

Wage Dispersion and Labor Turnover with Adverse Selection

Carlos Carrillo-Tudela ^{*} Leo Kaas [†]

Abstract

This paper analyzes the effects of adverse selection on worker turnover, wage dispersion, and resource allocation in labor markets. We consider a model of on-the-job search where firms offer long-term wage contracts to workers of different ability. Firms do not observe worker ability upon hiring but learn it gradually over time. Provided that information frictions are sufficiently strong, low-wage firms offer screening contracts and hire all types of workers in equilibrium, whereas high-wage firms specialize in attracting high-ability workers only. This implies that low-ability workers have higher turnover rates, earn lower wages on average and face an earnings distribution with a higher frictional component. Furthermore, positive sorting obtains in equilibrium.

Keywords: Adverse Selection, On-the-job search, Wage Dispersion, Sorting

JEL: D82; J63; J64

February 15, 2011

PRELIMINARY AND INCOMPLETE

^{*}Department of Economics, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. Email: cocarr@essex.ac.uk

[†]Department of Economics, University of Konstanz, 78457 Konstanz, Germany. Email: leo.kaas@uni-konstanz.de

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 as an impediment for labor turnover.¹ Taken together these frictions can present formidable barriers for efficient resource allocation and have profound effects on the distribution of wages. Lockwood (1991), for example, suggests that adverse selection exacerbates the negative effects of search frictions by reducing workers' mobility across firms and the re-employment chances of unemployed workers. With almost no exceptions, however, current contributions on labor search with adverse selection² abstract from job-to-job transitions, although these account for a sizable part of worker flows, and one would expect that asymmetric information has non-trivial implications for worker turnover between heterogeneous firms. Similarly, there is no analysis of the effects of adverse selection on the distribution of wages and in particular on dispersion that is attributed to market frictions (see Mortensen (2003) and Hornstein, Krusell, and Violante (2010)).

In this paper we consider a frictional labor market where workers search on the job and firms post wage contracts. Upon hiring, firms cannot observe the ability of their applicants, but they learn worker ability with delay during the employment spell. Firms are allowed to offer self-selection contracts, so they are free to decide to attract all types of workers or to attract only a subset of them. Using this framework we study three questions. (i) What characterizes job-to-job transitions in an environment of adverse selection and search frictions? (ii) What is the resulting allocation of workers among firms? (iii) What is the impact on the wage distribution? We argue that the combination of on-the-job search and adverse selection can have profound effects on the allocation of resources and on the distribution of wages, particularly when information frictions are rather severe.

Our model is based on the equilibrium search model proposed by Burdett and Mortensen (1998). As this model provides an elegant theory of worker turnover and wage dispersion under

¹Search models of the labor market are surveyed in Rogerson, Shimer, and Wright (2005). For labor market implications of adverse selection, see e.g. Salop and Salop (1976), Greenwald (1986), Gibbons and Katz (1991). One conclusion of this work is that low-ability workers tend to have higher turnover rates.

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

perfect information about worker ability, it is the natural benchmark for our work. In deviation from this benchmark, information is asymmetrically distributed in our model: while workers are perfectly informed about their ability, firms learn worker ability slowly over time. In this setting, firms can decide either to post screening contracts and target all types of workers, or to specialize in employing one type of worker only. In either case firms commit to pay their posted wages for as long as the workers remain employed in the firm. Similar to earlier work on adverse selection in labor markets, we demonstrate that low ability workers have higher turnover rates. Nonetheless, the model gives rise to positive sorting. High wage and high productivity firms end up employing a larger share of high ability workers. The explanation is that high-wage firms compete more strongly for high-ability workers and find it too costly to offer incentive-compatible contracts for workers of lower ability. Hence, these firms specialize in employing workers of higher ability. Although they also attract low ability workers, these workers leave soon after the employer learns their type. In contrast, firms with lower wages and lower productivity prefer to screen workers and hence offer stable wage contracts to both types of workers. Hence, these firms end up employing a larger fraction of low ability workers.

This sorting pattern has important consequences. It implies that the economy's total output is greater when firms face search and (sufficiently large) informational frictions than, for example, when firms face the same search frictions but are perfectly informed upon worker ability upon hiring. When informational frictions are small enough, all firms offer screening contracts. Random search then implies all firms obtain the same proportion of high and low skill workers and no sorting occurs. Regardless of the magnitude of asymmetric information, however, we demonstrate that there are always some firms that offer screening contracts. Thus an equilibrium where all firms pay the same wage to their new hires cannot occur.

The equilibrium sorting allocation that arises is consistent with recent empirical evidence showing that labor markets are characterized by positive sorting among workers and firms, or among workers and coworkers within firms (see Lopes de Melo (2009) and Bagger and Lentz (2008)). It is also consistent with the empirical evidence that documents the firm-size/wage-premium relation that is widely observed in many labor markets. Our model implies that high wage firms are not only bigger, but they also employ a more productive workforce. The workforce of a high wage firm is more productive because this firm is able to attract and retain a larger proportion of high ability workers. The model is therefore consistent with evidence demonstrating the importance of firm *and* worker characteristics in accounting for the positive relation between wages and firm size (see e.g. Brown and Medoff (1989), Abowd, Kramarz, and Margolis (1999), Haltiwanger, Lane, and Spletzer (1999), Idson and Oi (1999)). The (still substantial) part left

unexplained by these characteristics in those studies is attributed in this paper, as in Mortensen (2003), to search frictions.

The cross sectional variation in wages implied by the model is determined by (i) dispersion in worker ability, (ii) dispersion in firm productivity and (iii) frictional wage dispersion (workers of the same ability are paid differently in equally productive firms). As opposed to many previous studies that analyze wage dispersion using equilibrium search models (see e.g. Postel-Vinay and Robin (2002), Burdett, Carrillo-Tudela, and Coles (2009), and Hornstein, Krusell, and Violante (2010)), here the frictional component of the wage distribution combines the informational frictions faced by employers and the search frictions faced by both workers and firms. We show that the amount of frictional wage dispersion crucially differs among worker types: Low ability workers face a more dispersed earnings distribution and have a larger frictional component.

The associated wage dynamics and turnover patterns also differ decisively between workers: whilst the earnings of low ability workers undergo persistent upward and downward mobility, those of high ability workers typically rise along the life cycle. This property, coupled with the fact that low ability workers change jobs more often, implies that high job turnover is associated with lower average wages as found in several empirical studies (see, e.g., Mincer and Jovanovic (1981) and Light and McGarry (1998)). The main difference here is that this relationship arises due to firms' optimal wage policies in the presence of adverse selection and search frictions rather than from lower levels of firm-specific human capital of high mobility workers (Farber (1999)).

The paper is organized as follows. After a brief review of related literature, we set out the basic framework in Section 2. We focus first on the case where all firms are homogeneous which helps to derive a full equilibrium characterization in the most transparent way. In Section 3 we characterize equilibria with limited information. Particularly, we show that all firms screen their applicants when the firms' learning rate is high enough; but when informational frictions are sufficiently severe, a fraction of high-wage firms specializes in attracting high-ability workers. Implications for the firm-size–wage relation, for individual wage dynamics and wage dispersion are illustrated using numerical examples in Section 4. Section 5 introduces firm heterogeneity, it extends the key theoretical results for this setting and studies the sorting implications. Section 6 concludes. All proof and tedious derivations are relegated to the Appendix.

1.2 Related Literature

Besides a few earlier contributions (Lockwood (1991), Albrecht and Vroman (1992), Montgomery (1999)), a number of recent papers study the interrelation between search frictions and adverse

selection. Guerrieri, Shimer, and Wright (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. 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 his 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.

This paper also relates to the literature that analyze resource allocation in markets with search frictions. In particular, Lentz (2010) constructs a model based on the framework developed by Postel-Vinay and Robin (2002) in which workers of different abilities have different search intensities. He shows that in equilibrium more able workers search harder and hence have a higher chance of being employed in more productive firms when the production function is supermodular. We also assume a supermodular production function, but keep the search technology as simple as possible to stress the role adverse selection has on firms wage policies and generating positive sorting. Both papers share an important feature: firms operate under constant returns to scale and have no capacity constraints in hiring workers. This implies that in both cases the sorting process is driven by workers' ability to search on the job. This is in contrast to partnership models of sorting where both sides of the market engage in active search (e.g. Shimer and Smith (2000)).

This paper also contributes to the emerging literature that analyzes the joint implications of search frictions and workers' productivity differences on wage dispersion and wage dynamics. Although most of this literature allows for human capital accumulation (see Burdett, Carrillo-Tudela, and Coles (2009), Bagger et al. (2010), Fu (2010)), it also assumes that, upon a meeting, a firm is able to perfectly observe the productivity of its applicants. In our paper, workers do not accumulate human capital while employed, but firms learn the productivity of their applicants on-the-job. Asymmetric information thus generates a new source of frictional wage dispersion that has not been explored when analyzing the fundamental contributions to wage inequality.

2 Basic Framework

Consider a continuous time economy that is in steady state. There is a unit mass of risk neutral workers and 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. Hence, the objective of any worker is to maximize total expected lifetime utility, and the objective of any firm is to maximize expected the steady state profit flow.

Workers differ in their ability. In particular, we assume there are two types of workers. A fraction α_H has high ability ε_H and a fraction $\alpha_L = 1 - \alpha_H$ has low ability ε_L . Firms operate under a constant returns to scale technology and, for the main part of this paper, they all have the same productivity p . We consider the implications of firm heterogeneity in Section 5. An employed worker with ability ε_i generates flow output $\varepsilon_i p$ for $i = L, H$.

Once a firm and a worker meet, the productivity of the firm is common knowledge. The ability of the worker, however, remains the worker's private information. To simplify the learning process of firms, we assume that firms monitor the output of a particular worker at rate ρ . This parameter describes the firm's learning rate. Further, we assume that the monitoring technology is such that 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 do not observe an applicant's employment status or any other aspect of the worker's employment history.

Firms post fixed wage contracts and commit not to counter any outside offer. Particularly, we rule out that firms offer back-loading contracts, which implicates that contracts cannot be contingent on employer learning.³ To screen workers, firms offer self-selection contracts (w_L, w_H) , where w_i is the constant flow wage paid to a worker who reports type $i = L, H$ upon hiring. To enforce truth telling, firms announce to punish workers who misreport their type by paying a wage that makes them indifferent between staying in the firm or quitting to unemployment. Such an announcement is the strongest form of punishment that avoids inefficient job separation. It is also ex-post optimal for the punishing firm to employ a misreporting worker at his reservation wage. Let $F_i(w_i)$ denote the proportion of firms offering a wage no greater than w_i to workers of

³Stevens (2004) and Burdett and Coles (2003) show that optimal wage-tenure contracts exhibit an increasing wage-tenure profile. This paper restricts attention to constant wage contracts which allows to consider, in a simpler environment, the interaction between adverse selection and on-the-job search.

ability i , for $i = L, H$. Further, let $\underline{w}_i$ and $\overline{w}_i$ denote the infimum and supremum of the support of $F_i(\cdot)$. It is useful to restrict the analysis to rank-preserving wage policies: firms that offer higher wages to high-ability workers also offer higher wages to low-ability workers. That is, we use a strictly increasing function $\hat{w}(\cdot)$ that links the two wages offered by any particular firm such that $w_L = \hat{w}(w_H)$, and hence $F_L(\hat{w}(w_H)) = F_H(w_H)$ for all wages $w_H \in [\underline{w}_H, \overline{w}_H]$.⁴

Unemployed and employed workers meet firms according to a Poisson process with parameter $\lambda > 0$. Once a meeting takes place, a worker observes the menu of wages posted by the firm and can choose one of them, but nothing restricts the worker from choosing the contract the firm offers to workers of a different ability level. If both wages 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 wage offer is strictly preferred. Further, a worker reports his true type when indifferent between misreporting and truth-telling.

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 $b < \varepsilon_L p$ per unit of time. For simplicity we do not allow that workers of different abilities obtain different payoffs when unemployed. For example, 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).

2.1 Worker Strategies

Fix a pair of wage-offer distributions F_H, F_L and an associated function $\hat{w}$. Let U_i denote the expected 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. Let $V_{ij}(w)$ denote the maximum expected value of employment for a worker with ability i employed at a firm offering w after reporting type j . The function $\hat{w}$ is helpful to characterize these value functions as we can think of any worker randomly meeting firms by drawing high-ability wage offers from F_H . A worker that meets a firm offering w_H observes both w_H and $w_L = \hat{w}(w_H)$. The worker then decides which contract to choose (if any) to maximize expected lifetime utility. Using this insight and standard dynamic programming

⁴The restriction to rank-preserving wage policies implicitly constrains the set of equilibria that are considered. As we see later, however, rank preservation arises naturally in situations with binding incentive constraints.

arguments, the Bellman equation that describes U_i is given by

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

Next consider an employed worker of type i that reported his true type and is earning a wage w_i . Similar arguments as before imply that $V_{ii}(w_i)$ solves the following Bellman equation

$$\phi V_{ii}(w_i) = w_i + \lambda \int_{\underline{w}_H}^{\bar{w}_H} \max [V_{iL}(\hat{w}(x)) - V_{ii}(w_i), V_{iH}(x) - V_{ii}(w_i), 0] dF_H(x) + \delta(U_i - V_{ii}(w_i)). \quad (1)$$

If this worker misreported his type and is earning w_j , however, the value of employment $V_{ij}(w_j)$ takes into account that the worker is set back to his reservation wage (and thus attains unemployment utility) at rate ρ ; hence V_{ij} solves

$$\phi V_{ij}(w_j) = w_j + \lambda \int_{\underline{w}_H}^{\bar{w}_H} \max [V_{iL}(\hat{w}(x)) - V_{ij}(w_j), V_{iH}(x) - V_{ij}(w_j), 0] dF_H(x) + (\delta + \rho)(U_i - V_{ij}(w_j)). \quad (2)$$

It is straightforward to verify that any worker's optimal search strategy is characterized by a reservation wage. Let $R_{ijk}(x)$ denote the reservation wage of a worker: (i) currently receiving flow payoff x , (ii) is of type $i = L, H$, (iii) has reported (in the case of an employed worker) type $j = L, H$ and (iv) when meeting a firm decides to report type $k = L, H$. Thus, $R_{ijk}(x)$ is defined by $V_{ij}(x) = V_{ik}(R)$. The above value functions then imply that an unemployed worker of type i accepts a wage offer w' if and only if $w' \geq R_{ik}(b) = b$ for all $i, k = L, H$.⁵ Further, the reservation wage of any unemployed worker is the same whether he reports his true type or not.

Now consider an employed worker of type i that reported his true type and is earning a wage w_i . Given contact with a firm and revealing his true type once again (i.e. $k = i$), (1) implies that this worker accepts employment if and only if the firm offers him a wage $w'_i > R_{iii}(w_i) = w_i$. If the worker decides to misreport his type (i.e. $k \neq i$), however, (1) and (2) imply that the worker accepts employment if and only if the firm offers him a wage $w'_k > R_{iik}(w_i) = w_i + \rho[V_{ii}(w_i) - U_i]$. In this case, the worker must be compensated by the expected loss of misreporting his type.

Now suppose that an employed worker of type i misreported his true type $j \neq i$ and is earning a wage w_j . Given contact with a firm and reporting his true type ($k = i$), (1) and (2)

⁵Assuming that unemployed and employed workers meet firms at the same rate considerably simplifies the worker's problem as now *all* unemployed workers have the same reservation wage which is independent of firms' wage offer strategies. Although empirical evidence suggests that unemployed workers meet firms at a faster rate than employed workers (see Jolivet, Postel-Vinay and Robin, 2006), this restriction allows us to analyze, in a simpler environment, how firms' optimal wage policies are affected by adverse selection.

imply that the worker accepts employment if and only if the firm offers a wage $w'_i > R_{iji}(w_j) = w_j - \rho[V_{ij}(w_j) - U_i]$. In this case, the worker voluntarily accepts a wage cut as the punishment risk disappears with truth-telling. 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 $w'_j > R_{ijj}(w_j) = w_j$.

Note that a worker will 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 (1) and (2), it follows that this is equivalent to

$$w'_j - w'_i \leq \rho[V_{ij}(w'_j) - U_i] . \quad (3)$$

Namely the flow gain from misreporting on the left side may not exceed the expected loss of punishment on the right side.⁶

2.2 Firms' Profits

Consider a firm offering any pair of wages w_H, w_L . Recall that this firm does not know the type of its applicants and, for example, the posted wage w_H might attract both type of workers, while w_L does not attract any worker (or vice versa). We denote the firm's steady-state profit as $\Omega(w_H, w_L) = \sum_{i=L,H} \Omega_i(w_H, w_L)$, where $\Omega_i(w_H, w_L)$ describes the firm's profit from hires of type $i = H, L$ at the offered wages. These functions are described in more detail below; they are equilibrium objects that depend upon workers' search strategies and the wage-offer distributions. The firm's objective is to choose a pair (w_H, w_L) to maximize $\Omega(w_H, w_L)$. Equilibrium requires that the optimal choices of w_i must be consistent with the offer distributions $F_i(w_i)$ and the associated function $\hat{w}$. We define $\bar{\Omega} = \max \Omega(w_H, w_L)$ and now turn to formally define an equilibrium.

2.3 Market Equilibrium

Definition: A Market Equilibrium is a tuple $\{\hat{w}, F_i(\cdot), \Omega, R_{ijk}(\cdot), V_{ij}\}$ for each $i, j, k = L, H$ such that

(i) Firms maximize profits, i.e. $\Omega(w_H, w_L) \leq \bar{\Omega}$ for all (w_H, w_L) , and

$$\Omega(w_H, w_L) = \bar{\Omega} \text{ and } F_L(w_L) = F_H(w_H) \quad \text{for all } w_H \in \text{supp } F_H \text{ and } w_L = \hat{w}(w_H) .$$

⁶Note that it also follows from (1) and (2) that (3) is equivalent to $w'_j - w'_i \leq \rho[V_{ii}(w'_i) - U_i]$.

- (ii) Workers' search strategies are described by reservation wages $R_{ijk}(\cdot)$ and value functions V_{ij} satisfying (1), (2) and (3).

Before we characterize equilibrium we make some preliminary points. First note that $\varepsilon_L p > b$ implies that offering $w_i = b$ strictly dominates offering $w_i < b$ as it generates strictly positive profit. Hence in any equilibrium firms offer a set of wages such that $\min\{w_L, w_H\} \geq b$, $\bar{\Omega} > 0$ and $\underline{w}_i \geq b$ for $i = L, H$.

It is also useful to consider the equilibrium outcomes in the limiting cases where there is no possibility of learning a worker's type, $\rho = 0$, and when, upon a meeting, firms perfectly observe the worker's type, $\rho = \infty$. These limiting cases have the same structure as Burdett and Mortensen (1998) and are useful benchmarks against which we compare our equilibrium allocations.

2.3.1 Perfect Information

When $\rho = \infty$, firms are able to perfectly screen their applicants. As in Carrillo-Tudela (2009), this implies that firms segment their markets and choose w_L and w_H independently, each to maximize the corresponding steady state profit

$$\Omega_i(w_i) = \frac{\lambda(\phi + \delta)(\varepsilon_i p - w_i)\alpha_i}{[\phi + \delta + \lambda(1 - F_i(w_i))]^2}, \quad (4)$$

where $\Omega(w_L, w_H) = \Omega_L(w_L) + \Omega_H(w_H)$.⁷ Workers' reservation wage strategies are such that unemployed workers accept any wage above b and employed workers of type i earning a wage w_i accept any wage $w'_i > w_i$.

In this case, the offer distribution for each worker type is given by

$$F_i(w_i) = \left(\frac{\phi + \delta + \lambda}{\lambda} \right) \left[1 - \left(\frac{\varepsilon_i p - w_i}{\varepsilon_i p - \underline{w}_i} \right)^{1/2} \right]. \quad (5)$$

In this equilibrium $\underline{w}_i = b$ and $\bar{w}_i = \varepsilon_i p - [(\phi + \delta)/(\phi + \delta + \lambda)]^2(\varepsilon_i p - b)$ for $i = L, H$.

It is easy to verify that $\varepsilon_H > \varepsilon_L$ implies that $F_H(\cdot)$ first order stochastically dominates $F_L(\cdot)$, and in equilibrium more able workers are paid, on average, higher wages than low ability

⁷Each wage w_i attracts the correct worker type and hence the associated hiring flows are $h_i(w_i) = \lambda[u_i + G_i(w_i)(\alpha_i - u_i)]$, where u_i denotes steady state unemployment and $G_i(\cdot)$ is the steady state earnings distribution of type i workers. These measures follow from steady state turnover and are described in (12) and (13) below. A job filled with a worker of type i has value $J_i(w_i) = (\varepsilon_i p - w_i)/[\phi + \delta + \lambda(1 - F_i(w_i))]$. Then $\Omega_i(w_i) = h_i(w_i)J_i(w_i)$.

workers. At the level of an individual firm, however, low ability employees could potentially receive higher wages than their more able peers; i.e. $w_L > w_H$, which is a consequence of the constant profit condition. A firm, in equilibrium, is indifferent between posting any wage in the interval $w_i \in [\underline{w}_i, \bar{w}_i]$ for a given $i = L, H$. Our restriction on rank-preserving wage policies rules out these possibilities, however. That is, rank preservation implies that wage offers are linked according to

$$w_L = \hat{w}(w_H) = b + \left[\frac{\varepsilon_L p - b}{\varepsilon_H p - b} \right] (w_H - b) . \quad (6)$$

2.3.2 No Information

In the opposite scenario of no information, firms treat all worker as having the same average ability $\tilde{\varepsilon} = \varepsilon_H \alpha_H + \varepsilon_L (1 - \alpha_H)$. A firm cannot screen workers and offers the same wage w to any worker it meets. It follows that $\hat{w}$ is uniquely determined by $w_L = \hat{w}(w_H) = w_H = w$. The steady state profit of a firm is then given by⁸

$$\Omega(w) = \frac{\lambda(\phi + \delta)(\tilde{\varepsilon}p - w)}{[\phi + \delta + \lambda(1 - F(w))]^2}.$$

Workers' reservation wage strategies are such that unemployed workers accepts any wage above b and employed workers earning a wage w accept any wage $w' > w$.

Burdett and Mortensen (1998) establish that in this case there exists a unique equilibrium in which firms differentiate their wage policies such that

$$F(w) = \left(\frac{\phi + \delta + \lambda}{\lambda} \right) \left[1 - \left(\frac{\tilde{\varepsilon}p - w}{\tilde{\varepsilon}p - \underline{w}} \right)^{1/2} \right].$$

Similar to the perfect information case $\underline{w} = b$ and $\bar{w} = \tilde{\varepsilon}p - [(\phi + \delta)/(\phi + \delta + \lambda)]^2(\tilde{\varepsilon}p - b)$. Compared to the perfect-information case, low ability workers are paid on average higher wages, while high ability workers are paid lower wages on average.

3 Equilibria with Limited Information

The limiting cases explored above generate a positive relation between firm size and pay policies. This relationship is purely due to search frictions and arises from the Burdett and Mortensen

⁸A firm offering wage w hires a flow of $h(w) = \lambda[u + G(w)(1 - u)]$ workers, where u and G are the (unconditional) steady state unemployment and earnings distribution (again similar to (12) and (13) below). An employed worker generates profit value $J(w) = [\tilde{\varepsilon}p - w]/[\Phi + \delta + \lambda(1 - F(w))]$. It follows that $\Omega(w) = h(w)J(w)$.

(1998) framework. However, the ability composition in the workforce of high paying and bigger firms is the same as that of low paying and smaller firms. This feature seems to contradict the empirical evidence on the wage - firm size premium. Brown and Medoff (1989) and Idson and Oi (1999), for example, document that firm size is positively correlated with both wages and workforce quality.

In what follows we explore under what conditions there exists a positive relation between workers' productivity, firm size and pay policies, when (positive and finite) search and learning frictions coexist in the labor market. We show that when the learning rate of firms is sufficiently high, all firms offer screening contracts and have the same workforce productivity. For lower values of ρ , however, this type of equilibrium cannot be sustained as low ability workers have an incentive to misreport their type. In particular, we explore equilibria in which high wage firms offer wages to attract high ability workers only, while low wage firms offer screening contracts. We use the latter to analyze the relationship between firm size, wage policies and the ability composition of the workforce.

3.1 Non-binding Incentive Constraints

We start by showing that the perfect information equilibrium described in 2.3.1 can be sustained with limited information, provided that the learning rate ρ is sufficiently high. Consider such an equilibrium with wage offer distributions (5) and function $\hat{w}$ as in (6). Clearly, only low ability workers might have an incentive to misreport their type when firms cannot learn the worker type immediately. Indeed, the next result shows that low ability workers will not misreport their type if and only if firms learn the worker's type sufficiently fast.

Proposition 1: *The perfect information equilibrium where firms' wage offers are draws from distributions (5) is an equilibrium in the imperfect information economy if and only if, $\rho \geq \rho_1$ where*

$$\rho_1 \equiv (\phi + \delta + \lambda) \frac{(\varepsilon_H - \varepsilon_L)p}{\varepsilon_L p - b}.$$

It is intuitive that not only fast learning, but also small values of ϕ , δ , λ and $(\varepsilon_H - \varepsilon_L)$ are conducive to prevent misreporting: a small ability gap leads to small wage differentials and thus smaller gains from misreporting. A low layoff rate $(\phi + \delta)$ or a low job offer arrival rate reduces the incentive to misreport since workers cannot expect to escape punishment easily.⁹

⁹Proposition 1 does not extend to an environment in which high ability workers have a higher reservation wage. A differential in reservation wages would emerge if, for example, unemployment income or job arrival rates

Now consider values of $\rho < \rho_1$. The next result shows that there is another threshold $\rho_2 < \rho_1$, defined in (9) below, such that, for $\rho \in (\rho_2, \rho_1)$, incentive constraints bind for a fraction of firms but are slack for the remaining firms, and that all firms offer screening contracts. We fully characterize this type of equilibrium in the Appendix.

Proposition 2: *For values of $\rho \in [\rho_2, \rho_1)$, an equilibrium with wage dispersion exists in which all firms offer screening contracts. Incentive constraints bind for a fraction of (low wage) firms and they are slack for the remaining fraction of (high wage) firms if $\rho > \rho_2$.*

In the case $\rho = \rho_2$, where incentive constraints bind on all firms, one can calculate the wage offer distribution explicitly. Workers of ability $i = L, H$ earn wages $w_i \in [b, \bar{w}_i]$, with

$$\bar{w}_L = b + \frac{2(\tilde{\varepsilon}p - b)}{(\phi + \delta + \lambda)^2} \left\{ (\alpha_H \rho)^2 \log \left[\frac{\phi + \delta + \lambda + \alpha_H \rho}{\phi + \delta + \alpha_H \rho} \right] + \frac{1}{2} [(\phi + \delta + \lambda)^2 - (\phi + \delta)^2] - \alpha_H \lambda \rho \right\}, \quad (7)$$

$$\bar{w}_H = \frac{1}{\alpha_H} \left\{ \tilde{\varepsilon}p - \alpha_L \bar{w}_L - \frac{(\phi + \delta)^2}{(\phi + \delta + \lambda)^2} (\tilde{\varepsilon}p - b) \right\}. \quad (8)$$

To verify whether the incentive constraint indeed binds for all firms, we need to ensure that no firm has an incentive to reduce the wage for high ability workers while offering the same wage to low-ability workers (and hence continuing to screen workers at non-binding incentive constraints). In the proof of Proposition 2, we show that this is true if, and only if,

$$\varepsilon_{LP} - \bar{w}_L \leq (\varepsilon_{HP} - \bar{w}_H) \frac{\phi + \delta}{\phi + \delta + \rho}. \quad (9)$$

Intuitively, if the profit margin for high ability workers is large relative to the profit margin for low ability workers, firms have no incentive to reduce w_H (or to increase w_L), and hence incentive constraints must bind. The binding condition (9) is important as it implicitly defines the threshold value for parameter ρ_2 beyond which incentive constraints are slack for a fraction of firms.

If ρ is smaller than ρ_2 , incentive constraints continue to bind for all firms that offer screening contracts. We now turn to characterize equilibria for values of $\rho < \rho_2$ and derive conditions for existence. We show there exist equilibria in which all firms offer screening contracts and equilibria

differ between worker types. In this case, the full information outcome has wage offer distributions F_i whose lower bounds are the corresponding reservation wages. But then for any finite ρ (no matter how large), low ability workers employed at their reservation wage (or any wage close to to it) would misreport their type. They could earn the reservation wage of high ability worker temporarily, but they would not suffer from punishment which pays the same wage as under truth-telling. Hence any equilibrium with finite ρ must involve binding incentive constraints at low wages.

in which a fraction of high-wage firms decide not to screen workers. Instead, high wage firms specialize in employing high ability workers, whilst low wage firms employ all workers in different contracts. Intuitively, high wage firms desire to compete stronger for high ability workers and refuse to offer the associated incentive-compatible wage offers for low-ability workers. The latter contracts are too costly relative to the alternative of employing low-ability workers temporarily at high wages and punishing them after learning their true type.

3.2 Binding Incentive Constraints

Suppose that there is a fraction $\eta \leq 1$ of screening firms, where η is endogenously determined below. A screening firm offers “low wages” to both type of workers and screens its applicants by offering jobs to high ability workers at a wage $w_H \in [b, \tilde{w}_H]$ and to low ability workers at a wage $w_L = \hat{w}(w_H) \leq w_H$, with $\tilde{w}_L = \hat{w}(\tilde{w}_H)$. The function $\hat{w}$ here traces the binding incentive constraint; hence it follows from (3) that

$$w_L = \hat{w}(w_H) = w_H - \rho[V_{LH}(w_H) - U_L] , \text{ for all } w_H \in [b, \tilde{w}_H] . \quad (10)$$

That is, $\hat{w}(w_H)$ is the lowest wage that can be offered to low-ability workers so as to induce them not to misreport their type. It follows from the proof of Proposition 2 that screening firms indeed face binding incentive constraints when $\rho \leq \rho_2$.

The remaining fraction $1 - \eta$ of firms do not screen their applicants, but specialize in attracting high ability workers. To do so, a non-screening firm offers a “higher wage” $w_H > \tilde{w}_H$ only to high ability workers, while it offers workers of low ability a sufficiently low wage. Given the rank preservation property, we impose that these firms offer $w_L = \hat{w}(\tilde{w}_H)$ to any low-ability worker with a strictly increasing and continuous function $\hat{w}$ which is, however, below the incentive-compatible wage schedule that enforces truth-telling.¹⁰ That is, the incentive constraint is violated:

$$w_L = \hat{w}(w_H) < w_H - \rho[V_{LH}(w_H) - U_L] , \text{ for all } w_H \in [\tilde{w}_H, \bar{w}_H] . \quad (11)$$

Then, low-ability workers misreport their type when contacted by such a firm; they earn w_H until the firm learns their true type and pays them the reservation wage.

¹⁰It follows from differentiation of (10) that the right-hand side in (10) is strictly increasing (see also (18) below). Hence, a strictly increasing function $\hat{w}$ satisfying (11) exists. The exact shape of this function is clearly irrelevant for equilibrium because all low-ability workers misreport their type when contacted by a non-screening firm.

The above equilibrium structure has the following implications. First, as the reservation wage of all unemployed workers is given by b , they again accept any job offered. An employed worker of high ability always reports his true type, and hence accepts a job if it offers a wage w_H strictly above the one he is currently earning. Similarly, if a low ability worker employed in a screening firm earning $w_L \leq \tilde{w}_L$ meets another screening firm, he accepts the job offer if he is promised a wage $w'_L > w_L$. If this worker meets a non-screening firm offering (w'_H, w'_L) with $w'_H > \tilde{w}_H$, the worker will also accept the offer. This follows directly from the fact that accepting w'_H strictly dominates accepting $w'_L > \tilde{w}_L$ (because the incentive constraint fails), which strictly dominates the current job $w_L \leq \tilde{w}_L$ (hence $w_H > R_{LLH}(w_L)$). Lastly, consider a low-ability worker that is earning w_H in a non-screening firm. This worker then accepts any wage $w'_H > w_H$ from other non-screening firms. If this worker meets any screening firm offering $w'_L \leq \tilde{w}_L$, it again follows that the worker will not accept such an offer:¹¹ The current job strictly dominates truth-telling at the current employer (since the incentive constraint fails at non-screening firms) which itself strictly dominates the outside offer $w'_L \leq \tilde{w}_L < w_L$.

Note that the same arguments as in Burdett and Mortensen (1998) guarantee that any equilibrium wage offer distribution F_H is continuous and has connected support. In turn, F_L does not exhibit any mass points either and it also has connected support. Note, however, that the earnings distribution of low ability workers will have a mass point at the reservation wage when $\eta < 1$, which are those low-ability workers in non-screening firms that have been caught lying about their type. Furthermore, no low ability worker is employed at wages $w \in (\tilde{w}_L, \tilde{w}_H]$, and hence the earnings distribution of low ability workers has no connected support if $\eta < 1$.

3.2.1 Steady state measures

To simplify notation, it is useful to let the quit rate of high-ability workers earning x be denoted by $q(x) \equiv \phi + \delta + \lambda(1 - F_H(x))$. Given the reservation wage strategies described above, steady state turnover implies that unemployment of workers of type i is given by

$$u_i = \frac{(\phi + \delta)\alpha_i}{\phi + \delta + \lambda} , \quad (12)$$

and the proportion of employed workers of high type earning a wage $w'_H \leq w_H$ is given by

$$G_H(w_H) = \frac{(\phi + \delta)F_H(w_H)}{q(w_H)} . \quad (13)$$

¹¹Once a low ability worker is caught misreporting his type, he receives b and has the same reservation strategy as a low ability worker employed in a screening firm.

To construct the earnings distribution for low ability workers, first consider those workers earning wages no greater than $w_L \in [b, \tilde{w}_L]$. The flow into this category is given by $\lambda F_H(\hat{w}^{-1}(w_L))u_L + \rho[1 - G_L(\tilde{w}_L)](\alpha_L - u_L)$. The first term gives the number of unemployed workers that meet screening firms offering wages no greater than $w_L = \hat{w}(w_H)$; the second term gives the number of low ability workers employed in non-screening firms that earn wage b after the firm learned their true type. The outflow, on the other hand, is given by those workers that left the market, got displaced or found a better paying job, $[\phi + \delta + \lambda(1 - F_H(\hat{w}^{-1}(w_L)))]G_L(w_L)(\alpha_L - u_L)$. Steady state turnover and $F_H(\tilde{w}_H) = \eta$ then imply

$$G_L(w_L) = \frac{1}{q(\hat{w}^{-1}(w_L))} \left[(\phi + \delta)F_H(\hat{w}^{-1}(w_L)) + \frac{\rho(\phi + \delta + \lambda)(1 - \eta)}{\phi + \delta + \rho + \lambda(1 - \eta)} \right], \quad (14)$$

for all $w_L \in [b, \tilde{w}_L]$. Note that for $\eta < 1$, G_L exhibits a positive mass at $w_L = b$, which reflects those low ability workers who are punished after misreporting their type at non-screening firms.

Next consider the proportion of low ability workers employed at non-screening firms at wages $w_H \in [\tilde{w}_H, w_H]$, before the firm learns their type. Since any low ability worker will misreport his type when offered a wage $w'_H \in [\tilde{w}_H, w_H]$, the flow of workers into this category is given by $\lambda[u_L + G_L(\tilde{w}_L)(\alpha_L - u_L)][F_H(w_H) - F_H(\tilde{w}_H)]$. The worker flow out of this category is $[\phi + \delta + \rho + \lambda(1 - F_H(w_H))][G_L(w_H) - G_L(\tilde{w}_H)](\alpha_L - u_L)$. That is, low ability workers stop earning a wage $w \in [\tilde{w}_H, w_H]$ because they leave the market, get displaced, their employer learns their type or because they meet another non-screening firm offering a higher wage and they misreport their type once again. Noting that $G_L(\tilde{w}_L) = G_L(\tilde{w}_H)$, steady state turnover implies that the proportion of low ability workers earning a wage no greater than $w \in [\tilde{w}_H, \bar{w}_H]$ is

$$G_L(w_H) = \frac{(\phi + \delta)F_H(w_H) + \rho}{q(w_H) + \rho}. \quad (15)$$

3.2.2 Firms' payoffs

Consider a screening firm that offers $w_H \leq \tilde{w}_H$ to high ability workers and $w_L = \hat{w}(w_H) \leq \tilde{w}_L$ to low ability workers. Given that job applicants correctly report their type when meeting this firm, the hiring flows by posting w_H and w_L are $h_i(w_i) = \lambda u_i + \lambda G_i(w_i)(\alpha_i - u_i)$, for $i = L, H$. Using (12), (13) and (14), this firm's steady state profit is given by¹²

$$\Omega^S(w_H, w_L) = \frac{\lambda\theta(\eta)\alpha_L(\varepsilon_{LP} - w_L) + \lambda(\phi + \delta)\alpha_H(\varepsilon_{HP} - w_H)}{q(w_H)^2}. \quad (16)$$

where $\theta(\eta) \equiv [(\phi + \delta + \rho)(\phi + \delta + \lambda(1 - \eta))]/[\phi + \delta + \rho + \lambda(1 - \eta)]$.

¹²Appendix C contains a full derivation of the expressions in (16) and (17).

Now consider a non-screening firm that offers $w_H > \tilde{w}_H$ to high ability workers and $w_L = \hat{w}(w_H)$ to low ability workers. Since low ability workers will misreport their type when meeting this firm, the hiring flows associated with posting w_H equals $h_i(w_H) = \lambda u_i + \lambda G_i(w_H)(\alpha_i - u_i)$, for $i = L, H$. Also note that no worker accepts the wage w_L offered to low ability workers because of (11). Using (12), (13) and (15), implies that this firm's steady state profit is given by

$$\Omega^{NS}(w_H, w_L) = \frac{\lambda(\phi + \delta + \rho)\alpha_L[\varepsilon_{LP} - w_H + \rho J_L(b)]}{(q(w_H) + \rho)^2} + \frac{\lambda(\phi + \delta)\alpha_H(\varepsilon_{HP} - w_H)}{q(w_H)^2}, \quad (17)$$

where $J_L(b) \equiv (\varepsilon_{LP} - b)/(\phi + \delta + \lambda)$ is the value of a job filled by a low-ability worker earning b .

3.2.3 Wage-Offer Distributions

To solve for the equilibrium wage offer distributions, first consider a screening firm that offers a menu of wages $\{w_H, w_L\}$ such that low ability workers do not misreport their type, i.e. the binding incentive constraint (10) holds. Differentiation of (2) and (10) implies that $\hat{w}$ is described by the first-order differential equation

$$\hat{w}'(w_H) = \frac{\phi + \delta + \lambda(1 - F_H(w_H))}{\phi + \delta + \rho + \lambda(1 - F_H(w_H))}, \quad (18)$$

subject to $\hat{w}(b) = b$.

Further, in any equilibrium a screening firm offering w_H and $w_L = w(w_H)$ must be indifferent between this contract and the reservation wage contract such that $\Omega^S(w_H, \hat{w}(w_H)) = \Omega^S(b, b)$. Differentiation of this equation together with (18) gives the following result.

Lemma 1: *Given $\eta \leq 1$, the wage offer distribution $F_H(\cdot)$ solves the first-order differential equation*

$$F_H'(w_H) = \frac{(\phi + \delta + \lambda)^2}{2\lambda[(\phi + \delta)\alpha_H(\varepsilon_{HP} - b) + \theta(\eta)\alpha_L(\varepsilon_{LP} - b)]} \left[\frac{\rho(\phi + \delta)\alpha_H + [(\phi + \delta)\alpha_H + \theta(\eta)\alpha_L]q(w_H)}{q(w_H)[q(w_H) + \rho]} \right]$$

for all $w \in [b, \tilde{w}_H]$, subject to $F_H(b) = 0$.

The highest wage offered by screening firms $\tilde{w}_H$ is determined by $F_H(\tilde{w}_H) = \eta$, for any given $\eta \leq 1$. The corresponding wage $\tilde{w}_L = \hat{w}(\tilde{w}_H)$ then follows from integration of (18). Denote these solutions $\tilde{w}_H(\eta)$ and $\tilde{w}_L(\eta) = \hat{w}(\tilde{w}_H(\eta))$, respectively. In Appendix C we provide a closed-form solution for $\tilde{w}_L$.

Next consider non-screening firms offering a wage $w_H > \tilde{w}_H$. Equilibrium requires that the profits of any non-screening firm must satisfy $\Omega^{NS}(w_H, w_L) = \Omega^S(b, b)$. Substituting out the

corresponding expressions leads to the following characterization of the wage offer distribution at non-screening firms.

Lemma 2: *Given $\eta < 1$, the wage offer distribution $F_H(\cdot)$ solves the first-order differential equation*

$$F'_H(w_H) = \frac{q(w_H)(q(w_H) + \rho)[(\phi + \delta)\alpha_H(q(w_H) + \rho)^2 + (\phi + \delta + \rho)\alpha_L q(w_H)^2]}{2\lambda[(\phi + \delta)\alpha_H(\varepsilon_{Hp} - w_H)(q(w_H) + \rho)^3 + (\phi + \delta + \rho)\alpha_L(\varepsilon_{Lp} - w_H + \rho J_L(b))q(w_H)^3]}$$

for all $w \in (\tilde{w}_H, \bar{w}_H]$, subject to $F_H(\tilde{w}_H) = \eta$.

Similar to Lemma 1, we require $F_H(\bar{w}_H) = 1$, which then characterizes the upper bound $\bar{w}_H$. Let the solution to this upper bound be denoted $\bar{w}_H(\eta)$.

The distribution of wage offers for low ability workers follows directly from $F_L(w_L) = F_H(\hat{w}^{-1}(w_L))$ for $w_L \in [b, \bar{w}_L]$ with $F_L(\bar{w}_L) = 1$. Hence, the above characterizes the equilibrium solutions for $F_H(\cdot; \eta)$, $F_L(\cdot; \eta)$, $\tilde{w}_H(\eta)$, $\tilde{w}_L(\eta)$ and $\bar{w}_H(\eta)$, for a given $\eta \leq 1$.

3.2.4 Characterization and Existence

The final step is to derive the equilibrium fraction of screening firms, η^* . The next result shows that any equilibrium must have a positive measure of firms offering screening contracts. This follows since the lowest paying firms always find profitable deviations to incentive-compatible screening contracts if all firms were to offer non-screening contracts. For these firms, the screening contract to low-ability workers is more attractive compared to the alternative of offering a non-screening contracts.

Proposition 3: *In any Market Equilibrium $\eta > 0$.*

Given that equilibrium requires $\eta^* > 0$, there are two possible cases: (i) Equilibria in which $\eta^* \in (0, 1)$ and screening and non-screening firms coexist. (ii) Equilibria in which $\eta^* = 1$ and all firms offer screening contracts. We analyze each in turn. When screening firms co-exist with non-screening firms, firms must be indifferent between the two types of contract. In particular, at the threshold wage $\tilde{w}_H$, this necessitates $\Omega^S(\tilde{w}_H(\eta), \tilde{w}_L(\eta)) = \Omega^{NS}(\tilde{w}_H(\eta), b)$. Using (16) and (17), this condition implies that $\eta^* \in (0, 1)$ must solve the following fixed point problem

$$\eta = T(\eta) \equiv \frac{(\phi + \delta + \lambda)}{\lambda(\varepsilon_{Lp} - b)} \left[(\tilde{w}_L(\eta) - b) - [\tilde{w}_H(\eta) - \tilde{w}_L(\eta)] \left(\frac{\phi + \delta + \lambda(1 - \eta)}{\rho} \right) \right], \quad (19)$$

where $\tilde{w}_H(\eta)$ follows from Lemma 1 with $F_H(\tilde{w}_H(\eta)) = \eta$ and $\tilde{w}_L(\eta)$ from (18). In the case in which all firms offer screening contracts, however, equilibrium requires that $\Omega^S(\tilde{w}_H(\eta), \tilde{w}_L(\eta)) \geq$

$\Omega^{NS}(\tilde{w}_H(\eta), \tilde{w}_L(\eta))$ at $\eta = 1$. With T defined in (19), this is equivalent to $T(1) \leq 1$ being a necessary condition for existence of a pure screening equilibrium. The proof of Theorem 1 below shows that the function T is an increasing and convex function of η with $T(0) = T'(0) = 0$. This implies that, given $T(1) > 1$, a unique equilibrium exists in which $\eta^* \in (0, 1)$. Otherwise, a unique equilibrium exists in which $\eta^* = 1$.

The condition for $T(1) > 1$ that guarantees an equilibrium in which some (but not all) firms offer screening contracts can be related to the value of ρ . In particular, there exists a threshold value ρ_3 such that for $\rho < \rho_3$ any equilibrium implies $\eta^* \in (0, 1)$, while for $\rho \in [\rho_3, \rho_2]$ any equilibrium has $\eta^* = 1$. In the proof of Theorem 1 we show that ρ_3 is the implicit solution of equation

$$\frac{\lambda\rho(\varepsilon_L p - b)}{\phi + \delta + \lambda} = \rho[\bar{w}_L - b] - (\phi + \delta)[\bar{w}_H - \bar{w}_L] , \quad (20)$$

with $\bar{w}_L$ and $\bar{w}_H$ defined by (7) and (8).

Theorem 1: *A Market Equilibrium with screening firms, $\eta^* > 0$, exists and is unique. Moreover, there is a threshold value $\rho_3 \in (0, \rho_2)$ such that non-screening firms co-exist with screening firms if $\rho < \rho_3$.*

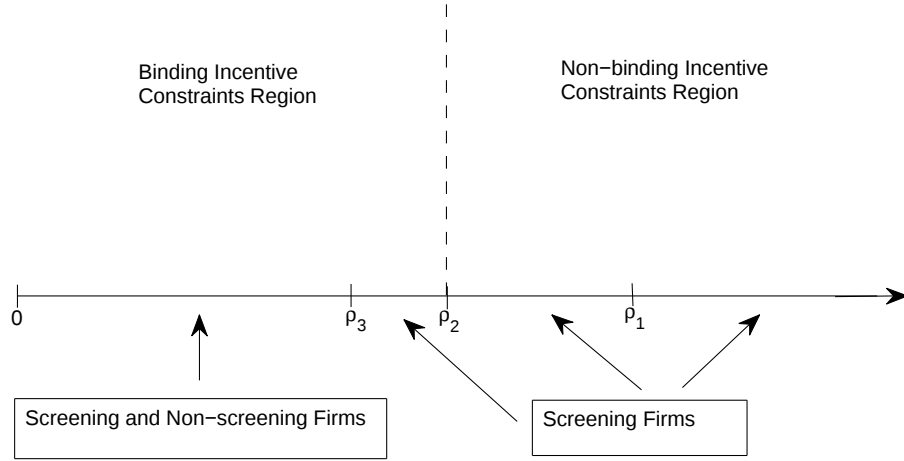


Figure 1: Set of Equilibria Parameterized by ρ

Note that Proposition 1, 2 and Theorem 1 together imply that the set of equilibria can be partitioned in terms of the firms' learning rate. Figure 1 depicts this partition. For all values

of $\rho \geq \rho_3$ firms offer screening contracts. There exists a positive relation between firm size and (average) wages paid that arises due to search frictions. However, all firms have the same workforce productivity. For values of $\rho < \rho_3$, those firms who offer the highest wages decide to offer non-screening contracts. These firms want to compete strongly for high ability workers and find it too costly to provide an appropriate incentive-compatible contract for low-ability workers. This induces low ability workers to cheat about their type, which entails higher separation rates for these workers. By implication, low-ability workers have a higher degree of turnover and they are under-represented in high-wage firms and over-represented in low-wage firms. In turn, this turnover pattern generates a positive relation between workers' types and firm size and wage policies. We summarize these findings as follows.

Corollary: *If $\rho \geq \rho_3$, all workers have the same turnover patterns, and all firms have the same ability composition of the workforce. If $\rho < \rho_3$, low-ability workers have higher turnover rates. High-wage firms have a higher workforce quality than low-wage firms.*

4 Implications

Given that the model cannot be fully characterized in closed form, we illustrate the implications about labor turnover and wage dispersion using a numerical example.

Consider the following parametrisation. Set the time period to a month and let $\phi = 0.0018$ to reflect an average working life of 45 years. Following Hornstein, Krusell, and Violante (2010), let $\delta = 0.036$ and $\lambda = 0.13$ to roughly match the average separation and job to job transition rates in the US economy. We choose $\varepsilon_L = 1$ and $\varepsilon_H = 2$ arbitrarily and let $\alpha_H = \alpha_L = 0.5$. We normalise $p = 1$ and set $b = p(\alpha_H \varepsilon_H + \alpha_L \varepsilon_L)/2 = 0.75$; this choice implies that unemployment income is at roughly 65% of the average wage. We set $\rho = 0.08$ as a benchmark. This number implies that on average firms learn their employees' true type after one year of employment.

Figure 2 depicts (19) and shows that in equilibrium 62.4 percent of firms offer screening contracts (i.e. $\eta^* = 0.624$). Figure 3 shows the associated earnings distribution of high ability and low ability workers, G_H and G_L . Screening firms compress the wages of high ability workers to offer sufficiently high wages and enforce truth-telling from low ability workers. Some screening firms even offer wages above the productivity of low ability workers, with $\tilde{w}_L(\eta^*) = 1.124$ and $\tilde{w}_H(\eta^*) = 1.357$.¹³ The wage policies of screening firms also affect the wages offered to high ability workers by non-screening firms. The set of wages offered to these workers is lower than in

¹³Recall, however, that in equilibrium all firms make strictly positive steady state profits, $\bar{\Omega} = 0.144$.

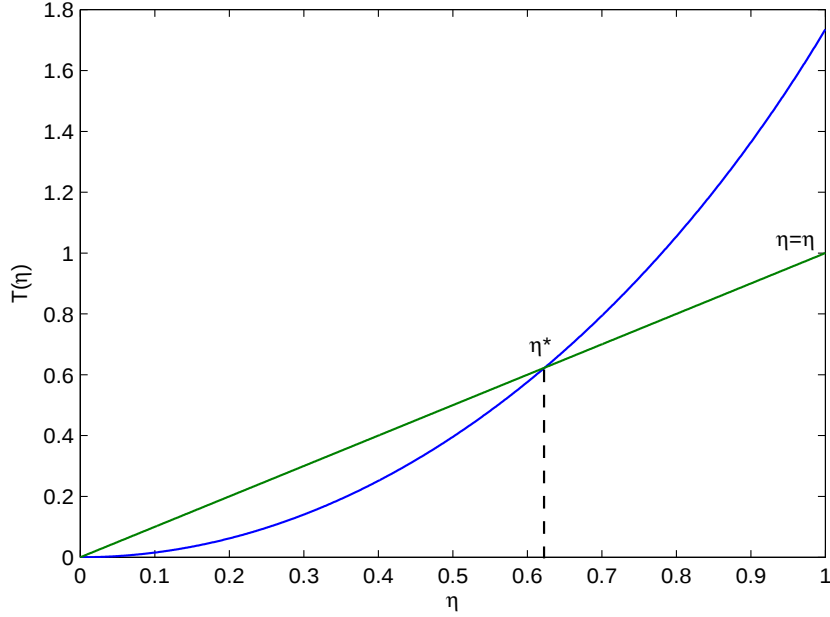


Figure 2: Existence of Equilibrium

the perfect information case with the highest wage paid to these workers being $\bar{w}_H(\eta^*) = 1.723$, while with perfect information $\bar{w}_H = 1.937$. Overall the earnings distribution faced by low ability workers is more dispersed than the one faced by high ability workers, implying that the former type of workers face more frictional wage dispersion. The benchmark parametrisation shows that, compared with the perfect information case, the interaction of search and learning frictions imply that average pay and wage dispersion for high ability workers fall, while average pay and wage dispersion rise for low ability workers.

The above parametrisation also implies that $\rho_1 = 0.671$, $\rho_2 = 0.431$ and $\rho_3 = 0.32$. Pure screening equilibria can only be sustained when firms learn the true type of their workers on average at the third month, $1/\rho_3$, of employment. Given that the latter number seems to require very fast learning from employers, the benchmark parametrisation suggests it is more reasonable to expect some firms not to screen their applicants, but target high ability workers. As we increase the learning rate of firms and move towards the set of pure screening equilibria, $\tilde{w}_L(\eta)$ decreases and $\tilde{w}_H(\eta)$ and $\bar{w}_H(\eta)$ increase. The mean and variance of G_H increases, while they decrease for G_L ; both moving in the direction of their perfect information values. Conversely, as $\rho \rightarrow 0$ the equilibrium converges to the no information case described in section 2.3.2. All firms offer an identical wage to both worker types such that $w_L = w_H = w$ for all $w \in [b, \bar{w}]$, where

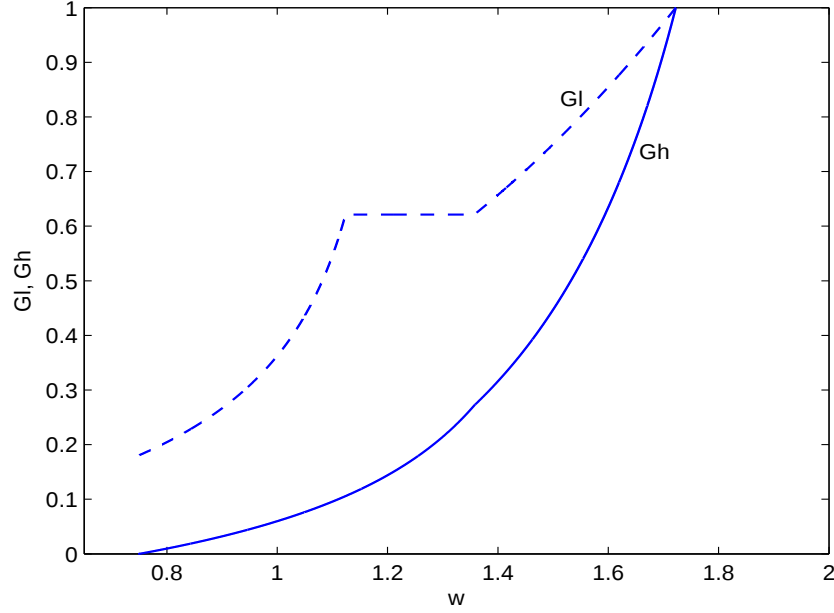


Figure 3: Earnings Distributions

$\bar{w} = \tilde{\varepsilon}p - [(\phi + \delta)/(\phi + \delta + \lambda)]^2(\tilde{\varepsilon}p - b) = 1.462$. In this equilibrium the firm never learns the worker true type and hence treats all workers as having the average productivity, $\tilde{\varepsilon}p$.¹⁴

An important implication of our model is that low ability workers have a higher degree of turnover than high ability workers. Such a result also obtains in other settings (as, for example, in Greenwald (1986)) where firms are restricted to offer the same contract to all workers. Here, in contrast, firms are free to offer screening contracts; nonetheless, a fraction of high-wage firms decides not to screen workers, and the higher turnover rate of low ability workers arises due to the punishment strategy of these firms. In turn, this turnover pattern generates a positive relation between workers' types and firms' size and wage strategies. First consider the relation between workers' type and firms' size. Using (31) and (32) in the Appendix, the workforce size of the low ability and high ability workers employed at a screening firm offering $w_H \leq \tilde{w}_H(\eta)$ are given by

$$n_L^s(\hat{w}(w_H)) = \frac{\lambda\theta(\eta)\alpha_L}{q(w_H)^2} \text{ and } n_H^s(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{q(w_H)^2}, \quad (21)$$

¹⁴Interestingly, when $\rho \rightarrow 0$, the solution to (19) is $\eta = 0.43$. However, since all "screening" firms offer the same wage to any worker they meet, and "non-screening" firms never (in the limit) punish workers, the equilibrium structure is identical to that of the no information case.

respectively; while for a non-screening firm offering $w_H > \tilde{w}_H(\eta)$ these measures are given by

$$n_L^{ns}(w_H) = \frac{\lambda(\phi + \delta + \rho)\alpha_L}{[q(w_H) + \rho]^2}, \quad n_L^{ns}(b) = \frac{\lambda\rho(\phi + \delta + \rho)\alpha_L}{[q(w_H) + \rho]^2(\phi + \delta + \lambda)} \quad \text{and} \quad n_H^{ns}(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{q(w_H)^2}, \quad (22)$$

where the total number of low ability workers in a non-screening firm offering w_H is $n_L^{ns}(w_H) + n_L^{ns}(b)$. It is then easy to verify that screening firms have a higher proportion of low ability workers in their workforces, while non-screening firms have a higher proportion of high ability workers. Furthermore, the average workforce size among screening firms is given by

$$E(n^s) = \frac{1}{\eta} \int_b^{\tilde{w}_H(\eta)} [n_L^s(\hat{w}(w_H)) + n_H^s(w_H)] dF_H(w_H);$$

while the average workforce size among non-screening firms is given by

$$E(n^{ns}) = \frac{1}{1 - \eta} \int_{\tilde{w}_H(\eta)}^{\bar{w}_H(\eta)} [n_L^{ns}(w_H) + n_L^{ns}(b) + n_H^{ns}(w_H)] dF_H(w_H).$$

The numerical solution for the above expressions indeed shows that the average size of the workforce in non-screening firms is greater than that of screening firms. Hence, firms that employ a more productive workforce are, on average, bigger.

Now consider the relation between firms' size and wages offered. The average wage earned in a screening firm offering $w_H \leq \tilde{w}_H(\eta)$ is given by $E^s(w_H) = w_H \tilde{n}_H^s(w_H) + \hat{w}(w_H) \tilde{n}_L^s(\hat{w}(w_H))$, where $\tilde{n}_i^s(w_H) = n_i^s(w_H)/n^s(w_H)$ and $n^s(w_H) = n_H^s(w_H) + n_L^s(\hat{w}(w_H))$; while the average wage earned in a non-screening firm offering $w_H > \tilde{w}_H(\eta)$ is $E^{ns}(w_H) = w_H [\tilde{n}_H^{ns}(w_H) + \tilde{n}_L^{ns}(w_H)] + b \tilde{n}_L^{ns}(b)$ and $\tilde{n}_i^{ns}$ is defined in the same way as above. Integrating across firms, yields the average wage earned in screening and non-screening firms, E_w^s and E_w^{ns} , respectively. The numerical solution of the latter expressions then yields $E_w^s = 1.006$ and $E_w^{ns} = 1.417$. Hence, firms that employ a more productive workforce are, on average, bigger and pay higher wages (see Brown and Medoff, 1989, for empirical evidence of this relationship).¹⁵

5 Firm Heterogeneity and Sorting

We now extend our basic model to include heterogeneity in firm productivity. The main aim is to analyze under what conditions there exists sorting by types. That is, equilibria in which more

¹⁵It is important to note that when screening and non-screening firms coexist, the relationship between firm size (workforce size) and wages is not monotonic as in the Burdett and Mortensen (1998) model. Although there is a positive relation between firm size and wages offered within each type of firm, it is easy to verify that the size of the highest paying screening firm is greater than the size of the lowest paying non-screening firm.

productive firms attract and retain a more productive workforce. The outcome of this type of equilibrium leads to one of the main results of the paper. In a labor market with search frictions and adverse selection, firms' wage policies can improve the allocation of resources and increase total output relative to an environment with perfect information. Further, such equilibria are consistent with the empirical evidence showing that the positive relation between firm size and wages is not only due to the fact that high wage firms have a more productive workforce, but that these firms are in itself more productive (see Haltiwanger et al. 1999).

The formal analysis of this case is very similar to that with homogeneous firms and is relegated to the Appendix. Here we present some important results and a numerical solution to such a model. Let β_H denote the fraction of firms with high productivity p_H and $\beta_L = 1 - \beta_H$ the fraction of firms with low productivity p_L . Lemma a.1 in the Appendix shows that in equilibrium more productive firms offer higher wages than the ones offered by less productive firms. Equilibria can then be of two types. Equilibria in which $\eta \in (0, \beta_L)$ such that low productivity firms offer screening and non-screening contracts and all high productivity firms offer non-screening contracts. Equilibria in which $\eta \in (\beta_L, 1]$ and all low productive firms offer screening contracts, while high productivity firms offer screening and non-screening contracts. More importantly, (21) and (22) imply the following result.

Proposition 4: *Given there exists an equilibrium in which fraction $\eta < 1$ of firms offer screening contracts, some (or all) high productivity firms employ a higher proportion of high ability workers than low productivity firms.*

Hence in any equilibrium with $\eta \in (0, 1)$ there is positive sorting of workers among firms. The degree of sorting depends on the value of η . For $\eta < \beta_L$ all high productivity firms have a higher proportion of high ability workers, while for $\eta > \beta_L$ only some high productivity firms have a higher proportion of these workers. In both cases, however, total output is higher than in any equilibria in which all firms screen workers.

5.1 Numerical Example

To solve the model we use the characterization of $\tilde{w}_H(\eta)$, $\tilde{w}_L(\eta)$, $\bar{w}_{Hk}(\eta)$, $F_H(\cdot | p_k)$, $F_H(\cdot)$, η for $k = L, H$ and the conditions for existence described in the appendix. We use the same parameter values as before, but set $p_L = 1$, $p_H = 2$. The value of β_L becomes important to determine which type of equilibrium is obtained. Given β_L is greater than the value of η computed for the homogeneous case, $\eta^{hm} = 0.624$, then the equilibrium is of the type in which a measure $\eta^* = \eta^{hm}$ of low productivity firms offer screening contracts and a measure $\beta_L - \eta^*$ offer

non-screening contracts; while all high productivity firms offer non-screening contracts. On the other hand, given $\beta_L < \eta^{hm}$, then all low productivity firms offer screening contracts; while a measure $\eta^{hm} - \beta_L$ of high productivity firms offer screening contracts and a measure $1 - \beta_L$ of high productivity firms offer non-screening contracts. We consider two values for $\beta_L = \{0.55, 0.75\}$ to show the properties of the model in each case and to reflect the decreasing probability mass function observed in empirical firm productivity distributions (see Lentz and Mortensen, 2006).

In the case of $\beta_L = 0.75$, we have that $\eta^* = \eta^{hm}$, $\bar{w}_{HL} = \underline{w}_{HH} = xxx$, and $\bar{w}_{HH} = xx$. Further, the average size of non-screening firms is greater than screening firms, $E(n^{ns}) > E(n^s)$; while the average size of high productivity firms is also greater than of low productivity firms, $E_H(n^{ns}) > E_L(n^{ns}) + E_L(n^s)$. The numerical solutions for the average wages show that the average wages earned in high productivity firms is higher than the average wage earned in low productivity firms. Hence this equilibrium implies that more productive firms not only employ a more productive workforce, but also offer, on average, higher wages and are bigger than less productive firms.

Now let $\beta_L = 0.55$. In this case we have that $\eta^* = xx$, $\bar{w}_{HL} = \underline{w}_{HH} = xxx$, and $\bar{w}_{HH} = xx$. Further, the average size of non-screening firms is greater than screening firms, $E(n^{ns}) > E(n^s)$; while the average size of high productivity firms is also greater than of low productivity firms, $E_H(n^{ns}) > E_L(n^{ns}) + E_L(n^s)$. The numerical solutions for the average wages show that the average wages earned in high productivity firms is higher than the average wage earned in low productivity firms. Hence this equilibrium implies that more productive firms not only employ a more productive workforce, but also offer, on average, higher wages and are bigger than less productive firms.

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 the ability of their applicants, but gradually learn it over time. Given this adverse selection problem, we show that when the learning rate is sufficiently low a unique equilibrium emerges in which low wage firms attempt to hire both low and high ability workers by offering incentive-compatible screening contracts. High wage firms offer contracts that intend to attract and retain only high ability workers. Moreover, in this equilibrium low ability workers have a higher degree of turnover and a more dispersed wage distribution. We also show, under reasonable parameter values, that there is a positive relation between wages, firm size and the productivity of the workforce, in line with the empirical

evidence presented by Brown and Medoff (1989) and others.

We extend our model and introduce firm heterogeneity to show that the previous type of equilibrium, implies positive sorting of workers among firms. High productivity firms attract and retain a more productive workforce. Total output is higher relative to an equilibrium where all firms offer screening contracts, as is the case under perfect information. We also show, again under reasonable parameter values, that more productive firms offer, on average, higher wages, are bigger and employ a more productive, in line with the empirical evidence presented by Abowd, Kramarz, and Margolis (1999) and Haltiwanger, Lane, and Spletzer (1999).

This paper restricts attention to flat wage contracts. However, Burdett and Coles (2003) and Stevens (2004) show that with on-the-job search firms benefit from offering upward-sloping wage-tenure contracts to reduce workers' quit probability. Our choice of contract space is purely motivated to preserve a tractable analysis while considering two-sided heterogeneity. As shown by Burdett and Coles (2010), wage-tenure contracts with firm productivity dispersion becomes highly intractable. Although the model could then generate richer wage dynamics we leave this extension for future research.

References

- ABOWD, J., F. KRAMARZ, AND D. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–334.
- 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.
- BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J.-M. ROBIN (2009): "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation," Discussion Paper No. 09/02, University of Bristol.
- BAGGER, J., AND R. LENTZ (2008): "An Empirical Model of Wage Dispersion and Sorting," Mimeo, University of Wisconsin.
- BROWN, C., AND J. MEDOFF (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1059.
- BURDETT, K., C. CARRILLO-TUDELA, AND M. COLES (2009): "Human Capital Accumulation and Labour Market Equilibrium," *International Economic Review*, forthcoming.

- BURDETT, K., AND M. COLES (2003): “Equilibrium Wage-Tenure Contracts,” *Econometrica*, 71, 1377–1404.
- (2010): “Wage/Tenure Contracts with Heterogeneous Firms,” *Journal of Economic Theory*, 145, 1408–1435.
- BURDETT, K., AND D. MORTENSEN (1998): “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review*, 39, 257–273.
- CARRILLO-TUDELA, C. (2009): “An Equilibrium Search Model when Firms Observe Workers’ Employment Status,” *International Economic Review*, 50, 485–506.
- FARBER, H. (1999): “Mobility and Stability: The Dynamics of Job Change in Labor Markets,” in *Handbook of Labor Economics Vol. 3*, ed. by O. Ashenfelter, and D. Card, pp. 2439–2483. Elsevier, Amsterdam.
- FU, C. (2010): “Training, Search and Wage Dynamics,” *Review of Economic Dynamics*, forthcoming.
- GIBBONS, R., AND L. KATZ (1991): “Layoffs and Lemons,” *Journal of Labor Economics*, 9, 351–380.
- 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.
- HALTIWANGER, J., J. LANE, AND J. SPLETZER (1999): “Productivity Differences across Employers: The Roles of Employer Size, Age, and Human Capital,” *American Economic Review*, 89, 94–98.
- HORNSTEIN, A., P. KRUSELL, AND G. VIOLANTE (2010): “Frictional Wage Dispersion in Search Models: A Quantitative Assessment,” *American Economic Review*, forthcoming.
- 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.

- 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.
- 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.
- LOPES DE MELO, R. (2009): “Sorting in the Labor Market: Theory and Measurement,” Mimeo, University of Chicago.
- 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*, ed. by S. Rosen, pp. 21–63. University of Chicago Press.
- MONTGOMERY, J. (1999): “Adverse Selection and Employment Cycles,” *Journal of Labor Economics*, 17, 281–297.
- MORTENSEN, D. (2003): *Wage Dispersion: Why are Similar Workers Paid Differently?* The MIT Press, Cambridge, MA.
- POSTEL-VINAY, F., AND J.-M. ROBIN (2002): “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, 70, 2295–2350.
- 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.
- SHIMER, R., AND L. SMITH (2000): “Assortative Matching and Search,” *Econometrica*, 68, 343–369.
- STEVENS, M. (2004): “Wage-Tenure Contracts in a Frictional Labour Market: Firms’ Strategies for Recruitment and Retention,” *Review of Economic Studies*, 71(2), 535–551.

VISSCHERS, L. (2007): “Employment Uncertainty and Wage Contracts in Frictional Labor Markets,” Mimeo, University of Pennsylvania.

A Proofs

Proof of Proposition 1:

First note that a low ability worker reports the correct type if and only if $V_{LH}(w_H) - V_{LL}(\hat{w}(w_H)) \leq 0$. Monotonicity of $\hat{w}$ and the Bellman equations (1) and (2) then imply that this condition can be expressed as

$$V_{LH}(w_H) - V_{LL}(\hat{w}(w_H)) = \frac{\hat{w}^{-1}(w_L) - w_L + \rho[U_L - V_{LL}(w_L)]}{\phi + \delta + \rho + \lambda(1 - F_L(w_L))} \leq 0.$$

Since the lowest paying firm offers $w_L = w_H = b$ and hence $U_L = V_{LL}(b) = V_{LH}(b)$, the above and (6) imply that low ability workers will self-select into the correct contract if and only if,

$$\varphi(w_L) \equiv V_{LL}(w_L) - V_{LL}(b) \geq \frac{(w_L - b)(\varepsilon_H - \varepsilon_L)p}{(\varepsilon_L p - b)\rho} \equiv \psi(w_L). \quad (23)$$

Equation (1) and the constant profit condition imply that

$$\begin{aligned} \varphi(w_L) &= \int_b^{w_L} V'_{LL}(x)dx = \int_b^{w_L} \frac{dx}{\phi + \delta + \lambda(1 - F_L(x))} \\ &= \frac{2(\varepsilon_L p - b)^{1/2}}{\phi + \delta + \lambda} [(\varepsilon_L p - b)^{1/2} - (\varepsilon_L p - w_L)^{1/2}]. \end{aligned}$$

Note that function ψ increases linearly in w_L . Since φ is strictly increasing and convex and $\varphi(b) = \psi(b) = 0$, it follows that condition (23) holds for all w_L if and only if $\varphi'(b) \geq \psi'(b)$. This is equivalent to the firm's learning rate satisfying the following condition

$$\rho \geq \frac{(\varepsilon_H - \varepsilon_L)p}{(\varepsilon_L p - b)\varphi'(b)}.$$

Substituting out for $\varphi'(b)$ then yields the condition stated in the proposition. This completes the proof of Proposition 1. $\square$

Proof of Proposition 2:

The proof of Proposition 1 reveals that the incentive constraint starts to bind at low wage firms when ρ is just below threshold ρ_1 . Thus we characterize an equilibrium in which the incentive constraint binds on a fraction $\gamma \leq 1$ of firms offering wages $w_H \in [b, \tilde{w}_H]$, and is slack for the remaining fraction $1 - \gamma$ of firms offering $w_H \in [\tilde{w}_H, \bar{w}_H]$, where γ is an equilibrium

object determined below. The associated wage offers for low-ability workers are described by a function $w_L = \hat{w}(w_H)$. When the incentive-constraint binds, it follows that function $\hat{w}$ obeys differential equation (18); the proof of this assertion is exactly as in Section 3.2 and follows from differentiation of (2) and (10). We denote by $\hat{w}^{IC}$ the unique solution of this differential equation with initial condition $\hat{w}^{IC}(b) = b$. Because the RHS of (18) is strictly decreasing in w_H , $\hat{w}^{IC}$ is a strictly concave function.

With $q(w_H) \equiv \phi + \delta + \lambda(1 - F_H(w_H))$, firms facing binding incentive constraints make constant profit if

$$\Omega^{IC}(w_H) = \frac{\lambda(\phi + \delta)}{q(w_H)^2} \left[\tilde{\varepsilon}p - \alpha_H w_H - \alpha_L \hat{w}^{IC}(w_H) \right] = \frac{\lambda(\phi + \delta)(\tilde{\varepsilon}p - b)}{q(b)^2}. \quad (24)$$

Differentiation of this equation yields a differential equation for the wage offer distribution F_H :

$$F'_H(w_H) = \frac{(\phi + \delta + \lambda)^2 \left[\alpha_L q(w_H) + \alpha_H (q(w_H) + \rho) \right]}{2\lambda q(w_H)(q(w_H) + \rho)(\tilde{\varepsilon}p - b)}. \quad (25)$$

Let F_H be the solution of this differential equation with $F_H(b) = 0$, and define $\tilde{w}_H(\gamma)$ by $F_H(\tilde{w}_H) = \gamma$. Further, define $\tilde{w}_L(\gamma) = \hat{w}^{IC}(\tilde{w}_H(\gamma))$.

For the remaining fraction of firms, the incentive constraint is slack (which will be verified below). These firms offer wages $w_H \geq \tilde{w}_H$ to maximize Ω_i as defined in (4). It follows from the constant-profit conditions $\Omega_i(w_i) = \Omega_i(\tilde{w}_i)$ and $F_i(\tilde{w}_i) = \gamma$ that the wage offer distributions satisfy

$$F_i(w_i) = \frac{1}{\lambda} \left\{ \phi + \delta + \lambda - (\phi + \delta + \lambda(1 - \gamma)) \left[\frac{\varepsilon_i p - w_i}{\varepsilon_i p - \tilde{w}_i} \right]^{1/2} \right\}, \text{ for } w_i \geq \tilde{w}_i,$$

for $i = H, L$. This defines $\bar{w}_i$ from $F_i(\bar{w}_i) = 1$ and it also implies that

$$w_L = \hat{w}^S(w_H) \equiv \varepsilon_L p + \frac{\varepsilon_L p - \tilde{w}_L}{\varepsilon_H p - \tilde{w}_H} (w_H - \varepsilon_H p), \quad w_H \in [\tilde{w}_H, \bar{w}_H]. \quad (26)$$

This shows that $\hat{w}$ is defined by the strictly concave function $\hat{w}(w_H) = \hat{w}^{IC}(w_H)$ for $w_H \in [b, \tilde{w}_H]$, and by the linear function $\hat{w}(w_H) = \hat{w}^S(w_H)$ on $w_H \in [\tilde{w}_H, \bar{w}_H]$. Evidently, $\hat{w}$ is continuous and strictly increasing. Because $\hat{w}^{IC}$ describes binding incentive constraints, the incentive constraint is slack at all wages $w_H > \tilde{w}_H$ if and only if $\hat{w}^{IC}(w_H) < \hat{w}^S(w_H)$ for $w_H > \tilde{w}_H$. Because $\hat{w}^{IC}$ is strictly concave and $\hat{w}^S$ is linear, this is the case iff

$$\hat{w}^{IC'}(\tilde{w}_H) \leq \hat{w}^{S'}(\tilde{w}_H)$$

holds. (18) and (26) imply that this is true iff

$$\frac{\varepsilon_L p - \tilde{w}_L}{\varepsilon_H p - \tilde{w}_H} \geq \frac{q(\tilde{w}_H)}{q(\tilde{w}_H) + \rho}. \quad (27)$$

This condition is necessary for an equilibrium with $\gamma < 1$. On the other hand, a binding incentive constraint implies that the firm offering $\tilde{w}_H$ (or any wage below) does not find it profitable to decrease w_H while keeping $w_L = \hat{w}(\tilde{w}_H)$ fixed. This is true if Ω_H , as defined in (4), has a non-negative lower derivative at $w_H = \tilde{w}_H$, which is true iff

$$q(\tilde{w}_H) \leq \frac{(\varepsilon_H p - \tilde{w}_H)(\phi + \delta + \lambda)^2(q(\tilde{w}_H) + \alpha_H \rho)}{(\tilde{\varepsilon} p - b)(q(\tilde{w}_H) + \rho)}.$$

Using the constant-profit condition $(\tilde{\varepsilon} p - b)/(\phi + \delta + \lambda)^2 = (\tilde{\varepsilon} p - \alpha_H \tilde{w}_H - \alpha_L \tilde{w}_L)/q(\tilde{w}_H)^2$, this condition is equivalent to

$$\frac{\varepsilon_L p - \tilde{w}_L}{\varepsilon_H p - \tilde{w}_H} \leq \frac{q(\tilde{w}_H)}{q(\tilde{w}_H) + \rho}. \quad (28)$$

These considerations show that in any equilibrium with $\gamma < 1$, (27) and (28) must hold with equality, whereas an equilibrium with $\gamma = 1$ (all firms face binding incentive constraints) must satisfy the weak inequality (28). At $\gamma = 1$, $\tilde{w}_L = \bar{w}_L$ and $\tilde{w}_H = \bar{w}_H$, and it follows that (28) coincides with (9). This condition, therefore, implicitly pins down threshold parameter ρ_2 . For any $\rho \in (\rho_2, \rho_1)$, the binding condition (28) then defines the equilibrium value of $\gamma \in (0, 1)$. This equilibrium exists because the RHS in (28) is larger than the LHS at $\gamma = 0$ (which follows from $\rho < \rho_1$) and since the RHS is smaller than the LHS at $\gamma = 1$ (which follows from $\rho > \rho_2$). Since all functions are continuous in γ , existence follows. To obtain a closed-form expression for this condition, calculate $\tilde{w}_L(\gamma)$ using (18) and (25):

$$\begin{aligned} \tilde{w}_L(\gamma) &= b + \int_b^{\tilde{w}_H} \frac{q(w_H)}{q(w_H) + \rho} dw_H = b + \int_{\phi + \delta + \lambda(1-\gamma)}^{\phi + \delta + \lambda} \frac{q}{(q + \rho)\lambda F_H'} dq \\ &= b + \frac{2(\tilde{\varepsilon} p - b)}{(\phi + \delta + \lambda)^2} \int_{\phi + \delta + \lambda(1-\gamma)}^{\phi + \delta + \lambda} \frac{q^2}{q + \alpha_H \rho} dq \\ &= b + \frac{2(\tilde{\varepsilon} p - b)}{(\phi + \delta + \lambda)^2} \left\{ (\alpha_H \rho)^2 \ln \left[\frac{\phi + \delta + \lambda + \alpha_H \rho}{\phi + \delta + \lambda(1-\gamma) + \alpha_H \rho} \right] \right. \\ &\quad \left. + \frac{(\phi + \delta + \lambda)^2 - (\phi + \delta + \lambda(1-\gamma))^2}{2} - \alpha_H \rho \lambda \gamma \right\}. \end{aligned}$$

Furthermore, $\tilde{w}_H(\gamma)$ can be calculated from the constant-profit condition (24):

$$\tilde{w}_H(\gamma) = \frac{1}{\alpha_H} \left\{ \tilde{\varepsilon} p - \alpha_L \tilde{w}_L(\gamma) - \frac{(\phi + \delta + \lambda(1-\gamma))^2}{(\phi + \delta + \lambda)^2} (\tilde{\varepsilon} p - b) \right\}.$$

For $\gamma = 1$, $\tilde{w}_L(1) = \bar{w}_L$ and $\tilde{w}_H(1) = \bar{w}_H$ coincide with (7) and (8). This completes the proof of Proposition 2. $\square$

Proof of Proposition 3:

We use a contradiction argument to show that there does not exist an equilibrium in which all firms offer non-screening contracts. If $\eta^* = 0$, no firm has an incentive to offer a screening

contract. In particular, consider a firm offering the lowest wages, $w_H = w_L = b$. If $\eta^* = 0$, this firm has no incentive to offer a screening contract. Namely,

$$\left[\frac{d\Omega^S(w_H, \widehat{w}(w_H))}{dw_H} \right]_{w_H=b} \leq 0. \quad (29)$$

Using (16), some algebra then establishes that condition (29) is the same as

$$2q(b)\overline{\Omega}^{NS}F'_H(b) \leq \theta(0)\alpha_L\widehat{w}'(b) + (\phi + \delta)\alpha_H,$$

where $\overline{\Omega}^{NS} = \overline{\Omega}_L^{NS} + \overline{\Omega}_H^{NS} = \Omega^{NS}(b, b)$ denotes the profits firms obtain in this equilibrium and $\overline{\Omega}_i^{NS}$ refers to the equilibrium profits obtained from workers with ability $i = L, H$.

Also note that in an equilibrium with $\eta^* = 0$, all firms offering wages $w_H \in [b, \overline{w}_H]$ obtain the same profits and hence,

$$\begin{aligned} 0 &= \left[\frac{d\Omega^{NS}(w_H, b)}{dw_H} \right]_{w_H=b} = -\frac{\lambda(\phi + \delta + \rho)\alpha_L}{(q(b) + \rho)^2} - \frac{\lambda(\phi + \delta)\alpha_H}{q(b)^2} + 2\lambda F'_H(b) \left[\frac{\overline{\Omega}_L^{NS}}{q(b) + \rho} + \frac{\overline{\Omega}_H^{NS}}{q(b)} \right] \\ &< -\frac{\lambda(\phi + \delta + \rho)\alpha_L}{(q(b) + \rho)^2} - \frac{\lambda(\phi + \delta)\alpha_H}{q(b)^2} + 2\lambda F'_H(b) \left[\frac{\overline{\Omega}^{NS}}{q(b)} \right]. \end{aligned}$$

Using (18) and re-arranging implies that

$$\begin{aligned} 2q(b)\overline{\Omega}^{NS}F'_H(b) &> \left(\frac{q(b)}{q(b) + \rho} \right)^2 (\phi + \delta + \rho)\alpha_L + (\phi + \delta)\alpha_H \\ &= \theta(0)\alpha_L\widehat{w}'(b) + (\phi + \delta)\alpha_H, \end{aligned}$$

which provides the required contradiction to condition (29). This completes the proof of Proposition 3. $\square$

Proof of Theorem 1:

Differentiating the function T with respect to η yields

$$T'(\eta) = \frac{\phi + \delta + \lambda}{\lambda(\varepsilon_L p - b)} \left[\frac{d\widetilde{w}_H}{d\eta} \left(\frac{d\widehat{w}(\widetilde{w}_H)}{dw_H} - \left(1 - \frac{d\widehat{w}(\widetilde{w}_H)}{dw_H} \right) \frac{\phi + \delta + \lambda(1 - \eta)}{\rho} \right) + \frac{\lambda[\widetilde{w}_H(\eta) - \widetilde{w}_L(\eta)]}{\rho} \right].$$

Using (18) to substitute out $d\widehat{w}(\widetilde{w}_H)/dw_H$ then simplifies the above expression to

$$T'(\eta) = \frac{(\phi + \delta + \lambda)[\widetilde{w}_H(\eta) - \widetilde{w}_L(\eta)]}{\rho(\varepsilon_L p - b)} > 0$$

for all $\eta > 0$. Further differentiation implies convexity,

$$T''(\eta) = \frac{\phi + \delta + \lambda}{(\varepsilon_L p - b)(\phi + \delta + \rho + \lambda(1 - \eta))} \frac{d\widetilde{w}_H}{d\eta} > 0$$

for all $\eta > 0$. Noting that $T(0) = T'(0) = 0$, continuity implies that T has at most one fixed point. A fixed point $\eta \in (0, 1)$ exists if and only if $T(1) > 1$. In this case, a proportion η of firms offer screening contracts, while $1 - \eta$ offer non-screening contracts. Otherwise, $\eta = 1$ and all firms offer screening contracts. The condition $T(1) = 1$ implicitly defines the threshold level ρ_3 at which the highest-wage firms are exactly indifferent between offering a screening or a non-screening contract. Since $\tilde{w}_i(1) = \bar{w}_i$ holds, with $\bar{w}_i$ defined in (7) and (8) (see also the proof of Proposition 2), condition (20) immediately follows.

Given the equilibrium value of $\eta \in (0, 1]$, note that by construction all firms are indifferent from offering wages in the support of $w \in [\bar{b}, w_H]$. Further, no firm would offer a wages below b as a worker would not accept such an offer, yielding zero profit for the firm. Further, no firm would offer a wage $w_H > \bar{w}_H$ as doing so strictly reduces profits. The firm does not increase hiring and retention rates and reduces its profit flow per worker. Hence given $\eta^* \in (0, 1]$, an equilibrium exists, is unique and is fully characterized by the wage-offer distributions characterized in the text. This completes the proof of Theorem 1. $\square$

B Firm Heterogeneity

Let β_H denote the fraction of firms with high productivity p_H and $\beta_L = 1 - \beta_H$ the fraction of firms with low productivity p_L . A worker with ability ε_i employed at a firm with productivity p_k then generates flow output $\varepsilon_i p_k$ for $i, k = L, H$. Let $F_i(w_i | p_k)$ denote the proportion of firms with productivity k offering a wage no greater than w_i to workers of ability i , for $i, k = L, H$. Further, let $\underline{w}_{ik}$ and $\bar{w}_{ik}$ denote the infimum and supremum of the support of $F_i(\cdot | p_k)$. Hence,

$$F_i(w_i) = \beta_H F(w_i | p_H) + (1 - \beta_H) F(w_i | p_L) \quad (30)$$

denotes the proportion of firms that offer a wage no greater than w_i to workers of ability i , for $i = L, H$, with $\underline{w}_i$ and $\bar{w}_i$ denoting the infimum and supremum of the support of F_i . Again we consider a candidate equilibrium with the rank-preservation property: Firms offer wages w_H and $w_L = \hat{w}(w_H)$, with an increasing function $\hat{w}$.

Given the specification for F_i , the worker's problem is the same as in the homogeneous case. A firm of type k maximizes $\Omega_k(w_H, w_L) = \sum_{i=L,H} \Omega_{ik}(w_H, w_L)$. Let $\bar{\Omega}_k = \max \Omega_k(w_H, w_L)$.

Finally, we use the same equilibrium concept as before, but require that the constant-profit condition (i) is satisfied for each firm type p_k ; i.e.

$$\Omega_k(w_H, w_L) = \bar{\Omega}_k \text{ and } F_L(w_L | p_k) = F_H(w_H | p_k) \quad \text{for all } w_H \in \text{supp } F_H(\cdot | p_k) \text{ and } w_L = \hat{w}(w_H).$$

B.1 Sorting Equilibrium

In what follows we focus on a candidate equilibrium in which a fraction $\eta \leq 1$ of firms offer screening contracts. As before screening firms offer jobs to high ability workers at a wage $w_H \in [b, \tilde{w}_H]$ and to low ability workers at a wage $w_L = \hat{w}(w_H) \leq w_H$ with $\tilde{w}_L = \hat{w}(\tilde{w}_H)$ satisfying (10). The remaining fraction $1 - \eta$ of firms specialize in attracting high ability workers by offering $w_H > \tilde{w}_H$ to high ability workers and $w_L = \hat{w}(w_H)$ to low ability workers, satisfying (11). Note that as in the homogeneous case, the arguments of Burdett and Mortensen (1998) imply that the wage offer distributions, $F_H(\cdot \mid p_k)$ for $k = L, H$, are continuous and exhibit connected supports. Further, given that the steady state measures u_i and G_i are the same as the ones derived in the homogeneous case, the following result shows that high productivity firms offer higher wages than low productivity firms.

Lemma a.1: *Given $\eta \neq \beta_L$, any equilibrium implies $\underline{w}_{iH} \geq \overline{w}_{iL}$ for $i = H, L$. Hence, firms with productivity p_H offer higher wages to both types of workers than firms with productivity p_L .*

Proof: To follow.

Given Lemma a.1 some non-screening firms will be of high productivity and hence there will be positive sorting between workers and firms as described in Proposition 4.

B.1.1 Characterization

To characterize equilibrium with firm heterogeneity we divide the analysis in two cases.

Case I: $\eta < \beta_L$

In this case, some low productivity firms offer screening contracts and some low productivity firms and all high productivity firms offer non-screening contracts. It is immediate that the arguments presented for the homogeneous case also apply here and imply that for a given η the wages offered to low ability workers by screening firms $w_L = \hat{w}(w_H)$ are described by (18) subject to the initial condition $\hat{w}(b) = b$. Further, Lemma 1 and Lemma 2 (with $p = p_L$) describe the offer distribution, F_H , for wages $w_H \in [b, \overline{w}_{HL}]$ such that $\tilde{w}_H(\eta)$ solves $F_H(\tilde{w}_H) = \eta$ using Lemma 1 and $\overline{w}_{HL}(\eta)$ solves $F_H(\overline{w}_{HL}) = \beta_L$ using Lemma 2. It then follows from Lemma a.1 and (30) that $F_H(w_H \mid p_L) = F_H(w_H)/\beta_L$ for all $w_H \in [b, \overline{w}_{HL}]$.

To obtain the offer distribution for wages $w_H \in [\underline{w}_{HH}, \overline{w}_{HH}]$ first note that optimality implies $\underline{w}_{HH}(\eta) = \overline{w}_{HL}(\eta)$. Further, since equilibrium requires that $\overline{\Omega}^{ns}(\underline{w}_{HH}, \underline{w}_{LH}) = \overline{\Omega}^{ns}(w_H, w_L)$ for all $w_H \in [\underline{w}_{HH}, \overline{w}_{HH}]$, $w_L = \hat{w}(w_H)$, distribution F_H is described by the differential equation in Lemma 2 with $p = p_H$ subject to the initial condition $F_H(\underline{w}_{HH}) = \beta_L$ and $\overline{w}_{HH}(\eta)$ solves

$F_H(\bar{w}_{HH}) = 1$. In this case, Lemma a.1 and (30) imply $F_H(w_H | p_H) = [F_H(w_H) - \beta_L]/[1 - \beta_L]$ for $w_H \in [\underline{w}_{HH}, \bar{w}_{HH}]$.

The last step to characterize the equilibrium is to solve for η . This can be done using the arguments of the homogeneous case by obtaining the fixed point of T in (19) with $p = p_L$. Note, however, that we must apply the restriction $\eta \in (0, \beta_L)$.

Case II: $\eta > \beta_L$

Now consider the case in which all low productivity firms and some high productivity firms offer screening contracts, while some high productivity firms offer non-screening contracts. Once again, the arguments presented in the homogeneous can be applied here and imply that for a given η the wages offered to low ability workers $w_L = \hat{w}(w_H)$ are described by (18) subject to the initial condition $\hat{w}(b) = b$. Further, the offer distribution, F_H , for wages $w_H \in [b, \bar{w}_{HL}]$ solves the differential equation in Lemma 1 with $p = p_L$ subject to the initial condition $F_H(b) = 0$ and $\bar{w}_{HL}$ solves $F_H(\bar{w}_{HL}) = \beta_L$. As before we have that $F_H(w_H | p_L) = F_H(w_H)/\beta_L$ for all $w_H \in [b, \bar{w}_{HL}]$.

Since optimality implies $\underline{w}_{HH}(\eta) = \bar{w}_{HL}(\eta)$, (18) describes $w_L = \hat{w}(w_H)$ for those screening firms with high productivity. The differential equation in Lemma 1 (with $p = p_H$) describes the corresponding offer distribution, $F_H(\cdot)$, for wages $w_H \in [\underline{w}_{HH}, \tilde{w}_H(\eta)]$ subject to the initial condition $F_H(\underline{w}_{HH}) = \beta_L$ and $\tilde{w}_H(\eta)$ solves $F_H(\tilde{w}_H) = \eta$. Lemma 2 with $p = p_H$ describes the offer distribution for wages $w_H \in (\tilde{w}_H(\eta), \bar{w}_{HH}]$, where $\bar{w}_{HH}$ solves $F_H(\bar{w}_{HH}) = 1$ and $F_H(w_H | p_H) = [F_H(w_H) - \beta_L]/[1 - \beta_L]$ for $w_H \in [\underline{w}_{HH}, \bar{w}_{HH}]$.

Finally, η is determined by the fixed point of T as described in (19) with $p = p_H$, given the restriction that $\eta \in (\beta_L, 1]$.

B.1.2 Existence

To follow.

C Omitted Derivations

Derivation of the steady state profits for screening and non-screening firms:

First consider a screening firm that offers $w_H \leq \tilde{w}_H$ to high ability workers and $w_L = \hat{w}(w_H) \leq \tilde{w}_L$ to low ability workers. This firm's steady-state profit is characterized by

$$\Omega^S(w_H, w_L) = h_L(w_L)J_L(w_L) + h_H(w_H)J_H(w_H) ,$$

where h_i are the hiring flows and J_i are profit values per hire of type $i = H, L$. Noting that $h_i(w_i) = \lambda u_i + \lambda G_i(w_i)(\alpha_i - u_i)$, for $i = L, H$ and using (12), (13) and (14) yields

$$h_H(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{\phi + \delta + \lambda(1 - F_H(w_H))}, \text{ and } h_L(w_L) = \frac{\lambda\theta\alpha_L}{\phi + \delta + \lambda(1 - F_H(\widehat{w}^{-1}(w_L)))}, \quad (31)$$

where θ is defined in the main text. Further, since all workers quit to a firm offering higher wages, the expected profit per new hire associated with each wage offer is given by

$$J_H(w_H) = \frac{\varepsilon_H p - w_H}{\phi + \delta + \lambda(1 - F_H(w_H))}, \text{ and } J_L(w_L) = \frac{\varepsilon_L p - w_L}{\phi + \delta + \lambda(1 - F_H(\widehat{w}^{-1}(w_L)))}.$$

Substituting out the above expressions in $\Omega^S(w_H, w_L)$ and some algebra establishes (16) in the text.

Next consider the a non-screening firm that offers $w_H > \widetilde{w}_H$ to high ability workers and $w_L = \widehat{w}(w_H)$ to low ability workers, satisfying (11). Since low ability workers will misreport their type when meeting this firm, its steady state profit is characterized by

$$\Omega^{NS}(w_H, w_L) = [h_L(w_H) + h_H(w_H)]J(w_H).$$

Noting that posting w_H yields a hiring rate $h_i(w_H) = \lambda u_i + \lambda G_i(w_H)(\alpha_i - u_i)$ for $i = L, H$ and using (12), (13) and (15), we have that

$$h_L(w_H) = \frac{\lambda(\phi + \delta + \rho)\alpha_L}{\phi + \delta + \rho + \lambda(1 - F_H(w_H))}, \text{ and } h_H(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{\phi + \delta + \lambda(1 - F_H(w_H))}. \quad (32)$$

Further, the expected profit per new hire by offering w_H is given by

$$J(w_H) = \frac{[\widetilde{\alpha}_H(w_H)\varepsilon_H + \widetilde{\alpha}_L(w_H)\varepsilon_L]p - w_H + \rho[\widetilde{\alpha}_H(w_H)J_H(w_H) + \widetilde{\alpha}_L(w_H)J_L(b)]}{\phi + \delta + \rho + \lambda(1 - F_H(w_H))},$$

where $\widetilde{\alpha}_i(w_H) = h_i(w_H)/[h_L(w_H) + h_H(w_H)]$ denotes the proportion of type $i = L, H$ workers the firm attracts by posting w_H ; and $J_H(w_H)$ and $J_L(b)$ denote the expected value to the firm of employing a worker *after* learning his true type. These value functions are given by

$$J_L(b) = \frac{\varepsilon_L p - b}{\phi + \delta + \lambda}, \text{ and } J_H(w_H) = \frac{\varepsilon_H p - w_H}{\phi + \delta + \lambda(1 - F_H(w_H))}.$$

Substituting out these expressions in $\Omega^{NS}(w_H, w_L)$ and some algebra establishes (17) in the text.

Derivation of $\widetilde{w}_L = \widehat{w}(\widetilde{w}_H)$:

First consider the differential equation (18) describing w_L subject to the initial condition $\widehat{w}(b) = b$, and note that this equation applies for values of $w_H \in [b, \widetilde{w}_H]$. Integration implies

$$\widetilde{w}_L = b + \int_b^{\widetilde{w}_H} \frac{\phi + \delta + \lambda(1 - F_H(w_H))}{\phi + \delta + \rho + \lambda(1 - F_H(w_H))} dw_H.$$

Now consider the following change of variable: $q = \phi + \delta + \lambda(1 - F_H(w_H))$ such that $dq = -\lambda F'_H(w_H)dw_H$. Using the expression for $F'_H(w_H)$ described in Lemma 1 we have that

$$\begin{aligned} \tilde{w}_L = & b + \frac{(\phi + \delta + \lambda)^2}{2[(\phi + \delta)\alpha_H(\varepsilon_H p - b) + \theta(\eta)\alpha_L(\varepsilon_L p - b)]} \times \\ & \int_{\phi + \delta + \lambda(1-\eta)}^{\phi + \delta + \lambda} \left[\frac{\rho(\phi + \delta)\alpha_H}{(q + \rho)^2} + \frac{[(\phi + \delta)\alpha_H + \theta(\eta)\alpha_L]q}{(q + \rho)^2} \right] dq. \end{aligned}$$

Integration then yields

$$\begin{aligned} \tilde{w}_L = & b + \frac{(\phi + \delta + \lambda)^2}{2[(\phi + \delta)\alpha_H(\varepsilon_H p - b) + \theta(\eta)\alpha_L(\varepsilon_L p - b)]} \times \\ & \left[[(\phi + \delta)\alpha_H + \theta(\eta)\alpha_L] \log \left(\frac{\phi + \delta + \rho + \lambda}{\phi + \delta + \rho + \lambda(1 - \eta)} \right) - \frac{\theta(\eta)\alpha_L \lambda \rho \eta}{(\phi + \delta + \rho + \lambda)(\phi + \delta + \rho + \lambda(1 - \eta))} \right]. \end{aligned}$$

A closed-form expression for $\tilde{w}_H$ can also be obtained using the constant-profit condition $\Omega^S(\tilde{w}_H, \tilde{w}_L) = \Omega^S(b, b)$.