DISCUSSION PAPER SERIES

DP11869

ADVERSE SELECTION AND ASSORTATIVE MATCHING IN LABOR MARKETS

Daniel Ferreira and Radoslawa Nikolowa

FINANCIAL ECONOMICS and LABOUR ECONOMICS

ISSN 0265-8003

ADVERSE SELECTION AND ASSORTATIVE MATCHING IN LABOR MARKETS

Daniel Ferreira and Radoslawa Nikolowa

Discussion Paper DP11869
Published 22 February 2017
Submitted 22 February 2017

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

ADVERSE SELECTION AND ASSORTATIVE MATCHING IN LABOR MARKETS

Abstract

We show that adverse selection in the labor market may generate negative assortative matching of workers and firms. In a model in which employers asymmetrically learn about the ability of their workers, high-productivity firms poach mediocre workers, whereas low-productivity firms retain high-ability workers. We show that this flipping property is caused by information asymmetry alone. Our model has a number of positive and normative predictions: External promotions are not an indication of high talent, within-job wage growth is higher in industries with more revenue dispersion, and non-compete clauses are inefficient in industries with significant firm heterogeneity.

JEL Classification: N/A

Keywords: N/A

Daniel Ferreira - d.ferreira@lse.ac.uk
London School of Economics and CEPR

Radoslawa Nikolowa - r.nikolowa@qmul.ac.uk
Queen Mary University of London

Acknowledgements

Adverse Selection and Assortative Matching in Labor Markets*

Daniel Ferreira

London School of Economics and CEPR

Radoslawa Nikolowa

Queen Mary University of London

February 2017

Abstract

We show that adverse selection in the labor market may generate negative assortative matching of workers and firms. In a model in which employers asymmetrically learn about the ability of their workers, high-productivity firms poach mediocre workers, whereas low-productivity firms retain high-ability workers. We show that this *flipping property* is caused by information asymmetry alone. Our model has a number of positive and normative predictions: External promotions are not an indication of high talent, within-job wage growth is higher in industries with more revenue dispersion, and non-compete clauses are inefficient in industries with significant firm heterogeneity.

Keywords: Adverse Selection, Matching, Labor Markets, Asymmetric Employer Learning

*We benefited from comments from and/or conversations with Ulf Axelson, William Fuchs, Christian Hellwig, Marco Ottaviani, and Mike Waldman, and from comments by seminar participants at Bocconi, Cass, ESCP, Essex, LSE, Munich, QMUL, Queen's University, and the Econometric Society European Meeting (Geneva). Contacts: d.ferreira@lse.ac.uk, r.nikolowa@qmul.ac.uk.

1. Introduction

If firms can only learn about the talent of their workers by observing them on the job, initial matches of firms and workers can be inefficient. After learning occurs, market forces may reallocate workers across firms. We show that if firms have an informational advantage when learning about their own workers, then worker flows tend to exacerbate inefficiencies associated with initial allocations. That is, adverse selection prevents the positive assortative matching of workers and firms: High-productivity firms poach mediocre workers, whereas low-productivity firms retain high-ability workers. We thus say that the equilibrium has a *flipping property.*

Our model belongs to the asymmetric employer learning literature, which was initiated by Waldman (1984) and Greenwald (1986). In such models, the current employer learns about the talent of her incumbent workers, while competing employers remain uninformed. This form of information asymmetry implies that competitors may learn about a worker's ability from the actions taken by the worker's employer, such as decisions involving promotion, retention, and termination. These models typically assume that workers accumulate firm-specific skills over time and are thus more valuable to their incumbent firms than to competing firms.[1]

We depart from the existing literature by introducing a specific form of firm heterogeneity: Some firms are more productive/profitable than others.[2] This feature allows us to study worker poaching and job mobility in equilibrium. Initial allocations of workers to firms may turn out to be inefficient, because they happen without any knowledge of workers' abilities. When initial allocations are inefficient, our question is the following: Do market forces eliminate or exacerbate inefficiencies? Under symmetric learning (and supermodularity of firm and worker qualities), we obtain the standard result that the best firms are indeed matched with the best workers in equilibrium, i.e., there is positive assortative matching of workers and firms. By contrast, in an equilibrium with asymmetric employer learning, the

[1] Earlier papers in this literature include Lazear (1986), Milgrom and Oster (1987), Waldman (1990), Laing (1994), Chang and Wang (1996), and Acemoglu and Pischke (1998, 1999), among others.

[2] Dispersion in profitability is widely documented, even within narrowly defined industries. A large body of strategy literature attributes profitability dispersion to monopoly profits, which are explained by barriers to entry or ownership of unique resources (McGahan and Porter, 1997; Rumelt, 1991). Even in industries with free entry, equilibrium (ex post) profitability dispersion can be explained by the accumulation of organizational capital (Atkeson and Kehoe, 2005). For a review of the literature on productivity dispersion, see Syverson (2011).

socially optimal allocation of talent cannot generally be attained: There is either too much or too little job mobility in equilibrium, or both.

To understand the consequences of asymmetric information for the matching of firms and workers, we first consider the problem of a social planner. The social planner decides whether a worker stays with her current employer or leaves for a more productive firm. The planner faces no constraints on transfers and actions, but does not know the worker's talent, which is only known by the current employer. To induce the employer to truthfully reveal information about the worker's talent, the planner designs a mechanism that forces the firm to pay a flat fee (which can be interpreted as a wage) for all types of workers that are retained. This flat-fee scheme is the only schedule that is incentive compatible.[3]

Because the firm pays the same fee regardless of which workers it retains, the firm retains only workers with talent above a given threshold. Therefore, only the very best workers can be retained. This in turn implies that only mediocre workers can be reassigned to high-profitability firms. This leads to a reversal of positive assortative matching. We call such a reversal a *flipping property* of the equilibrium. The social planner's solution makes it clear that the flipping property is a consequence of asymmetric information alone.

The flipping property implies that matching workers and firms on the basis of information revealed by firms is very inefficient. The planner thus prefers to make assignment decisions without using any information revealed by the mechanism. The solution then exhibits either excessive reassignment of mediocre workers or excessive retention of high-talent workers.

In the decentralized competitive equilibrium, which we consider next, both types of inefficiencies may coexist. In equilibrium, the incumbent employer offers the same wage to all retained workers, which implies that the incumbent retains only the most talented workers. Thus, high-profitability firms can only poach mediocre workers. It is rational for high-profitability firms to poach mediocre workers because these workers are better than the unemployed agents.

When solving for the decentralized equilibrium, we assume that contracts that fully bind workers to firms are not available. We make this assumption because the inalienability of labor is a key difference between labor markets and asset markets. "Excessive mobility" of mediocre workers occurs precisely because workers are free to move. We note, however, that

[3]It is assumed that outsiders cannot observe performance outcomes and, therefore, that the planner cannot design transfers contingent on those outputs.

the absence of bonding contracts does not cause the flipping property. The social planner's problem reveals that the flipping property is a consequence of informational asymmetries alone. The absence of bonding contracts simply allows the flipping property to manifest itself in the form of excessive mobility.

Our results may appear surprising in light of the original analysis of markets with asymmetric information by Akerlof (1970). In a lemons market in which the seller of an asset has private information, there is typically little or no trade. By analogy, one would expect that a labor market in which the current employer knows more about the quality of its worker than a competitor is likely to generate too little "trade," i.e., insufficient worker mobility. However, this analogy is imperfect for two reasons. First, matching considerations are important in labor markets, implying that low worker mobility is sometimes efficient. Second, workers are not like assets, which can be freely bought and sold. Assuming no slavery, a worker is free to work for the highest bidder, and the current employer typically receives no compensation if the worker is poached by another firm. In this context, and in contrast with traditional lemons market models, the flipping property typically leads to "too much trade."

Ours is not meant to be a general theory of labor markets. We expect our analysis to be relevant to those industries in which incumbent employers enjoy a natural advantage in discovering talent. Examples that fit such a description include innovative industries, such as information technology, and some sectors of the financial industry, such as asset management. In those cases, our model has distinctive empirical predictions. For example, in a model with symmetric learning and matching, people who are "externally promoted" (i.e., those who leave their jobs to move to higher-ranking positions in other firms) are typically the most talented workers. By contrast, in our model, because learning is asymmetric, externally promoted workers are not the most talented. Consistent with such a prediction, a recent paper by Berk, van Binsbergen, and Liu (forthcoming), which studies the labor market for mutual fund managers, finds that internal "promotions" (i.e., when a manager is given control over a larger value of assets) add value to the firm, whereas external promotions (i.e., external hires with an increase in assets under management) do not.[4]

Our model also predicts that employee earnings are higher in sectors with greater rev-

[4]Berk, van Binsbergen, and Liu (forthcoming) argue convincingly that mutual fund firms have private information about the skill of their managers. Their evidence shows that this information is used for promotion decisions. Outsiders are informationally disadvantaged because fund portfolio decisions are not disclosed in real time and fund performance alone is a very noisy measure of managerial skill.

enue dispersion and, more strikingly, that within-job earnings growth is also higher in such sectors. Consistent with these implications, in their study of the software industry, Andersson et al. (2009) show that the "rewards to loyalty" (i.e., within-job earnings growth for those employees retained by their firms) are greater in software sectors with high revenue dispersion. By contrast, between-job earnings growth is smaller in such sectors.

Our analysis also yields an important normative conclusion: The desirability of contracts that restrict worker mobility (e.g., non-compete clauses and bonding contracts) may vary across industries depending on the degree of firm heterogeneity and on the importance of firm-specific skills. Indeed, in some instances, facilitating employee mobility may be socially desirable. This normative conclusion also has some potentially positive predictions: Under the assumption that laws are chosen partly for efficiency reasons, non-compete clauses should be banned in industries where firm quality is very heterogeneous. Empirically, we observe variation in the enforceability of non-compete clauses across jurisdictions.[5]

After a brief discussion of the related literature, in Section 2 we present the setup of the model. Section 3 solves the problem of a financially unconstrained planner and shows our key result that any incentive-compatible mechanism exhibits a flipping property. In Section 4 we characterize the decentralized labor market equilibrium. We then provide a discussion of the key implications of the model and the empirical evidence supporting them in Section 5. Section 6 concludes.

Related Literature

Our paper belongs to the literature on adverse selection initiated by Akerlof (1970). Most papers in this literature consider decentralized trading situations in which the buyers' and sellers' valuations are not observable. The focus is usually on the impact of private information about the quality of a good on the occurrence of trade.[6] Here we adopt a similar approach, with similar assumptions. A key difference is that, in our setup, matching is an important consideration; thus, trading gains may not exist for a range of worker types.

[5]Non-compete clauses are controversial and have often been challenged in courts. For example, in California, non-compete clauses are considered void and non-enforceable, except in a small set of cases.

[6]For example, Ellingsen (1997) shows that there exists a separating equilibrium in which some trade of high-quality goods occurs in lemons markets. Levin (2001) studies how the degree of information asymmetry affects trade. Daley and Green (2012) and Fuchs and Skrzypacz (2015) develop dynamic models of adverse selection and its impact on trade. Bar-Isaac, Jewitt, and Leaver (2014) study how the degree of information asymmetry impacts efficiency when public and private information is multi-dimensional.

Therefore, in contrast to models of asset markets in which there is typically little or no trade,[7] our model shows that, in labor markets, the equilibrium often displays both excessive and insufficient worker mobility.

In the adverse selection literature, the closest paper to ours is Adriani and Deidda (2009). They consider a decentralized market in which a low-quality good is more valuable to the seller than to the buyer. They show that the unique equilibrium surviving the D1 refinement involves no trade. Market breakdown occurs because of an upward pressure on prices for signaling reasons. In our model, trade does not break down because the "asset" traded is an employee. An employee chooses to work for the firm with the highest wage offer. Because the wage offered to an employee is a cost for the incumbent firm, the incumbent retains only the best workers at the lowest wage that prevents poaching. However, the incumbent firm is not able to retain mediocre workers, and therefore, these workers choose to work for a poacher. Hence the decentralized equilibrium displays simultaneously too little trade of the best workers and too much trade of mediocre ones.

Our paper is also related to the literature on matching in labor markets, in particular to models in which information creates matching frictions. Franco, Mitchell, and Vereshchagina (2011) study the problem of designing work teams under moral hazard. They show that, even when technological complementarities favor positive assortative matching, a principal may choose to match high types with low types to save on incentive costs. Negatively sorted teams may thus be a consequence of the need to provide incentives for effort when individual contributions to team output are not observable. Related results can be found in Kaya and Vereshchagina (2015), who derive new results for teams organized as partnerships, and in Kaya and Vereshchagina (2014), who derive predictions concerning matching patterns in both corporations and partnerships. The reason for negative assortative matching in these models is very different from the economic forces in our model. Negative assortative matching arises as an optimal response to the problem of providing incentives under moral hazard. By contrast, in our model, which is based on hidden information about types, players do not truthfully reveal their types under any mechanism intended to implement a positively assortative matching allocation.

Hoppe, Moldovanu, and Sela (2009) and Kaya (2013) are examples of matching models

[7] Under private information and common knowledge, no-trade or insufficient-trade results can be found in a variety of different contexts and applications. Classic examples include Milgrom and Stokey (1982) and Myerson and Satterthwaite (1983).

with two-sided private information. Our model differs from theirs in many aspects, but crucially in that, in our setup, ex post payoffs are also private information, and thus contracting on payoffs is not possible.

More narrowly, our model belongs to the labor literature on asymmetric learning. In a seminal paper, Waldman (1984) considers internal job assignment as a signal of employee ability. Homogeneous firms attempt to poach workers assigned to higher-level jobs by making offers corresponding to workers' expected values. The resulting assignment of workers to jobs is inefficient; employees who would be more productive in higher-level jobs are not promoted. In a related paper, Milgrom and Oster (1987) show that employees whose abilities are observed only by their current employers tend to be promoted less often and paid lower wages than employees whose abilities are visible to other employers. Waldman's promotion model has also been extended by Bernhardt and Scoones (1993) and Bernhardt (1995) to analyze a number of issues, such as turnover, compensation, demotions, and other labor market outcomes.

As in our paper, Greenwald (1986) focuses on the role played by asymmetric information on employee mobility. Employees may leave a firm either for exogenous reasons or because they are not retained. A key result is that there are few layoffs among the best workers (who only quit for exogenous reasons), and the stream of people changing jobs thus disproportionately consists of "bad" employees. A key difference in our model is that mobility is always endogenous and driven by firm heterogeneity. In a related paper, Laing (1994) considers a model in which the decision to retain or fire an employee is a signal of employee ability; however, unlike our paper, firms are homogeneous and the focus is on the properties of the optimal contract for risk-averse employees. Mukherjee (2008) considers the related problem of designing an optimal disclosure policy when information about a worker's ability may trigger poaching attempts by more productive competitors.

Some papers show that variations of the key assumptions in these models can produce significantly different results. Ricart I Costa (1988) shows that if workers learn about their abilities and are able to choose from a menu of wage contracts, there is a separating equilibrium that resolves the "lemons" problem in Waldman's (1984) model. In our model, we assume that workers do not know their types, and our main results will hold as long as employers have *some* informational advantage about *some* aspects of their workers' abilities.

Golan (2005) challenges a different assumption in Waldman's model: the timing of wage offers. This study shows that if the incumbent always has the option of matching outside offers, efficiency can be restored.[8] Recent work by Waldman and Zax (2016) shows however that Golan's results are not robust; under reasonable assumptions, inefficient equilibria arise even when the incumbent is able to match outside offers.

Another application of asymmetric learning models involves the problem of investing in general and/or firm-specific skills; these models are developed by Waldman (1990), Chang and Wang (1996), Acemoglu and Pischke (1998, 1999), and Zabojnik and Bernhardt (2001), among others.[9]

More broadly, our paper is related to recent research that also emphasizes the welfare costs of fierce competition for talent, such as Bénabou and Tirole (2016) and Acharya, Pagano and Volpin (2016). These papers emphasize the deleterious effect that competition for talent can have on incentives (i.e., multitasking and/or project selection issues). By contrast, our focus is on the impact that competition has on how workers are allocated to firms. More closely related is Terviö (2009), who also shows that competition for talent creates inefficiencies. In his model, a worker's talent is revealed on the job but – unlike our model – this information is public. Terviö shows that in a competitive labor market, firms invest too little in talent discovery and over-recruit workers with mediocre abilities.

Finally, our analysis also shares certain ideas found in models of executive markets. As in firm-CEO assignment models, workers and firms are heterogeneous (Edmans, Gabaix, and Landier, 2009; Eisfeldt and Kuhnen, 2013; Gabaix and Landier, 2008; Terviö, 2008). As in Frydman (2013) and Murphy and Zabojnik (2004, 2006), workers are endowed with both firm-specific and general skills. As in Edmans and Gabaix (2011), the process of matching workers with firms is distorted by informational frictions.

[8] In an earlier paper, Lazear (1986) makes a similar point.

[9] There is also an important empirical literature on asymmetric employer learning. Gibbons and Katz (1991) provide empirical evidence that is compatible with the predictions of a model of layoffs with asymmetric employer learning. Pinkston (2009) constructs a model in which firms use bidding wars to compete for talent and finds empirical evidence of substantial asymmetric employer learning. Kahn (2013) also finds substantial evidence in favor of asymmetric learning. In contrast, Schönberg (2007) finds little evidence that employer learning is asymmetric.

2. Model Setup

The economy is populated with a continuum of firms that live for two periods; in the Internet Appendix, we present a model with infinitely-lived firms. Firms can be of one of two types, L or H, and these represent both the type and the mass of firms of each type. We denote a representative firm of each type by $i \in \{l, h\}$, which also denotes the value of a *profitability parameter*: Low-profitability – L firms – have parameter $l = 1$, and high-profitability firms – H firms – have parameter $h = \theta$, where $\theta > 1$. This is the only source of (exogenous) heterogeneity between firms.[10]

There is a continuum of agents (i.e., those who can become workers) who live for two periods: young age and old age. At $t = 0$, a mass M of young agents enter the labor market. Young agents are in excess supply: $M \gg H + L$. Young agents who are not employed at $t = 0$ remain available for hire at $t = 1$, when they are old. The outside option of an unemployed agent (young or old) is normalized to zero.

At $t = 0$, each firm (of either type, L or H) hires one young worker; all workers are observationally identical. At $t = 1$, each firm learns about the talent of its incumbent worker, which is given by $\tau \in [0, \overline{\tau}]$. Worker's talent is distributed according to a differentiable cumulative distribution function (c.d.f.) $F(.)$ with support $[0, \overline{\tau}]$.

Workers do not observe τ. Our interpretation of this assumption is that a firm has a better signal of its worker's ability than the worker herself. For example, τ may represent skills that are specific to the industry, and the employer may have more experience in assessing the value of such skills.[11] This assumption rules out the possibility of workers signaling their types to potential employers. It also rules out the possibility of potential employers screening workers through a menu of contracts. We choose to rule out these possibilities in order to focus on the role of asymmetric information among employers. Our approach has the advantage of making clear precisely what informational assumptions are required for the results. By contrast, the literature on employer learning typically adopts a different approach that imposes exogenous restrictions on actions – and sometimes on the space of contracts – to eliminate screening and employee signaling.

At $t = 1$, a firm of type $i \in \{1, \theta\}$ that successfully retains an incumbent worker of type

[10] A mass of firms with a known common productivity type is sometimes interpreted as "sector," see, e.g., Gibbons et al (2005).

[11] We assume that workers do not learn anything about τ for simplicity only. The important assumption here is that a worker learns less about τ than does her incumbent employer.

τ receives (expected) payoff $i\tau$. We assume that a firm's payoff is not directly observable, and thus remains private information to the firm. One possibility is that performance is only observed with noise. This could happen for a number of reasons, such as insufficient disclosure, imperfect measurement of the performance of complex tasks, difficulties in measuring a worker's individual contribution to the output of a team, or any other similar confounding effects. In all such cases, the firm may have an informational advantage over outsiders when estimating worker performance, because the firm can directly observe the worker's actions.

If the firm loses the incumbent worker, it may replace that worker with either a randomly selected unemployed agent or an incumbent worker poached from another firm. If firm i replaces an incumbent worker with an externally-hired worker of type τ, the firm's payoff is $i\gamma\tau$. Parameter $\gamma \in (0,1)$ represents the loss in firm-specific skills that results when the incumbent worker is replaced by an outsider. Higher levels of γ mean that firm-specific skills are less important. Our interpretation is that a worker acquires firm-specific skills after working for a particular firm in $t = 0$, and these skills remain valuable at $t = 1$.

We call the set of unemployed agents available for hire at $t = 1$ the *outside pool.* The outside pool is comprised only of old agents who were not employed at $t = 0$ (this is without loss of generality; in equilibrium, a firm with a vacancy would never hire a worker who was dismissed by another firm). If a firm of type $i \in \{1, \theta\}$ hires from the outside pool at $t = 1$, the firm's expected payoff is $i\gamma\mu$, where μ is the mean of $F(.)$.

At $t = 1$, some firms will have vacancies. For example, if $\tau < \gamma\mu$, a firm prefers a randomly selected unemployed agent to its incumbent worker, which means that the firm will fire its incumbent worker and open a vacancy. Firms with vacancies can either hire from the outside pool or try to poach a worker from another firm.

We are interested in how efficiently vacancies can be filled. If all firms were identical, all vacancies should be filled with workers from the outside pool; poaching workers from other firms would be inefficient because job mobility destroys firm-specific skills. However, since firms are heterogeneous – H firms are more productive than L firms – and there is a technological complementarity between τ and i, in some cases it is efficient for H firms to poach workers from L firms.

We make the following simplifying assumption:

Assumption A1 $\frac{H}{L} > \frac{1-F(\gamma\mu)}{2F(\gamma\mu)-1}$.

This is a sufficient, but not necessary, assumption to guarantee that poachable workers are always in "short supply," which is the most interesting case to analyze.

3. The Planner's Problem

We first consider the problem of a planner who wants to maximize social surplus. We place no exogenous restrictions on the set of mechanisms that the planner can choose.

3.1. Symmetric Information

As a benchmark, we start with the case in which information is symmetric, and thus the planner has the same information that firms do. In this case the planner can easily implement the first-best allocation.

At $t = 0$, there is no meaningful decision problem; each firm should hire one worker from the outside pool.

At $t = 1$, because of firm-specific skills, it is never efficient to reallocate workers from one firm to another when both firms are of the same type. Similarly, transferring workers from H firms to L firms is always inefficient. Thus, the planner only needs to consider the possibility of transferring workers from L firms to H firms.

The planner only needs to consider as potential poachers the set of H firms with workers with talent below $\gamma\mu$. This is because of Assumption A1. This implies that an H firm's incumbent worker is retained if $\tau \geq \gamma\mu$, and dismissed otherwise.

We now only need to consider the case of L firms with incumbent workers. Thus, to simplify the exposition, when solving the planner's problem we refer to an L firm with an incumbent worker at the beginning of $t = 1$ as an *incumbent firm*, and to H firms with vacancies (i.e., H firms with $\tau < \gamma\mu$) as potential *poachers*.

To formally study the planner's problem when allocating L firms' incumbent workers, we first introduce some notation and terminology. The planner's problem is to assign incumbent workers to one of three possible sets: P denotes the set of workers who are assigned to a poacher (i.e., an H firm with a vacancy), R denotes the set of workers who remain with the incumbent firm, and S denotes the set of workers who are unassigned (i.e., they are "sacked").

Definition 1 *An **allocation** is a function $a(\tau) : [0, \overline{\tau}] \to \Delta^2$, where Δ^2 is the standard 2-simplex, that maps a worker of type τ to a probability distribution $\{p_P, p_R, p_S\}$, where p_i is the probability that the worker is assigned to set $i \in \{P, R, S\}$.*

In other words, we define an allocation as a stochastic assignment rule. The allocation function determines which types of incumbent workers are allocated to L firms, to H firms, or to no firm. In the special case in which the assignment rule is deterministic, we can define an allocation in a more conventional way as $a(\tau) : [0, \overline{\tau}] \to \{P, R, S\}$. Definition 1 is however more general.

The *net surplus* created by a worker of type τ who is assigned to an incumbent firm is $\tau - \gamma\mu$. Similarly, the net surplus created by a worker of type τ who is assigned to a poacher is $\theta\gamma\tau - \theta\gamma\mu$. A social planner who wants to maximize social surplus should (i) replace all workers such that $\tau \leq \gamma\mu$ with a randomly selected replacement from the outside pool (whose expected type is μ) and (ii) assign worker $\tau \geq \gamma\mu$ to a poacher if and only if

$$\theta\gamma\tau - \theta\gamma\mu \geq \tau - \gamma\mu. \tag{1}$$

In other words, worker τ should be matched with a poacher when the incremental surplus to the poacher is larger than the net loss to the incumbent firm. Condition (1) implies that poaching should occur only if $\tau \geq \tau^{\#}$, where

$$\tau^{\#} = \begin{cases} \overline{\tau} & \text{if } \theta\gamma \leq 1 \\ \min\{(\theta - 1)\gamma\mu / (\theta\gamma - 1), \overline{\tau}\} & \text{if } \theta\gamma > 1 \end{cases}. \tag{2}$$

We call $\tau^{\#}$ the *critical type*: Workers who are more talented than the critical type should be assigned to a poacher, while those less talented than the critical type should be either retained or fired. Clearly, only those workers such that $\tau \leq \gamma\mu$ should be fired, which implies that all workers in $[\gamma\mu, \tau^{\#}]$ should be retained.[12] This completes the characterization of the first-best allocation $a^{FB}(\tau)$, which is given by

$$a^{FB}(\tau) = \begin{cases} (1, 0, 0) & \text{if } \tau \in [\tau^{\#}, \overline{\tau}] \\ (0, 1, 0) & \text{if } \tau \in [\gamma\mu, \tau^{\#}] \\ (0, 0, 1) & \text{if } \tau \in [0, \gamma\mu] \end{cases}. \tag{3}$$

[12] Because what happens in the (zero measure) boundary cases $\tau = \gamma\mu$ and $\tau = \tau^{\#}$ is not important, for simplicity we write all intervals as closed intervals.

Note that the first-best allocation is a deterministic assignment rule and displays positive assortative matching: Among those workers initially paired with L firms, the best workers – those in $\left[\tau^{\#}, \overline{\tau}\right]$ – are assigned to H firms, while the mediocre workers – those in $\left[\gamma\mu, \tau^{\#}\right]$ – remain with L firms.

3.2. Asymmetric Information

We now consider the case of an informationally constrained social planner: The planner cannot observe the incumbent workers' types. To make information asymmetries relevant, we maintain the assumption that outsiders (including the planner) cannot observe performance outcomes. Clearly, if the firms' payoffs were perfectly observable to all, the planner could implement the first-best outcome by forcing some L firms to "integrate" with some H firms (or write a contract that mimics the integrated solution), thus making both firms internalize the consequences of their actions. The interesting case is thus the one in which performance information is only partially revealed and integration is not viable.[13] We initially consider the simplest case of full asymmetric information and then, in Subsection 3.3, we analyze the case in which a (noisy) public signal is available.

We assume that the planner can force firms and workers to participate in any mechanism, and also that the planner can assign workers to firms in any way she chooses. This assumption implies that even slavery contracts are admissible.[14] Similarly, we assume that the planner faces no constraints on the transfers she can impose on players, e.g., there are no liquidity or budget-balance constraints. Our planner is thus completely unconstrained in her choices and actions; the only *endogenous* constraint the planner faces is incomplete information about the types of incumbent workers.

As in the case of symmetric information, transferring workers between firms of the same type, or from H firms to L firms, is inefficient. Thus again, here we only consider the possibility of transferring workers from L firms to H firms.

Because of Assumption 1, the planner wants to make sure that no H firm with $\tau \geq \gamma\mu$ dismisses its worker. This can be easily accomplished by setting the maximum payoff for H firms who dismiss workers at $\theta\gamma\mu$. Thus, as before, the planner only needs to consider as

[13] Technically speaking, integration is not viable under our assumptions because the managers of each type of firm have private information about the output produced by their own respective technologies.

[14] In other words, we don't require the mechanisms to satisfy individual rationality constraints. Our goal in this section is to show that incentive compatibility constraints are the main reason for our results.

potential poachers the set of H firms with workers with talent below $\gamma\mu$.

Let $\mathcal{A}$ denote the space of all allocations for workers initially paired with an L firm. For expositional simplicity, we restrict the space of admissible allocations to $\mathcal{A}^* \subset \mathcal{A}$, which is the space of all allocations for which, for a given $\widehat{\tau} \in [0, \overline{\tau}]$, $p_S = 1$ for all $\tau < \widehat{\tau}$ and $p_S = 0$ for all $\tau \geq \widehat{\tau}$. Although such a constraint substantially simplifies the presentation, it has no implications for the analysis, because this constraint is not binding when allocations are chosen optimally. In particular, note that any efficient allocation must belong to $\mathcal{A}^*$.

Without loss of generality, we can fully represent an allocation in $\mathcal{A}^*$ by a function $p(\tau) : [\widehat{\tau}, \overline{\tau}] \rightarrow [0, 1]$, where $p(\tau)$ is the probability that worker τ is assigned to set P. Henceforth, when there is no possibility of confusion, we will simply refer to $p(\tau)$ as an allocation.

Definition 2 *A* ***positive assortative matching (PAM) allocation*** *is an allocation in $\mathcal{A}^*$ such that, for any $\tau', \tau'' \in [\widehat{\tau}, \overline{\tau}]$, $\tau'' > \tau'$ implies $p(\tau'') \geq p(\tau')$, with strict inequality for at least one such pair.*

In other words, under a PAM allocation, more talented workers face a higher probability of being assigned to high-profitability firms. A *negative assortative matching* (*NAM*) allocation is defined analogously.

A *mechanism* $\langle p, t \rangle$ is an allocation rule $p(\tau^m)$ and a *transfer function* $t(\tau^m)$, where τ^m is a message sent by an L firm. We consider only symmetric mechanisms where the planner offers the same contract to all L firms. Thus, to simplify notation, we omit firm subscripts.

The timing is as follows. First, the planner offers (and commits to) a mechanism (i.e., a contract) to each incumbent L firm. Second, each incumbent firm sends a message $\tau^m \in [0, \overline{\tau}]$.[15] Third, the allocation is implemented (recall that all types $\tau^m < \widehat{\tau}$ are assigned to S with probability 1).

Let $U(\tau, \tau^m \mid p, t)$ denote the payoff of an incumbent firm with type τ from reporting τ^m under mechanism $\langle p, t \rangle$. An allocation p is *implementable* if there exists at least one transfer function t such that

$$\tau \in \arg \max_{\tau^m \in [0, \overline{\tau}]} U(\tau, \tau^m \mid p, t). \tag{4}$$

[15]Notice that, by appealing to the Revelation Principle, we can restrict the set of messages to the set of types.

In other words, p is implementable if there exists at least one transfer function such that truth-telling is incentive compatible.

Our first proposition restricts the set of implementable allocations:

Proposition 1 ***(The Flipping Property)*** *For any implementable allocation p, if $p(\tau') > p(\tau'')$ for some $\tau', \tau'' \in [\widehat{\tau}, \overline{\tau}]$, then it must be that $\tau' < \tau''$.*

Proof. The Revelation Principle implies that there is no loss of generality from focusing on truth-telling direct mechanisms. Define an incumbent firm's payoff function under mechanism $\langle p, t \rangle$ as

$$U(\tau, \tau^m \mid p, t) = \begin{cases} (1 - p(\tau^m))\,\tau + p(\tau^m)\,\gamma\mu + t(\tau^m) & \text{if } \tau^m \in [\widehat{\tau}, \overline{\tau}] \\ \gamma\mu + t(\tau^m) & \text{if } \tau^m \in [0, \widehat{\tau}) \end{cases}. \tag{5}$$

Note that an implicit assumption here is that a firm that loses its worker ends up employing a randomly selected worker from the outside pool.[16]

Suppose that an allocation p with $p(\tau') > p(\tau'')$ for some $\tau' > \tau''$ is implementable (i.e., it is incentive compatible for the firm to report $\tau^m = \tau$). Incentive compatibility requires

$$\begin{aligned} (1 - p(\tau'))\,\tau' + p(\tau')\,\gamma\mu + t(\tau') &\geq (1 - p(\tau''))\,\tau' + p(\tau'')\,\gamma\mu + t(\tau'') \\ t(\tau') - t(\tau'') &\geq [p(\tau') - p(\tau'')]\,(\tau' - \gamma\mu) \geq 0, \end{aligned}$$

but in this case, if τ'' deviates and reports τ', we have that $U(\tau'', \tau' \mid p, t) - U(\tau'', \tau'' \mid p, t)$ equals

$$t(\tau') - t(\tau'') - [p(\tau') - p(\tau'')]\,(\tau'' - \gamma\mu) \geq [p(\tau') - p(\tau'')]\,(\tau' - \tau'') > 0,$$

which implies that p is not incentive compatible. Finally, notice that we cannot have $p(\tau') > p(\tau'')$ for $\tau'' = \tau'$ because p must be a function. ■

Proposition 1 implies that, if the planner engages in matching based on types (i.e., if $p(\tau') \neq p(\tau'')$ for some $\tau' \neq \tau''$), only NAM can arise. In other words, incentive compatibility implies that an allocation that matches workers based on their types must exhibit NAM.

[16] We could have also included this possibility in the description of the mechanism, at the cost of some minor additional complexity.

This is our main result. The analysis in this section reveals that this *flipping property* is a consequence of asymmetric information alone.

Proposition 1 has a straightforward corollary:

Corollary 1 *There is no mechanism that implements a positive assortative matching allocation.*

Intuitively, Corollary 1 holds because, under a PAM allocation, the planner has to compensate a firm that risks losing a high-talent worker with a high monetary transfer to induce this firm to truthfully reveal the worker's type. If the planner does that, however, then a low-type firm, which cares less than the high-type firm about losing its worker, would prefer to pretend to have high type in order to receive a higher transfer.[17]

By contrast, it is easy to verify that all NAM allocations are implementable.

Another immediate consequence of the flipping property is that the first-best allocation is not implementable. From Subsection 3.1, we know that the first-best allocation implies

$$p^{FB}(\tau) = \begin{cases} 1 & \text{if } \tau \in \left[\tau^{\#}, \overline{\tau}\right] \\ 0 & \text{if } \tau \in \left[\gamma\mu, \tau^{\#}\right] \end{cases}, \tag{6}$$

which is a PAM allocation and thus not implementable.[18]

Although it is not surprising that non-monotonic allocations are not implementable, in our application this leads to an extreme form of inefficiency: Not only the first-best allocation is not implementable, but no allocation in which *some* better workers are more likely to match with better firms is implementable. That is, allocations in which PAM occurs for only a subset of types are also not implementable. There are no regions for which even imperfect matching can be efficiency enhancing.

Another class of implementable allocations is the set of allocations that exhibit no matching on types:

Definition 3 *A* ***matching-free allocation*** *is a function such that* $p(\tau) = c \in [0,1]$, *for all* $\tau \in [\widehat{\tau}, \overline{\tau}]$.

[17] Formally, Corollary 1 holds because all PAM allocations violate the typical monotonicity requirement for implementable decisions (here, for simplicity, we call a decision an allocation) under incomplete information (see, e.g., Fudenberg and Tirole, 1991, p. 260).

[18] More formally, for the non-generic case in which $\tau = \tau^{\#}$, $p^{FB}(\tau)$ can be any probability.

Under a matching-free allocation, the planner chooses to ignore the information revealed by the types in $[\widehat{\tau}, \overline{\tau}]$ when deciding to assign workers to firms. It is easy to see that matching-free allocations are also implementable.[19]

We now consider the optimal mechanisms. We postulate the planner's objective function as

$$\begin{aligned}\mathcal{S}(p, \widehat{\tau}) &= [F(\gamma\mu) H - L] \theta\gamma\mu + H \int_{\gamma\mu}^{\overline{\tau}} \theta\tau dF(\tau) + LF(\widehat{\tau})(\gamma\mu + \theta\gamma\mu) \\ &\quad + L \int_{\widehat{\tau}}^{\overline{\tau}} [p(\tau)(\theta\gamma\tau + \gamma\mu) + (1 - p(\tau))(\tau + \theta\gamma\mu)] dF(\tau). \end{aligned} \tag{7}$$

Note that, for simplicity, we assume that the planner only cares about the total surplus created by the allocation of talent, and not about the transfers. The planner maximizes $\mathcal{S}(p, \widehat{\tau})$ over all incentive-compatible mechanisms in $\mathcal{A}^*$.

Proposition 2 *The optimal mechanism implements a matching-free allocation $p^*(\tau) = c^*$ for $\tau \in [\widehat{\tau}, \overline{\tau}]$, such that*

$$\begin{aligned} &c^* = 1 \text{ and } \widehat{\tau} = \mu, \quad \text{if } E[\tau \mid \tau \geq \mu] \geq \tau^{\#} + k, \\ &c^* = 0 \text{ and } \widehat{\tau} = \gamma\mu, \quad \text{if } E[\tau \mid \tau \geq \mu] \leq \tau^{\#} + k, \end{aligned} \tag{8}$$

where

$$k \equiv \frac{\int_{\gamma\mu}^{\mu} (\tau - \gamma\mu) dF(\tau)}{(1 - F(\mu))(\theta\gamma - 1)}. \tag{9}$$

Proof. See the Appendix. ■

The economic intuition behind Proposition 2 is easier to grasp for the limiting case in which γ is close to 1 and $k \approx 0$. Ideally, the planner would like to assign all types in $[\mu, \overline{\tau}]$ that are higher than the critical type $\tau^{\#}$ to an H firm, and those below to an L firm. This is however not possible because PAM allocations are not feasible. The planner thus ignores the information revealed by incumbent firms and makes her decision by comparing the *expected type* $E[\tau \mid \tau \geq \mu]$ with the critical type $\tau^{\#}$. If the expected type is greater than the critical type, the planner assigns all workers in $[\mu, \overline{\tau}]$ to H firms. Similarly, if the expected type is

[19] To see this, suppose first that $c > 0$ and that the planner sets $t = 0$ for $\tau < \gamma\mu$ and $t = -\varepsilon$, with $\varepsilon > 0$, for $\tau \in [\gamma\mu, \overline{\tau}]$. All types less than $\gamma\mu$ report truthfully because they strictly prefer to replace the worker. All types such that $\tau \geq (\varepsilon/c) + \gamma\mu$ will also report truthfully. As we make $\varepsilon \to 0$, all types in $[\gamma\mu, \overline{\tau}]$ report truthfully. If $c = 0$ instead, then any flat transfer implements the allocation.

lower than the critical type, all workers in $[\mu, \overline{\tau}]$ are retained by incumbent L firms.[20]

Proposition 2 implies that the planner has to choose between the lesser of two evils: The planner either chooses to assign all incumbent workers with types greater than $\widehat{\tau}$ to L firms, or chooses to assign all such workers to H firms. Fine tuning the allocation of talent to efficiently match workers and firms is not possible. The first solution displays inefficient retention of the best workers – workers in $\left[\tau^{\#}, \overline{\tau}\right]$ are retained but should have been poached. The second solution displays inefficient poaching of the mediocre workers – workers in $\left[\mu, \tau^{\#}\right]$ are poached but should have been retained, and there is also inefficient firing of workers in $[\gamma\mu, \mu]$.

Remark 1. One straightforward implementation of the optimal mechanism is as follows. If the planner wants to set $c^* = 0$ (all workers in $[\gamma\mu, \overline{\tau}]$ retained), the planner should effectively forbid all poaching – workers should be forced to work for their incumbent firms. If instead the planner wants to set $c^* = 1$ (all workers in $[\mu, \overline{\tau}]$ poached), L firms should be heavily fined for retaining any worker.

Remark 2. Proposition 2 provides a justification for banning contracts in which firms own labor – i.e., quasi-slavery contracts. Even if workers voluntarily enter such contracts, these contracts generate externalities because there will be too much retention of high types. If the planner would like to set $c^* = 1$ but can only use regulatory tools, the planner may choose to ban non-compete clauses or other contracts that effectively give incumbent firms rights to retain their workers under most circumstances.

Remark 3. Under an optimal mechanism, the messages sent by incumbent firms are only useful for deciding which workers should be fired (i.e., assigned to set S). By assuming alternative technologies with PAM properties, we can construct a variation of the current model in which no one in an L firm should be fired. In such a model, the flipping property would force the planner to ignore all information revealed by the incumbent under a mechanism, making mechanism design irrelevant.

[20] The general (non-limiting) case of γ not close to 1 is slightly different because of an additional trade-off: If $c^* = 1$, there is inefficient firing of types in $[\gamma\mu, \mu]$, and thus the planner compares the expected type with the critical type *plus* some adjustment for the cost of inefficient firing, here measured by k.

3.3. Public Signals

We have thus far considered the case in which only the incumbent firm has any information about the talent of a worker. It is natural to ask whether our results, in particular the flipping property, change when the planner also has some information about workers.

Here we consider the case in which a publicly observable (and contractible) signal σ may reveal some information about τ. For example, σ could be a vector of worker characteristics (such as education, experience, etc.). This information could be produced in $t = 0$ and become publicly available at the beginning of $t = 1$. In this case, we say that σ is an *ex ante signal*: The planner observes the signal before offering a mechanism. Alternatively, σ could be an observable performance variable, such as a profit signal, which is observed only at the end of $t = 1$. In this case, σ is an *ex post signal*, which is observed only after the allocation is implemented.

Ex ante signals pose no additional complications to the analysis. After observing an ex ante signal σ, the planner will update her belief about the distribution of types; the planner now believes types are distributed according to F_σ. The analysis remains unchanged once F is replaced by F_σ.

Ex post signals, however, expand the set of options available to the planner. Although allocations cannot directly depend on the signal, the transfers can and they should now be written as $t(\tau^m, \sigma)$. With ex post signals, it is not obvious whether the planner should ignore the message τ^m when making matching decisions. This is because the planner can now impose large fines on firms that report types τ^m that are unlikely given σ, which should improve the planner's ability to induce incumbents to truthfully reveal information.

Do ex post signals eliminate the flipping property? This question cannot be answered without specifying the information structure induced by the signal σ. Consider for example an information structure under which firms' payoffs were perfectly observable. The planner could then propose a mechanism that mimics an "integrated firm" by offering incumbents a transfer that is identical to the poachers' payoff. This mechanism implements the first-best allocation. The interesting cases are thus those in which σ is not perfectly correlated with firms' payoffs, that is, cases in which the signal produces non-fully revealing information structures.

For arbitrary information structures, little can be said. Here we consider an example with a simple but rich information structure: a signal $\sigma \in \{1, ..., n\}$ that partitions the interval

$[\widehat{\tau}, \overline{\tau}]$ into n sets $\Omega_1, ..., \Omega_n$. Notice that even binary signals can be fully informative for matching purposes: if $n = 2$ with $\Omega_1 = [\gamma\mu, \tau^{\#}]$, knowledge of the public signal is sufficient for implementing the first-best allocation in $\mathcal{A}^*$: make $p = 0$ if $\sigma = 1$ and $p = 1$ if $\sigma = 2$. By contrast, if $\Omega_1 = [\gamma\mu, \overline{\tau}]$ and $\Omega_2 = ... = \Omega_n = \varnothing$, the signal is completely uninformative, and the analysis is then identical to the case in which no signal exists. All other cases represent some intermediate levels of informativeness.

We can now show the following:

Proposition 3 ***(The Conditional Flipping Property)*** *Suppose there is an ex post signal $\sigma \in \{1, ..., n\}$ that partitions the interval $[\gamma\mu, \overline{\tau}]$ into n sets $\Omega_1, ..., \Omega_n$. Then, the flipping property applies for each set Ω_σ, $\sigma \in \{1, ..., n\}$: For any implementable allocation p, if $p(\tau') > p(\tau'')$ for some $\tau', \tau'' \in \Omega_\sigma$, then it must be that $\tau' < \tau''$.*

Proof. The planner can always force an incumbent to reveal the set Ω_σ that contains the true τ by offering a transfer function such that $t(\tau^m, \sigma) = -\infty$ if $\tau^m \notin \Omega_\sigma$. Because an incumbent always knows the set Ω_σ that contains the true τ, the incumbent avoids such infinite fines by always reporting some τ' in Ω_σ. Conditional on observing σ, the planner can only distinguish between any two types $\tau', \tau'' \in \Omega_\sigma$ by relying on the incumbent's message. Thus, the planner's problem with an ex post signal σ is equivalent to choosing an allocation p_σ for each possible set Ω_σ, $\sigma \in \{1, ..., n\}$. By the same argument as in the proof of Proposition 1, if $p_\sigma(\tau') > p_\sigma(\tau'')$ then $\tau'' > \tau'$. Noticing that $p(\tau) = p_\sigma(\tau)$ for $\tau \in \Omega_\sigma$, $\sigma \in \{1, ..., n\}$, completes the proof. ■

We conclude that public signals do not eliminate the flipping property, which continues to hold conditionally for all those types that cannot be separated by the signal alone. As the signal becomes more informative, PAM allocations become possible precisely because information asymmetries are exogenously reduced: The signal eventually reveals some information that was initially private. However, in all cases in which information asymmetries persist and are not fully eliminated with time, the flipping property holds.

4. Decentralized Equilibrium

Here we consider a decentralized equilibrium version of the model. We make no important additional assumptions except for a *no-bonding assumption*: A worker is free to work for the highest bidder, and the current employer receives no compensation if the worker is poached by another firm. In such a market, an incumbent can only retain its worker by paying more than a competitor.

This assumption perhaps needs further elaboration. The solution to the planner's problem makes it clear that talent misallocation in general does not depend on this assumption, nor does the flipping property in particular. Allowing for bonding contracts will thus not restore efficiency. In fact, banning bonding contracts is sometimes efficient (see Remark 2 above).[21]

The reason that we rule out contracts that bind workers to firms is that we want our model to be different from an asset market model under private information. Note that, in the planner's problem, it makes no difference whether τ is an attribute of a worker or of a physical asset. However, in a decentralized equilibrium, this distinction is important: If the incumbent firm "owns" a worker of type τ, only excessive retention ("too little trade") can be observed. Since Akerlof (1970), this is a well understood result, to which we add by showing that too little trade is partly a consequence of the enforcement of ownership rights.

Our contribution, however, is to study the less well understood case of "too much trade." The assumption of unrestricted labor mobility distinguishes our model from models of trading of physical assets under asymmetric information. However, it should be clear that the reason for the lack of PAM is the incumbents' incentive-compatibility constraint, and that such a result holds even when bonding contracts are allowed, as illustrated by the planner's problem.

4.1. Setup

We retain the same setup as described in Section 2: a continuum of firms, with two types L and H, which differ from each other only because of their profitability parameter $i \in \{1, \theta\}$, and a continuum of agents in excess supply. All firms and agents live for two periods.

At $t = 0$, each firm hires a worker from the outside pool. Because all workers are ex ante identical, i.e. their types are distributed according to F, the initial pairing of firms and workers is random. For each type $i \in \{l, h\}$, we use subscripts ji to denote a unique firm

[21] Even when it is optimal to ban bonding contracts, incumbent firms may still choose to write such contracts. In the Internet Appendix, we present a setting in which a firm commits in $t = 0$ to a deferred compensation contract in which a worker is only paid at the end of the game, conditional on the worker not (voluntarily) quitting the firm. We show that such contracts, even when feasible, may not be voluntarily adopted by firms.

j of type i. Each firm ji offers a wage w_{ji} to a worker hired at $t = 0$. At the beginning of $t = 1$, this wage can be revised (upwards or downwards) based on the information that firms have at that time. Wages are always paid at the end of each period, conditional on the worker having stayed with the firm for that period.

Wage determination at $t = 0$ is a trivial problem. If there are no constraints on transfers from workers, firms will choose a negative wage to extract all future expected surplus from workers. If instead there is limited liability, for example, if wages cannot be negative, then this constraint will bind and wages will be set at the lowest level possible. Because there is nothing interesting happening in $t = 0$, here we focus on characterizing the equilibrium in $t = 1$ only. In the Internet Appendix, we solve a fully dynamic version in which, among other things, we characterize wages at all periods.

At the beginning of $t = 1$, all players face the following timing:

Date 1. Each firm ji learns the type $\tau_{ji} \in [0, \overline{\tau}]$ of its incumbent worker and independently offers the worker a wage w_{ji}.

Date 2. After observing all wage offers made by all firms in the sector, a firm j of type i that has a vacant position makes offers w^p_{ji} to incumbent workers; all firms act simultaneously.

Date 3. A worker who holds offers decides which offer, if any, to accept.

Date 4. All firms that do not have a worker at this date randomly select one agent from the outside pool. Outside pool agents accept to work for a wage of zero.

Date 5. Payoffs are realized.

We now explain the game in more detail.

At Date 1, after observing their incumbent workers' types, each firm j of type i simultaneously commits to a wage offer $w_{ji} \in \mathbb{R}$ to their incumbent workers. We permit strictly negative wage offers, as these offers will not be accepted, which implies that a negative wage offer is equivalent to dismissing the incumbent worker.

Define the set of retention wages for firms of type L as $W_l = \{w : w = w_{jl} \text{ for some } jl \in L\}$, and define W_h analogously. The set of all possible W_i is the set $\mathcal{W}_i$.

At Date 2, firms who have dismissed their workers (i.e., they have made negative wage offers to their incumbent workers) now have vacancies, and will thus make poaching offers w^p_{ji} after observing some $W \equiv W_l \cup W_h \in \mathcal{W}_l \cup \mathcal{W}_h$. Without loss of generality, we restrict the analysis to the case in which only H firms make poaching offers. This restriction is not binding in equilibrium, because, for the same worker, H firms would always make better offers than L firms. Thus, we impose this restriction only to simplify the notation.

Importantly, poachers do not observe the incumbent workers' types. Instead, they form beliefs regarding these types after observing W. Poachers believe that the unconditional distribution of τ is $F(.)$.[22] We assume that all poachers share the same beliefs, whether on or off any equilibrium path, which is a usual assumption in sequential games with incomplete information that use Perfect Bayesian Equilibrium (PBE) as a solution concept. Thus, we denote by $F^W(\tau \mid w, i)$ the common belief about τ that poachers hold after observing a worker who has an offer of w from a firm of type i, when the set of all offers to workers is W.

Let $\omega_i(w)$ denote the set of incumbent workers in firms of type i who hold offer w. We assume that each poacher can make an offer (a *poaching wage*) $w^p_{jh}(w, i)$ to all workers in set $\omega_i(w)$. We will also write the poaching wage as $w^p_{jh}(w, i, W)$ whenever we wish to emphasize that poaching wages are equilibrium strategies, and as such, they depend on the set of all observed wages W.

Because of Assumption A1, "poachable" workers are in short supply, thus if we assume – as we do – that poachers compete among themselves in Bertrand fashion, no poacher can have a payoff larger than the outside payoff $\theta\gamma\mu$ (recall that only type H firms are poachers). After observing the set of offers, W, a poacher offers

$$w^p_{jh}(w, i, W) = \theta\gamma \left(\int_0^{\overline{\tau}} \tau dF^W(\tau \mid w, i) - \mu \right) \tag{10}$$

to all workers in the set $\omega_i(w)$.[23] As above, if $w^p_{jh}(w, i, W) < 0$, the offer is not accepted, which means that a negative poaching wage offer is equivalent to no offer. Because the right-hand side of (10) doesn't depend on jh, for simplicity we now drop this subscript from function w^p.

At Date 3, workers always agree to work for the maximum non-negative wage that is offered to them. We make the following assumption:

Assumption A2 A worker in $\omega_i(w)$ accepts all offers where $w^p(w, i) > w$ and rejects all

[22] This assumption is made for the sake of simplicity; nothing important changes if the unconditional c.d.f. for incumbent workers is $G(.) \neq F(.)$.

[23] It is easy to see from (10) that if L firms were allowed to make offers, these offers would be dominated by the offers made by H firms.

offers where $w^p(w,i) \leq w$.

In other words, if indifferent, a worker stays with her current employer, which is a standard assumption in the literature (see, e.g., Waldman, 1984). However, this assumption entails some loss of generality, because it eliminates a number of equilibria in mixed strategies. Thus, we consider Assumption A2 as an equilibrium selection criterion with intuitive properties: Workers may have a small bias against changing jobs because of unmodeled costs.[24]

Comments on timing and assumptions. The timing assumes that incumbent firms move before poachers. Changing the timing such that incumbent firms move after poachers and make the final offer makes retention easier, but does not fundamentally affect the qualitative properties of the equilibrium.[25]

There are only two dates when meaningful decisions are made: Dates 1 and 2, i.e., when incumbent firms and poachers, respectively, choose their actions.[26] We only consider pure strategies at Date 1, but this is without loss of generality; the continuum assumption allows for mixing at the population level. The assumption that poachers only play pure strategies at Date 2 is also without loss of generality because of Assumption A2.

The assumption that an incumbent firm makes an offer to its worker is meant to imply that workers have no bargaining power vis-à-vis incumbent firms; workers either accept their firms' offers or move elsewhere. Alternatively, there could also be multiple rounds of offers and counter-offers by incumbents and poachers. We assume a single round as a simple way of introducing costs of delayed negotiations.[27]

We assume that the continuation game beginning at Date 3 must be in equilibrium regardless of the history of play, and that if $w^p(w,i,W) > w$, each poacher is matched with a worker with equal probability (i.e., we assume random rationing).

[24] Relaxing this assumption makes mixed-strategy equilibria possible. A complete characterization and discussion of mixed-strategy equilibria can be found in the Internet Appendix.

[25] A complete analysis of the case in which incumbents move last can be found in the Internet Appendix.

[26] At Date 4, there is no meaningful choice because both incumbent firms and poachers are strictly better off by hiring a worker from the outside pool than by keeping a job post unfilled. Date 3 is when workers who hold (potentially multiple) offers choose whether to stay or leave. This decision is not strategic.

[27] Alternatively, we could also consider a situation in which there are potentially infinite rounds of offers and counter-offers, in which each additional round introduces a cost paid by the incumbent (equivalently, the incumbent discounts the future). Because poachers are competitive, the incumbent may face a different bidder for its worker in each round. In this modified game, the incumbent would immediately offer either the wage that would retain the worker or any wage that would not lead to retention.

4.2. Symmetric Information

In this subsection, we briefly discuss the benchmark case of symmetric information.

Suppose that all firms have the same information. Suppose a type H firm has a vacancy at Date 2 (that is, at Date 1, it dismissed its incumbent worker). We call such a firm a poacher. Poachers compete à la Bertrand for each type $\tau \in [0, \overline{\tau}]$. Their profits must equal their outside payoff $\theta\gamma\mu$. Thus, the poaching wage offered to type τ is given by

$$w^{pS}(\tau) = \theta\gamma(\tau - \mu), \tag{11}$$

where the superscript S denotes symmetric information.

In a subgame perfect equilibrium, incumbent firm ji solves $\max_{w \in \mathbb{R}} \pi_{ji}(w)$, where

$$\pi_{ji}(w) = \begin{cases} \tau_{ji} - w & \text{if } w \geq \max\{\theta\gamma(\tau_{ji} - \mu), 0\} \\ i\gamma\mu & \text{otherwise} \end{cases}. \tag{12}$$

Suppose first that $\tau_{ji} \leq \mu$. In this case, firm i does not have to worry about poaching and will pay $w_{ji} = 0$ if $\tau_{ji} \in [\gamma\mu, \mu]$ and some $w_{ji} < 0$ if $\tau_{ji} < \gamma\mu$ (in other words, it dismisses the worker).

If instead $\tau_{ji} > \mu$ and firm ji wants to retain a worker, then the firm must offer at least as much as a poacher, that is, w_{ji} must be equal to or greater than $\theta\gamma(\tau_{ji} - \mu) > 0$. Then, ji's payoff is $\pi_{ji} = i\tau_{ji} - \theta\gamma(\tau_{ji} - \mu)$, which implies that this is an optimal choice if and only if $i\tau_{ji} - \theta\gamma(\tau_{ji} - \mu) \geq i\gamma\mu$. If $i = h$, this condition holds always, thus implying that, in equilibrium, no worker is poached from an H firm. An H firm's optimal strategy regarding its incumbent worker is summarized by:

$$w^S_{jh} = \begin{cases} \text{any } w < 0 & \text{if } \tau_{jh} \leq \gamma\mu \\ 0 & \text{if } \tau_{jh} \in [\gamma\mu, \mu] \\ \theta\gamma(\tau_{jh} - \mu) & \text{if } \tau_{jh} \in [\mu, \overline{\tau}] \end{cases}. \tag{13}$$

Now the analysis that follows refers to L firms only. If $\theta\gamma \leq 1$, condition $\tau_{jl} - \theta\gamma(\tau_{jl} - \mu) \geq \gamma\mu$ is true for any $\tau_{jl} > \mu$ (recall that $\theta > 1$). If $\theta\gamma > 1$, this condition holds for any $\tau_{jl} \leq (\theta - 1)\gamma\mu / (\theta\gamma - 1)$.

This reasoning implies that an L firm's optimal strategy is to offer

$$w_{jl}^{S}=\begin{cases}\text{any } w<0 & \text{if } \tau_{jl}\leq\gamma\mu\\ 0 & \text{if } \tau_{jl}\in[\gamma\mu,\mu]\\ \theta\gamma\left(\tau jl-\mu\right) & \text{if } \tau_{jl}\in\left[\mu,\tau^{\#}\right]\\ \text{any } w<w^{pS}\left(\tau_{jl}\right) & \text{if } \tau_{jl}\geq\tau^{\#}\end{cases}, \tag{14}$$

where $\tau^{\#}$ is the critical type as defined in (2). Recall that, for simplicity, we always use closed intervals to denote the equilibrium sets of types.

In equilibrium, L firms retain all types in $\left[\gamma\mu,\tau^{\#}\right]$. Types that are lower than $\gamma\mu$ are fired and not poached. Types higher than $\tau^{\#}$ are poached in equilibrium. We conclude that the equilibrium of this game implements the first-best allocation as defined in (3).

4.3. Asymmetric Information

4.3.1. Equilibrium: Definition

We now define the equilibrium conditions under asymmetric information. Poachers' strategies are given by the function $w^{p}\left(w,i,W\right)$, as defined in (10). We denote an incumbent firm's strategy by $w_{ji}\in\mathbb{R}$, and a given set of such strategies for each firm type is denoted by $\tilde{w}_{l}\equiv\{w_{jl}:jl\in L\}$ and $\tilde{w}_{h}\equiv\{w_{jh}:jh\in H\}$.

Recall that we defined the c.d.f. $F^{W}\left(\tau\mid w,i\right)$ as the common belief of poachers about the type of worker who is offered w in firm i when W is observed. Beliefs are given by a family of functions $F^{W}\left(\tau\mid w,i\right)$ defined for each $W\in\mathcal{W}$. We denote such a family of functions simply by F^{W}. Workers' beliefs about their own types do not influence equilibrium outcomes, because (i) workers do not know more than outsiders and (ii) optimal worker behavior depends only on wage offers, regardless of their beliefs. Thus we do not include workers' beliefs in the definition of the equilibrium.

Let $\pi_{ji}(w_{ji},w^{p}\left(w_{ji},i,W\right))$ denote the expected payoff to firm i if it chooses to offer w_{ji} to its worker, while poachers play strategy $w^{p}\left(w,i,W\right)$. Note first that this payoff does not depend directly on the strategies of other incumbent firms or on poachers' beliefs F^{W}; knowledge of the poaching wage $w^{p}\left(w_{ji},i,W\right)$ is sufficient for firm ji to forecast its payoff. Second, note that firm ji can compute $\pi_{ji}(w_{ji},w^{p}\left(w_{ji},i,W\right))$ with no ambiguity because we assume that poaching wages $w^{p}\left(w_{ji},i,W\right)$ are given by (10) and are common knowledge.

Finally, because there is a continuum of firms $ji \in L \cup H$, and a continuum of types $\tau \in [0, \overline{\tau}]$, for each set of pure strategies $\tilde{w} \equiv \tilde{w}_l \cup \tilde{w}_h$, there is a unique W, which occurs with probability 1. The difference between $\tilde{w}$ and W is that the former keeps track of which firms $ji \in L \cup H$ made which offer, while the latter only contains those offers made by each type of firm, without distinguishing, within each type, among the firms that made such offers. We denote the set of wages W induced by strategy $\tilde{w}$ by $W(\tilde{w})$.

Definition 4 *A strategy profile $(\tilde{w}, w^p(w, i, W))$ and a family of belief functions F^W constitute an equilibrium of the game if*

(i) for each $ji \in L \cup H$, $w_{ji} \in \tilde{w}$ only if $w_{ji} \in \arg\max_{w \in \mathbb{R}} \pi_{ji}(w, w^p(w, i, W(\tilde{w})))$;

(ii) poaching wages $w^p(w, i, W)$ are given by (10); and

(iii) all poachers hold identical beliefs $F^W(\tau \mid w, i)$ for all $w \in W$ and all $W \in \mathcal{W}$. These beliefs must be consistent with Bayes's rule for all $w \in W(\tilde{w})$. Poachers believe that the incumbent firms behave independently of one another, which specifically implies that, if $ji \neq j'i$, $F^W(\tau_{ji}, \tau_{j'i} \mid w_{ji}, w_{j'i}, i) = F^W(\tau_{ji} \mid w_{ji}, i) \cdot F^W(\tau_{j'i} \mid w_{j'i}, i)$ for all $W \in \mathcal{W}$.

This definition is equivalent to a Perfect Bayesian Equilibrium (PBE). Conditions (i)-(ii) are the standard requirement that the equilibrium strategies are best responses to one another.

Condition (iii) not only requires that beliefs are updated by Bayes's rule whenever possible, but also imposes some additional weak restrictions on beliefs off the equilibrium path. As usual in PBE definitions with many players, we require all poachers to hold the same beliefs, both on and off the equilibrium path. We also require that beliefs depend only on W, which is mostly for tractability. This is a slightly stronger restriction because it implies $F^W(\tau_{ji} \mid w, i) = F^W(\tau_{j'i} \mid w, i)$ for any $ji, j'i \in L \cup H$. One interpretation is that all incumbent firms of the same type i are observationally identical; thus, they cannot be differentiated by poachers when these firms play the same wage w_i in equilibrium.

4.3.2. Equilibrium: Characterization

To characterize the equilibium, we begin by making two additional simplifying assumptions to deal with equilibrium multiplicity. We first assume that an incumbent would never make an offer that is weakly dominated by making no offer:

Assumption E1 Incumbent ji offers $w_{ji} \geq 0$ only if $i\tau_{ji} - w_{ji} \geq i\gamma\mu$.

We also assume the following:

Assumption E2 (*Divinity*) After observing an off-the-equilibrium-path wage w', poachers believe that the probability that type $\tau' \geq \frac{w'}{i} + \gamma\mu$ deviates is no less than the probability that type $\tau'' > \tau'$ deviates.

Assumption E2 is a technical assumption that restricts the set of admissible off-the-equilibrium-path beliefs. This assumption is an adaptation to our setup of the Divinity Criterion of Banks and Sobel (1987). Assumption E2 is not particularly restrictive and is compatible with (infinitely) many off-the-equilibrium beliefs; thus, it does not eliminate equilibrium multiplicity. None of our main conclusions depends on this assumption.[28]

The role of Assumptions E1 and E2 is to restrict the set of equilibria; thus, they may be interpreted as equilibrium selection criteria. They simplify the analysis significantly, although they do not eliminate equilibrium multiplicity.

We now state some preliminary results.

Lemma 1 *In any equilibrium, all workers retained by a type-i firm are offered the same wage.*

This important result has a very simple proof. Suppose that there are two types, τ' and τ'', where $\tau'' > \tau'$. Suppose that the incumbent wishes to retain both types. Suppose also that $w'' > w'$ (the argument is analogous if $w'' < w'$). This situation cannot be an equilibrium because there is a profitable deviation for an incumbent with worker τ''. Indeed, the incumbent prefers to offer w' to a worker of type τ''. Type τ'' would nonetheless be retained, but at a lower wage.

Lemma 2 *Any equilibrium must have a threshold property: If type τ' is retained by an incumbent in equilibrium, type $\tau'' > \tau'$ must also be retained in equilibrium.*

[28]The intuition for Assumption E2 is as follows. For concreteness, suppose that type τ'' is retained by an L firm in an equilibrium with wage w'', while type $\tau' \in [w' + \gamma\mu, \tau'')$ is not retained (the intuition for the other cases is analogous to this example). An incumbent with a worker of type τ'' that deviates and offers this type wage w' can only benefit from the deviation if poachers offer $w^p(w') \leq w'$. However, for this set of poaching wages, type τ' would also benefit from a deviation. On the other hand, type τ'' would be worse off if $w^p(w') > w'$, whereas type τ' would not be worse off. Thus, the logic of Banks and Sobel's Divinity Criterion requires that the probability of τ' deviating should be no less than that of τ'' deviating.

This is again easily proven: For a given retention wage, w, if it is optimal to retain τ' ($i\tau' - w \geq i\gamma\mu$), then it is also optimal to retain any τ such that $\tau \geq \tau'$.

Lemma 2 is a manifestation of the flipping property and, as such, is a consequence of the incumbent's informational advantage.

We now prove the following proposition:[29]

Proposition 4 *An equilibrium exists. All equilibria have the following properties:*

1. *There is a unique $\tilde{\tau}_i \in [\gamma\mu, \overline{\tau}]$ such that, for each firm type $i \in \{l, h\}$, all types $\tau \geq \tilde{\tau}_i$ are retained. Threshold $\tilde{\tau}_i$ is the same for all equilibria and is either $\overline{\tau}$ or the least element of the set of fixed points of*

$$G_i(\tau_i) \equiv \frac{w_i^*(\tau_i)}{i} + \gamma\mu, \tag{15}$$

   where

$$w_i^*(\tau_i) = \theta\gamma\left(\int_{\tau_i}^{\overline{\tau}} x dF(x \mid x \geq \tau_i) - \mu\right) \tag{16}$$

   is the wage offered to retained workers whose types are greater than τ_i.

2. *All types $\tau \in [0, \gamma\mu]$ are fired in equilibrium (wages are negative).*

3. *There is a subset of types $P_i \subseteq [\gamma\mu, \tilde{\tau}_i]$ that are poached in equilibrium and a subset of types $S_i \subseteq [\gamma\mu, \tilde{\tau}_i]$ that are fired in equilibrium (wages are negative), with $S_i \cup P_i = [\gamma\mu, \tilde{\tau}_i]$.*

4. *If $\tau_i \in P_i$, then the incumbent offers some $w_i' \in [0, w^p(w_i'))$, where*

$$w^p(w) = \theta\gamma\left(\int_0^{\overline{\tau}} x dF^W(x \mid w) - \mu\right) \tag{17}$$

   and $F^W(\tau \mid w_i', i) = F(\tau \mid \tau_i \in P_i)$.

Proof. See the Appendix. ■

[29] In what follows, for simplicity, we define all equilibrium sets of types as closed intervals. That is, we refrain from specifying what happens in equilibrium in the knife-edge cases in which an incumbent is indifferent between retaining or not retaining a type. The equilibrium is unaffected by what happens in these cases.

This proposition shows that, in equilibrium, incumbent workers will find themselves in one of the following three situations: unemployed, employed by their incumbent firm, or employed by a high productivity poacher. Part 1 of Proposition 4 implies that the very best workers will typically be retained by the incumbent firm, which implies that if workers are retained at all, they must be the best workers. Because this is true even when the incumbent firm has low productivity, the equilibrium will be inefficient due to the lack of PAM.

Part 2 of Proposition 4 implies all types $\tau < \gamma\mu$ are fired because the unemployment replacement value is higher. Part 3 implies that worker types in $[\gamma\mu, \tilde{\tau}_i]$ are not retained by an incumbent i, and will be either fired or poached. Sets P_i and S_i are not pinned down because the incumbent is indifferent to how workers leave the firm.

Only those workers with talent in the interval $[\gamma\mu, \tilde{\tau}_i]$ may be poached in equilibrium. With some abuse of language, we call these workers *mediocre workers*, although in some cases this interval will also contain the very best workers (e.g., if $\tilde{\tau}_i$ is close to or equal to $\overline{\tau}$).

Proposition 4 also reveals that equilibria differ from one another (meaningfully) only because the sets P_i and S_i may differ.[30] Because we focus on the efficiency properties of the equilibria, it is natural to select the most efficient equilibrium as the focal equilibrium:[31]

Corollary 2 *There is a most-efficient equilibrium in which* $P_i = [\mu, \tilde{\tau}_i]$ *and* $S_i = [\gamma\mu, \mu]$.

We prove the existence of this equilibrium in the proof of Proposition 4. This equilibrium is the most efficient one because any other equilibrium must have either some $\tau' < \mu$ who is poached or some $\tau' > \mu$ who is fired, or both. In the former case, allocational efficiency can be improved by firing τ'. In the latter case, allocational efficiency can be improved by letting poachers hire τ'.

In the most efficient equilibrium, the equilibrium outcome changes monotonically with τ: As τ increases, outcomes change from unemployment to poaching, and then from poaching to retention. Note that the most-efficient decentralized equilibrium implements a NAM

[30] Two observationally equivalent equilibria with the same P_i and S_i may also differ from one another because they are sustained by different beliefs off the equilibrium path and may display different wages offered by incumbent firms for types in P_i.

[31] An alternative refinement would be to focus on equilibria in which the incumbent firm only fires workers with abilities less than $\gamma\mu$. This requirement has intuitive properties; the incumbent offers zero wage to mediocre workers and retains them with a vanishingly small probability (i.e., only if poachers "tremble" and do not make offers). Although our results would remain basically unchanged, except for the elimination of inefficient firing, under such a refinement, equilibrium existence is not always guaranteed and may depend on distributional assumptions.

allocation, whenever workers from L firms are poached by H firms, and is thus a third-best allocation.

To focus on the most interesting case, in the remainder of this paper, we assume that the following condition holds:

Condition G $\max_{\tau_l \in [0,\overline{\tau})} \tau_l - G_l(\tau_l) > 0$.

Under Condition G, there is an interior $\tilde{\tau}_i < \overline{\tau}$ such that all types $\tau_i \in [\tilde{\tau}_i, \overline{\tau}]$ are retained by i incumbents. Condition G always holds for any set of parameters if $\overline{\tau} \to \infty$. If Condition G does not hold, incumbent firms never retain any worker in equilibrium, i.e., $\tilde{\tau}_i = \overline{\tau}$.

4.3.3. Equilibrium: Efficiency

The most-efficient equilibrium implies that types $\tau_l \in [\tilde{\tau}_l, \overline{\tau}]$ are retained by incumbents of type L and $\tau_l \in [\mu, \tilde{\tau}_l]$ are poached. That is, the most-efficient decentralized equilibrium implements a (deterministic) NAM allocation. By contrast, allocational efficiency requires a PAM allocation, as in (3). Thus, the most efficient equilibrium does not lead to an efficient allocation of talent, which is formally stated in the next corollary.

Corollary 3 *The most-efficient decentralized equilibrium under asymmetric information is talent-allocation inefficient. In particular, there are three different sources of misallocation of talent:*

1. ***Excessive firing**: Types $\tau_i \in [\gamma\mu, \mu]$ are fired but should have been retained.*

2. ***Excessive poaching of mediocre types:** Types $\tau_l \in [\mu, \min\{\tau^{\#}, \tilde{\tau}_l\}]$ and $\tau_h \in [\mu, \tilde{\tau}_h]$ are poached but should have been retained.*

3. ***Insufficient poaching of high types**: Types $\tau_l \in [\max\{\tau^{\#}, \tilde{\tau}_l\}, \overline{\tau}]$ are retained but should have been poached.*

The corollary above shows that the possibility of poaching creates three distortions relative to the first-best scenario. Incumbent firms do not attempt to retain some workers who are potential poaching targets, which leads to excessive turnover. Such turnover results in misallocation of talent because some workers who have acquired firm-specific skills are either inefficiently fired (Case 1) or inefficiently poached by high-profitability firms (Case 2). Thus, the equilibrium displays a "Peter Principle Property": Mediocre workers are "promoted" to positions in better firms, whereas the best workers stay with their current employers. This is a key empirical prediction of the model, and an illustration of the flipping property. Finally, low-productivity firms might be too successful in retaining workers who would otherwise be matched with better firms in the first-best allocation. In other words, there might be too little poaching in equilibrium (Case 3).

4.3.4. Comparative Statics

To perform comparative statics, we focus on two parameters with intuitive interpretations. The first is θ, which could be interpreted as the (cross-sectional) dispersion in firm profitability (or revenue). Because firm profitability in reality may be positively related to firm survival and growth, θ can also be interpreted as a measure of heterogeneity in firm size. The second parameter, γ, measures the importance of general skills relative to firm-specific skills. Alternatively, an increase in γ can also be interpreted as a decrease in the cost of recruiting a new worker (e.g., search costs).

It's immediate from (15) and (16) that θ has no effect on $\widetilde{\tau}_h$. However, θ does affect $\tilde{\tau}_l$. Under Condition G, $\tilde{\tau}_l$ is the least fixed point of $G_l(\tau)$. Because $G_l(0) < 0$, it follows that $1 - G_l'(\tilde{\tau}_l) > 0$. By the implicit function theorem, we have

$$\frac{d\tilde{\tau}_l}{d\theta} = \gamma \frac{\int_{\tilde{\tau}_l}^{\overline{\tau}} \tau f(\tau)\, d\tau - [1 - F(\tilde{\tau}_l)]\,\mu}{[1 - F(\tilde{\tau}_l)]\,[1 - G_l'(\tilde{\tau}_l)]} > 0. \tag{18}$$

In other words, the retention threshold for L firms increases with the profitability dispersion parameter θ. Intuitively, as L and H firms become more heterogeneous, L firms find it increasingly more difficult to retain workers, and are thus only able to retain the very best workers. This result also means that job mobility increases with θ, because the set of poached workers $[\mu, \tilde{\tau}_l]$ increases with $\tilde{\tau}_l$. This increase in mobility can be either efficient or inefficient. For example, if $\tilde{\tau}_l < \tau^{\#}$, then increasing θ leads to more inefficient poaching.

The effect of the importance of general skills relative to firm-specific skills is also easily inferred from

$$\frac{d\tilde{\tau}_i}{d\gamma} = \frac{\theta}{i} \frac{\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau f(\tau)\, d\tau - [1 - F(\tilde{\tau}_i)]\,\mu\left(\frac{\theta - i}{\theta}\right)}{[1 - F(\tilde{\tau}_i)]\,[1 - G_i'(\tilde{\tau}_i)]} > 0. \tag{19}$$

Both retention thresholds increase with the relative importance of general skills γ. Again, this result is intuitive: There is more poaching when general skills are more important (i.e., if skills are more portable). Therefore, an increase in γ also increases job mobility. An increase in poaching from H firms is always inefficient. An increase in poaching from L firms may be either efficient or inefficient.

Our model also has predictions for wages. Consider for example $w_i^*(\tilde{\tau}_l)$, which is the wage paid to workers retained by L firms. From (16) and (18) we have

$$\frac{dw^*(\tilde{\tau}_l)}{d\theta} = \gamma\left(E\left(\tau \mid \tau \geq \tilde{\tau}_l\right) - \mu\right) + \theta\gamma \frac{dE\left(\tau \mid \tau \geq \tilde{\tau}_l\right)}{d\tilde{\tau}_l}\frac{d\tilde{\tau}_l}{d\theta} > 0. \tag{20}$$

If a worker is first hired with a zero wage (as it would happen if, for example, they could not be paid negative wages), then $w_i^*(\tilde{\tau})$ measures the increase in earnings for those workers who are retained by their firms. This result shows that within-job earnings growth increases in the dispersion of firm payoffs.

5. Applications and Implications

Here we discuss some applications and novel positive and normative implications of our model.

Mutual fund managers One of the key predictions of our analysis is that internally promoted workers (i.e., workers who are retained by their firms) are more talented than externally promoted workers (i.e., workers who are poached and promoted to higher-level positions). This implication is unique to models in which employers are better informed about the talent of their workers. Ours is the first model in the asymmetric employer learning literature in which firms may rationally poach mediocre workers from other firms.

This prediction should only apply to industries in which asymmetric information about talent is important. An industry that fits this description is the mutual fund industry. The main task of mutual fund managers is to construct and manage stock portfolios. As a regulated industry, US mutual funds are required to disclose their portfolio holdings quarterly. This means that all changes in portfolio holdings between two disclosure dates remain private information to the firm.

It is difficult for outsiders to assess the skill of mutual fund managers. Fund performance is a very noisy measure of stock-picking and market timing skill because of the difficulty of separating true skill from luck. Nevertheless, there is evidence that some mutual fund managers have skill and are thus able to command higher fees (see, e.g., Berk and van Binsbergen, 2015). Mutual fund executives have an advantage over outsiders when evaluating their managers because, unlike outsiders, they can observe all of their managers' choices (i.e., portfolio changes) and their explanations.

Berk, van Binsbergen, and Liu (forthcoming) find that mutual fund firms are able to identify their best managers, who are then internally "promoted" (i.e., they are allocated more assets to manage). Such internal promotions appear to add value to the company. By contrast, managers who are externally promoted – i.e., they are poached by other mutual funds and experience an increase in assets under management – do not appear to add much value. The authors conclude that their evidence indicates that mutual fund executives have private information about their managers, and that this information is useful for allocating capital to managers.

Our model provides a possible explanation for the most puzzling aspect of this evidence, which is the fact that mutual fund firms hire outsiders who deliver mediocre results. In our model, poacher firms understand that workers who are let go are mediocre, but these workers are still preferable to workers of whom we know very little. A normative implication of our analysis is that external promotions are socially inefficient in these cases, which suggests that mutual fund manager mobility is excessive.

The software industry The assumptions of our model fit well with innovative industries, where talent is crucial but difficult to measure. Andersson et al. (2009) study compensation patterns in a number of sectors of the software industry. They find that sectors in which there is more dispersion in potential payoffs offer higher earnings growth for those employees who are retained by their firms. In our model, the dispersion in potential payoffs is measured by the profitability differential parameter θ. From (20), we see that within-job earnings growth is higher in sectors with larger θ.

Non-compete clauses Contractual clauses that forbid workers to quit and work for competitors are common in a number of high-skill occupations. Such non-compete clauses typically impose a quarantine period before a worker can join another firm in the same industry and/or geographical region.

Non-compete clauses offer an imperfect solution to the problem of excess mobility, because they are of limited duration, and because the definitions of both the applicable industry and geographical area may be fuzzy in practice. Furthermore, non-compete clauses are controversial and have often been challenged in courts. For example, in California non-compete clauses are considered void and non-enforceable, except in a small set of cases.

Our analysis reveals that non-compete clauses can be socially inefficient, justifying their restriction in some cases. In sectors or industries with substantial firm heterogeneity (i.e., high θ), both $\tau^{\#}$ (the critical type) and k are low, making it more likely that the second-best allocation involves poaching of all types (see Proposition 2). Intuitively, when firm heterogeneity is high, preventing workers from moving from low-productivity firms to high-productivity firms is very inefficient. However, if non-compete clauses are available, low-productivity firms may use them to avoid poaching. Although it might be rational for workers and firms to write such contracts, the parties involved do not internalize the negative effects of mobility restrictions on high-productivity firms that have vacancies. Information asymmetry is the friction that makes non-compete agreements inefficient.

6. Conclusions

We show that adverse selection in the labor market may generate (inefficient) negative assortative matching of workers and firms. To the best of our knowledge, this is a new result. It is also a somewhat unexpected result. Although it is not surprising that informational asymmetries generate inefficiencies, it is not obvious that such inefficiencies can fully offset the tendency for the best firms to hire only the best workers, especially if there are strong technological complementarities between talent and firm quality. Previous research has emphasized the fact that adverse selection in the labor market implies insufficient job mobility for high-ability workers (e.g., Greenwald, 1986). We show that there can be too much job mobility for mediocre workers.

Our results also illustrate how labor markets differ from asset markets. A simple asset market under adverse selection never displays too much trade. Our model makes the reason for this result clear: The institution of asset ownership forces buyers to compensate sellers fairly in asset exchanges, which prevents the poaching of assets against the seller's will. In contrast, too much (i.e., inefficient) "trade" in labor markets may arise precisely because labor cannot be owned by firms, and because contracts that mimic ownership of labor are rare.

References

Acemoglu, D., and Pischke, J-S.1998. Why Do Firms Train? Theory and Evidence. *Quarterly Journal of Economics* 113: 79–119.

Acemoglu, D., and Pischke, J-S. 1999. The Structure of Wages and Investment in General Training. *Journal of Political Economy* 107: 539–572.

Acharya, V., Pagano, M., and Volpin, P. 2016. Seeking Alpha: Excess Risk Taking and Competition for Managerial Talent. *Review of Financial Studies* 29: 2565-2599.

Adriani, F., and Deidda, L. 2009. Price Signaling and the Strategic Benefits of Price Rigidities. *Games and Economic Behavior* 67: 335–350.

Akerlof, GA. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics* 84: 488–500.

Andersson, F., Freedman, M., Haltiwanger, J., Lane, J., and Shaw, K. 2009. Reaching for the Stars: Who Pays for Talent in Innovative Industries? *Economic Journal* 119: F308-F332.

Atkeson, A. and P.J. Kehoe. 2005. Modeling and Measuring Organization Capital. *Journal of Political Economy* 113: 1026-53.

Banks, J. and Sobel, J. 1987. Equilibrium Selection in Signaling Games. *Econometrica* 55: 647–661.

Bar-Isaac, H., Jewitt, I., and Leaver, C. 2014. Asymmetric Information and Adverse Selection. *Working Paper*

Bernhardt, D. 1995. Strategic Promotion and Compensation. *Review of Economic Studies* 62: 315–339.

Bernhardt, D., and D. Scoones. 1993. Promotion, Turnover, and Preemptive Wage Offers. *American Economic Review* 83: 771–791.

Bénabou, R. and Tirole, J. 2016. Bonus Culture: Competitive Pay, Screening, and Multitasking. *Journal of Political Economy.* 124: 305–370.

Berk, J. B., van Binsbergen, J. H. 2015. Measuring Skill in the Mutual Fund Industry. *Journal Financial Economics* 118, 1–20.

Berk, J. B., van Binsbergen, J. H., and Liu, B. Matching Capital and Labor. *Journal of Finance*, forthcoming.

Chang, C., and Y. Wang. 1996. Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited. *Journal of Labor Economics* 14: 505–519.

Daley B. and B. Green. 2012. Waiting for News in the Market for Lemons. *Econometrica* 80: 1433–1504.

Edmans, A., and X. Gabaix. 2011. The Effect of Risk on the CEO Market. *Review of Financial Studies* 24: 2822-2863.

Edmans, A., Gabaix, X., and A. Landier. 2009. A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium. *Review of Financial Studies* 22: 4881-4917.

Eisfeldt, A. L., and C. M. Kuhnen. 2013. CEO Turnover in a Competitive Assignment Framework. *Journal of Financial Economics* 109: 351-372.

Ellingsen, T. 1997. Price Signals Quality: The Case of Perfectly Inelastic Demand. *International Journal of Industrial Organization* 16: 43–61.

Franco, A., Mitchell, M., and G. Vereshchagina, 2011. Incentives and the Structure of Teams. *Journal of Economic Theory* 146: 2307-2332.

Frydman, C. 2013. Rising Through the Ranks: The Evolution of the Market for Corporate Executives, 1936-2003. Working paper

Fuchs, W. and A. Skrzypacz. 2015. Costs and Benefits of Dynamic Trading in a Lemons Market. Working paper

Fudenberg, D., and J. Tirole. 1991. Game Theory. MIT Press.

Gabaix, X., and A. Landier. 2008. Why Has CEO Pay Increased so Much? *Quarterly Journal of Economics* 123: 49–100.

Gibbons, R., and L. Katz. 1991. Layoffs and Lemons. *Journal of Labor Economics* 9: 351-380.

Gibbons, R., Katz, L., Lemieux, T., and D. Parent, 2005. Comparative Advantage, Learning, and Sectoral Wage Determination. *Journal of Labor Economics* 23: 681-724.

Golan, L. 2005. Counteroffers and Efficiency in Labor Markets with Asymmetric Information. *Journal of Labor Economics* 23: 373-393.

Greenwald, B. 1986. Adverse Selection in the Labour Market. *Review of Economics Studies* 53: 325–47.

Hoppe, H., Moldovanu B., and A. Sela. 2009. The Theory of Assortative Matching Based on Costly Signals. *Review of Economics Studies* 76: 253–281.

Kahn, L. 2013. Asymmetric Information between Employers. *American Economic Journal: Applied Economics*, 5: 165-205.

Kaya, A. 2013. Two-sided Matching with Private Information. Working Paper

Kaya, A. and G. Vereshchagina, 2014. Partnerships versus Corporations: Moral Hazard, Sorting and Ownership Structure. *American Economic Review*, 104: 291–307.

Kaya, A. and G. Vereshchagina, 2015. Moral Hazard and Sorting in a Market for Partnerships. *Economic Theory*, 60: 73–121.

Laing, D. 1994. Involuntary Layoffs in a Model with Asymmetric Information Concerning Worker Ability. *Review of Economics Studies* 61: 375–92.

Lazear, E. 1986. Raids and Offer Matching. *Research in Labor Economics* 8:141–65.

Levin, J. 2001. Information and the Market for Lemons. *Rand Journal of Economics* 32:657–666.

McGahan, A. M., and M. E. Porter. 1997. How Much Does Industry Matter, Really? *Strategic Management Journal* 18: 15-30.

Milgrom, P., and S. Oster. 1987. Job Discrimination, Market Forces, and the Invisibility Hypothesis. *Quarterly Journal of Economics* 102: 453-476.

Milgrom, P., and N. Stokey. 1982. Information, Trade, and Common Knowledge. *Journal of Economic Theory* 26: 17-27.

Mukherjee, A., 2008. Career Concerns, Matching, and Optimal Disclosure Policy. *International Economic Review* 49:1211–1250.

Murphy, K. J., and J. Zabojnik. 2004. CEO Pay and Appointments: A Market Based Explanation for Recent Trends. *American Economic Review* 94: 192–196.

Murphy, K. J., and J. Zabojnik. 2006. Managerial Capital and the Market for CEOs. Queen's University, working paper.

Myerson, R. B., and M. A. Satterthwaite. 1982. Efficient Mechanisms for Bilateral Trade. *Journal of Economic Theory* 29: 265-281.

Pinkston, J. C. 2009. A Model of Asymmetric Employer Learning with Testable Implications. *Review of Economic Studies* 76: 367-394.

Ricart i Costa, J. 1988. Managerial Task Assignment and Promotions. *Econometrica* 56: 449–466.

Schönberg, U. 2007. Testing for Asymmetric Employer Learning. *Journal of Labor Economics* 25: 651-692.

Rumelt, R. 1991. How Much Does Industry Matter? *Strategic Management Journal* 12: 167–185.

Syverson, C. 2011. What Determines Productivity? *Journal of Economic Literature* 49: 326-365.

Terviö, M. 2008. The Difference that CEOs Make: An Assignment Model Approach. *American Economic Review* 98: 642–668.

Terviö, M. 2009. Superstars and Mediocrities: Market Failure in The Discovery of Talent. *Review of Economic Studies* 72: 829–850.

Waldman, M. 1984. Job Assignments, Signaling, and Efficiency. *Rand Journal of Economics* 15: 255–267.

Waldman, M. 1990. Up-or-Out Contracts: A Signaling Perspective. *Journal of Labor Economics* 8: 230–250.

Waldman, M., and O. Zax, 2016. An Exploration of the Promotion Signaling Distortion. *Journal of Law, Economics, and Organization* 32: 119–149.

Zabojnik, J., and D. Bernhardt, 2001. Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation. *Review of Economic Studies* 68: 693–716.

A. Appendix: Omitted Proofs

Proposition 2.

Proof. If $\theta\gamma \leq 1$, or $\theta\gamma > 1$ and $(\theta-1)\,\gamma\mu/\,(\theta\gamma-1) \geq \overline{\tau}$, (3) implies that the first-best outcome can be achieved by a matching-free allocation with $\widehat{\tau} = \gamma\mu$ and $p(\tau) = 0$ for all $\tau \in [\gamma\mu, \overline{\tau}]$.

If $\theta\gamma > 1$ and $(\theta-1)\,\gamma\mu/\,(\theta\gamma-1) < \overline{\tau}$, the first best-outcome requires a PAM allocation, which is not feasible, because any feasible allocation must be non-increasing in $\tau \in [\widehat{\tau}, \overline{\tau}]$. To solve for the optimal mechanism, we proceed in two steps. First, we find the set of optimal mechanisms for a given $\widehat{\tau}$; $m(\widehat{\tau})$ denotes the set of all such mechanisms. Second, we find the $\widehat{\tau}$ that maximizes surplus among all mechanisms in $\{m(\widehat{\tau}) : \widehat{\tau} \in [\gamma\mu, \overline{\tau}]\}$.

Take $\widehat{\tau}$ as given and consider an implementable allocation p. For any given τ' we have

$$p(\tau')(\theta\gamma\tau' + \gamma\mu) + (1 - p(\tau'))(\tau' + \theta\gamma\mu) = p(\tau')\left[(\theta\gamma-1)\,\tau' - (\theta-1)\,\gamma\mu\right] + \theta\gamma\mu + \tau', \quad (21)$$

If $\tau' \in \left[\widehat{\tau}, \tau^{\#}\right]$, (21) is decreasing in $p(\tau')$ because $\tau' \leq (\theta-1)\,\gamma\mu/\,(\theta\gamma-1)$. Thus, $\mathcal{S}(p, \widehat{\tau})$ can be weakly increased by (pointwise) replacing $p(\tau')$ with $p\left(\tau^{\#}\right)$ for all $\tau' \in \left[\widehat{\tau}, \tau^{\#}\right]$ (recall that p must be non-increasing because of Proposition 1). By the same argument, if $\tau'' \in \left[\tau^{\#}, \overline{\tau}\right]$, the planner can increase surplus by replacing $p(\tau'')$ with $p\left(\tau^{\#}\right)$. Thus the optimal allocation must be a matching-free allocation $p(\tau) = c$, with surplus

$$\mathcal{S}(p, \widehat{\tau}) = Q + LF(\widehat{\tau})(\gamma\mu + \theta\gamma\mu) + L\int_{\widehat{\tau}}^{\overline{\tau}} (\theta\gamma\mu + \tau)\, dF(\tau) + cL\int_{\widehat{\tau}}^{\overline{\tau}} [(\theta\gamma - 1)\tau - (\theta - 1)\gamma\mu]\, dF(\tau), \tag{22}$$

where Q is a constant given by

$$Q \equiv [F(\gamma\mu) H - L]\theta\gamma\mu + H\int_{\gamma\mu}^{\overline{\tau}} \theta\tau dF(\tau). \tag{23}$$

The optimal choice of c will depend on the last term of function (22), which can be rewritten as

$$cL\int_{\widehat{\tau}}^{\overline{\tau}} [(\theta\gamma - 1)\tau - (\theta - 1)\gamma\mu]\, dF(\tau) = cL(1 - F(\widehat{\tau}))[(\theta\gamma - 1)E(\tau \mid \tau \geq \widehat{\tau}) - (\theta - 1)\gamma\mu], \tag{24}$$

which implies that the optimal choice of c is

$$c^* = \begin{cases} 0 & \text{if } E(\tau \mid \tau \geq \widehat{\tau}) \leq \tau^{\#} \\ 1 & \text{if } E(\tau \mid \tau \geq \widehat{\tau}) \geq \tau^{\#} \end{cases}. \tag{25}$$

Now, if $c^* = 0$, the optimal $\widehat{\tau}$ is $\gamma\mu$, because an incumbent is better off retaining any type above $\gamma\mu$ than hiring from the outside pool. If $c^* = 1$, the optimal $\widehat{\tau}$ is μ, because an H firm with a vacancy is better off employing any type above μ than hiring from the outside pool. Thus, the optimal mechanism requires either $c^* = 0$ and $\widehat{\tau} = \gamma\mu$ or $c^* = 1$ and $\widehat{\tau} = \mu$. The mechanism that implements $c^* = 1$ (all workers above μ poached) is optimal if

$$\int_{\mu}^{\overline{\tau}} (\theta\gamma\tau + \gamma\mu) dF(\tau) + \int_{\gamma\mu}^{\mu} (\gamma\mu + \theta\gamma\mu) dF(\tau) \geq \int_{\gamma\mu}^{\overline{\tau}} (\tau + \theta\gamma\mu) dF(\tau), \tag{26}$$

which can be rewritten as

$$\int_{\mu}^{\overline{\tau}} [(\theta\gamma - 1)\tau - \gamma\mu(\theta - 1)]\, dF(\tau) \geq \int_{\gamma\mu}^{\mu} (\tau - \gamma\mu) dF(\tau), \tag{27}$$

$$(1 - F(\mu))\left[(\theta\gamma - 1)E(\tau \mid \tau \geq \mu) - \gamma\mu(\theta - 1) - \frac{\int_{\gamma\mu}^{\mu} (\tau - \gamma\mu) dF(\tau)}{1 - F(\mu)}\right] \geq 0, \tag{28}$$

The result then follows by defining

$$k = \frac{\int_{\gamma\mu}^{\mu} (\tau - \gamma\mu) dF(\tau)}{(1 - F(\mu))(\theta\gamma - 1)}. \tag{29}$$

■

Proposition 4.

Proof. *Part 1.* Lemma 2 implies that an equilibrium with retention must have thresholds $\tilde{\tau}_i$ as in Part 1 of the proposition (If no type τ_i for an i firm is retained in equilibrium, we set $\tilde{\tau}_i = \overline{\tau}$ and this part is trivially proved). Lemma 1 implies that all types τ_i in $[\tilde{\tau}_i, \overline{\tau}]$ are paid the same wage. To prevent poaching, this wage must be such that $w_i^* \geq w^p(w_i^*)$. Because poachers know that all types in $[\tilde{\tau}_i, \overline{\tau}]$ are offered w_i^*, their beliefs must be given by $F(\tau \mid \tau \geq \tilde{\tau}_i)$ upon observing w_i^*, which implies

$$w^p(w_i^*) = \theta\gamma\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau \mid \tau \geq \tilde{\tau}_i) - \mu\right). \tag{30}$$

Because $w_i^* \geq w^p(w_i^*)$, a necessary condition for an incumbent firm of type i who has a worker of type $\tau_i \in [\tilde{\tau}_i, \overline{\tau}]$ not to deviate and fire the worker is

$$\tau_i - \frac{w^p(w_i^*)}{i} \geq \gamma\mu, \tag{31}$$

a condition that is equivalent to

$$\tilde{\tau}_i = \frac{w^p(w_i^*)}{i} + \gamma\mu. \tag{32}$$

We now show that $w_i^* = w^p(w_i^*)$ if the equilibrium threshold is $\tilde{\tau}_i$. Suppose first that $w_i^* > w^p(w_i^*)$ and consider a deviation from an incumbent with type $\tau_i > \tilde{\tau}_i$ who chooses to offer $w^p(w_i^*)$ instead. For this not to constitute a profitable deviation, we must have that

$$w^p(w^p(w_i^*)) > w^p(w_i^*) = \theta\gamma\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau \mid \tau \geq \tilde{\tau}_i) - \mu\right), \tag{33}$$

which can only happen if

$$\int_0^{\overline{\tau}} \tau dF^W(\tau \mid w^p(w_i^*)) > \int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau \mid \tau \geq \tilde{\tau}_i).$$

This condition requires the existence of at least one $\tau'' > \tilde{\tau}_i \geq \frac{w^p(w_i^*)}{i} + \gamma\mu$ such that its probability of deviation is strictly greater than that of some type $\tau' \in (\tilde{\tau}_i, \tau'')$. However, this is ruled out by Assumption E2. Thus, $w_i^* = w^p(w_i^*)$.

Next, we show that $\tilde{\tau}_i$ is unique. Define the function

$$\tau_i - G_i(\tau_i) = \tau_i - \frac{\theta\gamma}{i}\left(\int_{\tau_i}^{\overline{\tau}} x dF(x \mid x \geq \tau_i) - \mu\right) - \gamma\mu. \tag{34}$$

Clearly, the existence of an equilibrium with retention requires this function to be non-negative for some τ_i (to see this, insert (33) in (32)). Because $G_i(\tau_i)$ is continuous and $-G_i(0) = -\gamma\mu < 0$, at least one fixed point of G_i exists if and only if

$$\max_{\tau_i \in [0,\overline{\tau})} \tau_i - G_i(\tau_i) \geq 0. \tag{35}$$

(If (35) does not hold, the unique equilibrium displays no retention). This condition always holds for $i = \theta$, but it may or may not hold for $i = 1$. Assuming that (35) holds, we define the threshold $\tilde{\tau}_i$ as the least element of the set of fixed points of $G_i(\tau)$:

$$\tilde{\tau}_i = \min_{\{x: G_i(x) = x\}} x. \tag{36}$$

Clearly, $\tilde{\tau}_i \geq \gamma\mu$. At $\tilde{\tau}_i$, the incumbent is just indifferent between retaining the worker for w_i^* or firing the worker:

$$\tilde{\tau}_i = \frac{\theta\gamma}{i}\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau \mid \tau \geq \tilde{\tau}_i) - \mu\right) + \gamma\mu. \tag{37}$$

To show that this threshold is part of an equilibrium, notice first that because $G_i(0) > 0$, $\tau_i - G_i(\tau_i)$ crosses zero from below at $\tilde{\tau}_i$, which is also a necessary condition for an equilibrium. We only need to show that no other $\tau > \tilde{\tau}_i$ can be an equilibrium. To see this, suppose that there is $\tau' > \tilde{\tau}_i$ such that only types $\tau > \tau'$ are retained at wage

$$w' = \theta\gamma\left(\int_{\tau'}^{\overline{\tau}} \tau dF(\tau \mid \tau \geq \tau') - \mu\right). \tag{38}$$

Then, an incumbent with type $\tilde{\tau}_i + \varepsilon$, with $\varepsilon > 0$ arbitrarily small, could deviate and offer

$w_i^* < w'$, with

$$w_i^* = \theta\gamma \left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau \mid \tau \geq \tilde{\tau}_i) - \mu \right). \tag{39}$$

If type $\tilde{\tau}_i + \varepsilon$ is successfully retained after this deviation, then the incumbent is strictly better off. For such a deviation not to be profitable, poachers' beliefs must be such that $w^p(w_i^*) > w_i^*$. However, again, this could only be true if there were some type $\tau'' > \tilde{\tau}_i$ whose probability of deviation is strictly greater than that of a type $\tau \in (\tilde{\tau}_i, \tau'')$. This is ruled out by Assumption E2. Thus, $\tilde{\tau}_i$ is uniquely determined as the least fixed point of $G_i(\tau_i)$, and the retention wage is given by w_i^* as in (39).

Part 2. It follows trivially from Assumption E1.

Part 3. Suppose that there is some type τ_i' in $[\gamma\mu, \tilde{\tau}_i]$ that is retained in equilibrium. Lemma 2 implies that all types in $[\tau_i', \tilde{\tau}_i]$ are also retained, and Lemma 1 implies that all types in $[\tau_i', \overline{\tau}]$ must be paid the same wage. However, because $\tau_i' \leq \tilde{\tau}_i$, then by the definition of $\tilde{\tau}_i$ in (36), we have $\tau_i' - G_i(\tau_i') \leq 0$. Thus, type τ_i' cannot be profitably retained. Thus, all types in $[\gamma\mu, \tilde{\tau}_i]$ must be either poached (and thus included in set P_i) or fired (and thus included in set S_i). Thus, if an equilibrium exists, Part 3 must hold.

Part 4. If $\tau_i \in P_i$, then the incumbent must offer some wage w_i' that is lower than the poaching wage $w^p(w_i')$. Because poachers' beliefs must be Bayesian on the equilibrium path, then

$$w^p(w) = \theta\gamma \left(\int_0^{\overline{\tau}} x dF^W(x \mid w) - \mu \right), \tag{40}$$

and poachers' beliefs are given by $F^W(\tau \mid w_i', i) = F(\tau \mid \tau_i \in P_i)$

To complete the proof, we only need to show that at least one equilibrium exists. Suppose first that $\max_{\tau_l \in [0,\overline{\tau})} \tau_l - G_l(\tau_l) > 0$. In this case, we know that there exists a unique pair $\{\tilde{\tau}_l, \tilde{\tau}_h\} < \{\overline{\tau}, \overline{\tau}\}$. The following fully characterizes one possible equilibrium:

Consider the retention wages

$$w_i(\tau_i) = \begin{cases} w_i^* & \text{if } \tau_i \in [\tilde{\tau}_i, \overline{\tau}] \\ 0 & \text{if } \tau_i \in [\mu, \tilde{\tau}_i] \\ -1 & \text{if } \tau_i \in [0, \mu] \end{cases}, \tag{41}$$

the poaching wages on the equilibrium path

$$w^p(w_i) = \begin{cases} w_i^* & \text{if } w_i = w_i^* \\ \theta\gamma(\int_{\mu}^{\tilde{\tau}_i} \tau dF(\tau \mid \tau \in [\mu, \tilde{\tau}_i]) - \mu) & \text{if } w_i = 0 \\ -1 & \text{if } w_i = -1 \end{cases}, \tag{42}$$

and beliefs such that $F(\tau \mid \tau \geq \frac{w_i}{i} + \gamma\mu)$ for any w_i that is off the equilibrium path. In this equilibrium, $P_i = [\mu, \tilde{\tau}_i]$ and $S_i = [\gamma\mu, \mu]$.

If we instead have $\max_{\tau_l \in [0,\overline{\tau})} \tau_l - G_l(\tau_l) \leq 0$, nothing is change for h firms. For l firms, no type τ_l is retained, and an equilibrium in which all types $\tau_l \geq \mu$ are offered $w_l = 0$, and types below μ are fired, exists and is sustained by beliefs such that $F(\tau \mid \tau \geq w_l + \gamma\mu)$ for any w_l that is off the equilibrium path. This equilibrium implies $P_l = [\mu, \overline{\tau}]$ and $S_l = [\gamma\mu, \mu]$.

■