_H$, $\overline{w}_L = \overline{w}_H - x$, and the wage offer distributions characterize the separating equilibrium with binding incentive constraints.

Segmented equilibrium with pooling at top wages

As in the main text, a pooling equilibrium is characterized by the variables $(\varphi, \xi, \eta, R_L, R_H)$. All firms offer separating initial wages (R_H, R_L) to unemployed workers and dispersed initial wages to employed workers. Fraction $\varphi > 0$ of firms offer separating contracts (w_H, w_L) with $w_L < p_L$, fraction

$\eta > 0$ offer the same separating contract $(p_L + x, p_L)$, and fraction $1 - \varphi - \eta \geq 0$ offer dispersed pooling contracts (w_H) with $w_H > p_L + x$. Low-ability workers who are employed at initial wage $w_L < p_L$ and quit to a firm offering $(p_L + x, p_L)$ misreport their type with probability $\xi > 0$. One possibility is that the mass point at $w_H = p_L + x$ is the highest offered contract; this happens when the learning rate is sufficiently large (greater than some threshold $\rho_3 < \rho_2$). The other possibility, when $\rho < \rho_3$, is that positive mass of firms offer pooling contracts with $w_H > p_L + x$.

To obtain steady-state measures, write $G_i(w)$ for the earnings distribution of initial wages for those workers who were hired from another employer. Write h_i for the measure of workers who were hired from unemployment and currently earn initial wage R_i, and write $g_i^*(p_i)$, $i = H, L$, for the measures of employed workers after promotion. Since the distribution of initial wages has a mass point at (w_H^*, p_L), we write $g_L(p_L)$, $g_L(w_H^*)$ and $g_H(w_H^*)$ for the measures of workers earning p_L (low ability) or w_H^* (high and low ability) before promotion/layoff. These earnings distribution depend on equilibrium variables $(\varphi, \xi, \eta, R_L, R_H)$ as follows.

Low-ability workers: Write $g_1 = G_{L-}(p_L)$, $g_2 = G_L(p_L) - G_{L-}(p_L) = g_L(p_L)$, $g_3 = G_L(w_H^*) - G_{L-}(w_H^*) = g_L(w_H^*)$, $g_4 = G_L(\overline{w}_H) - G_L(w_H^*)$, $g_6 = g^*(p_L)$. Thus, fraction $G_0 \equiv g_1 + g_2 + g_3 + g_4$ of employed low-ability workers receive initial wages and were hired from another employer. Fraction g_6 have been promoted, and h_L were hired from unemployment. Hence, $G_0 + g_6 + h_L = 1$. Mass $\alpha_L u_L$ of workers is unemployed and mass $\alpha_L(1 - u_L)$ is employed. As in the benchmark model, $G_{L-}(w_H^*) = G_L(p_L)$ because no low-ability worker earns a wage in the interval (p_L, w_H^*). G_0, g_1, g_2, g_3, g_6, h_L and u_L satisfy the following steady-state equations

$$
\begin{aligned}
g_1[\phi + \delta + \rho + \lambda(1 - \varphi)] &= h_L \lambda \varphi \,, \\
g_2[\phi + \delta + \rho + \lambda(1 - \varphi - \eta)] &= (h_L + g_1)\lambda(1 - \xi)\eta \,, \\
g_3[\phi + \delta + \rho + \lambda(1 - \varphi - \eta)] &= (h_L + g_1)\lambda\xi\eta \,, \\
G_0[\phi + \delta + \rho] &= h_L\lambda + \lambda g_6(1 - \eta - \varphi) \,, \\
g_6[\phi + \delta + \lambda(1 - \varphi - \eta)] &= \rho(h_L + g_1 + g_2) \,, \\
(1 - u_L)h_L[\phi + \delta + \rho + \lambda] &= u_L\lambda \,, \\
u_L[\phi + \lambda] &= \phi + (1 - u_L)\delta + (1 - u_L)(g_3 + g_4)\rho \,.
\end{aligned}
$$

We can solve these as follows:

$$
\begin{aligned}
g_1 &= \frac{\lambda\varphi}{\phi + \delta + \rho + \lambda(1 - \varphi)} h_L \equiv A h_L \,, \\
g_2 &= \frac{(1 + A)\lambda(1 - \xi)\eta}{\phi + \delta + \rho + \lambda(1 - \varphi - \eta)} h_L \equiv B h_L \,, \\
g_3 &= \frac{(1 + A)\lambda\xi\eta}{\phi + \delta + \rho + \lambda(1 - \varphi - \eta)} h_L \,, \\
g_6 &= \frac{(1 + A + B)\rho}{\phi + \delta + \lambda(1 - \varphi - \eta)} h_L \equiv D h_L \,, \\
h_L &= \frac{\phi + \delta + \rho}{\lambda + \phi + \delta + \rho + D(\phi + \delta + \rho + \lambda(1 - \eta - \varphi))} \,, \\
u_L &= \frac{(\phi + \delta + \rho + \lambda)h_L}{\lambda + h_L(\phi + \delta + \rho + \lambda)} \,.
\end{aligned}
$$

The stationary earnings distribution of initial wages is

$$
G_L(w) = \frac{\lambda h_L F_L(w)}{\phi + \delta + \rho + \lambda(1 - F_L(w))} \quad \text{if} \quad w < p_L \,,
$$

$$
G_L(w) = \frac{\lambda h_L F_H(w) + \lambda g_6(F_H(w) - \varphi - \eta)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} \quad \text{if} \quad w > w_H^* \,.
$$

High-ability workers: Here the same equations for G_H as in the proof of Proposition 2 apply. The unemployment rate is $u_H = (\phi + \delta)/(\phi + \delta + \lambda)$, and mass $h_H = u_H \lambda/(1 - u_H)/(\phi + \delta + \rho + \lambda)$ of employed workers were hired from unemployment and earn initial wage R_H.

Profit Maximization: To find wage-offer distributions from the firms' profit-maximization conditions, define $\hat{\alpha} \equiv \alpha u_L/u_H$ as the measure of low-ability unemployed workers per unemployed high-ability worker. Hence, a random worker hired from unemployment has high ability with probability $1/(1 + \hat{\alpha})$ and low ability otherwise. Further define parameter A_0 as above.

Wage-offer distributions in the four possible cases are similar to those in the proof of Proposition 2 with the following differences:

$w_H \in [R_H, \tilde{w}_H)$: **Separating contracts with slack incentive constraints.** Here the wage offer distributions are again given by (8), and the threshold wage $\tilde{w}_H$ is given by (65).

$w_H \in [\tilde{w}_H, w_H^*)$: **Separating contracts with binding incentive constraints.** With binding incentive constraints $w_L = w_H - x$, the firms' profit is

$$\Omega(w_H, w_H - x) = \frac{A_0}{(\phi + \delta + \rho + \lambda(1 - F_H(w_H)))^2} \left[p_H - w_H + \hat{\alpha}(p_L - w_H + x) \right].$$

The constant profit condition $\Omega(w_H, w_H - x) = \Omega(R_H, R_L)$ yields the same wage offer distribution as in (44), where the term $(b - R_L)$ is replaced by x, and $\hat{p} \equiv p_H + \hat{\alpha} p_L$ and $\hat{R} \equiv R_H + \hat{\alpha} R_L$ are similarly defined. With the same definition of $C(x)$ as in the proof of Proposition 2, the restriction $F_{H-}(w_H^*) = F_{L-}(p_L) = \varphi$ implies the equilibrium condition for φ:

$$C(\varphi)^2 = \frac{p_H - p_L - x}{\hat{p} - \hat{R}}. \tag{67}$$

Mass point at $w_H = w_H^*$ with some pooling and layoffs of low-ability workers. Fraction ξ of low-ability workers misreport their type while the rest accept p_L. On each misreporting worker of low ability hired at w_H^* the firm makes expected (negative) profit

$$J_L(w_H^*) = -\frac{x}{\phi + \delta + \rho + \lambda(1 - \varphi - \eta)}.$$

The rate at which low-ability workers are hired into this contract is

$$h_L(w_H^*) = \alpha \lambda \xi (1 - u_L) \left[h_L + G_{L-}(p_L) \right] = \frac{\lambda^2 \hat{\alpha} \xi (\phi + \delta)}{(\phi + \delta + \lambda)[\phi + \delta + \rho + \lambda(1 - \varphi)]}.$$

Therefore the firm's profit on low-ability workers at wage w_H^* is

$$\Omega_L(w_H^*) = h_L(w_H^*) J_L(w_H^*) = -\frac{A_0 \hat{\alpha} \xi x}{[\phi + \delta + \rho + \lambda(1 - \varphi)][\phi + \delta + \rho + \lambda(1 - \varphi - \eta)]}.$$

Firms' profits on high-ability workers $\Omega_H(w_H^*)$ is the same as in the proof of Proposition 2. The constant-profit condition $\Omega_L(w_H^*) + \Omega_H(w_H^*) = \Omega(R_H, R_L)$ then implies the equilibrium condition for variable ξ:

$$C(\varphi)C(\varphi + \eta) = \frac{p_H - p_L - x - \xi \hat{\alpha} x}{\hat{p} - \hat{R}}. \tag{68}$$

$w_H > w_H^*$: **Pooling contracts with layoffs of low-ability workers.**

This is only an equilibrium strategy if $1 - \varphi - \eta > 0$. Otherwise (i.e. $1 - \varphi - \eta = 0$), the highest initial wage is at w_H^*. Each worker of low ability hired at $w_H > w_H^*$ yields (negative) profit

$$J_L(w_H) = \frac{p_L - w_H}{\phi + \delta + \rho + \lambda(1 - F_H(w_H))}.$$

The rate at which low-ability workers are hired into this contract is

$$h_L(w_H) = \lambda\alpha(1 - u_L)\Big[h_L + G_L(w_H) + g_6\Big] = \lambda\alpha(1 - u_L)\frac{(\phi + \delta + \rho + \lambda)(h_L + g_6) - \lambda g_6(\varphi + \eta)}{\phi + \delta + \rho + \lambda(1 - F_H(w))} \ .$$

With the redefinition

$$\hat{\hat{\alpha}} \equiv \alpha(1 - u_L)\frac{[(\phi + \delta + \rho + \lambda)(h_L + g_6) - \lambda g_6(\varphi + \eta)](\phi + \delta + \lambda)}{\lambda(\phi + \delta)} \ ,$$

the firm's expected (negative) profit on low-ability workers at wages $w_H > w_H^*$ is

$$\Omega_L(w_H) = h_L(w_H)J_L(w_H) = \frac{A_0\hat{\hat{\alpha}}(p_L - w_H)}{[\phi + \delta + \rho + \lambda(1 - F_H(w_H))]^2} \ .$$

Profit on high-ability workers $\Omega_H(w_H)$ is the same as in the proof of Proposition 2. Then the constant-profit condition $\Omega_L(w_H) + \Omega_H(w_H) = \Omega(R_H, R_L)$ yields the wage offer distribution

$$F_H(w_H) = \frac{\phi + \delta + \rho + \lambda}{\lambda}\left[1 - \left(\frac{\hat{\hat{p}} - w_H(1 + \hat{\hat{\alpha}})}{\hat{p} - \hat{R}}\right)^{1/2}\right], \ w_H \in (w_H^*, \overline{w}_H] \ ,$$

where $\hat{\hat{p}} \equiv p_H + \hat{\hat{\alpha}}p_L$. The top wage follows from $F_H(\overline{w}_H) = 1$:

$$\overline{w}_H = \frac{1}{1 + \hat{\hat{\alpha}}}\left[\hat{\hat{p}} - \left(\frac{\phi + \delta + \rho}{\phi + \delta + \rho + \lambda}\right)^2(\hat{p} - \hat{R})\right] \ .$$

From the limiting condition $F_{H+}(w_H^*) = \lim_{w_H \searrow w_H^*} F_H(w_H) = \varphi + \eta$ follows the condition for variable η (provided that $\eta < 1 - \varphi$):

$$C(\varphi + \eta)^2 = \frac{p_H - p_L - x - \hat{\hat{\alpha}}x}{\hat{p} - \hat{R}} \ . \tag{69}$$

Reservation Wages

We first calculate the value $V_L(p_L)$:

$$\begin{aligned}(\phi + \delta)V_L(p_L) &= p_L + \delta U + \lambda\int_{w_H^*}^{\overline{w}_H}[V_{LH}(w) - V_L(p_L)] \, dF_H(w) \ , \\ &= p_L + \delta U + \int_{w_H^*}^{\overline{w}_H}\frac{\lambda(1 - F_H(w))}{\phi + \delta + \rho + \lambda(1 - F_H(w))} \, dw \ . \end{aligned} \tag{70}$$

The last integral is

$$\int_{w_H^*}^{\overline{w}_H}\frac{\lambda(1 - F_H(w))}{\phi + \delta + \rho + \lambda(1 - F_H(w))} \, dw = \overline{w}_H - w_H^* - \frac{\phi + \delta + \rho}{\phi + \delta + \rho + \lambda}\int_{w_H^*}^{\overline{w}_H}\left(\frac{\hat{p} - \hat{R}}{\hat{\hat{p}} - (1 + \hat{\hat{\alpha}})w}\right)^{1/2}dw$$

$$= \overline{w}_H - w_H^* - 2\frac{C(1)[C(\varphi + \eta) - C(1)]}{1 + \hat{\hat{\alpha}}}(\hat{p} - \hat{R}) \ .$$

This equation together with (70), $x = \rho[V_L(p) - U]$ and $w_H^* = p_L + x$ can be used to obtain x:

$$x = \frac{\rho}{\phi + \delta + \rho}\left\{\overline{w}_H - b - 2\frac{C(1)[C(\varphi + \eta) - C(1)]}{1 + \hat{\hat{\alpha}}}(\hat{p} - \hat{R})\right\} \ . \tag{71}$$

The reservation wage for low-ability workers satisfies

$$R_L = b - x - \lambda \int_{R_H}^{\overline{w}_H} \max[V_{LL}(\hat{w}(w)), V_{LH}(w)] - U \; dF_H(w) \; . \tag{72}$$

The integral is

$$\lambda \left[\int_{R_L}^{p_L} V_{LL}(w) - U dF_L(w) + \eta[V_L(p_L) - U] + \int_{w_H^*}^{\overline{w}_H} V_{LH}(w) - V_{LH}(w_H^*) dF_H(w) + (1 - \varphi - \eta)[V_L(p_L) - U] \right]$$

$$= \int_{R_L}^{p_L} \frac{\lambda(1 - F_L(w))}{\phi + \delta + \rho + \lambda(1 - F_L(w))} dw + \frac{\lambda(1 - \varphi)x}{\rho} + \frac{(\phi + \delta)x}{\rho} + b - p_L$$

$$= (p_L - R_L)(1 - 2C(\varphi)) + \frac{2C(\varphi)^2}{1 + \hat{\alpha}}(\hat{p} - \hat{R}) - \frac{2C(\varphi)}{1 + \hat{\alpha}}(p_H - R_H - p_L + R_L)^{1/2}(p_H - p_L - x)^{1/2}$$

$$+ \frac{(\lambda(1 - \varphi) + \phi + \delta)x}{\rho} + b - p_L \; .$$

Here the last equation makes use of the same integral formula as in the proof of Proposition 2. Substitution of this expression and (71) into (72) gives the reservation-wage equation for R_L.

Regarding the reservation-wage equation for R_H, we can also make use of the integral formula in the proof of Proposition 2, which gives

$$b - R_H - \frac{\rho}{\phi + \delta}(p_H - b) = \lambda \int_{R_H}^{\overline{w}_H} V_{HH}(w) - V_{HH}(R_H) dF_H(w) = \tag{73}$$

$$\overline{w}_H - R_H - 2C(1)(p_H - R_H) + \frac{2C(1)C(\varphi)}{1 + \hat{\alpha}}(\hat{p} - \hat{R})$$

$$+ \frac{2C(1)\hat{\alpha}}{1 + \hat{\alpha}}(p_H - p_L - R_H + R_L)^{1/2}(p_H - p_L - x)^{1/2} + \frac{2C(1)[C(1) - C(\varphi + \eta)]}{1 + \hat{\alpha}}(\hat{p} - \hat{R}) \; .$$

The equilibrium variables $(\varphi, \xi, \eta, R_L, R_H)$ are determined by the five equations (67), (68), (69), (72) and (73), when $\eta + \varphi < 1$. Otherwise, equation $\eta + \varphi = 1$ replaces (69).