

IZA DP No. 5936

**Wage Dispersion and Labor Turnover
with Adverse Selection**

Carlos Carrillo-Tudela
Leo Kaas

August 2011

Wage Dispersion and Labor Turnover with Adverse Selection

Carlos Carrillo-Tudela

*University of Essex
and IZA*

Leo Kaas

*University of Konstanz
and IZA*

Discussion Paper No. 5936
August 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Wage Dispersion and Labor Turnover with Adverse Selection^{*}

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. With sufficiently strong information frictions, low-wage firms offer separating contracts and hire all types of workers in equilibrium, whereas high-wage firms offer pooling contracts designed to retain high-ability workers only. Low-ability workers have higher turnover rates, they are more often employed in low-wage firms and face an earnings distribution with a higher frictional component. Furthermore, positive sorting obtains in equilibrium.

JEL Classification: D82, J63, J64

Keywords: adverse selection, on-the-job search, wage dispersion, sorting

Corresponding author:

Leo Kaas
Department of Economics
University of Konstanz
78457 Konstanz
Germany
E-mail: leo.kaas@uni-konstanz.de

^{*} We would like to thank Jim Albrecht, Melvyn Coles, Javier Fernandez-Blanco, Miltos Makris, Espen Moen, Peter Norman, Fabien Postel-Vinay, Ludo Visschers and Susan Vroman for their comments and insights. We also thank participants at the Essex Economics and Music workshop 2011, SAET 2011, SED 2011, and at seminars in BI Oslo, Mainz and Carlos III. The usual disclaimer applies.

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 the re-employment chances of unemployed workers. With almost no exceptions, however, current contributions on labor search with adverse selection abstract from job-to-job flows,² although these transitions account for a sizable part of worker flows. Furthermore, the rate at which workers change jobs is an important determinant of wage dispersion among similar workers (see, e.g., Mortensen (2003) and Hornstein, Krusell, and Violante (2010)). Thus one would expect that asymmetric information not only has non-trivial implications for worker job-to-job turnover, but also for the distribution of wages and in particular on dispersion that is attributed to market frictions.

In this paper we consider a frictional labor market where workers search on the job and firms post wages. Firms commit to pay their posted wages for as long as the workers remain employed in the firm. Upon hiring, firms cannot observe the ability of their applicants, but they learn the worker's ability with delay during the employment spell. 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 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 workers' ability slowly over time. Firms compete for workers by offering long-term contracts which specify a flat wage, specific for a worker type and promised to be paid for the duration of the employment relation.³ To separate workers, firms can commit to fire workers who

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).

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

³We motivate our focus on flat-wage contracts below.

misreport their type upon hiring.

Firms follow one of two strategies in equilibrium. Either they decide to offer separating contracts or they offer pooling contracts. Separating contracts provide all workers with the same retention rates, while pooling contracts provide higher retention rates for more able workers. We show that the set of equilibria can be parameterized by the degree of information frictions. When firms learn sufficiently fast the type of a worker, only separating contracts are offered in equilibrium. Otherwise, equilibria are segmented: low-wage firms offer separating contracts, while high-wage firms offer pooling contracts.

In any segmented equilibrium, low-ability workers have higher job-turnover rates. Precisely this feature 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 aim to compete more strongly for high-ability workers and find it too costly to provide the necessary information rents to low-ability workers in a menu of separating contracts. Hence, they offer pooling contracts which retain a larger share of workers of higher ability. Although these firms also attract low-ability workers, these workers are laid off after the employer learns their type. In contrast, firms with lower wages and lower productivity prefer to separate workers and hence offer stable wage contracts. These firms end up employing a larger fraction of low-ability workers.

These turnover and sorting patterns have important consequences. They imply that the economy's total output is smaller when firms face search and (sufficiently large) information frictions than, for example, when firms face the same search frictions but are completely uninformed (or perfectly informed) about worker ability. Indeed, both in the absence of information and under full information, random search implies that all firms employ the same proportion of high- and low-ability workers, and all workers have the same employment rates. With asymmetric information, however, low-ability workers are more likely to get fired which reduces aggregate output, even when all firms are equally productive. But if firms differ in productivity, the positive sorting pattern gives rise to an additional loss of output: this is because low-ability workers are more likely to find employment in low-productivity firms. As a consequence, low-productivity firms are bigger and total output is further depressed relative to the no-sorting benchmark.

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 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 part left unexplained by these characteristics in those studies is attributed in this paper, as in Mortensen (2003), to labor market 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). 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 information frictions faced by employers and the search frictions faced by both workers and firms. We show that when information frictions are sufficiently strong, frictional wage dispersion is higher for low-ability than for high-ability workers. We also show that the amount of frictional wage dispersion faced by low-ability workers follows a hump-shaped relation with the firms' learning rate. That is, wage dispersion is highest for intermediate informational asymmetries.

The associated wage dynamics and turnover patterns also differ decisively between workers. Low-ability workers change jobs and experience unemployment more often than high-ability workers. In turn, the earnings of low-ability workers are characterized by more frequent upward and downward mobility. This property implies that high job turnover is associated with lower average wages as found in 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)).

Our restriction to constant-wage contracts is motivated by the wage-posting model described in Burdett and Mortensen (1998). Under this specification, we assume that firms commit not to counter any outside offer. We also rule out that firms offer back-loading wage schedules. Stevens (2004) and Burdett and Coles (2003) show that optimal wage-tenure contracts exhibit an increasing wage-tenure profile. By restricting attention to constant wages we are able to consider the implications of adverse selection, on-the-job search and firm heterogeneity on wage dispersion, job turnover and the allocation of resources in a simple and tractable environment.⁴ This restriction is also motivated by recent evidence showing that there is very little or even no returns to firm-specific tenure, implying that in reality firms offer mostly flat-wage contracts, conditional on labor market experience. Indeed, Kambourov and Manovskii (2009) for the US and Williams (2009) for the UK show that returns to firm tenure greatly diminishes or even disappears when controlling for experience in an industry and/or occupation. We believe this modeling restriction is a good starting point to understand wage dispersion and labor

⁴For example, Stevens (2004) shows that the dispersion in contract offers and job-to-job turnover disappears when considering wage-tenure contracts and risk-neutral workers. Burdett and Coles (2003) show that one can regain these features when workers are risk averse, but a model with risk-averse workers and firm heterogeneity becomes highly intractable (see Burdett and Coles (2010)).

turnover in an environment with search frictions and adverse selection which generates predictions consistent with the empirical evidence on sorting patterns, the firm-size/wage premium relation and job turnover.

The rest of 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 separate their applicants when the firms' learning rate is high enough; but when information frictions are sufficiently severe, a fraction of high-wage firms offer pooling contracts and end up employing more 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 discusses the robustness of the main results. Section 7 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. Inderst (2005) analyzes existence of separating equilibria in a model of random search with adverse selection. In his model the composition of the pool of searching individuals evolves over time. However, once a productive match is formed and a contract agreed, the pair leaves the market. To the best of our knowledge, there are only two papers with on-the-job search under adverse selection. Kugler and Saint-Paul (2004) analyze the effects of firing cost on different types of workers in a model with search on-the-job, assuming however an ad-hoc wage schedule. This is very different from this paper which is interested in optimal wage policies under adverse selection. Visschers (2007) considers a model with random search based on Stevens (2004) and assumes that both the worker and 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.

A few papers consider the interaction of search frictions and adverse selection to study firms' decisions to offer a take-it-or-leave-it wage offer or to engage in bilateral bargaining with their job applicants. Camera and Delacroix (2004), for example, consider a random search model, while Michelacci and Suarez (2006) consider a directed search model to address this issue. As in our paper, firms choose between different types of contracts which impacts the type of workers they employ. Michelacci and Suarez (2006) shows that when firms are indifferent between the two, the market segments and firms

posting wages attract workers with low productivity, while the firms that bargain attract high productivity workers.⁵ In our paper we restrict attention to wage posting and let firms choose between offering separating contracts to hire both types of workers at different wages or posting a pooling contract that provides a higher retention rate for high-ability workers.

This paper also 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 are constrained in match formation (e.g. Shimer and Smith (2000)).

Finally, this work 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, Fontaine, Postel-Vinay, and Robin (2009), 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.

There are two types of workers who differ in their innate ability. 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

⁵Interestingly, there is no segmentation in the random search model proposed by Camera and Delacroix (2004).

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. We assume that firms monitor the output of a particular worker at rate ρ . This parameter describes the firm's learning rate.⁶ 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. In Section 6 we explore the implications of this assumption.

Unemployed and employed workers meet firms according to a Poisson process with parameter $\lambda > 0$. Once a meeting takes place, the firm offers a menu of contracts to the worker. A contract consists of two elements: (i) a wage and (ii) a firing policy. We assume that a firm can fully commit to both parts of the contract. More specifically, a firm offers the wage w_i to workers of ability $i = L, H$. If a worker truthfully reports his type upon hiring, the firm commits not to fire the worker and to pay w_i for the rest of the employment relation. If a worker misrepresented his type, he is paid w_i until the firm learns his true type. At that point, the firm issues the punishment and fires the worker. As we restrict the analysis to constant-wage contracts and rule out any further transfers between workers and firms, a layoff is the only possible form of punishment to a deviating worker. In Section 6 we explore the implications of relaxing our constant wage assumption, allowing the firm to cut the wage of the misreporting worker. We also discuss the assumption of commitment to the firing policy.

In the following, we identify contracts by the wage paid to a worker who reports type i . Let $F_i(w_i)$ denote the proportion of firms offering a wage no greater than w_i to workers of type 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 $w_H \in [\underline{w}_H, \overline{w}_H]$.⁷

When a worker meets a firm, the worker observes the posted contracts 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 contracts are rejected, however, the worker remains in his current state with no option to recall previously met firms. We make the following tie-breaking assumptions: an unemployed worker accepts a wage offer if indifferent to accepting it or remaining unemployed, while an employed

⁶The implicit assumption here is that the firm observes total output, but since it employs a mass of workers it is too costly to observe the output of each individual worker immediately.

⁷The 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.

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 $b < \varepsilon_{LP}$ 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 arguments, the Bellman equation that describes U_i is given by

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

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

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

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 unemployment at rate ρ ; hence V_{ij} solves

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

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 who (i) currently receives 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 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 $i, k = L, H$.⁸

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 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 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) 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 layoff 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 a layoff on the right side.⁹

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 and truth-telling 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.

⁸Assuming that unemployed and employed workers meet firms at the same rate and have the same flow value of unemployment considerably simplifies the worker's problem because *all* unemployed workers have the same reservation wage which is independent of firms' wage offer strategies. In Section 6 we discuss the implications of different reservation wages.

⁹Note that it also follows from (1) and (2) that (3) is equivalent to $w_j - w_i \leq \rho[V_{ii}(w_i) - U_i]$.

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

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

(ii) Workers' search and truth-telling 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 when 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 the Burdett and Mortensen (1998) model 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 profit¹⁰

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

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 equilibrium 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 $i = L, H$.

¹⁰Each 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 derived in a similar way as (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 $\Omega_i(w_i) = h_i(w_i)J_i(w_i)$.

It is easy to verify that $\varepsilon_H > \varepsilon_L$ implies that $F_H(\cdot)$ first order stochastically dominates $F_L(\cdot)$. In equilibrium more able workers face more frictional wage dispersion and are paid, on average, higher wages than low-ability 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 by¹¹

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

Workers' reservation wage strategies are such that unemployed workers accept 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. In this case, of course, all workers face the same frictional wage dispersion.

3 Equilibria with Limited Information

We now explore the case in which (positive and finite) search and information frictions coexist in the labor market. We show that when the learning rate of firms is sufficiently high, all firms offer separating contracts. Both types of workers truthfully reveal their type and self-select into the appropriate wage.

¹¹A 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 expected profit value $J(w) = [\tilde{\varepsilon}p - w]/[\phi + \delta + \lambda(1 - F(w))]$. It follows that $\Omega(w) = h(w)J(w)$.

For lower values of ρ , however, we show there exist segmented equilibria in which low-wage firms offer separating contracts, while high-wage firms offer pooling contracts.

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 drawn from distributions (5) satisfying (6) is an equilibrium in the imperfect information economy if and only if,*

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

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 separation rate $(\phi + \delta)$ or a low job offer arrival rate reduce the incentive to misreport since workers are more likely to be laid off once the firm learns ability.

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 separating contracts. We fully characterize this type of equilibrium in Appendix A.

Proposition 2: *For values of $\rho \in [\rho_2, \rho_1)$, an equilibrium with wage dispersion exists in which all firms offer separating 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 (see Appendix A). 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 separate 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.

Conversely, if ρ is smaller than ρ_2 , incentive constraints must bind for all firms offering separating contracts. However, not all firms may prefer to offer separating contracts because it can be too costly to provide the necessary information rents to low-ability workers. We now characterize equilibria for values of $\rho < \rho_2$ and derive conditions for existence.

3.2 Binding Incentive Constraints

Suppose that $\rho \leq \rho_2$ and consider a firm which offers w_H to high-ability workers. Regarding the wage offered to low-ability workers, $w_L = \hat{w}(w_H)$, this firm has two options. Either, it offers a high enough wage that is incentive-compatible and prevents low-ability workers from misreporting. Given that incentive constraints are binding for $\rho \leq \rho_2$, equation (3) implies that

$$w_L = \hat{w}(w_H) = w_H - \rho[V_{LH}(w_H) - U_L] . \quad (10)$$

Such a firm offers *separating contracts* to the two workers. Alternatively, the firm may find the contract to low-ability workers too costly and instead decides to violate the incentive constraint:

$$w_L = \hat{w}(w_H) < w_H - \rho[V_{LH}(w_H) - U_L] . \quad (11)$$

In this case, low-ability workers misreport their type when they meet this firm; they earn w_H until the firm learns their ability and lays them off. Equivalently, we can interpret this wage policy as a *pooling contract* since this firm achieves the same outcome by simply offering one contract: a wage w_H in combination with a layoff commitment to low-ability workers. Both worker types accept this contract, although their expected income patterns differ ex-post. Without loss of generality and to keep the notation consistent throughout, we will specify the analysis in terms of the equivalent menu of wages (w_H, w_L) where $w_L = \hat{w}(w_H)$ and \hat{w} satisfies (11).

We can prove that separating contracts always dominate pooling contracts at the lower end of the wage offer distribution. But at higher wages, pooling contracts can possibly dominate separating contracts.

Proposition 3: *Consider any given distribution of wage offers to high-ability workers F_H with support $[b, \bar{w}_H]$. Then there is a threshold wage $\tilde{w}_H \in (b, \bar{w}_H]$ such that a firm offers a separating contract if $w_H \leq \tilde{w}_H$ and a pooling contract if $w_H > \tilde{w}_H$.*

To understand why low-wage firms always prefer to offer separating contracts, consider a firm offering w_H close to the reservation wage b . For this firm, it is not very costly to prevent low-ability workers from misreporting because the worker's gain from doing so, $V_{LH}(w_H) - U_L$, is rather small. Conversely, when w_H is high, the incentive-compatible wage w_L is also high and may even exceed the worker's marginal product ε_{LP} . Then, in some situations, the firm may decide to offer a pooling contract which brings about a lower job-retention rate for low-ability workers.

Given the structure of wage policies described in Proposition 3, we characterize an equilibrium where a fraction $\eta \in (0, 1]$ of firms offer separating contracts and fraction $1 - \eta$ offering pooling contracts, with η to be endogenously determined below. Separating firms offer $w_H \in [b, \tilde{w}_H]$ to high-ability workers and the incentive-compatible wage $w_L = \hat{w}(w_H) \leq \tilde{w}_L$ to low-ability workers, satisfying (10) with $\tilde{w}_L = \hat{w}(\tilde{w}_H)$. Pooling firms target high-ability workers by offering $w_H > \tilde{w}_H$ and $w_L = \hat{w}(w_H)$, violating the incentive-compatibility constraint (i.e. satisfying (11)). Given the rank-preservation property, we maintain that function \hat{w} is strictly increasing and continuous.¹²

This 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 strictly above the one he is currently earning. Similarly, if a low-ability worker employed in a firm earning $w_L \leq \tilde{w}_L$ meets another firm offering $w'_L \leq \tilde{w}_L$, he accepts the job offer if he is promised a wage $w'_L > w_L$. If this worker meets a firm offering a wage $w'_H > \tilde{w}_H$, the worker will also accept the offer and misreport the type. Lastly, consider a low-ability worker that is earning $w_H \geq \tilde{w}_H$ before the firm learns ability. This worker then accepts any wage $w'_H > w_H$ from another firm. If the worker meets any firm offering $w'_L \leq \tilde{w}_L$, it follows that the worker will not accept such an offer.¹³

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. Also note that 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 $\eta < 1$. We now proceed to solve for equilibrium.

¹²It 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 this type of firm.

¹³The current job strictly dominates truth-telling at the current employer which itself strictly dominates the outside offer $w'_L \leq \tilde{w}_L < w_L = \hat{w}(w_H)$ and $V_{LL}(w_L) < V_{LH}(w_H)$.

3.2.1 Steady-state Measures

To simplify notation, it is useful to let the quit rate of high-ability workers earning w_H be denoted by $q(w_H) \equiv \phi + \delta + \lambda(1 - F_H(w_H))$ and note that $\eta = F_H(\tilde{w}_H)$. Given the reservation wage strategies described above, we show in Appendix C that steady-state turnover implies that unemployment of workers of both types is given by

$$u_H = \frac{(\phi + \delta)\alpha_H}{\phi + \delta + \lambda} \quad \text{and} \quad u_L = \frac{q(\tilde{w}_H)}{q(\tilde{w}_H) + \rho} \frac{(\phi + \delta + \rho)\alpha_L}{\phi + \delta + \lambda}, \quad (12)$$

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

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

Further, for all $w_L \in [b, \tilde{w}_L]$, the earnings distribution of low-ability workers is given by

$$G_L(w_L) = \frac{q(\tilde{w}_H)(\phi + \delta + \rho)F_H(\hat{w}^{-1}(w_L))}{(q(\tilde{w}_H) + \rho\eta)q(\hat{w}^{-1}(w_L))}. \quad (14)$$

The earnings distribution for all $w \in [\tilde{w}_H, \bar{w}_H]$ is given by

$$G_L(w_H) = \frac{(\phi + \delta + \rho)[\eta\rho + q(\tilde{w}_H)F_H(w_H)]}{(q(w_H) + \rho)(q(\tilde{w}_H) + \rho\eta)}. \quad (15)$$

3.2.2 Firms' Payoffs

Consider a separating 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 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 by¹⁴

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

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

Now consider a pooling firm that offers $w_H > \tilde{w}_H$ to high-ability workers and a non-incentive-compatible wage to low-ability workers. Since low-ability workers 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 $i = L, H$. Using (12), (13) and (15), implies that this firm's steady-state profit is given by

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

¹⁴Appendix C contains a full derivation of the expressions in (16) and (17).

3.2.3 Wage-Offer Distributions

To solve for the equilibrium wage offer distributions, first consider a 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

$$\hat{w}'(w_H) = \frac{q(w_H)}{q(w_H) + \rho}, \quad (18)$$

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

Further, in any equilibrium a firm offering $w_H \leq \tilde{w}_H$ with associated wage $w_L = \hat{w}(w_H)$ 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 wage \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 and \tilde{w}_H .

Next consider a firm offering a wage $w_H > \tilde{w}_H$. Equilibrium requires that the profits of this firm must satisfy $\Omega^P(w_H, w_L) = \Omega^S(b, b)$. Substituting out the corresponding expressions leads to the following characterization of the wage offer distribution for $w_H > \tilde{w}_H$.

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

$$F_H'(w_H) = \frac{q(w_H)(q(w_H) + \rho)[(\phi + \delta)\alpha_H(q(w_H) + \rho)^2 + (\phi + \delta + \rho)\alpha_L q(w_H)^2]}{2\lambda[(\phi + \delta)\alpha_H(\varepsilon_{HP} - w_H)(q(w_H) + \rho)^3 + (\phi + \delta + \rho)\alpha_L(\varepsilon_{LP} - w_H)q(w_H)^3]}$$

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

Similar to Lemma 1, we require $F_H(\bar{w}_H) = 1$ to characterize 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 η^* . Given that equilibrium requires $\eta^* > 0$, there are two possible cases: (i) Segmented equilibria in which $\eta^* \in (0, 1)$ and separating and pooling contracts

coexist. (ii) Equilibria in which $\eta^* = 1$ and all firms offer separating contracts. We analyze each in turn. When separating and pooling contracts coexist, firms must be indifferent between the two types of contract. In particular, at the threshold wage \tilde{w}_H , this necessitates $\Omega^S(\tilde{w}_H(\eta), \tilde{w}_L(\eta)) = \Omega^P(\tilde{w}_H(\eta), \tilde{w}_L(\eta))$. Using (16) and (17), this condition implies that $\eta^* \in (0, 1)$ must solve the following fixed point problem

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

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 separating contracts, equilibrium requires that $\Omega^S(\tilde{w}_H(\eta), \tilde{w}_L(\eta)) \geq \Omega^P(\tilde{w}_H(\eta), \tilde{w}_L(\eta))$ at $\eta = 1$. With T defined in (19), this is equivalent to $T(1) \geq 1$ being a necessary condition for existence of a pure separating equilibrium.

The proof of Theorem 1 below shows that the function T has at most one fixed point and that a unique equilibrium exists. We also prove that equilibrium is segmented, i.e. some firms offer pooling contracts, provided that the learning rate ρ is sufficiently low and a parameter condition is satisfied.

Theorem 1: *A Market Equilibrium with $\eta^* \in (0, 1]$ exists and is unique. Moreover, if*

$$\lambda^2(\tilde{\varepsilon}p - b) > (\phi + \delta + \lambda)^2(\varepsilon_{LP} - b) \quad (20)$$

holds, there is a threshold value $\rho_3 \in (0, \rho_2)$ such that pooling and separating contracts coexist if $\rho < \rho_3$. Otherwise all firms offer separating contracts.

Condition (20) is a necessary condition for pooling contracts to be profitable for the highest-wage firms. Intuitively, if productivity of low-ability workers is relatively low and the wage-offer distribution is sufficiently dispersed (λ is sufficiently large), it is too costly for high-wage firms to provide the necessary information rents to separate low-ability workers, so that a pooling contract with firing of low-ability workers is the more attractive option. In the proof of Theorem 1 we show that the threshold value ρ_3 is the implicit solution of equation

$$(\phi + \delta)(\bar{w}_H - \bar{w}_L) = \rho(\bar{w}_L - \varepsilon_{LP}), \quad (21)$$

with \bar{w}_L and \bar{w}_H defined by (7) and (8). The proof also reveals that at the threshold wage and beyond, firms make negative profits on low-ability workers: $\varepsilon_{LP} < \tilde{w}_L(\eta) < \tilde{w}_H(\eta)$. This implies that the firing of low-ability workers at high-wage firms is ex-post optimal.

Propositions 1, 2 and Theorem 1 together imply that the set of equilibria can be partitioned in terms of the degree of information frictions through the firms' learning rate. Figure 1 depicts this partition. For values of $\rho < \rho_3$, those firms who offer the highest wages find it too costly to offer incentive-compatible contracts to low-ability workers. They instead decide to offer pooling contracts. These contracts are accepted by both worker types, but low-ability workers have lower job-retention rates. For values of $\rho \geq \rho_3$ all firms offer separating contracts. Incentive constraints bind for all firms if $\rho < \rho_2$ and for a fraction of firms if $\rho \in [\rho_2, \rho_1]$; incentive constraints are slack for all firms if $\rho \geq \rho_1$.

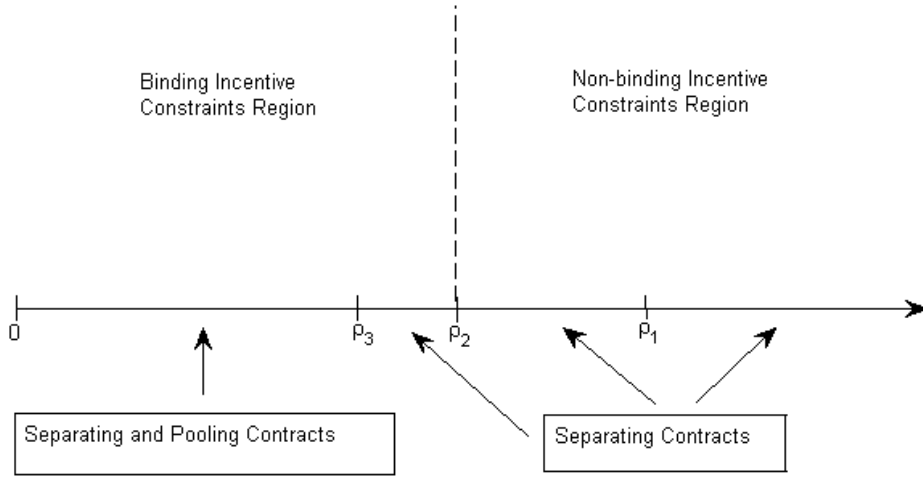


Figure 1: Set of equilibria parameterized by ρ

4 Implications

In this section we analyze some of the implications for labor turnover and wage dispersion. First, we show that in a segmented equilibrium low-ability workers have a higher degree of turnover and they are underrepresented in high-wage firms (offering pooling contracts) and overrepresented in low-wage firms (offering separating contracts). Formally, using (36) and (37) in Appendix C, the workforce sizes of low-ability and high-ability workers employed at a separating firm offering $w_H \leq \tilde{w}_H(\eta)$ are given by

$$n_L^S(\hat{w}(w_H)) = \frac{\lambda\theta(\eta)\alpha_L}{q(w_H)^2} \quad \text{and} \quad n_H^S(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{q(w_H)^2}, \quad (22)$$

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

$$n_L^P(w_H) = \frac{\lambda(\phi + \delta + \rho)\alpha_L}{[q(w_H) + \rho]^2} \quad \text{and} \quad n_H^P(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{q(w_H)^2}. \quad (23)$$

It is then easy to verify that firms offering separating contracts have a higher proportion of low-ability workers in their workforces, while firms offering pooling contracts have a higher proportion of high-ability workers. Furthermore, since (23) implies that $n_H^P(w_H)/(n_L^P(w_H) + n_H^P(w_H))$ is increasing in w_H , the proportion of high-ability workers is increasing in w_H among pooling firms. The intuition is that high-wage firms are able to attract and retain more workers of both types, while they detect misreporting low-ability workers at the same rate ρ , independent of the offered wage. We summarize these findings as follows.

Proposition 4: *If $\rho \geq \rho_3$, both worker types 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. Firms offering pooling contracts (high-wage firms) have a more productive workforce than firms offering separating contracts (low-wage firms). Among high-wage firms, the workforce productivity is increasing in w_H .*

A further immediate consequence of our model is that low-ability workers have higher unemployment as they are more likely to be laid off at firms offering pooling contracts. This implies that total output is lower in the economy with small information frictions ($0 < \rho < \rho_3$), both relative to the full-information benchmark and relative to the no-information case ($\rho = 0$).

4.1 Numerical Example

Since our model cannot be fully solved in closed form, we use a numerical example to illustrate how wage dispersion changes with information frictions and to study the relation between wages, firm size and workforce ability.

Consider the following parametrization. Set the time period to a month and let $\phi = 0.0018$ reflect an average working life of 45 years. Following Hornstein, Krusell, and Violante (2010), set $\delta = 0.036$ 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 $\alpha_H = 0.75$, $\alpha_L = 0.25$. We normalize $p = 1$ 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 $\rho = 0.16$ as a benchmark. This number implies that on average firms learn their employees' true type after six months of employment.

The above parametrization implies that in equilibrium 69.2 percent of firms offer separating contracts (i.e. $\eta^* = 0.692$). It also implies that $\rho_1 = 1.34$, $\rho_2 = 0.86$ and $\rho_3 = 0.52$. That is, pure separating equilibria can only be sustained when firms learn the true type of their workers on average at the second month, $1/\rho_3$, of employment. Given that the latter number seems to require very fast learning from employers, a segmented equilibrium is quite plausible in our benchmark parametrization.

4.1.1 Wage Dispersion

In equilibrium separating firms compress the wages offered to the two types so as to enforce self-selection of low-ability workers. As implied by Theorem 1, some firms offer wages above the productivity of low-ability workers, with $\tilde{w}_L(\eta^*) = 1.205$ and $\tilde{w}_H(\eta^*) = 1.625$. In turn, the wage policies of these firms affect the wages offered to high-ability workers by pooling firms. In particular, the highest wage offered to high-ability workers is $\bar{w}_H(\eta^*) = 1.883$, which is smaller than the upper bound in the perfect information case where $\bar{w}_H = 1.943$.

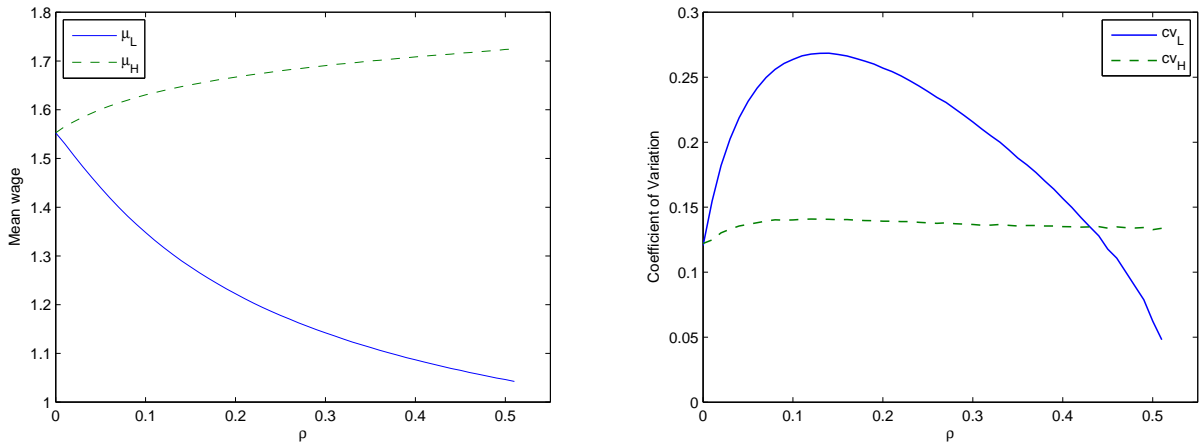


Figure 2: Mean wages and coefficients of variation for G_L and G_H and varying ρ .

Note that the earnings distributions G_H and G_L reflect wage dispersion that arises purely due to search and information frictions. By computing their standard deviations, the benchmark parametrization implies that low-ability workers face a more dispersed distribution and hence more frictional wage dispersion than high-ability workers. Figure 2 depicts the mean and coefficient of variation (cv) of these distributions as we change the firms' learning rate. As we increase ρ from zero (no information) to $\rho_3 = 0.52$ (where all firms offer separating contracts), $\tilde{w}_L(\eta)$ decreases and $\tilde{w}_H(\eta)$ and $\bar{w}_H(\eta)$ increase. While the mean and the cv of G_H increase and the mean of G_L decreases, the cv of G_L follows a non-monotonic pattern. When ρ is increased further, all these moments move in the direction of their perfect information values.¹⁵ 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 $w_L = w_H = w \in [b, \bar{w}]$, where $\bar{w} = \tilde{\varepsilon}p - [(\phi + \delta)/(\phi + \delta + \lambda)]^2(\tilde{\varepsilon}p - b) = 1.706$. In this equilibrium the firm never learns the worker's true type, so it is unable to separate workers or to fire low-ability workers later on.¹⁶

For values of $\rho < 0.43$, where separating and pooling contracts coexist in the market, low-ability workers face higher frictional wage dispersion than high-ability workers. For higher values of ρ the opposite holds. Intuitively, as ρ increases away from zero, the wage offer distribution to high-ability workers becomes a bit more dispersed. At the same time, there is a group of low-ability workers who earn temporarily rather high wages ($w > \tilde{w}_H$) while many others earn low wages at separating firms ($w < \tilde{w}_L$). This segmentation dramatically drives up wage dispersion for low-ability workers. As ρ

¹⁵Figure 2 shows the mean and the cv for $\rho \leq \rho_3 = 0.52$. For $\rho > \rho_3$ both measure converge monotonically to their perfect information values which are $E(w_H) = 1.75$, $E(w_L) = 0.97$, $cv(w_H) = 0.137$, $cv(w_L) = 0.027$.

¹⁶Interestingly, in the limit $\rho \rightarrow 0$, equilibrium is segmented with a share $\eta = 0.49$ of firms offering separating contracts. However, since all "separating" firms offer nearly the same wage to any worker they meet, and "pooling" firms never (in the limit) fire workers, the equilibrium structure is identical to the one of the no-information case.

increases further, however, the fraction of separating firms η converges to unity and almost all workers are employed in separating firms which reduces dispersion.¹⁷

4.1.2 Ability, Firm Size and Wages

Proposition 4 shows that low-ability workers have a higher degree of turnover than high-ability workers. This turnover pattern generates a positive relation between workforce productivity, firm size, and wage strategies. First recall that firms offering pooling contracts have a higher proportion of high-ability workers. Next note that the average workforce size among firms offering separating contracts 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 firms offering pooling contracts is given by

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

The numerical solution for the above expressions shows that the average size of the workforce in pooling firms is greater than that of separating firms; $E(n^S) = 0.443$ and $E(n^P) = 1.389$. This result shows that, in our benchmark parametrization, firms that employ a more productive workforce are bigger on average. Obviously, pooling firms offer higher wages than separating firms; in our parametrization we compute the average wages offered by the two types of firms as $E(w^S) = 1.21$ and $E(w^P) = 1.76$. 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).¹⁸

5 Firm Heterogeneity and Sorting

We now extend our basic model to include heterogeneity in firm productivity. The main aim is to analyze whether there exists sorting by types. That is, do more productive firms attract and retain a more productive workforce? We show that such a sorting pattern obtains in our adverse selection model, although it does not obtain in the corresponding perfect information or no information benchmarks. Sorting equilibria are consistent with empirical evidence showing that the positive relation between firm

¹⁷The general properties of these graphs remain unchanged when we alter the share of low-ability workers α_L . For example, increasing α_L implies that cv_L intersects cv_H at lower values of ρ , while reducing α_L generates the opposite effect.

¹⁸It is important to note that when separating and pooling contracts coexist, the relationship between firm 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 a firm offering \tilde{w}_H is greater than the size of the a firm offering $\tilde{w}_H + \varepsilon$, $\varepsilon > 0$ is sufficiently small. Among pooling firms, the relation between wages, firm size and workforce productivity is monotonic, however (see Proposition 4).

size and wages is not only due to the fact that high-wage firms have a more productive workforce, but that these firms are themselves more productive (see Haltiwanger et al. 1999).

The formal analysis of this case is very similar to the one with homogeneous firms and is relegated to Appendix B. 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 . 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 \overline{w}_{ik} denote the infimum and the supremum of the support of $F_i(\cdot | p_k)$. Lemma a.1 in Appendix B shows that in equilibrium more productive firms offer higher wages than less productive firms, and that separating firms offer lower wages than pooling firms. There are then two equilibrium configurations of interest: Equilibria in which $\eta \in (0, \beta_L)$ such that low-productivity firms offer separating and pooling contracts and all high-productivity firms offer pooling contracts; and equilibria in which $\eta \in (\beta_L, 1]$ and all low-productivity firms offer separating contracts, while high-productivity firms offer separating and pooling contracts. Again, equations (22) and (23) imply that pooling firms have a more productive workforce than separating firms. Furthermore, among pooling firms, the workforce productivity is increasing in the offered wage. This directly yields the following result:

Proposition 5: *Consider a segmented equilibrium in which one type of firms offer either separating or pooling contracts. Then some (or all) high-productivity firms employ a higher proportion of high-ability workers than low-productivity firms.*

Hence 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, positive sorting obtains.

As in the homogeneous case, total output is reduced because low-ability workers have lower employment rates than high-ability workers in any segmented equilibrium. With firm heterogeneity, the positive sorting entails an additional loss of output relative to the no-sorting benchmarks that obtain under either full information or no information. The explanation is as follows. The turnover pattern of high-ability workers does not depend on the amount of information: their relative employment shares do not vary with the firms' learning rate. Low-ability workers, however, are more likely to be employed in low-productivity firms when information frictions are sufficiently strong. Hence, low-productivity firms employ a larger share of the total labor force, which ultimately reduces aggregate output. We note that this loss of output due to sorting also emerges in the variation of this model where firms cut wages instead of firing misreporting workers and where all workers have the same employment rates (see Section 6.1).

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 equilibrium conditions described in Appendix B. We use the same parameter values as before, but set $p_L = 1$ and $p_H = 1.1$. The value of β_L becomes important to determine which type of equilibrium is obtained. We consider two values for $\beta_L \in \{0.5, 0.9\}$ 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 (2008)).

In the case of $\beta_L = 0.9$, we have that $\eta^* = 0.744$, $\bar{w}_{HL} = \underline{w}_{HH} = 1.615$, and $\bar{w}_{HH} = 1.797$. All high-productivity firms offer pooling contracts, whereas low-productivity firms offer both types of contracts. As should be expected, the average size of high-productivity firms is greater than that of low-productivity firms, $E_H(n^P) = 2.073 > E_L(n^P + n^S) = 0.592$. Further, low-productivity firms employ a larger share of low-ability workers (27%) compared to high-productivity firms (7.4%). Under perfect information (or with zero information), in contrast, both types of firms have a balanced workforce. Relative to these no-sorting situations, low-productivity firms employ a larger number of low-ability workers but the same number of high-ability workers. Hence total output is lower under asymmetric information (1.361 relative to 1.4 without sorting).

Now let $\beta_L = 0.5$. In this case we have that $\eta^* = 0.843$, $\bar{w}_{HL} = \underline{w}_{HH} = 1.495$, and $\bar{w}_{HH} = 2.008$. Then all low-productivity firms separate workers whereas high-productivity firms offer both pooling and separating contracts. Again, the average size of high-productivity firms is greater than that of low-productivity firms, $E_H(n^P + n^S) = 1.192 > E_L(n^S) = 0.313$, and low-productivity firms employ a larger share of low-ability workers than high-productivity firms (31.8% relative to 20.4%). The output loss of asymmetric information is tiny in this example (1.441 compared to 1.466 without sorting).

Note that in both cases the average wages earned in high-productivity firms are higher than the average wages earned in low-productivity firms. Hence, more productive firms not only employ a more productive workforce, but they also offer, on average, higher wages and are bigger than less productive firms.

6 Robustness

In this section we relax several of the assumptions made earlier. We argue that the qualitative features of the benchmark model still hold in each of the following extensions.

6.1 Wage Cuts

In the benchmark model we assume that a layoff is the only form of punishment for a misreporting worker. An alternative is to allow the firm to cut the wage of these workers, while maintaining the

assumption that firms offer flat-wage contracts for all workers who report truthfully. In this case, it is easy to see that cutting the wage of the misreporting worker to his reservation wage dominates the firing of this worker. By using the alternative punishment, the firm continues to extract rents out of this worker (at least for some time), while it is an equally strong threat as it provides the same expected payoff to the worker as a layoff. We now briefly explore the implications of this variation. Details are available from the authors upon request.

First note that Propositions 1 and 2 are exactly the same in this case. It is only when incentive constraints bind for all firms that the analysis is modified. In particular, we can extend Proposition 3 to show that low-wage firms always offer separating contracts, and we can also show that segmented equilibria exist where a fraction of high-wage firms offer pooling contracts, particularly when the learning rate is sufficiently low. In any segmented equilibrium, the equilibrium fraction of separating firms η solves the fixed point problem

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

where $\tilde{w}_L(\eta)$ and $\tilde{w}_H(\eta)$ are obtained in the same way as in the benchmark model. Furthermore, \widehat{T} is an increasing and convex function of η with $\widehat{T}(0) = \widehat{T}'(0) = 0$, such that $\widehat{T}(1) > 1$ implies the existence of a segmented equilibrium; conversely if $\widehat{T}(1) \leq 1$, all firms offer separating contracts.

In a segmented equilibrium, pooling firms cut wages of low-ability workers. Again, these workers have higher turnover rates than high-ability workers, and pooling (high-wage) firms have a more productive workforce than separating (low-wage) firms. All workers have the same employment rates, however. Numerical examples show that the main implications for wage dispersion and for the relation between firm size, workforce productivity and wages still hold in this case. An interesting difference, however, is that pooling firms offer higher wages to high-ability workers than in the benchmark model. This follows because these firms gain more from low-ability worker once they are paid their reservation wage after being caught misreporting. Hence they can bid more aggressively for high-ability workers. When firms have heterogeneous productivity, the analysis presented in the benchmark model remains basically unchanged. In particular, the sorting pattern described in Proposition 5 extends to this case.

6.2 Limited Commitment on Firing

In the benchmark model we assume that firms are able to commit to fire misreporting workers. Likewise, they are able to commit not to fire workers who report their ability truthfully. As we show in the proof of Theorem 1, in equilibrium only those workers are fired who actually yield losses to the firm; hence equilibrium firing is ex-post optimal in the benchmark model. Nonetheless, on the equilibrium path, some separating firms retain low-ability workers even though they yield losses (the firms offering $w_L \in (\varepsilon_{LP}, \tilde{w}_L]$). Ex-post these firms would prefer to fire these workers. Furthermore, all off-equilibrium

punishment strategies rest on the commitment assumption. What would happen if firms can only commit to a flat wage but are unable to commit to a firing policy?

Any limitation on commitment clearly makes it more difficult for firms to separate workers. Specifically, at the lowest wages firms are not able to credibly announce that they will fire misreporting workers. Those low-ability workers who misreport high ability and earn $w_H < \varepsilon_L p$ would still yield positive profit, so firms would like to retain these workers ex-post. Hence, low-wage firms will only be able to offer pooling contracts; these are accepted by both types of workers and no worker will be fired. Conversely, at the highest wages firms are unable to separate workers because they cannot commit to retain low-ability workers earning wages $w_L = \hat{w}(w_H) > \varepsilon_L p$. These firms, again, can only offer pooling contracts, but they would fire low-ability workers (as they do in the benchmark model if $\rho < \rho_3$). The only firms that are able to separate workers are those offering $w_H \geq \varepsilon_L p \geq w_L = \hat{w}(w_H)$: these firms credibly fire misreporting workers and they credibly retain truthful workers of low ability. Therefore, the equilibrium patterns are very similar to those of the benchmark model. As long as a positive fraction of high-wage firms offer pooling contracts and fire low-ability workers, the main qualitative predictions remain the same.

6.3 Different Reservation Wages

Our assumptions that unemployment income is the same for all workers and that job arrival rates are the same for employed and for unemployed workers imply that workers of both types have the same reservation wage. If any of these two assumptions is violated, reservation wages would differ. For example, if unemployed workers have a higher job-arrival rate than employed workers, the reservation wage exceeds unemployment income since unemployed workers want to be compensated for giving up the higher option value of search. Since high-ability workers draw higher wage offers from firms offering separating contracts, they also have a higher reservation wage. How does the equilibrium structure change when high-ability workers have a higher reservation wage?

A first observation is that the full-information outcome cannot be an equilibrium, even when the firms' learning rate is very large (that is, Proposition 1 fails). To see this, note that the full-information outcome has wage offer distributions F_i whose lower bounds are the corresponding reservation wages $R_H > R_L$. But then for any finite ρ , no matter how large, low-ability workers employed at their reservation wage (or at any wage close to it) would misreport their type: They could earn the higher reservation wage of high-ability worker temporarily, but they would not suffer from a layoff which gives the same utility as truth-telling.

This implies that any equilibrium with finite ρ must involve binding incentive constraints, at least at the lowest wages. In particular, the firm with the lowest wage offer to high-ability workers $w_H = R_H$ must provide some rents to low-ability workers, $w_L = \hat{w}(R_H) > R_L$, if it wants to enforce truth-telling. At higher wages, as in our benchmark model, some firms possibly find it profitable to offer pooling

contracts for high-ability workers. At the same time, however, there can also be a fraction of firms that decide to offer pooling contracts exclusively for low-ability workers at the lowest wages $w_L \in [R_L, w_L^0]$ with some $w_L^0 < R_H$. Such an outcome would resemble features of the Albrecht and Axell (1984) model where a fraction of firms hire only low-reservation-wage workers while others hire all types of workers at higher wages. As long as some firms in the higher wage range offer pooling contracts to high-ability workers, our main implications remain intact, however: low-ability workers have higher turnover, and they are underrepresented in high-wage firms and overrepresented in low-wage firms (and even more so when pooling contracts for low-ability workers are offered).

6.4 Information on Workers' Employment Status

There is a large literature that consider the implications of adverse selection for unemployment (see Greenwald (1986), Gibbons and Katz (1991), and Kugler and Saint-Paul (2004), among others). In these models the composition of unemployed workers becomes biased towards low-ability ones as firms lay off these workers more frequently. As a consequence the re-employment wages of all unemployed workers decrease due to adverse selection.¹⁹ In our model, a segmented equilibrium implies that low-ability workers experience unemployment more often than high-ability workers and hence the unemployment rate is also biased towards low-ability workers. This can imply that firms have an incentive to further differentiate their contracts based on the employment status of their applicants.

Even without adverse selection, however, firms in the Burdett and Mortensen (1998) framework have an incentive to differentiate their wage offers by employment status so as to extract additional rents; see Carrillo-Tudela (2009) for a formal analysis. If all unemployed workers have the same reservation wage, information on employment status perfectly reveals the lowest wage the firm has to offer to hire from the unemployment pool. In equilibrium, the offer distribution faced by unemployed workers degenerates to a mass point at the reservation wage. For employed workers, the equilibrium offer distribution is characterized by the Burdett-Mortensen distribution, where the infimum of this distribution equals the reservation wage. When unemployed workers have different reservation wages, the offer distribution faced by these workers is similar to the one described in Albrecht and Axell (1984). The offer distribution for employed workers continues to be described by the one of the Burdett-Mortensen model.

The arguments in Carrillo-Tudela (2009) can also be applied under adverse selection. In particular, in the benchmark model all firms offer unemployed workers their reservation wage, but they differentiate their contracts (either offering separating or pooling contracts) when hiring employed workers as described in Section 3.2. Hence our main implications are not altered in this case. When workers have different reservation wages the analysis becomes more cumbersome. The arguments in Carrillo-Tudela (2009), however, suggest that the offer distribution faced by unemployed worker will be described by

¹⁹It is important to note that in these models firms are (by assumption) not allowed to offer separating contracts.

two mass points each at the corresponding reservation wages. Our conjecture is that the higher proportion of low-ability workers in the unemployment pool will mainly affect the dispersion of the offer distribution faced by unemployed workers.²⁰

7 Conclusions

In this paper we consider a model of the labor market in which search frictions coexist with information frictions. The latter arise as firms do not observe worker ability upon hiring but gradually learn it over time. Given this adverse selection problem, we show that when the learning rate is sufficiently low, a unique equilibrium emerges in which low-wage firms attempt to hire both low- and high-ability workers by offering incentive-compatible separating contracts. High-wage firms offer contracts that intend to retain only high-ability workers. In such a segmented equilibrium, low-ability workers have a higher degree of turnover and a more dispersed earnings distribution, and they are underrepresented in high-wage firms. 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 empirical results.

We extend our model and introduce firm heterogeneity to show that a segmented equilibrium implies positive sorting of workers among firms. High-productivity firms employ a more productive workforce, and low-ability workers are overrepresented in low-productivity firms. Total output is lower both relative to the no-information outcome where firms are unable to separate workers and relative to an equilibrium where all firms offer separating contracts, as is the case under perfect information. We also show that more productive firms offer, on average, higher wages, they are bigger and employ a more productive workforce, in line with the empirical evidence of 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 motivated to preserve a tractable analysis while considering two-sided heterogeneity; it is also based on evidence related to the lack of wage-tenure effects found in empirical studies. Nevertheless, adding this feature to the model is of interest as it would allow us to analyze the joint effects of adverse selection and search frictions in a more general contract space. We leave this extension for future research.

²⁰All these arguments rely on equilibria in which the supports of the offer distributions faced by unemployed workers do not overlap with the offer distributions faced by employed workers. As shown in Carrillo-Tudela (2009), it seems difficult to rule out equilibria with non-overlapping supports when firms are heterogeneous.

References

- ABOWD, J., F. KRAMARZ, AND D. MARGOLIS (1999): “High Wage Workers and High Wage Firms,” *Econometrica*, 67, 251–334.
- ALBRECHT, J., AND B. AXELL (1984): “An Equilibrium Model of Search Unemployment,” *Journal of Political Economy*, 92, 824–840.
- 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.
- CAMERA, G., AND A. DELACROIX (2004): “Trade Mechanism Selection in Markets with Frictions,” *Review of Economic Dynamics*, 4, 851–868.
- 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 Dispersion,” *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.
- KAMBOUROV, G., AND I. MANOVSKII (2009): “Occupational Specificity of Human Capital,” *International Economic Review*, 50(1), 63–115.
- 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.
- LENTZ, R., AND D. MORTENSEN (2008): “An Empirical Model of Growth through Product Innovation,” *Econometrica*, 76, 1317–1373.
- 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, 535–551.
- VISSCHERS, L. (2007): “Employment Uncertainty and Wage Contracts in Frictional Labor Markets,” Mimeo, University of Pennsylvania.
- WILLIAMS, N. (2009): “Seniority, experience, and wages in the UK,” *Labour Economics*, 16, 272–283.

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,

$$\varphi(w_L) \equiv V_{LL}(w_L) - V_{LL}(b) \geq \frac{(w_L - b)(\varepsilon_H - \varepsilon_L)p}{(\varepsilon_{LP} - b)\rho} \equiv \psi(w_L). \quad (24)$$

Equation (1) and the constant profit condition imply that

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

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 (24) 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

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

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

$$\Omega^{IC}(w_H) = \frac{\lambda(\phi + \delta)}{q(w_H)^2} \left[\tilde{\varepsilon}p - \alpha_H w_H - \alpha_L \hat{w}^{IC}(w_H) \right] = \frac{\lambda(\phi + \delta)(\tilde{\varepsilon}p - b)}{q(b)^2}. \quad (25)$$

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

$$F'_H(w_H) = \frac{(\phi + \delta + \lambda)^2 [\alpha_L q(w_H) + \alpha_H (q(w_H) + \rho)]}{2\lambda q(w_H)(q(w_H) + \rho)(\tilde{\varepsilon}p - b)}. \quad (26)$$

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

$$w_L = \hat{w}^S(w_H) \equiv \varepsilon_{LP} + \frac{\varepsilon_{LP} - \tilde{w}_L}{\varepsilon_{HP} - \tilde{w}_H} (w_H - \varepsilon_{HP}), \quad w_H \in [\tilde{w}_H, \bar{w}_H]. \quad (27)$$

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 (27) imply that this is true iff

$$\frac{\varepsilon_{LP} - \tilde{w}_L}{\varepsilon_{HP} - \tilde{w}_H} \geq \frac{q(\tilde{w}_H)}{q(\tilde{w}_H) + \rho}. \quad (28)$$

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

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

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

$$\frac{\varepsilon_{LP} - \tilde{w}_L}{\varepsilon_{HP} - \tilde{w}_H} \leq \frac{q(\tilde{w}_H)}{q(\tilde{w}_H) + \rho}. \quad (29)$$

These considerations show that in any equilibrium with $\gamma < 1$, (28) and (29) must hold with equality, whereas an equilibrium with $\gamma = 1$ (all firms face binding incentive constraints) must satisfy the weak inequality (29). At $\gamma = 1$, $\tilde{w}_L = \bar{w}_L$ and $\tilde{w}_H = \bar{w}_H$, and it follows that (29) coincides with (9). This

condition, therefore, implicitly pins down threshold parameter ρ_2 . For any $\rho \in (\rho_2, \rho_1)$, the binding condition (29) then defines the equilibrium value of $\gamma \in (0, 1)$. This equilibrium exists because the RHS in (29) 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 (26):

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

Furthermore, $\tilde{w}_H(\gamma)$ can be calculated from the constant-profit condition (25):

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

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:

Suppose that $\rho \leq \rho_2$ and consider a firm which offers w_H to high-ability workers. Such a firm makes the same profit on high-ability workers, irrespective of its wage offer to low-ability workers. This is because high-ability workers' acceptance and quit rates are independent of the wage offer $w_L \leq w_H$. Hence, the decision of what wage to offer to low-ability workers only depends on a firm's profit on low-ability workers. Let \hat{w}_{IC} denote the incentive-compatibility relation between w_H and w_L satisfying (10). Also denote by $q(w_H) \equiv \phi + \delta + \lambda(1 - F_H(w_H))$ the separation rate of a firm offering w_H .

If the firm decides to offer $w_L = \hat{w}_{IC}(w_H)$ to these workers (i.e. a separating contract), its profit per hire is $(\varepsilon_{LP} - \hat{w}_{IC}(w_H))/q(w_H)$. But if the firm offers $w_L = \hat{w}(w_H) < \hat{w}_{IC}(w_H)$ to low-ability workers (i.e. a pooling contract), all these workers accept wage w_H and are laid off back to unemployment at rate ρ . The firm's expected profit per low-ability hire is then $(\varepsilon_{LP} - w_H)/(q(w_H) + \rho)$. Now in both cases, the hiring rate of low-ability workers is the same: The firm hires all low-ability workers whose current lifetime utility is smaller than $V_{LH}(w_H)$, because of the binding incentive compatibility constraint, the worker would also obtain $V_{LL}(w_L) = V_{LH}(w_H)$ in a separating contract. Thus hiring rates are the same regardless of the type of contract offered. It follows that offering separating contracts dominates offering a pooling contract iff

$$\frac{\varepsilon_{LP} - \hat{w}_{IC}(w_H)}{q(w_H)} \geq \frac{\varepsilon_{LP} - w_H}{q(w_H) + \rho}.$$

Rewrite this condition as

$$\Phi(w_H) \equiv \rho(\varepsilon_{LP} - \hat{w}_{IC}(w_H)) + q(w_H)(w_H - \hat{w}_{IC}(w_H)) \geq 0. \quad (30)$$

It is easy to see that $\Phi(b) > 0$ and $\Phi'(w) < 0$. Hence, there is a unique threshold wage \tilde{w}_H satisfying

$$\Phi(\tilde{w}_H) \geq 0, \quad \tilde{w}_H \leq \bar{w}_H,$$

with complementary slackness. This implies that firms offer separating contracts when $w_H \leq \tilde{w}_H$ and they offer pooling contracts when $w_H > \tilde{w}_H$. $\Phi(b) > 0$ directly implies that $\tilde{w}_H > b$, so that separating dominates pooling at the lowest-wage firms. This completes the proof of Proposition 3. \square

Proof of Theorem 1:

From the discussion in the text, the equilibrium fraction of firms offering separating contracts $\eta > 0$ satisfies the complementary slackness condition

$$T(\eta) \geq \eta, \quad \eta \leq 1. \quad (31)$$

Conversely, any η solving this condition defines a market equilibrium. Because of $\lim_{\eta \rightarrow 0} \tilde{w}_i(\eta) = b < \varepsilon_{LP}$ for $i = L, H$, it follows that $\lim_{\eta \rightarrow 0} T(\eta) = \infty$. Because T is a continuous function, existence and uniqueness follows under the provision that T intersects the 45-degree line at most once: then the complementary-slackness condition (31) must have a unique solution $\eta \in (0, 1]$.

To prove that T has at most one fixed point, we claim that $T(\eta) = \eta$ implies $T'(\eta) = 0$. To show this property, consider any candidate $\eta = T(\eta)$ and differentiate function T :

$$\begin{aligned} T'(\eta) &= -\frac{\rho}{\lambda(\tilde{w}_H(\eta) - \tilde{w}_L(\eta))} \tilde{w}'_L(\eta) - \frac{\rho(\varepsilon_{LP} - \tilde{w}_L(\eta))}{\lambda(\tilde{w}_H(\eta) - \tilde{w}_L(\eta))^2} (\tilde{w}'_H(\eta) - \tilde{w}'_L(\eta)) \\ &= -\frac{\rho \tilde{w}'_L(\eta)}{\lambda(\tilde{w}_H(\eta) - \tilde{w}_L(\eta))} \left\{ \hat{w}'(\tilde{w}_H(\eta)) + \frac{\varepsilon_{LP} - \tilde{w}_L(\eta)}{\tilde{w}_H(\eta) - \tilde{w}_L(\eta)} (1 - \hat{w}'(\tilde{w}_H(\eta))) \right\}. \end{aligned}$$

Here the last line uses that $\tilde{w}'_L = \hat{w}' \tilde{w}'_H$. From (18) follows that

$$\hat{w}'(\tilde{w}_H(\eta)) = \frac{q(\tilde{w}_H(\eta))}{q(\tilde{w}_H(\eta)) + \rho},$$

and $T(\eta) = \eta$ implies that

$$\frac{\varepsilon_{LP} - \tilde{w}_L(\eta)}{\tilde{w}_H(\eta) - \tilde{w}_L(\eta)} = -\frac{q(\tilde{w}_H(\eta))}{\rho}.$$

Taken together, this implies that the expression in braces is zero, and hence $T'(\eta) = 0$. This proves that equilibrium exists and is unique.

To prove the second claim, note that an equilibrium with $\eta < 1$ exists if, and only if, $T(1) < 1$. Furthermore, note that $\tilde{w}_i(1) = \bar{w}_i$, defined in (7) and (8); cf. also the expressions for \tilde{w}_i in Appendix C. Then, $T(1) < 1$ gives rise to the equivalent condition

$$(\phi + \delta)(\bar{w}_H - \bar{w}_L) < \rho(\bar{w}_L - \varepsilon_{LP}). \quad (32)$$

Denote the dependence on parameter ρ by $\bar{w}_i(\rho)$, and define

$$\Psi(\rho) \equiv (\phi + \delta)(\bar{w}_H(\rho) - \bar{w}_L(\rho)) - \rho(\bar{w}_L(\rho) - \varepsilon_L p) .$$

Note that $\bar{w}_L(0) = \bar{w}_H(0) = \bar{w} = \tilde{\varepsilon}p - [(\phi + \delta)/(\phi + \delta + \lambda)]^2(\tilde{\varepsilon}p - b)$, which is the highest wage offer in the no-information case. Because of $\Psi(0) = 0$ the condition for a segmented equilibrium (32) is satisfied for sufficiently low values of ρ iff $\Psi'(0) < 0$. Differentiate Ψ at $\rho = 0$:

$$\Psi'(0) = \varepsilon_L p - \bar{w} + (\phi + \delta)(\bar{w}'_H(0) - \bar{w}'_L(0)) .$$

Now differentiate (7) and (8) at $\rho = 0$:

$$\bar{w}'_L(0) = -\frac{2(\tilde{\varepsilon}p - b)\lambda\alpha_H}{(\phi + \delta + \lambda)^2} , \quad \bar{w}'_H(0) = -\frac{\alpha_L}{\alpha_H}\bar{w}'_L(0) .$$

Substitute this into the above to obtain

$$\Psi'(0) = \varepsilon_L p - \bar{w} + \frac{2\lambda(\phi + \delta)(\tilde{\varepsilon}p - b)}{(\phi + \delta + \lambda)^2} .$$

Then, some simple manipulation shows that $\Psi'(0) < 0$ is equivalent to the inequality condition stated in Theorem 1. Therefore, under this condition, there exists a threshold value $\rho_3 > 0$ such that $\Psi(\rho) < 0$ (and hence a segmented market equilibrium exists) for $\rho \in (0, \rho_3)$. This completes the proof of Theorem 1. \square

B Firm Heterogeneity

Let

$$F_i(w_i) = \beta_H F_i(w_i | p_H) + (1 - \beta_H) F_i(w_i | p_L) \tag{33}$$

denote 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: Wages offered by any firm satisfy $w_L = \hat{w}(w_H)$ 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 expected profit $\Omega_k(w_H, w_L)$. 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 k ; i.e.

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

Before we characterize the relevant sorting equilibria, we prove a few results on the optimal wage policies of heterogeneous firms. Note that optimal worker behavior is exactly the same as in the model

with homogeneous firms. Particularly, workers select into wage contracts according to the same incentive constraint (3) as before. We focus on the case where the firms' learning rate is sufficiently low so that incentive constraints are binding for all firms.

Consider any offer distribution F_H with support $[\underline{w}_H, \bar{w}_H]$ such that $\underline{w}_H \geq b$ and $\bar{w}_H < \infty$ and recall that $q(w_H) = \phi + \delta + \lambda(1 - F_H(w_H))$. It is convenient to define the following constants,

$$\bar{\Phi}_k = \rho(\varepsilon_L p_k - b) + (\phi + \delta)(\bar{w}_H - b) - (\phi + \delta + \rho) \int_b^{\bar{w}_H} \frac{q(x)}{q(x) + \rho} dx ,$$

for $k = L, H$.

Lemma a.1:

(i) Given $\bar{\Phi}_k < 0$, there exist threshold wages $\hat{w}_k > 0$, $k = H, L$, such that a firm of type k offering w_H to high-ability workers prefers to offer separating contracts if $w_H < \hat{w}_k$ and prefers to offer pooling contracts if $w_H > \hat{w}_k$.

(ii) Given such thresholds exist, then $\hat{w}_H > \hat{w}_L$. That is, if a low-productivity firm offering w_H to high-ability workers prefers to offer separating contracts, a high-productivity firm would strictly prefer to offer separating contracts when it offers w_H to high-ability workers.

(iii) If two pooling firms of type $k = H, L$ offer wages w_{Hk} to high-ability workers, it must be that $w_{HH} > w_{HL}$.

(iv) If two separating firms of type $k = H, L$ offer wages w_{Hk} to high-ability workers, it must be that $w_{HH} > w_{HL}$.

Proof: To prove the first two parts, consider a firm of type k offering w to high-ability workers. This firm then makes the same expected profit from high-ability workers, irrespective of its contract choice (cf. the profit expressions (16) and (17)). To determine whether offering separating contracts is better than offering pooling contracts, we need to compare the corresponding profits from hiring of low-ability workers. If the firm offers separating contracts, its profit from low-ability workers is

$$h(w) \frac{\varepsilon_L p_k - \hat{w}(w)}{q(w)} , \tag{34}$$

where $h(w)$ is the hiring rate of low-ability workers (see Appendix C), $1/q(w)$ is expected job duration, and $\hat{w}(w)$ is the separating wage, implicitly defined from (10). If the firm offers pooling contracts, its profit from hiring of low-ability workers is

$$h(w) \frac{\varepsilon_L p_k - w}{q(w) + \rho} , \tag{35}$$

where $h(w)$ is the same hiring rate as in (34).

The firm decides to offer separating contracts if (34) is larger than (35), i.e.

$$(q(w) + \rho)(\varepsilon_L p_k - \hat{w}(w)) \geq q(w)(\varepsilon_L p_k - w) .$$

Define

$$\Phi_k(w) \equiv \rho \varepsilon_L p_k - (q(w) + \rho) \hat{w}(w) + q(w)w$$

and note that the above inequality corresponds to $\Phi_k(w) \geq 0$. It is easy to verify that $\Phi_k(b) > 0$ and that $\Phi'_k(w) = -\lambda F'_H(w)[w - \hat{w}(w)] < 0$ for all $w \in [\underline{w}_H, \bar{w}_H]$ and $\Phi'_k(w) = 0$ for $w > \bar{w}$. Given $\bar{\Phi}_k = \Phi_k(\bar{w}_H) < 0$, continuity implies that there exists a unique threshold wage $\hat{w}_k > b$ such that $\Phi(\hat{w}_k) = 0$. In this case, the firm prefers to offer separating contracts if $w < \hat{w}_k$ and it prefers to offer pooling contracts if $w > \hat{w}_k$. Otherwise all firms offer separating contracts. This completes the proof of part (i).

Part (ii) follows directly because Φ is strictly increasing in p_k .

To prove part (iii), consider a low- and high-productivity firm offering a pooling contract with wages w_{HL} and w_{HH} , respectively. The aim is to show that $w_{HH} \geq w_{HL}$ in equilibrium. Consider equation (17), which describes the profits of a firm offering a pooling contract. Let $L_H^P(w_H) = \lambda(\phi + \delta)\alpha_H/q(w_H)^2$ and $L_L^P(w_H) = \lambda(\phi + \delta + \phi)\alpha_L/(q(w_H) + \rho)^2$ and note that both expressions are increasing in w_H . Using a similar argument as in Burdett and Mortensen (1998), it holds that in equilibrium

$$\begin{aligned} & L_L^P(w_{HH})(\varepsilon_{LPH} - w_{HH}) + L_H^P(w_{HH})(\varepsilon_{HPH} - w_{HH}) \\ & \geq L_L^P(w_{HL})(\varepsilon_{LPH} - w_{HL}) + L_H^P(w_{HL})(\varepsilon_{HPH} - w_{HL}) \\ & > L_L^P(w_{HL})(\varepsilon_{LPL} - w_{HL}) + L_H^P(w_{HL})(\varepsilon_{HPL} - w_{HL}) \\ & \geq L_L^P(w_{HH})(\varepsilon_{LPL} - w_{HH}) + L_H^P(w_{HH})(\varepsilon_{HPL} - w_{HH}), \end{aligned}$$

which implies $L_L^P(w_{HH})\varepsilon_L + L_H^P(w_{HH})\varepsilon_H \geq L_L^P(w_{HL})\varepsilon_L + L_H^P(w_{HL})\varepsilon_H$. Since this inequality holds for any equilibrium offers w_{HL} and w_{HH} , it follows from the monotonicity of L_k^P that $\underline{w}_{HH} \geq \bar{w}_{HL}$. This completes the proof of part (iii).

To prove (iv), consider a low-productivity firm offering w_{HL} and $w_{LL} = \hat{w}(w_{HL})$ and a high-productivity firm offering w_{HH} and $w_{LH} = \hat{w}(w_{HH})$. Recall that \hat{w} is increasing in w_H . Now consider equation (16), which describes the profits of firms offering separating contracts. Let $L_H^S(w_H) = \lambda(\phi + \delta)\alpha_H/q(w_H)^2$ and $L_L^S(w_H) = \lambda\theta(\eta)\alpha_L/q(w_H)^2$ and note that both expressions are increasing in w_H . Using the same arguments as above it follows that $L_L^S(w_{HH})\varepsilon_L + L_H^S(w_{HH})\varepsilon_H \geq L_L^S(w_{HL})\varepsilon_L + L_H^S(w_{HL})\varepsilon_H$ which implies $\underline{w}_{HH} \geq \bar{w}_{HL}$. This completes the proof of Lemma a.1. \square

B.1 Sorting Equilibrium

The previous Lemma shows that in any market equilibrium: (i) conditional on productivity, firms offering separating contracts pay lower wages than firms offering pooling contracts, and (ii), high-productivity firms pay higher wages than low-productivity firms. In what follows we focus on a candidate equilibrium in which a fraction $\eta \leq 1$ of firms offer separating contracts. As before firms offering separating contracts post wages $w_H \in [b, \tilde{w}_H]$ and $w_L = \hat{w}(w_H) \leq w_H$ with $\tilde{w}_L = \hat{w}(\tilde{w}_H)$ satisfying

(10). The remaining fraction $1 - \eta$ of firms offers $w_H > \tilde{w}_H$ 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, $F_H(\cdot | p_k)$ for $k = L, H$, are continuous and exhibit connected supports.

Given Lemma a.1, there are two natural equilibrium candidates. First, if $\eta < \beta_L$, low-productivity firms offer both separating and pooling contracts and all high-productivity firms offer pooling contracts; in this equilibrium the threshold wages of Lemma a.1 satisfy $\hat{w}_k < \tilde{w}_H$ for $k = L, H$. Second, if $\eta > \beta$, all low-productivity firms offer separating contracts and high-productivity firms offer separating and pooling contracts; here we have $\hat{w}_k > \tilde{w}_H$, $k = L, H$.²¹ We now turn to characterize these two types of equilibria.

Case I: $\eta < \beta_L$

In this case, some low-productivity firms offer separating contracts and some low-productivity firms and all high-productivity firms offer pooling 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 separating 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 $w_H \in [b, \bar{w}_{HL}]$ such that $\tilde{w}_H(\eta)$ solves $F_H(\tilde{w}_H) = \eta$ using Lemma 1 and $\bar{w}_{HL}(\eta)$ solves $F_H(\bar{w}_{HL}) = \beta_L$ using Lemma 2. It then follows from Lemma a.1 and (33) that $F_H(w_H | p_L) = F_H(w_H)/\beta_L$ for all $w_H \in [b, \bar{w}_{HL}]$.

To obtain the offer distribution for wages $w_H \in [\underline{w}_{HH}, \bar{w}_{HH}]$ first note that optimality implies $\underline{w}_{HH}(\eta) = \bar{w}_{HL}(\eta)$. Further, since equilibrium requires that $\Omega^P(\underline{w}_{HH}, \underline{w}_{LH}) = \Omega^P(w_H, w_L)$ for all $w_H \in [\underline{w}_{HH}, \bar{w}_{HH}]$, $w_L = \hat{w}(w_H)$, 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 $\bar{w}_{HH}(\eta)$ solves $F_H(\bar{w}_{HH}) = 1$. In this case, Lemma a.1 and (33) 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 separating contracts, while some high-productivity firms offer pooling 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

²¹There is actually a third equilibrium candidate where both types of firms offer both types of contract ($\hat{w}_L < \tilde{w}_H < \hat{w}_H$). However, Lemma a.1 part (ii) rules out equilibria in which all low-productivity firms offer separating contracts and all high-productivity firms offer pooling contracts.

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 firms with high productivity offering separating contracts. 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 $\eta \in (\beta_L, 1]$.

C Omitted Derivations

Derivation of the steady-state unemployment rates and the earnings distribution of workers:

Consider unemployment of high-ability workers, u_H . The inflow into this category is $(\phi + \delta)(\alpha_H - u_H)$, while the outflow is given by λu_H . Steady-state turnover then yields the expression in the main text. Similarly consider the proportion of high-ability workers earning a wage no greater than w_H , $G_H(w_H)$. The inflow into this category is given by $\lambda F_H(w_H)u_H$, while the outflow is given by $q(w_H)G_H(w_H)(\alpha_H - u_H)$. Steady-state turnover then implies (13).

Now consider unemployment of low-ability workers. The inflow into this category is given $[\phi + \delta + \rho(1 - G_L(\tilde{w}_L))](\alpha_L - u_L)$, while the outflow is given by λu_L . Steady-state turnover yields

$$u_L = \frac{[\phi + \delta + \rho(1 - G_L(\tilde{w}_L))]\alpha_L}{\phi + \delta + \lambda + \rho(1 - G_L(\tilde{w}_L))}.$$

Next consider those low-ability workers earning wages no greater than $w_L \in [b, \tilde{w}_L]$. The inflow into this category is given by $\lambda F_H(\hat{w}^{-1}(w_L))u_L$. The outflow is given by $[\phi + \delta + \lambda(1 - F_H(\hat{w}^{-1}(w_L)))]G_L(w_L)(\alpha_L - u_L)$. Steady-state turnover and $F_H(\tilde{w}_H) = \eta$ then imply

$$G_L(w_L) = \frac{\lambda F_H(\hat{w}^{-1}(w_L))u_L}{[\phi + \delta + \lambda(1 - F_H(\hat{w}^{-1}(w_L)))](\alpha_L - u_L)}.$$

Evaluating $G_L(w_L)$ at $w_L = \tilde{w}_L$ and solving the above two equations gives the corresponding measures described in the main text.

Finally consider the proportion of low-ability workers employed in firms offering pooling contracts at wages $w'_H \in [\tilde{w}_H, w_H]$. Since any low-ability worker will misreport his type when offered such a wage, 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 $[\phi + \delta + \rho + \lambda(1 - F_H(w_H))][G_L(w_H) - G_L(\tilde{w}_H)](\alpha_L - u_L)$.

Noting that $G_L(\tilde{w}_L) = G_L(\tilde{w}_H)$ and using the above expressions imply that the steady-state proportion of low-ability workers earning a wage no greater than $w_H \in [\tilde{w}_H, \bar{w}_H]$ is given by (15).

Derivation of the steady-state profits for firms offering separating and pooling contracts:

First consider a 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

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

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

$$h_H(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{q(w_H)} \text{ and } h_L(w_L) = \frac{\lambda\theta(\eta)\alpha_L}{q(\hat{w}^{-1}(w_L))} , \quad (36)$$

where $\theta(\eta)$ 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

$$J_H(w_H) = \frac{\varepsilon_{HP} - w_H}{q(w_H)} \text{ and } J_L(w_L) = \frac{\varepsilon_{LP} - w_L}{q(\hat{w}^{-1}(w_L))} .$$

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

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

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

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

$$h_L(w_H) = \frac{\lambda(\phi + \delta + \rho)\alpha_L}{\rho + q(w_H)} \text{ and } h_H(w_H) = \frac{\lambda(\phi + \delta)\alpha_H}{q(w_H)} . \quad (37)$$

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

$$J(w_H) = \frac{[\tilde{\alpha}_H(w_H)\varepsilon_H + \tilde{\alpha}_L(w_H)\varepsilon_L]p - w_H + \rho[\tilde{\alpha}_H(w_H)J_H(w_H)]}{\rho + q(w_H)} ,$$

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

$$J_H(w_H) = \frac{\varepsilon_{HP} - w_H}{q(w_H)} .$$

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

Derivation of $\tilde{w}_L(\eta)$ and $\tilde{w}_H(\eta)$:

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

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

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

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

Integration then yields

$$\begin{aligned} \tilde{w}_L(\eta) &= b + \frac{2[(\phi + \delta)\alpha_H(\varepsilon_{HP} - b) + \theta(\eta)\alpha_L(\varepsilon_{LP} - b)]}{(\phi + \delta + \lambda)^2 [(\phi + \delta)\alpha_H + \theta(\eta)\alpha_L]^3} \times \\ &\quad \left[[(\phi + \delta)\alpha_H\rho]^2 \log \left(\frac{(\phi + \delta)\alpha_H\rho + ((\phi + \delta)\alpha_H + \theta(\eta)\alpha_L)(\phi + \delta + \lambda)}{(\phi + \delta)\alpha_H\rho + ((\phi + \delta)\alpha_H + \theta(\eta)\alpha_L)(\phi + \delta + \lambda(1 - \eta))} \right) \right. \\ &\quad \left. + \lambda\eta [(\phi + \delta)\alpha_H + \theta(\eta)\alpha_L]^2 \left(\phi + \delta + \lambda \left(1 - \frac{\eta}{2} \right) - \frac{(\phi + \delta)\alpha_H\rho}{(\phi + \delta)\alpha_H + \theta(\eta)\alpha_L} \right) \right]. \end{aligned}$$

The closed-form expression for $\tilde{w}_H(\eta)$ is obtained using the constant-profit condition $\Omega^S(\tilde{w}_H, \tilde{w}_L) = \Omega^S(b, b)$. This yields

$$\tilde{w}_H(\eta) = \varepsilon_{HP} + \frac{\theta(\eta)\alpha_L(\varepsilon_{LP} - \tilde{w}_L(\eta))}{(\phi + \delta)\alpha_H} - \left(\frac{\phi + \delta + \lambda(1 - \eta)}{\phi + \delta + \lambda} \right)^2 \left[\varepsilon_{HP} - b + \frac{\theta(\eta)\alpha_L(\varepsilon_{LP} - b)}{(\phi + \delta)\alpha_H} \right],$$

where $\tilde{w}_L(\eta)$ is given above. Further, note that when $\eta = 1$, these expression collapse to the ones in (7) and (8).