# Recruiting Talent*

Simon Board[†], Moritz Meyer-ter-Vehn[‡], and Tomasz Sadzik[§]

July 31, 2016

**Abstract**

We propose a model of firm dynamics in which a firm's primary asset is the talent of its workforce. Firms compete in wages to attract applicants, and managers seek to identify the most talented. Over time, a firm's quality evolves as today's recruits become tomorrow's managers. If talent is scarce, firm-applicant matching is positive assortative, with better firms posting higher wages and attracting better applicants. As a result, the economy converges to a steady state featuring persistent dispersion in talent, wages and productivity. Along the path, if firms are initially similar, then high-wage firms incur short term losses while they accumulate the talent that guarantees a sustainable competitive advantage. We also show that equilibrium leads to an inefficient selection of talent into the industry, and can be improved by policies that reduce wage dispersion.

## 1 Introduction

The success of a firm is typically built upon thousands of decisions made by hundreds of employees, making it critical for firms to identify and recruit the best talent. In addition to their direct contribution to productivity, talented workers identify and recruit the next generation of employees. In this paper, we propose a model of firm dynamics in which long-run success depends on building and maintaining a talented workforce. We show that differences in human capital can serve as a source of sustainable competitive advantage, and study the trajectories of talent, wages and productivity over time, and the dispersion across firms.

Our model is based on two key assumptions. First, talent is critical for many firms' success, meaning that recruiting is a top priority. This is self-evident in our own profession, where university rankings depend on the quality of professors rather than the number of buildings. Using data from Nazi Germany, Waldinger (2016) shows that a loss of human capital had a large, persistent effect on university quality over 50 years, whereas a loss of physical capital had a small, temporary effect.

*We are particularly grateful to Steve Tadelis for a conversation that inspired this project. We have received useful comments from Dirk Bergemann, V Bhaskar, Johannes Hörner, Peter Kondor, Alessandro Lizzeri, Benny Moldovanu, Volker Nocke, Kevin Roberts, and Larry Samuelson. We thank seminar audiences at Berlin, Bonn, Cambridge, Edinburgh, Games 2016, Gerzensee 2016, Leicester, Mannheim, Oxford, Queen Mary, UBC, UCL, UCLA, and Yale. Keywords: labor market competition, adverse selection, assortative matching, firm dynamics, productivity dispersion, wage dispersion.

†UCLA, http://www.econ.ucla.edu/sboard/

‡UCLA, http://www.econ.ucla.edu/mtv/

§UCLA, https://www.econ.ucla.edu/faculty/regular/Sadzik.html

Talent is also critical for technology firms, and in service industries such as consulting, law, or sales. For example, the Netflix HR manual states that "one outstanding employee gets more done and costs less than two adequate employees; we endeavor to have only outstanding employees" (Hastings and McCord, 2009). As as result, the leaders "continually told managers that building a great team was their most important task" (McCord, 2014).[1]

Second, we assume that talented employees are better at identifying talented applicants. At universities, an expert can more accurately assess the quality of an applicant than a novice. In addition, more talented employees typically have a richer network of contacts that provides them with more information about potential applicants. This view is consistent with Waldinger's (2012, 2016) finding that the loss of star faculty led to a permanent reduction in the quality of hires and thus the quality of the department, but did not affect the productivity of current faculty. The literature on referrals provides further evidence about this channel. Looking at three industries, Burks et al. (2015) find that employee referrals provide useful information about applicants, and that more productive employees make more frequent and higher quality referrals. In their field experiment, Beaman and Magruder (2012) find similar results and, in addition, show that productive agents have a better ability to predict the performance of their referrals than unproductive agents, implying that more talented agents have better information.

Based on these two assumptions, we show that when talent is scarce, firms with a talented workforce post high wages, attract better applicants, and hire better recruits. Employee excellence thus perpetuates across generations, meaning that the economy converges to a unique steady state featuring persistent heterogeneity in talent, productivity and wages. Intuitively, the positive assortative matching in the labor market counters the regression of mediocrity present in all firms, inducing a natural degree of dispersion to which the economy converges over time. Our results can thus account for the large, persistent differences in productivity and wages across firms documented by Foster, Haltiwanger, and Syverson (2008) and Abowd, Kramarz, and Margolis (1999), and the positive assortative matching of workers found by Card, Heining, and Kline (2013).

In Section 2, we analyze a static model of labor market competition. A continuum of firms competes for a continuum of workers that have high or low ability. Firms attract applicants by posting wages, and then receive a noisy signal about each applicant, helping them screen the wheat from the chaff.[2] Firms with more skilled recruiters receive more accurate signals. We suppose the market is frictionless: The highest-paying firm attracts all applicants, and hires the first applicant who produces a good signal (e.g. passes the firm's test). Other firms hire from the remaining, adversely selected pool of workers. The quality of a firm's applicant pool thus depends on its ordinal rank in the wage distribution, giving rise to a continuous equilibrium wage distribution.

If talent is scarce, firm-worker matching is *positive assortative*. That is, firms with more skilled recruiters post higher wages and attract better applicants. Intuitively, at high wages the applicant pool is balanced, and skilled recruiters have a comparative advantage at screening; but at low wages the remaining applicants are so bad that recruiting skill yields little benefit. As a result, firms with skilled recruiters hire better workers, meaning that quality tends to persist.

In Section 3, we turn to the full, dynamic model. Workers retire at an exogenous rate and, in order to fill the resulting vacancies, firms compete to recruit from a pool of new workers, as in

[1] There are many other studies documenting the heterogeneity of individual level productivity, ranging from senior managers of major firms (Bertrand and Schoar, 2003) to shop floor workers (Lazear, 2000).

[2] This "direct screening" differs from the "self-selection" use of the word "screening", e.g. Stiglitz (1975).

the static model. Crucially, we suppose that firms' recruiting skill corresponds to the number of talented workers. This means that both the productivity and talent of the firm evolve as the firm hires recruits who, in turn, become the new recruiters.

As in the static model, firms with more talent pay higher wages because of the complementarity between screening skills and applicant quality. Thus, equilibrium is unique with talented firms posting high wages, attracting good applicants and hiring talented recruits, reinforcing their initial advantage. These dynamics converge to a unique steady state, irrespective of the initial distribution of talent, meaning that dispersion arises endogenously over time. Steady-state dispersion balances two forces: positive assortative matching which amplifies differences across firms, and imperfect screening which leads to mean regression and equalizes firms. Human capital is thus a source of sustainable competitive advantage. While low-quality firms could in principle catch up by posting higher wages and hiring more talented workers, it is not profitable for them to do so.

To get a sense of the equilibrium dynamics, suppose all firms start off with (almost) identical talent. Over time the distribution of talent converge to steady state, with the high-talent firms hiring from good applicant pools and accumulating more talent, while the low firms hire from poor, deteriorating pools, and lose talent. Top firms thus initially pay high wages and lose money; they later recoup this investment when their accumulated talent raises their productivity and provides a competitive edge in the labor market. Thus wages tend to fall over time as firms become differentiated and the lower firms are no longer able to effectively compete for talent.

In Section 4, we enrich the dynamic model in two directions. First, we introduce heterogeneity in firm fundamentals so that firms differ in both their (exogenous) marginal product of talent, called their "technology", and their (endogenous) stock of talent. Equilibrium wages rise in both technology and talent, so a low-technology firm may initially outbid a high-technology firm if it has significantly more talent. Nevertheless, as talent evolves over time, the economy converges to a unique steady state in which matching is *stratified*, meaning that firms with high technology also accumulate more talent. Hence, the initial two-dimensional matching problem ultimately collapses to a single dimension. This suggests that universities with good faculty but poor fundamentals are inherently unstable in the long-run. We then use the model to look at how a firm adjusts to an exogenous increase in technology. We show that while such a firm raises its wages, it optimally hires recruits below its new steady-state level. This is because the firm is both poor at screening and posts wages below the steady-state level. Intuitively, skilled management complements high-quality applicant pools, so it is a mistake to pay high wages before the management can use the extra resources wisely.

Our second application is to investigate the impact of peer effects on the distribution of talent and wages. We suppose that all workers prefer to work at a firm with more talented colleagues, and show that equilibrium matching is unchanged, as is the evolution of talent. Wages, however, may be quite different. Since workers care about total utility, high-talent firms may be able to post lower wages than low-talent firms and still attract better applicants. Moreover, if firms start with similar talent, then the introduction of peer effects increases initial wage dispersion, but reduces it in the long run. Intuitively, firms have an extra incentives to invest in their stock of talent which they can later harvest in terms of lower wages.

In Section 5 we study welfare. Since every worker is equally productive at any firm, this depends solely on the aggregate sorting of talent into the industry. In contrast to classic matching models (e.g. Shapley and Shubik (1971), Becker (1973)), equilibrium in our model is inefficient. In particular, if a social planner could determine the order in which firms screen applicants, she would have low-quality firms screen first, so that matching would be negative assortative. Intuitively, when a high-quality firm picks first, this introduces more adverse selection in the applicant pool than when a low-quality firm chooses first. This negative externality outweighs the private gain from positive assortative matching. One implication is that if the planner cannot contract on firm types, but only choose a set of admissible wages, then her optimal mechanism is to have the firms choose in a random order; this can be implemented by only allowing trade at a single wage, or by imposing wage cap. Hence, while productivity dispersion is not *prima facie* evidence of misallocation in our model, our equilibrium does exhibit more dispersion than is socially optimal (cf. Hopenhayn (2014)).

### 1.1 Literature

Our paper is most closely related to the literature on matching with transfers. This was started by Shapley and Shubik (1971) and Becker (1973) who showed that equilibrium is efficient and that, for complementary production functions, matching is positive assortative. In the context of firms, Lucas (1978) and Kremer and Maskin (1996) considered a model where agents are divided into managers and workers, who then interact via a multiplicative production function. Garicano (2000) and Garicano and Rossi-Hansberg (2006) endogenize the production function and number of layers in the hierarchy, while Levin and Tadelis (2005) consider how a large mass of agents form partnerships.

A number of recent papers introduced dynamics into matching models. Anderson and Smith (2010) and Anderson (2016) suppose agents match each period and evolve as a function of the match. They prove that equilibrium is efficient, and derive sufficient conditions for matching to be positive assortative. Jovanovic (2014) considers an OLG model where young and old workers work together, enhancing the human capital of the young worker. He shows that noise in the match leads to misallocation, lowering the growth of the economy.

In contrast to this literature, our paper focuses on the hiring process, rather than production complementarities. This is important since "the literature has been less successful at explaining how firms can find the right employees" (Oyer and Schaefer, 2011). We propose that hiring is a skill that firms possess in different quantities, study how this interacts with firms' compensation policies, and analyze how the allocation of talent across firms evolves over time. We show that hiring provides a micro-foundation for the complementarities that give rise to positive assortative matching. And, unlike classic matching models, exhibits inefficiencies due to the presence of adverse selection.

Our static screening model is related to a variety of models of adverse selection.[3,4] The closest is Kurlat's (2016) model of financial markets with adverse selection in which buyers receive heterogeneous signals about sellers' assets. Our model assumes that firms' signals are independent, and show that firms post different wages, matching is positive assortative and inefficient. In comparison, Kurlat assumes that firms' signals are nested, meaning that a more informed firm knows everything that a less informed firm knows. He shows that firms post the same price with ties broken in favor of the less-informed agents, and equilibrium is efficient. Intuitively, when signals are nested, a less

[3]Such direct screening has been used before, in different contexts. Guasch and Weiss (1980) observed that tests can be used to separate workers with private information about their types. Lockwood (1991) uses a random search model with screening to argue that the duration of unemployment serves as a signal of worker quality.

[4]There is also a literature on matching with incomplete information and interdependent values. For stability notions, see Chakraborty, Citanna, and Ostrovsky (2010) and Liu et al. (2014).

informed firm imposes no externality on a more informed firm when they move first.

Our static model also relates to the literature on wage dispersion. Albrecht and Vroman (1992) consider a random-matching model in which agents differ in their reservation wage and show that firms are tempted to raise their wage to appeal to less enthusiastic workers. Burdett and Mortensen (1998) propose a model with homogeneous workers and on-the-job search, in which firms pay more to poach workers from their competitors. In these papers, dispersion derives from firms competing for more workers in an economy with matching frictions, whereas our dispersion derives from firms competing for better workers in an economy with adverse selection. Moreover, we focus on how differences in talent and productivity evolve over time as a result of selective hiring.

Our paper provides a theory of firm dynamics in which a firm's talent as its key strategic asset. Models of firm dynamics were first studied by Lucas and Prescott (1971) and Hopenhayn (1992), where firms' competitive advantage is determined by their technology and demand state. Jovanovic (1982) and Board and Meyer-ter-Vehn (2013) suppose firms' self-esteem and reputation serve as a state variable. Hopenhayn and Rogerson (1993) and Fuchs, Green, and Papanikolaou (2016) suppose that firms differ in their labor and capital stock, respectively, which are slow moving due to firing costs and adverse selection. By focusing on talent, our paper provides a new channel through which firms can sustain a competitive advantage that is particularly relevant for service industries (e.g. universities, technology) which make up over 70% of GDP. It also gives rise to predictions concerning the dispersion and persistence of productivity, wages and employee quality.

Finally, there are themes in our paper that echo those in dynamic models of political economy. As in our paper, Dewan and Myatt are interested in the evolution of talent within organizations. Dewan and Myatt (2010) focus on firing standards, arguing that as the government ages, its talent pool depletes and it standards fall. Dewan and Myatt (2014) focus on recruitment, supposing that a government that can recruit better talent as it becomes more successful, creating a feedback loop. Our model also captures the idea that when a firm hires a worker, it also hire his tastes. This is related to the literature on dynamic clubs, where today's members must decide who will make decisions tomorrow. Acemoglu and Robinson (2000), Lizzeri and Persico (2004) and Jack and Lagunoff (2006) consider the extension of the voting franchise, while Sobel (2001), Barberà, Maschler, and Shalev (2001) and Roberts (2015) consider general club goods.

# 2 Static Model

In this section we introduce the static competitive screening model. Section 2.1 describes how heterogeneous firms compete to attract and identify talent. Section 2.2 shows that wages are dispersed in equilibrium. Section 2.3 provides sufficient conditions for positive assortative matching.

## 2.1 Model

A unit mass of firms, each with one vacancy, competes for a unit mass of workers by posting wages. Workers and firms are both heterogeneous. Workers differ in their talent, with proportion $\bar{q} \in (0, 1)$ being talented and the remainder being untalented. Firms differ in their recruiting skill. In particular, each firm has a unit mass of recruiters, of which proportion $r$ is skilled, $\theta = H$, and $1 - r$ is unskilled, $\theta = L$. Firms' screening skills $r$ are distributed with pdf $f$, cdf $F$, and support

$[\underline{r}, \bar{r}]$.[5]

Firms simultaneously post wages; workers then apply to firms in order of wages. As in the Gale-Shapley mechanism, the firm with the highest wage picks a worker; the remainder then apply to the "second" firm, and so on until all firms and workers are matched. Firms select among applicants by picking a random recruiter to administer a pass/fail tests to each applicant. Talented workers pass the test, while untalented workers are screened out with probability $p_\theta$, where $0 < p_L < p_H < 1$. The firm randomly hires one applicant who passes the test, which is guaranteed to happen since there are a continuum of applicants.[6] This description assumes firms post different wages, as happens in equilibrium. For concreteness, assume that in case of a tie, all workers break the tie in the same way, as if the firms were infinitesimally differentiated.

Payoffs are as follows. Workers only care about wages, and so accept any job with positive wage. For firms, talented workers have productivity $\mu$ and untalented workers productivity 0. Thus, when a firm hires a worker with expected talent $\lambda$ at wage $w$, its expected profits are $\pi = \mu\lambda - w$.[7]

We solve for Nash equilibrium in wages, $w(r)$.

## 2.2 Preliminary Analysis

We first derive the quality of the recruits hired by firm $r$. When a recruiter of skill $\theta$ screens an applicant pool $q$, Bayes' rule implies that the expected talent of its recruit equals

$$\lambda(q;\theta) = \frac{q}{1-(1-q)p_\theta}. \tag{1}$$

The numerator is the proportion of applicants who are talented; the denominator is the proportion of applicants who pass the test. When proportion $r$ of the firm's recruiters have type $\theta = H$, the expected talent of its recruit is thus $\lambda(q;r) = r\lambda(q;H) + (1-r)\lambda(q;L)$, where we identify $H = 1$ and $L = 0$.

Next, we wish to understand how the quality of the applicant pool depends on the firm's rank in the wage distribution, as high-wage firms cherry-pick the applicants that pass their test. Suppose all firms post different wages,[8] denoted by $w(r)$, let $x(w)$ be the rank of a firm posting wage $w$, and $r(x)$ be the expected recruiter skill at wage quantile $x$. The highest ranked firm has $r(1)$ skilled recruiters and faces applicant pool $q(1) = \bar{q}$; thus, proportion $\lambda(q(1); r(1))$ of its recruits are talented. Since firms select talented workers disproportionately, lower-ranked firms have an adversely selected applicant pool, meaning $q(x)$ falls as the firm rank $x$ declines. Specifically, at rank $x$ there is a total of $xq(x)$ talented workers, of which firms $[x, x+dx]$ hire $\lambda(q(x); r(x))dx$;

[5] We assume that firms are distributed continuously. This is for notational convenience, so we can associate each firm with its screening skill, and avoid spurious equilibrium multiplicity. But Lemma 1, below, also applies if there were a mass of identical firms; that is, wage dispersion would arise endogenously.

[6] Since many workers pass the test, the random prioritization can occur before or after the test. That is, the firm could make all applicants take the test and then choose one, or it could randomly order the applicants and choose the first that passes the test.

[7] In the current model, the number of firms and workers is identical, and there is full employment. In Section 5 we endogenize the number of firms, and thus employment, by introducing an entry cost. This introduces a welfare margin, namely the aggregate allocation of talent in the industry.

[8] If an atom of firms posts the wage $w$ and workers need to break the tie, our tie-break rule implies that each firm's wage rank $x$ is drawn uniformly from $[\underline{x}(w), x(w)]$ where $\underline{x}(w) = \Pr(r : w(r) < w) < x(w)$. Lemma 1, below, implies that this complication does not arise on-path.

hence $d[xq(x)] = \lambda(q(x); r(x))dx$. Rearranging, the talent pool evolves according to,

$$q'(x) = \frac{\lambda(q(x); r(x)) - q(x)}{x} \tag{2}$$

Since screening is imperfect, some talent remains, $q(x) > 0$, for all $x > 0$. However, at the bottom, firms pick over the workers so many times, that no talent remains $q(0) = 0$. That is, if $q(0) > 0$ then $\lambda(q(0); r(0)) > q(0)$ and, using equation (1), the derivative $q'(0)$ behaves as $1/x$, contradicting the finiteness of $\int_0^1 q'(x)dx$. In addition, observe that the rate of decrease depends on which firms are choosing: a high-skilled firm picks out more talented workers than a low-skilled firm, and introduces more adverse selection.

Now consider equilibrium wages. Since the lowest paying firm $x = 0$ only recruits untalented workers, its productivity is zero, and thus it's wage must be zero, too. Moreover, the equilibrium wage distributions can have no atoms or gaps. If an atom of firms $x \in [\underline{x}, \bar{x}]$ offered the same wage $w$, then - recalling the our tie-break rule - a firm could attract discretely better job applicants with a marginal wage raise. On the other hand, if there were a gap so that no firms offered wages $w \in [\underline{w}, \bar{w}]$, then the firm offering $\bar{w}$ could profitably cut its wage to $\underline{w}$ and attract the same applicants. To summarize:

**Lemma 1.** *The equilibrium wage distribution has no atoms or gaps with minimum* 0*.*

### 2.3 Equilibrium

We have shown that no two firms pay the same wage, but do firms with skilled recruiters pay more or less than those with unskilled recruiters? In Becker (1973), positive assortative matching arises when high types of both market sides are complementary. Similarly, in Lucas (1978) better managers hire more workers since the production function is supermodular. Here, the question is whether skilled managers have a comparative advantage in screening applicants with higher expected talent. Formally, we say there is *positive assortative matching* (PAM) between firms and applicants if the equilibrium wage function $w(r)$ is increasing. Conversely, we say there is *negative assortative matching* (NAM) if $w(r)$ is decreasing.

Any wage profile $w(r)$ induces a firm-applicant matching $Q(r) := q(x(w(r)))$. This is an equilibrium if no firm $r$ wishes to change its wage, in particular mimic another firm $\tilde{r}$, by paying wage $\tilde{w} = w(\tilde{r})$ and attracting applicants $Q(\tilde{r})$. The incentive compatibility constraint is then:

$$\mu\lambda(Q(r); r) - w \geq \mu\lambda(Q(\tilde{r}); r) - \tilde{w}.$$

Adding to this firm $\tilde{r}$'s IC constraint not to mimic firm $r$, the wages cancel and we get

$$\lambda(Q(r); r) + \lambda(Q(\tilde{r}); \tilde{r}) \geq \lambda(Q(\tilde{r}); r) + \lambda(Q(r); \tilde{r})$$

Thus, in equilibrium, $\lambda(Q(r); \tilde{r})$ is supermodular in $(r, \tilde{r})$. Define the *benefit of skilled screeners* by

$$\Delta(q) := \frac{\partial}{\partial r}\lambda(q; r) = \lambda(q, H) - \lambda(q, L).$$

In equilibrium, $\Delta(Q(r))$ increases in $r$ meaning that the most skilled firms must be assigned to applicant pools where their benefit is the largest.

To understand the form of the equilibrium observe that, as shown in Appendix A.1, the benefit function $\Delta(\cdot)$ is single-peaked with peak at $\hat{q} \in (0,1)$. Intuitively, when all the applicants are talented, $q = 1$, then a firm will hire a talented recruit, independent of its screening ability, meaning that $\Delta(1) = 0$. Similarly, when all applicants are untalented, $q = 0$, a firm will always hire an untalented recruit, meaning that $\Delta(0) = 0$. In the middle, a firm's recruiting skills enable it to separate the wheat from the chaff.

Motivated by industries with relatively few highly productive individuals, such as universities of technology companies, the body of the paper hereafter focuses on the case of scarce talent. We characterize matching when talent is abundant in Appendix A.2.

**Assumption.** *Talent is scarce, $\bar{q} \leq \hat{q}$.*

Since applicant talent is highest at the highest-paying firm, where $q(1) = \bar{q}$, this assumption ensures that the benefit function $\Delta(\cdot)$ is increasing at all $q(x)$. That is, skilled firms have a comparative advantage at screening better applicant pools, meaning that equilibrium exhibits PAM. More formally, in any equilibrium $\Delta(Q(r))$ increases in $r$; since $\Delta(\cdot)$ is increasing, $Q(r)$ must also increase. We thus have:

**Theorem 1.** *Equilibrium exists and is unique; matching is positive assortative.*

*Proof.* The above discussion implies that matching must be positive assortative. In particular, firms employ pure strategies $w(r)$ that increase in $r$. We can then construct the equilibrium. Given PAM, the quality $r(x)$ of a rank-$x$ firm is given by the inverse of the distribution function $F(r)$. The applicant quality $q(x)$ of the rank-$x$ firm is then given by the sequential screening equation (2), and the recruit quality $\lambda(q(x); r(x))$ by the updating equation (1).

Denote the equilibrium wage required to attract applicants of quality $q$ by $W(q)$. At the bottom, $W(q(0)) = 0$. Firm $x$'s first-order condition

$$W'(q(x)) = \mu\lambda'(q(x); r(x)) \tag{3}$$

determines marginal wages, where $\lambda'(q; r)$ denotes the partial derivative $\partial\lambda(q; r)/\partial q$. Conversely, given these wages and the supermodularity of $\lambda(q; r)$, firm $r(x)$ optimally offers wage $W(q(x))$. Thus equilibrium exists and is unique. □

When talent is scarce, firms with skilled recruiters pay higher wages than those with unskilled recruiters. Using the first-order condition (3), wages are determined by the marginal benefit of a better applicant pool, $w(r) = \int_{\underline{r}}^{r} \mu\lambda'(Q(\tilde{r})), \tilde{r})Q'(\tilde{r})d\tilde{r}$. Then, using the envelope theorem, equilibrium profits are determined by the marginal screening ability, $\pi(r) = \int_{\underline{r}}^{r} \mu\Delta(Q(\tilde{r}))d\tilde{r}$. Intuitively, total value $\mu\lambda(q, r)$ depends on both the applicant pool quality and the screening ability; the workers captures the marginal benefit of the former, while the firms capture the marginal benefit of the latter.

Our analysis has interesting predictions for wage and productivity dispersion. Figure 1 illustrates our benchmark simulation and exhibits a decreasing density of productivity (which is the sum of wages and profits). Intuitively, when the test is relatively accurate and talent is scarce, a few top firms hire the bulk of the talent, and the remaining firms are left with untalented workers. Similarly, the density of wages is decreasing since firms compete fiercely for the top-ranks, but see little differences among the bottom ranks.

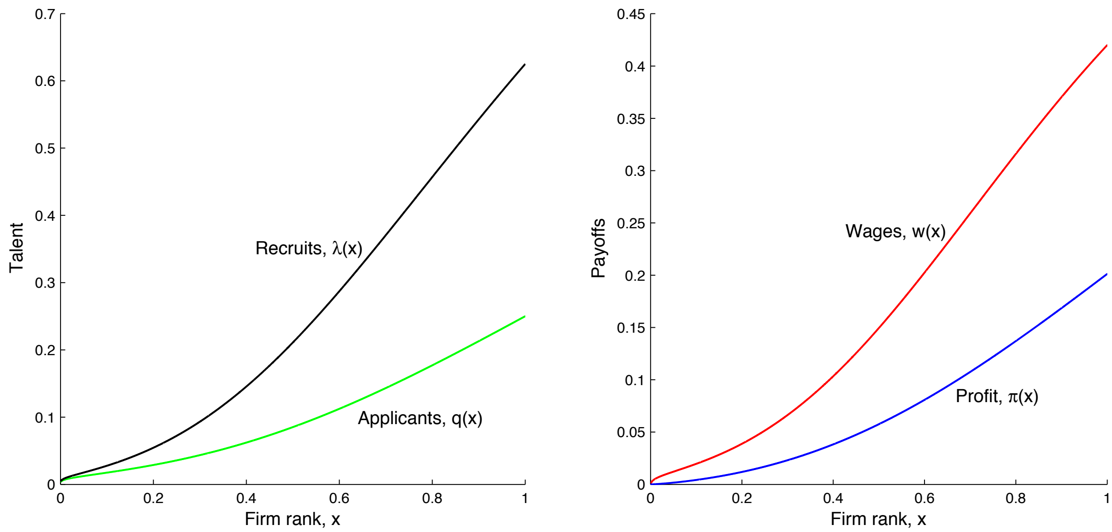


Figure 1: **Equilibrium with Scare Talent.** The **left** panel shows the quality of applicants and recruits. The **right** panel shows the resulting wages and profits. These figures assume $p_H = 0.8$, $p_L = 0.2$, $r \sim U[0,1]$, $\mu = 1$ and $\bar{q} = 0.25$.

Figure 2(a) shows the effect of change in screening technology (e.g. the information content of networks, or development of new assessment techniques) that alters the distribution of recruiter skills $F(r)$. First, suppose we go from the benchmark simulation with $r \sim U[0,1]$ to a setting where all firms have $r = 0.5$. Since the firms become undifferentiated, profits vanish, and wages rise to equal productivity. In addition, as screening ability shifts from top firms to bottom firms, productivity drops for top-ranked firms; this reduces adverse selection and productivity rises for middle and lower-ranked firms. Second, suppose the screening ability of firms rises from $r = 0.5$ to $r = 0.8$. The productivity/wages of the top firms rise as they select better workers; this lowers applicant quality for lower firms and reduces talent among their recruits, even though their screening skills have improved just the same, and thereby increases inequality. We formally establish a variant of this comparative static in Appendix A.3.

Figure 2(b) shows that skilled-biased technological change also leads to an increase in inequality. Specifically, we suppose an innovation doubles productivity $\mu$, but halves the number of productive agents $\bar{q}$, such that aggregate production $\mu\bar{q}$ remains unchanged. For the top firm, as the number of talented workers halves, the number of applicants passing the test falls, meaning that $\lambda(\bar{q})$ drops by less than half, and productivity $\mu\lambda(\bar{q})$ increases. The top firm thus hires proportionally more of the available talent, and proportionally less talent remains for lower firms, raising inequality. For the case of identical recruiters, Appendix A.3 formalizes this argument. The figure also shows that profits fall as it becomes more important for firms to be at the top of the distribution, making them fight harder from these top slots.

# 3 Dynamic Model

We now embed our static labor market into a model of firm dynamics, allowing us to endogenize the distribution of talent within and across firms. The key premise is that today's recruits become tomorrow's managers, and hire the next generation of workers. Firms thus desire talented workers

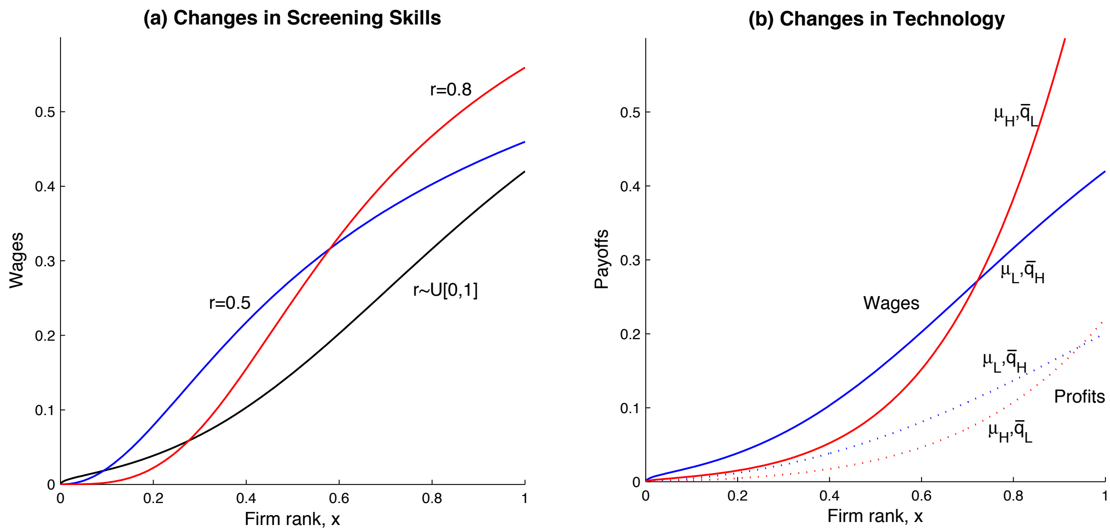


Figure 2: **Comparative Statics.** The **left** panel illustrates changes in the distribution of skill, $r$. The **right** panel illustrates changes in technology, with $(\mu_L \bar{q}_H) = (1, 0.25)$ and $(\mu_H, \bar{q}_L) = (5, 0.05)$. The other parameters are the same as Figure 1.

both for the immediate increase in productivity, and for the benefit of having skilled recruiters in the future.

We show that, as in the static model, equilibrium is unique with high-talent firms offering high wages that complement their superior screening skills; they thus attract superior applicants, and sustain their talent advantage. We characterize firm values, wages, and profits over time and simulate this equilibrium for initially identical firms. Eventually, the talent distribution converges to a steady state, where the positive assortative firm-applicant matching offsets the regression-to-mediocrity that results from imperfect screening.

## 3.1 Model

Time $t \geq 0$ is continuous. There is mass 1 of firms, each with mass 1 of jobs. At time $t$, a firm is described by its proportion of talented workers $r_t$; initially, the distribution of $r_0$ is exogenous.

At every instant $[t, t+dt]$, proportion $\alpha dt$ workers retire, leaving firms with vacancies. In the job market, there are then $\alpha dt$ open jobs and $\alpha dt$ applicants, of whom fraction $\bar{q}$ are talented. Analogous to the static model, firms compete for these applicants by posting life-time wages $w_t$;[9] applicants who do not find a job immediately leave the industry.

We assume that talented workers become skilled recruiters, untalented workers become unskilled recruiters, so the firm's skill coincides with its stock of talent, $r_t$. If a firm with talent $r_t$ posts a wage $w_t$ with rank $x_t(w_t)$, then it attracts applicants $q_t = q_t(x_t(w_t))$ and hires recruits of quality $\lambda(q_t; r_t)$. Writing $r_t(x)$ for the talent of the firm with wage rank $x$ at time $t$, the quality of the applicant pool $q_t(x)$ is determined by

$$q_t'(x) = \frac{\lambda(q_t(x); r_t(x)) - q_t(x)}{x} \text{ and } q_t(1) = \bar{q} \tag{4}$$

[9] Equivalently, and more naturally, one can think of worker compensation as constant flow wage $(\alpha + \rho)w_t$ until retirement, but it simplifies the accounting to have firms incur these costs up-front.

as in Section 2. The evolution of a firm's talent $r_t$ is then given by

$$\dot{r}_t = \alpha(\lambda(q_t; r_t) - r_t). \tag{5}$$

As for payoffs, workers maximize lifetime wages $w_t$, while firms' flow profits equal revenue $\mu r_t$ minus wages. A firm's problem is to choose wages to maximize total discounted profits. Denoting the discount rate by $\rho \geq 0$, its value function is

$$V_s(r_s) = \max_{\{w_t\}_{t\geq s}} \int_s^\infty e^{-\rho(t-s)}(\mu r_t - \alpha w_t)dt \tag{6}$$

where $r_t$ evolves according to (5).

An equilibrium is given by a wage path $\{w_t\}_{t\geq 0}$ for every firm,[10] so that given the induced wage ranks $x_t(w)$ and applicant qualities $q_t(x)$, every firm's wage path is optimal. We say an equilibrium is *essentially unique*, if the induced distribution over equilibrium trajectories $\{r_t\}_{t\geq 0}$ is unique.[11]

## 3.2 Firm's Problem

First, we study a firm's optimal wage path $\{w_t\}_{t\geq 0}$ for any given applicant function $q_t = q_t(x_t(w))$, that is, without imposing equilibrium restrictions on other firms. As in Section 2, it is convenient to write $W_t(q)$ for the wage required to attract applicants $q$ at time $t$, and let the firm optimize directly over the applicant pool $q_t$. After this change of variable, the firm's Bellman equation becomes

$$\rho V_t(r) = \max_q \{\mu r - \alpha W_t(q) + \alpha(\lambda(q; r) - r)V_t'(r) + \dot{V}_t(r)\}. \tag{7}$$

Intuitively, firm value is determined by its flow profits plus appreciation due to talent acquisition or a secular trend. Assuming wages are differentiable, the first-order condition is

$$W_t'(q) = \lambda'(q; r)V_t'(r). \tag{8}$$

Intuitively, the cost of attracting better applicants (the LHS) must balance the gains of a higher quality pool which increases the recruit quality and thereby firm value (the RHS).

Compared to the first-order condition in the static model (3), the additional factor $V_t'(r)$ on the RHS captures the fact that the firm in our dynamic model is only replacing an infinitesimal part of its workforce at every instant. Importantly, both terms on the RHS rise in $r$: The marginal benefit of better applicants $\lambda'(q; r)$ because of the complementarity as in Section 2, and the marginal value of higher talent $V_t'(r)$ because the value function $V_t(r)$ is convex in $r$. Convexity follows since discounted payoffs (6) for given wages $\{w_t\}_{t\geq s}$ are linear in current talent $r_s$, and the upper envelope of linear functions is convex.

Firms with more talent thus have a higher marginal benefit from attracting better applicant pools, yielding positive assortative matching. Hence, the firm's wages are dynamic complements:

[10] As in the static model, the restriction to deterministic wages is without loss. In principle, a firm might mix between two wages by switching between them arbitrarily fast. To avoid measurability issues associated with such strategies, we allow for "distributional wage strategies" but show in Section 3.3 that equilibrium strategies are almost always pure.

[11] This definition avoids two spurious notions of multiplicity. First, in continuous time, any firm's optimal strategy $\{w_t\}_{t\geq 0}$ can be unique only almost always. Second, if two or more firms are initially identical but then drift apart, only the distribution of trajectories can be determined uniquely.

an increase in today's wage, raises tomorrow's talent and raises tomorrow's optimal wage.

To further analyze the marginal value of talent and the optimal strategy $\{\hat{q}_s\}_{s\geq t}$, we apply the envelope theorem to get

$$V_t'(r_t) = \frac{\partial}{\partial r_t}\int_t^\infty e^{-\rho(s-t)}[\mu r_s - \alpha W(\hat{q}_s)]ds = \mu \int_s^\infty e^{-\rho(s-t)}\frac{\partial r_s}{\partial r_t}ds \tag{9}$$

where the partial derivative $\partial r_s/\partial r_t$ takes $\{\hat{q}_s\}_{s\geq t}$ as given. Since $r_t$ evolves according to (5) we can calculate this derivative to obtain[12]

$$V_t'(r_t) = \mu \int_t^\infty e^{-\int_t^s \rho+\alpha(1-\Delta(\hat{q}_u))du}ds. \tag{10}$$

Intuitively, the future benefit of better employees is discounted both at the interest rate $\rho$ and the retirement rate $\alpha$; but selective recruiting raises the persistence of firm talent, or equivalently, reduces the talent decay rate by a factor $(1-\Delta(q))$.

### 3.3 Equilibrium

Given the single-firm analysis, it is straight-forward to characterize equilibrium. The benefit of attracting talented applicants on the RHS of (8) increases in the firm's talent $r$, so firms with more talent post higher wages and attract better applicants. More strongly, even if firms share the same talent $r_0$ initially they post different wages, as in the static model, recruit different types of workers, and diverge immediately (see Appendix B.1). Thus, in equilibrium, each firm is characterized by a rank $x$, which describes the firm's position in the talent, applicant, and wage distribution at all times $t > 0$.

Equilibrium is then characterized in two simple steps

(1) Allocations. At any point in time, applicant quality $q_t(x)$ is determined by sequential screening (4). The evolution of a firm's talent $r_t(x)$ is then given by equation (5).

(2) Payoffs. Firms marginal value of talent is determined by (10), with $\hat{q}_u = q_u(x)$. Using this, wages $W_t(q)$ are given by the first-order condition (8), with $r = r_t$, $\hat{q}_u = q_u(x)$, and $W_t(0) = 0$.

Given these wages, the resulting value function solves (7) for all $r$ and $t$. To see that the first-order condition (8) implies global optimality, consider two firms $x, \tilde{x}$ at time $t$ with talent $r = r_t(x), \tilde{r} = r_t(\tilde{x})$ and applicants $q = q_t(x), \tilde{q} = q_t(\tilde{x})$. By firm $\tilde{x}$'s first-order condition, $W_t'(\tilde{q}) = \lambda'(\tilde{q};\tilde{r})V_t'(\tilde{r})$, firm $x$'s net benefit from attracting marginally better applicants than $\tilde{q}$ equals $\lambda'(\tilde{q};r)V_t'(r) - \lambda'(\tilde{q};\tilde{r})V_t'(\tilde{r})$. Since $\lambda$ is supermodular, $V$ is convex, and $q_t(x), r_t(x)$ increase in $x$, this difference is negative for $\tilde{q} < q$ and positive for $\tilde{q} > q$, and so firm $x$ optimally targets applicants $q$. Standard verification theorems then imply that the policy functions are indeed optimal. To summarize:

**Theorem 2.** *Equilibrium exists and is essentially unique. Firm-applicant matching is positive assortative and the distribution of talent has no atoms at $t > 0$.*

[12] To see this, write the solution of the ODE $\dot{r}_t = \phi(r_t) = \alpha(\lambda(q_t;L) - r_t(1-\Delta(q_t)))$ as a function of the initial value $r_0 = z$. Then, $\frac{\partial^2}{\partial t \partial z}r_t(z) = \frac{\partial}{\partial z}\phi(r_t(z)) = -(1-\Delta(q_t))\frac{\partial}{\partial z}r_t(z)$, or $\frac{\partial}{\partial t}\log\frac{\partial}{\partial z}r_t(z) = -(1-\Delta(q_t))$. Integrating up and de-logging we get $\frac{\partial}{\partial z}r_t(z) = \exp(-\int_0^t(1-\Delta(q_s))ds)$ since $\frac{\partial}{\partial z}r_0(z) = 1$. See, for example, Hartman (2002).

This result shows that even if firms start off with identical talent, some will post higher wages than others, attract better applicants and hire better recruits. These firms will then accumulate talent, continue to pay high wages, causing the distribution of talent to disperse over time.

The evolution of talent, wages and profits are illustrated in Figure 3. This simulation assumes that, at $t = 0$, all firms employ average workers, with quality $\bar{q} = 0.25$, pay a wage equal to average productivity, $\mu\bar{q}$. Panel (a) illustrates the convergence of $r_t(x)$ and $q_t(x)$. The "vertical" lines represent the cross-sectional distribution of $(r, q)$ at different times, while the "horizontal" lines represent the sample-paths of selected firms. The top-ranked firm recruits from the constant applicant pool $q_t(1) = \bar{q}$, and so (5) implies that its talent grows monotonically and converges to steady state. For lower-ranked firms, the dynamics are more subtle. For example, firm $x = 0.45$ initially improves as its recruits are more talented than retirees. However, as the firms above become better at identifying talent, its applicant pool deteriorates and its quality eventually falls back.

Panel (b) shows the evolution of firm value net of legacy wages.[13] Initially all firms are identical and have zero value. However, as some post high wages and acquire more talented workers, they become differentiated and the highly ranked firms have positive value. On the flip-side, these firms make a loss as they invest in talent, as shown in Panel (c). In particular, firms initially all make zero profits as their wage bill is determined by their legacy employees. As highly-ranked firms pay high wages to acquire talent, profits fall. However, this talent raises productivity and gives the firms a comparative advantage in recruiting talented workers in the future, so eventually profits become positive. Panel (d) then shows that wages fall over time. Initially, firms are identical and compete aggressively for talent to gain a comparative advantage in the future. Later, the differentiation between firms lessens the competition. Moreover, wages tend to fall proportionately more for lower ranked firms as top firms grow more productive, while the productivity of low firms falls.[14]

Figure 3 indicates that the economy converges to a steady state. Steady-state talent and applicant pool $\{r_*(x), q_*(x)\}$ is easily characterized. First, the talent of each firm's recruits and retirees balance:

$$\lambda(q; r) = r. \tag{11}$$

Since $\partial\lambda(q; r)/\partial r = \Delta(q) < 1$, (11) has a unique fixed point, which we denote by $r = \xi(q) := \lambda(q; L)/(1-\Delta(q))$. Naturally, firms with better applicants have higher talent; formally, $\xi(q)$ increases since $\lambda(q; r)$ rises in $q$. Second, substituting $r(x) = \xi(q(x))$ into the sequential screening equation (4) we obtain:

$$q'(x) = \frac{\lambda(q(x); \xi(q(x))) - q(x)}{x}. \tag{12}$$

Together, replacement (11) and sequential screening (12) determine equilibrium talent and applicants $\{r_*(x), q_*(x)\}$.

We now show that from any initial condition, the economy converges to the steady state.

**Theorem 3.** *The steady state talent distribution $r_*(x)$ is unique and has no gaps or atoms. For any distribution of initial talent $r_0$, firm $x$'s equilibrium talent $r_t(x)$ converges to $r_*(x)$.*

*Proof.* In Appendix B.2. □

[13] At time $t$, the legacy wage bill $\omega_t$ evolves according to $\dot{\omega}_t = \alpha(w_t - \omega_t)$, with initial condition $\omega_0 = \mu r_0 = \mu\bar{q}$. The firm's net value is then given by $V(r_t) - \omega_t/(\alpha + \rho)$. Similarly, flow profits are $\mu r_t - \omega_t$.

[14] This relates to the forces in Figure 2(a), which examines the impact of a mean-preserving spread in recruiting skills.

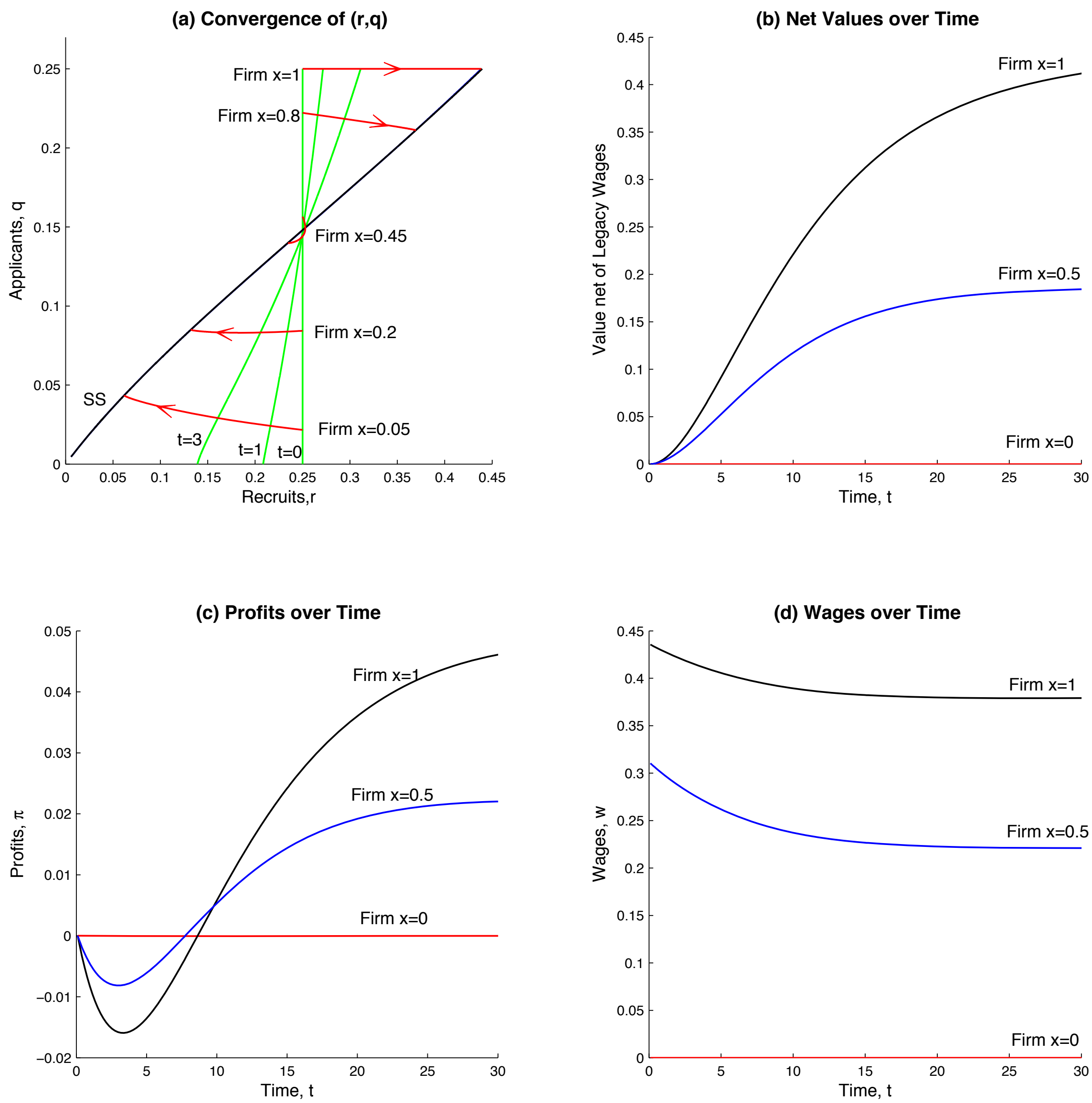


Figure 3: **Equilibrium Dynamics.** At $t = 0$, all firms have talent $\bar{q}$ and pay wages $\mu\bar{q}$. After, they choose wages optimally as characterized in the text. As in Figure 1, we assume $p_H = 0.8$, $p_L = 0.2$, $\mu = 1$, and $\bar{q} = 0.25$. Thinking of a time period as a year, we assume the interest rate is $\rho = 0.1$, and turnover is $\alpha = 0.2$.

In steady-state, the distribution of talent and productivity $r_*(x)$ is dispersed. This is perhaps surprising since firms' talent is subject to regression to mediocrity. With imperfect matching, when a skilled worker retires there is chance they will be replaced by an unskilled worker, and similarly an unskilled worker may be replaced by a skilled worker. Thus, if each firm hired from the same pool, then firms' talent levels would converge at the talent decay rate, $\alpha(1 - \Delta(q))$, as reflected by the exponent in (10). Countering this, is the effect of positive assortative matching: firms with skilled employees post higher wages and recruit from a better pool. As a result, the steady state supports permanent heterogeneity in firm quality, productivity and profits.[15]

Talent thus generates a sustainable competitive advantage. One might wonder why a firm with untalented workers doesn't choose to post high wages and increase its talent over time. While this is a feasible strategy, it is simply too expensive. A high-talent firm obtains higher marginal benefit from raising wages both because of the complementarity between a firm's recruiting skills and the applicant quality in the job market, and because firm value is convex.

To show convergence, Figure 3(a) suggests a proof by "induction". The top firm recruits from a pool of constant quality, so equation (5) implies that its talent converges exponentially to steady state. Then consider firm $x < 1$. If the talent of higher firms converges, then firm $x$'s applicant pool converges and, equation (5) implies that firm $x$'s talent converges. Since $x$ is continuous, the actual proof shows that the steady state satisfies a contraction property, and then applies the contraction mapping theorem over a small interval, akin to the proof of the Picard-Lindelof theorem.

The model generates predictions about how value is shared between workers and firms. Using (10), the steady-state value is given by

$$V_*'(r_*(x)) = \frac{\mu}{\rho + \alpha(1 - \Delta(q_*(x)))}.$$

Intuitively, an increase in talent raises productivity by $\mu$ discounted at the interest rate and the talent decay rate. Using (8), marginal flow wages are then given by

$$(\rho + \alpha)W_*'(q) = \frac{\rho + \alpha}{\rho + \alpha(1 - \Delta(q))}\mu\lambda'(q; r). \tag{13}$$

Steady state profits equal output minus flow wages, $\Pi_*(q) = \mu\xi(q) - (\alpha + \rho)W_*(q)$. Differentiating (11) and substituting into (13), the ratio of returns to capital and labor are given by

$$\frac{\Pi_*'(q)}{(\alpha + \rho)W_*'(q)} = \frac{\rho\Delta(q)}{(\rho + \alpha)(1 - \Delta(q))} \tag{14}$$

In the static model, firm differentiation is largely driven by the distribution of skills $F$ and the benefit of skilled screeners $\Delta$. In the dynamic model the distribution of skills is endogenous and is partly determined by $\Delta$. As a consequence, an increase in $\Delta$ leads to an increase the share of output that goes to firms both because it raises the comparative advantage of high-talent firms in the labor market, and because it raises the persistence of the stock of talent.

[15] For simplicity, we assume that talent and recruiting skill are identical. One can allow for imperfect correlation by assuming that a firm with productive talent $r$ has recruiting skills $\beta r + (1 - \beta)\bar{q}$, for $\beta > 0$. Again, equilibrium is unique and converges to a unique steady state with nondegenerate distribution of talent, wages and profits. While steady state dispersion is lower than for $\beta = 1$, it does not collapse as $\beta$ approaches 0 because of positive assortative matching. In contrast, if talent and skills are uncorrelated, $\beta = 0$, equilibrium is indeterminate with all firms being indifferent over all wages at all times.

To help us understand the forces at work, we consider two comparative statics. First, suppose there is increase in turnover, such as moving from history to economics departments, or manufacturing to technology sectors. The increase in $\alpha$ has no impact on the distribution of talent, but does raise the rate the economy converges to the steady state. This means that, in steady-state, profits fall, while wages and wage dispersion rise.[16] Intuitively, the comparative advantage of a firm is its stock of talent; when turnover is high, this stock depletes quickly, and firms get a small share of the (constant) output. Indeed, as turnover increases, a low-talent firm can mimic a high-talent firm and get almost the same profits as the past quickly becomes irrelevant; this intensifies competition and drives up wages.

Second, consider an increase in skill-biased technical change that doubles productivity $\mu$ but halves the number of qualified workers $\bar{q}$. For the top firm, if its recruiting ability stayed constant, then the quality of its recruits falls by less than half, as in the static model. However, since talent and recruiting ability are aligned, then the reduction in talent can make it harder to identify talented applicants, meaning that the firm's steady-state quality, $\xi(\bar{q})$, may fall by more than half. Thus, productivity can become less dispersed. Moreover, as $\Delta(q)$ shrinks because of the talent reduction, equation (14) implies workers gain more of the surplus.

# 4 Applications

In this section we consider two applications of the dynamic model. First, we suppose firms differ in both exogenous technology and endogenous talent. Second, we allow for peer effects, so that workers care about both wages and the average talent of the firm.

## 4.1 Heterogeneous Technology

We extend the basic dynamic model and suppose that firms differ in both their initial talent $r_0$ and also in their technology, $\mu$. While both jointly determine productivity, the former is endogenous, and evolves over time as the firm hire new recruits, while the latter is exogenous. In the case of universities, we interpret $\mu$ as the value a university places on research, perhaps because of its charter; in the case of tech firms, $\mu$ reflects the idea that talent has a higher return at Google than Pets.com.

Initially, talent and technology may differ in arbitrary ways across firms, and equilibrium must map this two-dimensional heterogeneity into one-dimensional wage rankings. However, we show that over time the two dimensions of heterogeneity collapse into one, and the economy converges to a unique, stratified steady state with positive assortative technology-talent matching. We then use this steady state to study how an individual firm optimally adjusts its talent pool after a technology shock.

Assume that technology is binary, $\mu \in \{\mu_L, \mu_H\}$, and that fraction $\nu$ firms have the low technology, $\mu_L < \mu_H$. We first argue that the target applicant pool $\hat{q}_t$ is increasing in both technology and talent. Fixing an aggregate wage schedule, $W_t(q)$, the firm's first-order condition is given by

[16] Formally, (13) implies that $W_*(q)$ rise in $\alpha$, while (13) and (14) imply that $\Pi_*(q)$ increases. Moreover, $W_*(q)$ is log-supermodular and supermodular in $(q, \alpha)$ which means wage dispersion rises in both absolute and proportional senses. To see this observe that $\frac{\partial^2}{\partial\alpha\partial q}\log[W_*'(q)] \overset{\text{sgn}}{=} \Delta'(q) > 0$ which means that $W_*'(q_H)/W_*'(q_L)$ increases in $\alpha$, for $q_H > q_L$, and hence implies $\frac{\partial^2}{\partial\alpha\partial q}\log[W_*(q)] = \frac{\partial}{\partial\alpha}[W_*'(q)/\int_0^q W_*'(\hat{q})d\hat{q}] > 0$. Finally note that a log-supermodular function that is increasing in both arguments is also supermodular.

(8). As before, an increase in talent $r$ increases both the marginal value of talent $V_t'(r;\mu)$ and the marginal quality of recruits $\lambda'(q,r)$, and thus increases the RHS of (8). Similarly, an increase in technology $\mu$ increases the marginal value of talent $V_t'(r;\mu)$ and the RHS of (8). We prove this in Appendix C.1; intuitively, it follows because technology and talent are complements in the profit function.

When comparing two firms, if $A$ has higher technology and talent than $B$, then it pays higher wages and attracts a better applicant pool. However, if $A$ has better technology, while $B$ has better talent, the ranking is ambiguous. Moreover, if $A$ posts higher wages than $B$, then it will hire from a better applicant pool, and may hire better recruits, leading it to accumulate more talent than $B$. Thus, in contrast to the homogeneous technology case in Section 3.3, firms' wage rank may change over time.

We now characterize the steady state. Firms have a constant wage rank $x$. As above, the steady state talent $r_*(x)$ and applicant quality $q_*(x)$ are thus determined by the constant firm quality (11) and sequential screening (12). These distributions are the same as in the homogenous model, and are thus independent of the distribution of technology. We then need to specify the technology level $\mu(x)$ of firm $x$. We say that the steady state is *stratified* if $\mu(x)$ equals $\mu_H$ for high wage firms $x > \nu$, and $\mu_L$ for low wage firms $x < \nu$. Anticipating the below result, we denote this function as $\mu_*(x)$. As before, the steady-state marginal value of talent is given by

$$V_*'(r;\mu) = \frac{\mu_*(x)}{\rho + \alpha(1 - \Delta(q_*(x)))} \tag{15}$$

with initial condition $V(0,\mu_L) = 0$. Finally, wages are given by the first-order condition

$$W_*'(q) = \lambda'(q_*(x); r_*(x)) V_*'(r_*(x); \mu_*(x)) \tag{16}$$

with initial condition $W_*(0) = 0$.

**Theorem 4.** *The stratified steady state $\{r_*(x), q_*(x), \mu_*(x), W_*(q)\}$ is the unique steady-state equilibrium. Any equilibrium converges to this steady state.*

*Proof.* We first verify that the stratified steady state indeed constitutes an equilibrium. Consider a firm with technology $\mu_L$, say, and rank $x \le \nu$. Observe that $W_*(q)$ has an upward kink at $q_*(\nu)$ since competition is fiercer among the high technology firms above $q_*(\nu)$. Now, define $W_L(q)$ by replacing $\mu_*(x)$ with $\mu_L$ in (16). This wage function coincides with actual wages $W_*(q)$ for $q \le q_*(\nu)$, but it is flatter for $q \ge q_*(\nu)$. For wages $W_L(q)$, the firm's steady state strategy is optimal by the same argument as for homogeneous firms in Section 3.3. Since actual wages $W_*(q)$ are higher than $W_L(q)$ above $q_*(\nu)$, a low-technology firm with talent rank $x \le \nu$ is thus best-responding by choosing $w = W_*(q_*(x))$. A similar argument shows that the postulated steady state strategy is optimal for a high firm with productivity $\mu_H$ and steady state talent rank $x \ge \nu$.

To show uniqueness of the steady-state, suppose technology-talent matching is not stratified, i.e. $\mu(x)$ has a downward jump at some $\tilde{x}$. Observe that workforce and applicant talent, $r_*(x), q_*(x)$, are continuous in $x$. Thus, the marginal benefit of attracting better applicants, $\lambda(q; r_*(x)) V'(r_*(x); \mu(x))$ also jumps downwards at $\tilde{x}$. Hence, high-technology firms just below $\tilde{x}$ want to offer higher wages than low-technology firms just above, contradicting steady-state equilibrium.

The convergence proof in Appendix C.2 proceeds in three steps. First, we argue that the equilibrium talent ranking across firms eventually settles down since high-technology firms' ranks always rise, while low-technology firms' ranks fall. Next, we argue that the wage ranking across firms eventually settles down, too. Then, by the uniqueness of steady state equilibrium and continuity arguments, the equilibrium converges to the stratified steady state. □

Theorem 4 shows that differences in exogenous technology eventually prevail over differences in endogenous talent. A high-tech firm has a higher marginal benefit from talent and posts higher wages than nearby low-tech firms, overcoming its talent deficit. Hence the high-tech firm recruits better workers and, in the long run, ends up with both better technology and talent.

One often sees universities and firms that have particularly good employees relative to their "fundamentals", often due to historic accidents. Theorem 4 shows that talent provides a temporary competitive advantage, but ultimately fundamentals prevail. For example, Arai (2003) shows that capital-intensive firms tend to pay higher wages; our model predicts that such firms will accumulate talent and provide them with a further comparative advantage in the job market.

We next consider how a single firm adjusts in response to an increase in technology. For example, a university obtains a new donor who particularly values quality, or Netflix changes business models from shipping DVDs to streaming.[17] We then wish to study how the firm's wage, recruits and talent change over time.

Formally, we suppose the economy is in steady state and a low-technology firm, "firm $A$", receives a unanticipated shock that makes it high-technology. Denote firm $A$'s initial talent by $r_0 < r_*(\nu)$, and its post-shock talent by $r_t$. Write $\tilde{W}(r)$ for steady-state wages of a firm with talent $r$, i.e. $\tilde{W}(r_*(x)) := W_*(q_*(x))$; also recall that $r_*(\nu)$ is the lowest steady-state talent among high-technology firms.

**Theorem 5.** *After its technology shock, firm $A$ posts wages $w_t \in (\tilde{W}(r_t), \tilde{W}(r_*(\nu))]$. Talent $r_t$ increases over time, and converges to (but never reaches) its steady state level $r_*(\nu)$.*

*Proof.* Since the optimal wage rises in a firm's talent and technology, firm $A$'s wage exceeds the wage of a firm with the same talent and low-technology, $w_t > \tilde{W}(r_t)$, but is less than the wage of a high-technology firm with talent $r_*(\nu)$, $w_t \leq \tilde{W}(r_*(\nu))$. The first inequality is strict since equilibrium wages are smooth at $\tilde{W}(r_t)$, but the second may be weak because the wage function has a kink at $r_*(\nu)$. Thus, firm $A$'s talent $r_t$ is bounded above $r_*(\nu)$, but rises over time, and hence converges. By the proof of convergence in Theorem 4, its limit exceeds the steady-state talent of all low-technology firms. However, since its recruits are weakly worse that that of firm $r_*(\nu)$, the convergence is only asymptotic. □

When a low-technology firm experiences an increase in productivity, its wages jump up. However, if its talent is sufficiently below steady state, its wages do not jump all the way up to its steady-state level. Intuitively, the firm does not initially have the talent to spend a wage boost wisely. The firm thus hires people below its aspired steady-state level of talent both because it pays (weakly) lower wages, and because of its inferior recruiting skills. In the long-run, the wage and applicant pool rises to the steady state, and the talent level follows along. These dynamics differ notably from the "burst hiring" whereby a firm spends extravagantly to build a complementary group of employees.

[17]The Netflix HR manual states "In procedural work, the best are 2× better than average. In creative/inventive work, the best are 10× better than average." Consistent with the above analysis, it goes on to say paying "top of market is core to high performance culture" (Hastings and McCord, 2009).

## 4.2 Peer Effects

In the baseline model, a worker chooses the firm that pays the highest wages. In this section, we suppose his preferences also depend on the fraction of talented workers at the firm, $r$. Formally, we assume that a hired worker's utility is given by $w + \gamma r$, for $\gamma > 0$. This could represent the prestige effect of being hired by a highly informed firm, or the increase in human capital when a new worker first joins the firm. We show that the introduction of peer effects has no affect on equilibrium matching or the evolution of talent, but does lower steady state wages.[18]

First, let us consider the static model of Section 2. As before, incentive compatibility implies that $\Delta(Q(r))$ must increase in $r$, so matching is positive assortative, in that firms with higher skills offer higher utility and attract better applicants. Intuitively, peer effects make high-skill firms more attractive but, unlike Becker (1973), do not make them relatively more attractive for talented agents. Since workers care about total compensation, equilibrium wages simply decrease by $\gamma r$, with slope $w'(r) = \mu\lambda'(Q(r), r)Q'(r) - \gamma$. This means that if firms are sparsely distributed, such as top universities, then $Q'(r)$ is low and wages can decrease in quality.

Now, let us consider the dynamic model of Section 3. In equilibrium, a firm must post utility $U_t(q)$ to attract applicant pool of quality $q$. To do this, firm $r$ must pay a wage $W_t(q;r) = U_t(q) - \gamma r$. The Bellman equation is again given by (7), with first-order condition

$$U_t'(q) = \lambda'(q;r)V_t'(r), \tag{17}$$

meaning that more talented firms attract better applicants, hire better recruits and stay ahead of less talented firms. Hence, talent and applicant quality $(r_t(x), q_t(x))$ converge to the same steady state as before, $r_*(x) = \xi(q_*(x))$.

Peer effects increase the marginal value of talent. Applying the Envelope Theorem, the steady-state value is given by,

$$V_*'(r_*(x)) = \frac{\mu + \alpha\gamma}{\rho + \alpha(1 - \Delta(q_*(x)))}. \tag{18}$$

Intuitively, a new worker that increases talent by $\Delta r$ allows the firm to reduce its wage of new recruits by $\gamma\Delta r$ while this worker is employed. In addition, the new worker helps identify future talent who can be used to lower wages even further.

We now turn to wages. If we start with all firms being identical, the introduction of peer effects increases initial wage dispersion.[19] Intuitively, talent allows a firm to lower future wages, increasing incentives to accumulate it in the first place. However, in steady state, peer effects lower the slope of the wage function meaning that, if the equilibrium wage function is increasing, then peer effects lower dispersion.[20] Intuitively, peer effects enable talented firms to pay less to their employees, but also increases the incentive to bid for future talent. The latter effect is discounted, so the former dominates. Hence, if firms start off with similar levels of talent, the introduction of peer effects shifts utility from later recruits to those who enter the industry earlier, magnifying the decreasing wages seen in Section 3.

[18]This should be viewed as a reduced form that shows the robustness of our results, and yields some insight about the form of wages. In a full reputation model, the agent might care about the market's inference of his quality, $\lambda$. In a full peer effects model, the worker might care about the talent over time, not just when hired.

[19]See Appendix C.3.

[20]See Appendix C.3.

## 5 Welfare

In order to keep the model as simple as possible, we have so far abstracted from the question of firm entry and the level of employment. We now address these questions, and in particular whether equilibrium matching sorts talent into the industry efficiently, by assuming that untalented workers should not enter the industry

Aggregate sorting is an important issue in our applications. In the academic context, some students may be a better fit for a professional capacity than for academic research. Similarly, Netflix's HR policy states "One outstanding employee gets more done and costs less than two adequate employees; we endeavor to have only outstanding employees", (Hastings and McCord, 2009).

Formally, assume that employment has a fixed cost $k \geq 0$, implying that a talented employee has value $\mu - k$, while an untalented employee has value $-k$. Equivalently, one can interpret $k$ as an opportunity cost for the employee (adjusting wages appropriately). Under either interpretation, untalented workers should not enter the industry and the organization of the market can affect how this sorting takes place.

### 5.1 Static (In)efficiency

We first consider the static model of Section 2 with a fixed cost $k \geq 0$. In equilibrium, profits are increasing in $r$, so firms above some cutoff $\tilde{r} > \underline{r}$ enter the market; matching is then positive assortative. Intuitively, if all firms entered the market, the lowest applicant pool would consist solely of untalented workers, $q(0) = 0$, and the lowest-wage firm would not recover its fixed cost. The cutoff is then determined by the zero profit condition, $\mu\lambda(Q(\tilde{r}); \tilde{r}) = k$, and mass $F(\tilde{r})$ of workers are unemployed, of whom fraction $Q(\tilde{r})$ are talented.

We now consider the surplus maximization problem of a social planner, who can decide which firms enter the market, and the order in which she selects applicants. Optimal entry is characterized by some threshold $\check{r}$ with quantile $\check{x} = F(\check{r})$.[21] The planner then chooses a measure-preserving map $r$ from $[\check{x}, 1]$ to $[\check{r}, \bar{r}]$ to maximize social surplus $\int_{\check{x}}^{1} [\mu\lambda(q(x); r(x)) - k]dx$, where $q(x)$ is given by sequential screening (2). For example, PAM corresponds to $r(x) = F^{-1}(x)$, and NAM to $r(x) = F^{-1}(1 + \check{x} - x)$.

Equilibria in continuous matching markets with transferable utility are typically efficient (e.g. Gretsky, Ostroy, and Zame (1992)). Surprisingly, this welfare theorem fails in our model. One can see the downside of PAM in a two firm variant of our model. Suppose there is a good firm with some skilled recruiters, and a bad firm with only unskilled recruiters who hire randomly, i.e. $p_L = 0$. Under PAM, the good firm has applicants $\bar{q}$ while the low firm faces an adversely selected pool $q < \bar{q}$. But under NAM, both firms face the original pool quality $\bar{q}$ since the bad firm hires randomly, generating no adverse selection. Thus, firms jointly hire better recruits under NAM than under PAM.

The key difference to standard matching models with exogenous types is that the quality of low applicant pools $q(x)$ is endogenous, subject to adverse selection induced by the screening of better paying firms. This adverse selection is maximized by PAM and minimized by NAM, so the

[21] Intuitively, the planner wants the most skilled firms in the market to help screen the wheat from the chaff. Formally, this follows from the fact $\zeta > 0$ in the proof of Theorem 6.

supermodularity of $\lambda(q;r)$ and adverse selection work at cross purposes. Surprisingly, we can show that adverse selection unambiguously dominates, and NAM is efficient.

**Theorem 6.** *In the static model with positive costs $k > 0$, the planner's solution is characterized by some entry threshold $\check{r} > \underline{r}$ and NAM for firms $r \in [\check{r}, \bar{r}]$.*

*Proof.* Observe that for a fixed entry threshold $\check{r}$ and unemployment $\check{x} = F(\check{r})$, maximizing surplus is equivalent to maximizing employed talent $m(\check{x})$, where $m(x) = \bar{q} - q(x)x$. Now, relax the planner's problem by letting her allocate additional recruiting skills to the firms. Specifically, for any threshold $\check{x}$ and screening skills $r(x)$ for $x \geq \check{x}$, we define $\zeta(x)$ as the derivative of employment $m(\check{x})$ with respect to $r(x)$. Marginally better screening skills at firm $x$ increase employed talent at rank $x$ by $dm(x) := \Delta(q(x))$. However, firms with lower ranks $\tilde{x} \in [\check{x}, x]$ now face additional adverse selection, reducing their talent intake.

To compute the effect on total employed talent $m(\check{x})$, note first that, by definition, total employed talent evolves according to $m'(x) = -\lambda(q(x); r(x)) = -\lambda((\bar{q} - m(x))/x; r(x))$. Thus, by footnote 12, incremental employed talent $dm(\tilde{x})$ depreciates at rate $\lambda'(q(\tilde{x}); r(\tilde{x}))/\tilde{x}$ as $\tilde{x}$ decreases from $x$ to $\check{x}$, and so

$$\zeta(x) := \frac{dm(\check{x})}{dr(x)} = \frac{dm(\check{x})}{dm(x)} \frac{dm(x)}{dr(x)} = \exp\left( -\int_{\check{x}}^{x} \frac{\lambda'(q(\tilde{x}); r(\tilde{x}))}{\tilde{x}} d\tilde{x} \right) \Delta(q(x)). \tag{19}$$

We claim that for fixed $\check{x} > 0$, the aggregate effect of screening skills, $\zeta(x)$, is strictly decreasing in the screening rank $x$. If we differentiate (19) and multiply through with $x \exp(-\int_{\check{x}}^{x} [\lambda'(q(\tilde{x}); r(\tilde{x}))/\tilde{x}] d\tilde{x})$, to get

$$\Delta'(q(x))q'(x) - \lambda'(q(x); r(x)) \frac{\Delta(q(x))}{x}. \tag{20}$$

In Appendix D.1, we show this is strictly negative. In this equation, the first, positive term captures the idea that allocating screening skills to higher-ranked firms is efficient because of the complementarity, as in Becker (1973). That is, shifting screening skills up by $\epsilon$ ranks, where applicant talent is $q'(x)\epsilon$ higher, increases talent intake by $\Delta'(q(x))q'(x)\epsilon$. The second, negative term captures the idea that better screening at higher ranked firms exacerbates adverse selection for lower ranked firms. That is, better skills at rank $x + \epsilon$ reduces quality of the applicant pool of the $x$-ranked firm by $\Delta(x)\epsilon/x$, and talent intake by $\lambda'(q(x); r(x))\Delta(x)\epsilon/x$. These two terms are illustrated in Figure 4. As we show in the Appendix, the negative, adverse selection effect unambiguously dominates the positive, complementarity effect.

Next, we argue NAM is efficient given any entry threshold $\check{r} > \underline{r}$. If the matching function is anything other than NAM, we can find $r' > r$ with wage ranks $x' > x$. Swapping the ranks $x$ and $x'$ effectively shifts screening skills from rank $x'$ to rank $x$. Since $\zeta(x)$ decreases in $x$, this raises aggregate surplus.

Finally, we argue that $\check{r}$ exceeds $\underline{r}$. If all firms did enter, $\check{r} = \underline{r}$, then all workers are employed and the lowest-ranked firms contributes negative social surplus, $\lambda(q(x); r(x)) - k$. Welfare is thus increased by having some firms exit (ideally, the lowest skilled firms). $\square$

Theorem 6 is interesting for two reasons. First, it shows that equilibrium exhibits excessive dispersion in productivity and talent. Second, it shows how this inefficiency comes from the externalities high-skill firms exert on low-skill firms by poaching talented workers. There is no comparable

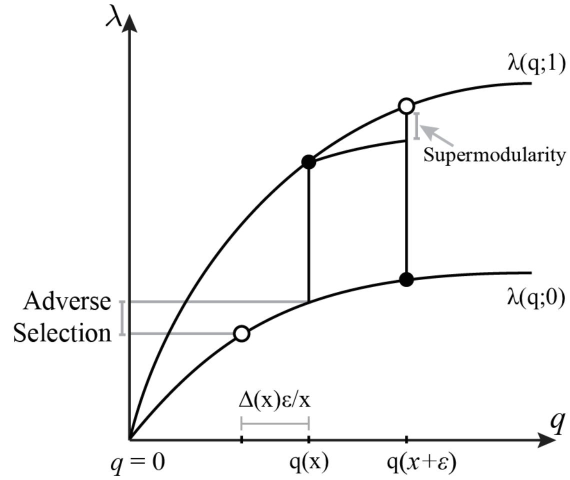


Figure 4: This figure illustrates the supermodularity and adverse selection forces arising in Theorem 6. The solid dots represent talent intake under NAM, where the rank-$x$ has high skill and the rank-$x+\epsilon$ has low skill. When shifting skills to the rank-$x+\epsilon$ firm, we get PAM with talent intake represented by the empty dots.

inefficiency in classical matching models because the choice set of a firm depends only on its rank in the wage distribution, and not the identity of the firms above it.

The theorem and its proof have practical implications. If the planner cannot observe firms' actual skill but can only choose a set of admisible wages, then it will pool all firms in a single atom. In equilibrium, high-skill firms always offer weakly higher wages than low-skill firms. Given this constraint, and the fact that $\zeta(x)$ is decreasing, the planner's optimal policy is a fixed wage so that firms choose in a random order.[22] This optimum can also be implemented by a wage cap.[23]

As an illustration, this argument suggests that the NCAA's ban on paying college athletes raises talent in competitive college athletics. It does not do so by lowering wages - the marginal athletes would be paid zero wages anyhow - but rather by preventing colleges with the best scouts bidding away the best athletes and lowering the quality of the marginal programs, causing them to exit.

Another possible policy to increase quality in the market, namely entry barriers, is not effective in our model. In models where the informed party makes the entry decision (e.g. Atkeson, Hellwig, and Ordoñez (2015)) such barriers can increase welfare by cutting the incentives of low-quality firms to enter. However, in our model, given the positive assortative matching, equilibrium entry is optimal. Intuitively, the marginal firm does not impose an externality on others and enters if and

[22]The "mechanism design" version of this result is to let all firms report their skill to the planner. Incentive compatibility implies applicant quality is weakly increasing in skill, so the optimal mechanism is to pool the firms into a single atom.

[23]When using a fixed wage $\check{w}$, the planner must choose $\check{w}$ so the marginal firm breaks even. That is, $\check{w} = \mu E_x[\lambda(q(x), \check{r}) | x \geq F^{-1}(\check{r})] - k$, where the expectation is taken over ranks in the atom, $x \in [F^{-1}(\check{r}), 1]$. If $\check{w}$ is a not a fixed wage, but rather a wage cap that firms are free to underbid, we need to check additionally that no firm wants to do so. It suffices to check that the marginal firm $\check{r}$, which is most tempted to cut wages, does not want to cut its wage to 0, which is the most profitable deviation. To see this, note that at $w = 0$ the firm exerts no externality on other firms or workers, so captures its full contribution to social surplus. By definition of $\check{r}$, this social surplus is zero when the firm is placed inside the atom, as instructed by the planner. When posting $w = 0$, the contribution to social surplus is thus negative, meaning the firm's profits from setting $w = 0$ are also negative.

only if entry is efficient.[24]

We have modeled welfare by studying aggregate sorting into the industry, but the same economic forces can lead to inefficiency in related environments. First, suppose that firms can invest in their screening technology, simultaneously choosing their screening ability $r$ at cost $c(r)$ before posting wages. For example, this might represent hiring more HR professionals, conducting more interviews or more background checks. Indeed, many firms use external headhunters to find executives, paying 20-30% of the annual salary (Coverdill and Finlay, 1998). In equilibrium, firms post different wages, and pick different skill levels to complement these choices, with high-wage firms investing more in their skill. However, such an equilibrium is inefficient: adverse selection means that firms have excessive incentive to invest, since this worsens the quality of the application pool for firms posting lower wages. This result contrasts to the finding of efficient investment found in classic matching models, e.g. Cole, Mailath, and Postlewaite (2001).

Second, suppose we introduce complementarity or substitutability between firms and workers, analogous to Becker (1973). For example, if high-skill firms have lower marginal products then output is maximized by distributing talent equally across firms. However, if this submodular effect is small enough, equilibrium will still exhibit PAM, again meaning that equilibrium is inefficient.

## 5.2 Dynamic (In)efficiency

We now turn to the dynamic model from Section 3. Suppose that each potential firm starts off with exogenous talent $r_0$ and chooses whether or not to enter at cost $k/\rho$. Labeling firms by their talent rank $x$, flow social surplus $\mu R_t - (1-\check{x})k$ then depends on the entry threshold $\check{x}$ and aggregate employed talent $R_t := \int_{\check{x}}^1 r_t(x)dx$. Under PAM, firm $x$ has wage rank $x$, and there is steady state dispersion of talent, as in Section 3. Under NAM, firm $x$ has wage rank $1-x+\check{x}$, so talent at any firm converges to the same constant in finite time; once a set of firms has reached this constant talent level, we interpret NAM as all those firms posting the same wage distribution.

**Theorem 7.** *Consider the dynamic model with positive costs $k>0$. For any entry threshold $\check{x}>0$, social surplus is strictly higher under NAM than under PAM at all times $t>0$.*

*Proof.* We wish to show that aggregate employed talent is higher under NAM than under PAM, $R_t^N > R_t^P$, at all times $t>0$. Let $q_t^P(x)$ be applicant pool quality and $m_t^P := \bar{q} - q_t^P(\check{x})\check{x}$ be aggregate talent intake under PAM, and similarly for NAM using "$N$" superscripts. At $t=0$, we have $R_0^N = R_0 = R_0^P$. For all $t \geq 0$, we claim that whenever $R_t^N \geq R_t^P$ then $m_t^N > m_t^P$. To see this observe that, by assumption, there is more recruiting talent under NAM, and it places more weight on lower ranks of the wage distribution. Since $\zeta(x)$ in equation (19) is positive and decreasing in $x$, this means that more talented workers are employed under NAM than PAM, $m_t^N \geq m_t^P$. Moreover, since $\zeta(x)$ is strictly decreasing, this inequality is strict unless the talent distribution under PAM and NAM is both identical and degenerate, equal to one atom of identical firms, which can only arise at $t=0$. Since aggregate employed talent evolves according to $\dot{R}_t^P = \alpha(m_t^P - R_t^P)$, and similarly for NAM, the above claim and the single-crossing lemma implies $R_t^N > R_t^P$ for all $t>0$, as required. □

[24] With this said, if the planner imposes NAM or pools the firms together, then an entering firm does have a negative externality on other firms. This is why the optimal single wage must be strictly positive to prevent excessive entry. Indeed, this suggests that a low, fixed wage is optimal for the NCAA, such as paying for tuition.

As in the static model, this result has practical implications. Suppose the planner cannot observe the firms' initial talent and impose a screening order directly, but can choose a set of admissible wages at each point in time. In equilibrium, high-talent firms pay higher wages than low-talent firms, so the planner maximizes aggregate employed talent at all times $t > 0$ by forcing all firms to offer the same wage $w_t$ at time $t$.

# 6 Conclusion

This paper proposed a new model of firm dynamics in which talent takes center-stage. The model is based on the idea that talented workers are both more productive and are better at identifying talented applicants. In the static labor market, matching is positive assortative when talent is scarce, with high-skill firms posting higher wages and attracting better applicants. We then embedded this in a dynamic model in which today's recruits become tomorrow's managers and hire the next generation of workers. The economy converges to a unique steady state in which talent, wages and productivity are dispersed, helping us understand how talent can serve as a sustainable competitive advantage. Finally, we showed that equilibrium is inefficient and can be improved by policies that reduce wage dispersion.

The paper raises a number of questions related to the competition between firms, and the organization within firms that we do not address. Currently, we assume that all agents are involved in the recruiting process, meaning that the recruiting skill of a firm is a linear combination of all of its workers. This is reasonable if firms never learn the talent of their workers, or if the identity of the recruiter is defined by their job title (e.g. theorists hire theorists, while macroeconomists hire macroeconomists). However, in practice, firms do learn about the talent of their workers, and may be able to place more authority in the hands of those with more talent. Moreover, if the firm is trying to improve, as in Section 4.1, then it can place more authority in the hand of the recent, high-quality recruits. Such organizational issues raise even more questions: What if workers differ in their taste as well as their information? What if workers take actions to influence the allocation of authority in their favor? How should the firm design recruiting committees to aggregate information inside the organization?

We also assume that all jobs offer lifetime employment. It is straightforward to allow agents to leave at some exogenous rate and re-enter the job market. Assuming firms cannot observe the age of agents, the applicant pool would then consist of both new workers and those who recently quit. One can go a step further, and allow firms to learn the values of their workers (e.g. at a Poisson rate) and fire the bad ones. These workers would then re-enter the pool, lowering the overall quality $\bar{q}$. Finally, one might wish to allow for on-the-job search. This is particularly interesting because job-to-job moves help firms aggregate information and can be welfare improving. Any such analysis would depend on whether employers know the quality of their workers (e.g. Greenwald (1986)), and whether poaching firms can treat new and old workers differently (e.g. Board and Meyer-ter Vehn (2015))

Another important step is to analyze the quantitative implications of the model for the dispersion and persistence of talent and productivity. For such an exercise, it would be natural to suppose that firm technology $\mu$ follows a Markov process. Firms would then switch ranks over time as a function of shocks to the firm and industry, and for Brownian or Poisson technology, steady state productivity should satisfy Zipf's law. One would also like to introduce product market competition and endogenize dynamic entry and exit. While entry is simple, one would need to think carefully about exit since flow wages would no longer be a sufficient statistic for the value of a job.

To sum up, we hope that our model provides a benchmark to study the impact of talent management strategies on firm dynamics. That is, how individual firms organize in order to recruit and retain the best talent, and how these strategies interact with features of the labor market to co-determine the persistence and dispersion of talent, wages and productivity across an industry.

# Appendix

## A Analysis from Section 2

### A.1 Properties of $\Delta(\cdot)$ Function

Here we show that $\Delta(q)$ is single-peaked, with maximum $\hat{q} \in (0,1)$. Differentiating, we have

$$\lambda'(q;\theta) = \frac{1-p_\theta}{[1-(1-q)p_\theta]^2}$$
$$\lambda''(q;\theta) = \frac{-2(1-p_\theta)p_\theta}{[1-(1-q)p_\theta]^3} = \lambda'(q;\theta)\frac{-2p_\theta}{[1-(1-q)p_\theta]}$$

and so $\lambda''(q;\theta)/\lambda'(q;\theta)$ decreases in $p_\theta$. Thus, if $\Delta'(\hat{q}) = \lambda'(\hat{q};H) - \lambda'(\hat{q};L) = 0$, then

$$\Delta''(\hat{q}) = \lambda''(\hat{q};H) - \lambda''(\hat{q};L) = \lambda'(\hat{q};H)\left[\frac{\lambda''(\hat{q};H)}{\lambda'(\hat{q};H)} - \frac{\lambda''(\hat{q};L)}{\lambda'(\hat{q};L)}\right] < 0,$$

and so $\Delta$ is single-peaked.

We can compute $\hat{q}$. Since $\lambda'(\hat{q};H) = \lambda'(\hat{q};L)$, we have $[1-(1-\hat{q})p_L]\sqrt{1-p_H} = [1-(1-\hat{q})p_H]\sqrt{1-p_L}$, which implies $\hat{q} == \frac{(1-p_L)\sqrt{1-p_H}-(1-p_H)\sqrt{1-p_L}}{p_H\sqrt{1-p_L}-p_L\sqrt{1-p_H}}$.

### A.2 General Matching

In the body of the paper we assume that talent is scarce, $\bar{q} < \hat{q}$, implying PAM by Theorem 1. Here, we characterize the equilibrium if talent is abundant, $\bar{q} \in (\hat{q},1)$.

In equilibrium the advantage of skilled screeners $\Delta(Q(r))$ increases in $r$. Since $\Delta(q)$ is single-peaked around $\hat{q}$ , equilibrium has the following form. First, the most skilled firm $\bar{r}$ is matched with an applicant pool of quality $\hat{q}$. Second, less-skilled firms over the range $[r^*,\bar{r})$ post two wages and are matched with both worse applicant pools $Q_P(r) < \hat{q}$, and better applicant pools $Q_N(r) > \hat{q}$ such that $\Delta(Q_N(r)) = \Delta(Q_P(r))$. As we describe below, this indifference condition, determines the weight that firms put on the two wages. Finally, the least-skilled firms $[\underline{r},r^*)$ are matched with bad applicant-pools $Q_P(r) < \hat{q}$, where the cutoff $r^*$ satisfies $\Delta(Q_P(r^*)) = \Delta(\bar{q})$. Since $\Delta(\cdot)$ increases for low talent $q < \hat{q}$, the lower branch $Q_P(r)$ increases, representing PAM. Conversely, $\Delta(\cdot)$ decreases for high talent $q > \hat{q}$ and so the upper branch $Q_N(r)$ increases, representing NAM. These functions are illustrated in Figure 5. More formally:

**Theorem 8.** *Suppose talent is abundant, $\bar{q} \in (\hat{q},1)$. There exists an equilibrium. Equilibrium is characterized by a threshold $r^* \in (\underline{r},\bar{r})$, an increasing PAM function $Q_P(r)$ for $r \in [\underline{r},\bar{r}]$, and a decreasing NAM function $Q_P(r)$ for $r \in [r^*,\bar{r}]$.*

*Proof.* Unlike in the case with scarce talent, $\bar{q} < \hat{q}$, where equilibria must be in pure strategies, the preamble to this theorem shows that we must now consider mixed strategies. Any mixed strategy profile induces a joint distribution over screening skills $r$ and wages $w$, which in turn induces a joint distribution over $r$ and wage ranks $x$. Writing $r(x)$ for the expected skills at rank $x$, applicant pools $q(x)$ are determined by sequential screening (2) as usual. The resulting firm-applicant matching $Q(r)$ may now be set-valued. The optimality requirement $\Delta(Q(r)) \le \Delta(Q(r'))$ then means $\Delta(q) \le \Delta(q')$ for all $q \in Q(r), q' \in Q(r')$ with $r \le r'$.

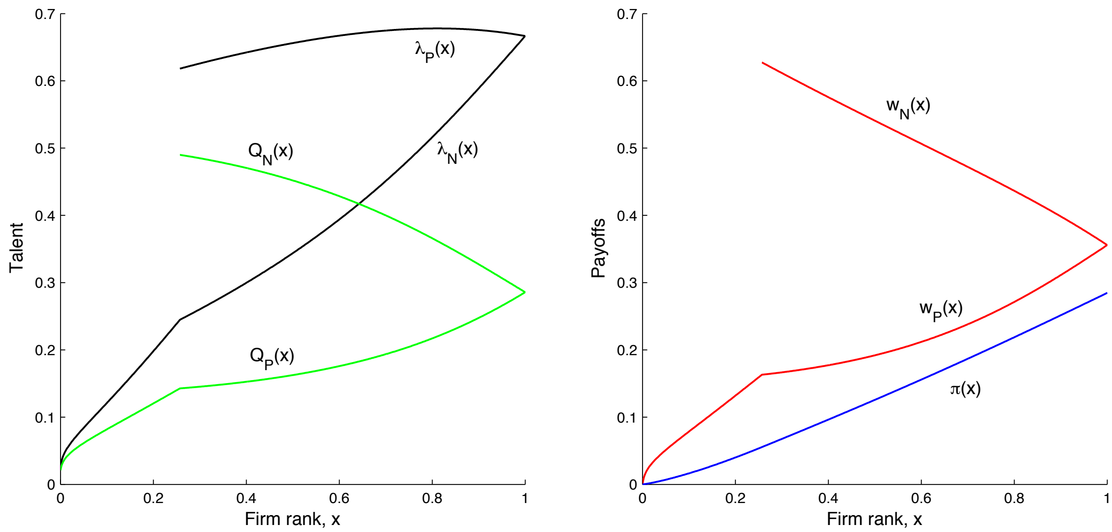


Figure 5: **Equilibrium with Abundant Talent.** The **left** panel shows the quality of applicants and recruits. The **right** panel shows the resulting wages and profits. The parameters are the same as Figure 1, except that $\bar{q} = 0.5$.

We first argue that only one firm is screening at a given wage rank $x$, and thus the skills of this firm equals $r(x)$. Assume otherwise, that two firms $r < r'$ both screen at rank $x \in Q(r), Q(r')$. Then $\Delta(Q(r)) = \Delta(Q(r'))$, and so incentive compatibility implies that any intermediate firm $\tilde{r} \in (r, r')$ must share the same value of $\Delta$. However, a mass of firms cannot share the same value of $\Delta$ since since $q(x)$ strictly increases and $\Delta(\cdot)$ has no flat spots.

Next, we characterize the form of matching. The monotonicity of $\Delta(Q(r))$ implies that $r(x)$ is single-peaked, increasing for $q(x) < \hat{q}$ and decreasing for $q(x) > \hat{q}$. Along the increasing branch, let firm $r$ be matched with rank $X_P(r)$, and let $Q_P(r) = q(X_P(r))$ be the corresponding applicant pool. Along the decreasing path, let firm $r$ be matched with rank $X_N(r)$ and let $Q_N(r) = q(X_N(r))$ be the corresponding applicant pool.

We construct equilibrium as follows. The top firm is matched with $\hat{q}$, meaning that $Q_P(\bar{r}) = Q_N(\bar{r}) = \hat{q}$. *A priori* we don't know the rank of firm $\bar{r}$, so we denote it by $\hat{x} := X_P(\bar{r}) = X_N(\bar{r})$. For lower firms, we have the indifference condition

$$\Delta(q(X_P(r))) = \Delta(q(X_N(r))) \tag{21}$$

and the availability of firms with type $r$

$$\left|X_N'(r)\right| + X_P'(r) = f(r). \tag{22}$$

We interpret $X_P'(r)$ as the weight firm $r$ places on the PAM branch. Since $\Delta(q)$ and $q(x)$ are continuous and the distribution of $r$ has no gaps, these supports are intervals with upper bound $\bar{r}$. Moreover, since $\Delta(q(0)) = \Delta(0) = 0 < \Delta(\bar{q}) = \Delta(q(1))$, the support of $X_P$ is all of $[\underline{r}, \bar{r}]$, while the support of $X_N$ is truncated below, at some $r^* \in (\underline{r}, \bar{r})$. We have thus established the equilibrium characterization of Theorem 8.

To establish equilibrium existence (and also show how to compute an equilibrium), we now show that there exist functions $\{Q_N(r), Q_P(r), X_N(r), X_P(r)\}$ satisfying (21) and (22), and wages $W(q)$

so that firm $r$ finds it optimal to attract applicants $Q_N(r)$ and $Q_P(r)$. For the PAM branch, we have to adjust the sequential screening equation (2) by the weight that the firms place on the PAM branch,

$$Q_P'(r) = \frac{\lambda(Q_P(r), r) - Q_P(r)}{X_P(r)} X_P'(r). \tag{23}$$

and similarly for the NAM branch. Differentiating (21), and substituting for $Q_P'(r)$ and $Q_N'(r)$, and then using (22) yields

$$\Delta'(Q_P(r)) \left[\frac{\lambda(Q_P(r), r) - Q_P(r)}{X_P(r)}\right] X_P'(r) = \Delta'(Q_N(r)) \left[\frac{\lambda(Q_N(r), r) - Q_N(r)}{X_N(r)}\right] \left(X_P'(r) - f(r)\right). \tag{24}$$

This equation then yields the weight $X_P'(r)$ and, integrating from $\hat{x}$, yields the mass functions $X_N, X_P$. One can then derive the applicant functions $Q_N, Q_P$ from (23).

To ensure this ODE solves the original problem, it needs to satisfy the boundary condition $Q_N(r(1)) = \bar{q}$, which says that the top ranked firm recruits from the unadulterated pool. We therefore study $Q_N(r(1))$ as a function of the free variable $\hat{x}$, and apply the intermediate value theorem. At the top, when $\hat{x} = 1$, $\bar{r}$ is the highest ranked firm by assumption implying $Q_N(r(1)) = \hat{q} < \bar{q}$. At the bottom, as $\hat{x} \to 0$ we claim that $Q_N(r(1)) \to 1 < \bar{q}$. To see why, observe that as $X_P(\bar{r}) = \hat{x} \to 0$ firms place all their weight on the NAM branch of the equilibrium. For this to be true, the indifference equation (21) means that $Q_N(r) > \hat{q}$ for all $r < \bar{r}$ which requires $\bar{q} \to 1$ so that $Q_N(r) \to 1$ pointwise. Thus by continuity, for some $\hat{x} \in (0,1)$, there exists a solution to the ODE with the original boundary conditions and $X_P(\bar{r}) = X_N(\bar{r}) = \hat{x}$.

Finally, we turn to wages. Given the functions $\{Q_N, Q_P, X_N, X_P\}$, wages $W(q)$ are given by the first-order condition. For the PAM branch, this means

$$W'(Q_P(r)) = \lambda'(Q_P(r); r)$$

with boundary condition $W(0) = 0$.

Now, to see that $Q_N, Q_P$, along with wages $W(q)$ constitutes an equilibrium, consider firm $r$, say in the joint support of $X_P$ and $X_N$. The marginal value of attracting better applicants $\lambda'(q;r) - W'(q) = \lambda'(q;r) - \lambda'(q;r(q)) = \Delta'(q)(r - r(q))$ is (1) positive for $q < Q_P(r)$ since $\Delta' > 0$ and $r > r(q)$; (2) negative for $q \in (Q_P(r), \hat{q})$ since $\Delta' > 0$ and $r < r(q)$; (3) positive for $q \in (\hat{q}, Q_N(r))$ since $\Delta' < 0$ and $r < r(q)$; and (4) negative for $q > Q_N(r)$ since $\Delta' < 0$ and $r > r(q)$. Thus, the two local maxima are $Q_P(r)$ and $Q_N(r)$. To see that these are equally profitable, we write their difference as

$$\begin{aligned}\int_{Q_P(r)}^{Q_N(r)} \Delta'(\tilde{q})(r - r(\tilde{q}))d\tilde{q} &= \int_{Q_P(r)}^{Q_P(\bar{r})} \Delta'(\tilde{q})(r - r(\tilde{q}))d\tilde{q} + \int_{Q_N(\bar{r})}^{Q_N(r)} \Delta'(\tilde{q})(r - r(\tilde{q}))d\tilde{q} \\ &= \int_r^{\bar{r}} \Delta'(Q_P(\tilde{r}))(r - \tilde{r})Q_P'(\tilde{r})d\tilde{r} + \int_{\bar{r}}^{r} \Delta'(Q_N(\tilde{r}))(r - \tilde{r})Q_N'(\tilde{r})d\tilde{r}\end{aligned}$$

which is zero by (24). □

It is worth noting that this model is highly tractable. In general, models with PAM and NAM regions are hard to characterize (e.g. Chiappori, McCann, and Nesheim (2010)). For us, the key to tractability is the fact that the screening advantage $\Delta(q)$ is independent of $r$, which follows from our assumption that a firm's screening skills are a linear combination of its recruiters' skills.

## A.3 Comparative Statics

Here we show that wage and productivity dispersion rise when either screening skills rise (due to improved technology), or aggregate talent falls (due to skill-biased technological change). In order to simplify the analysis, we assume that all recruiters are identical, screening out untalented workers with probability $p$. This means that firms are also identical, so profts are zero and wages, productivity and talent all coincide.

We say that *dispersion rises* when the talent ratio between the $x'$-ranked firm and the $x$-ranked firm rises for all $x' > x$. Writing the recruit quality as a function of rank $x$, screening skills $p$, and aggregate talent $\bar{q}$, we wish to show$\lambda(x', p, \bar{q})/\lambda(x, p, \bar{q})$ is decreasing in $\bar{q}$ and increasing in $p$. That is, a fall in $\bar{q}$ or an increase in $p$ allows the top firms to acquire a proportionately larger share of the talent. This is formalized by the following result:

**Lemma 2.** *$\lambda(x, p, \bar{q})$ is log-submodular in $(x, \bar{q})$ and, if $p(1 + \bar{q}) \leq 1$, log-supermodular in $(x, p)$.*

*Proof.* We first consider the effect of changing $\bar{q}$. Let us write the applicant quality by $q(x, p, \bar{q})$ and it's logarithm by $\vartheta(x, p, \bar{q}) = \log q(x, p, \bar{q})$. We claim that $q(x, p, \bar{q})$ is log-submodular in $(p, \bar{q})$. By (1) and (2), log-applicant quality obeys $\vartheta_x(x, p, \bar{q}) = \frac{q_x(x,p,\bar{q})}{q(x,p,\bar{q})} = \frac{1}{x}(\frac{1}{1-p(1-e^{\vartheta})} - 1) =: \psi(x, \vartheta, p)$. Differentiating gives us the result

$$\vartheta_{x\bar{q}} = \frac{d}{d\bar{q}}\psi(x, \vartheta(x, p, \bar{q}), p) = \psi_\vartheta \vartheta_{\bar{q}} < 0$$

since $\psi_\vartheta < 0$ and $\vartheta_{\bar{q}} > 0$. Intuitively, when applicant talent $\bar{q}$ drops by half, the talent of top firms drops by less than half since these firms screen out some of the newly untalented applicants, meaning they hire proportionally more of the available talent.

Next, we claim that $\lambda(x, p, \bar{q})$ is log-submodular in $(x, \bar{q})$. To see this, write log recruit quality, $\Lambda := \log \lambda$, as a function of screening skills $p$ and log applicant quality $\vartheta$, $\Lambda(\vartheta, p) = \log \frac{e^{\vartheta}}{1-p(1-e^{\vartheta})} = \vartheta - \log(1 - p(1 - e^{\vartheta}))$. First note that $\Lambda(\vartheta, p)$ increases in $\vartheta$ with $\Lambda_\vartheta = 1 - \frac{qp}{1-p(1-q)} = \frac{1-p}{1-p(1-q)} > 0$, and increases in $p$ with $\Lambda_p = \frac{1-q}{1-p(1-q)} = \frac{1}{1/(1-q)-p} > 0$. The submodularity in $(x, \bar{q})$ then follows easily because both factors in

$$\frac{d}{d\bar{q}}\Lambda(\vartheta(x, p, \bar{q}), p) = \Lambda_\vartheta(\vartheta(x, p, \bar{q}), p)\vartheta_{\bar{q}}(x, p, \bar{q})$$

are positive and fall in $x$. Intuitively, the larger proportional decline of applicant quality at lower-ranked firms is aggravated by the concavity of recruit quality $\Lambda$ in applicant quality $\vartheta$.

We now consider the effect of changing $p$. We first show that $q(x, p, \bar{q})$ is decreasing in $p$ and log-supermodular in $(x, p)$. To verify monotonicity, observe that for $x = 1$, $q(x, p, \bar{q}) = \bar{q}$, independent of $p$. Fixing $p' > p$, (1) and (2) imply $q_x(x, p, \bar{q}) = \frac{1}{x}\left(\frac{(1-q)p}{1-(1-q)p}\right)$, so $q_x(x, p', \bar{q}) > q_x(x, p', \bar{q})$ when $q(x, p', \bar{q}) = q(x, p', \bar{q})$. The single-crossing lemma then implies $q(x, p', \bar{q}) < q(x, p, \bar{q})$ for all $x < 1$. To verify log-supermodularity, differentiating yields

$$\vartheta_{xp} = \frac{d}{dp}\psi(x, \vartheta(x, p, \bar{q}), p) = \psi_\vartheta \vartheta_p + \psi_p > 0 \tag{25}$$

since $\psi_\vartheta, \vartheta_p < 0$ and $\psi_p > 0$, and thus falls by a greater amount at lower-ranked firms when screening skills $p$ increase.

Next, we claim that $\lambda(x, p, \bar{q})$ is log-supermodular in $(x, p)$ if $p(1 + \bar{q}) \leq 1$. This is harder to establish since the greater proportional drop of applicant quality at low-ranked firms is counteracted by a greater benefit of improved screening skills at those firms, as witnessed by $\Lambda_{\vartheta p} < 0$. For the aggregate effect to favor high-ranked firms, we need

$$\frac{d}{dp}\Lambda(\vartheta(x, p, \bar{q}), p) = \Lambda_\vartheta \vartheta_p + \Lambda_p = \frac{(1-p)\vartheta_p + (1-q)}{1-p(1-q)} \tag{26}$$

to rise in $x$. Indeed, differentiating (26) with respect to $x$, multiplying through with $\frac{x}{1-q}(1-p(1-q))^3 > 0$, recalling from (25) that $\vartheta_{xp} \geq \psi_p$ and $\vartheta_p \leq 0$, and then substituting $\psi_p = \frac{1}{x}\frac{1-q}{(1-p(1-q))^2}$ and $q_x = \frac{1}{x}(\frac{q}{1-p(1-q)} - q) = \frac{1}{x}\frac{pq(1-q)}{1-p(1-q)}$, we get

$$\begin{aligned}
\tfrac{d^2}{dxdp}\Lambda(\vartheta(x, p, \bar{q}), p) &\overset{\text{sgn}}{=} \frac{x(1-p(1-q))^2}{1-q}\frac{\partial[(1-p)\vartheta_p + (1-q)]}{\partial x} - \frac{x(1-p(1-q))}{1-q}pq_x[(1-p)\vartheta_p + (1-q)] \\
&\geq \frac{x(1-p(1-q))^2}{1-q}[(1-p)\psi_p - q_x] - x(1-p(1-q))pq_x \\
&= 1 - p - pq(1-p(1-q)) - p^2q(1-q) \\
&= 1 - p(1+q) \geq 0
\end{aligned}$$

where the last inequality follows by the assumption that $p(1 + \bar{q}) \leq 1$. □

# B Proofs from Section 3

## B.1 Proof that Talent Distribution is Smooth

Here we argue that if there is an atom of initially identical firms, these firms diverge immediately. Assume to the contrary, that at time $t > 0$ an atom of firms has the same worker quality $r_t$. Writing $x$ for the rank of $r_t$, let this atom be $[\underline{x}, \bar{x}]$. Since optimal wages rise in talent and hence talent differences never vanish, firms in the atom must have identical talent $r_s$ for all $s \in [0, t]$. At any time $s \in [0, t]$ the wage distribution must be smooth by the arguments in Section 2. If firms in the atom post different wages, they drift apart. Hence the firms must employ non-degenerate distributional strategies,[25] posting both high and low wages to attract good and bad applicants; they must thus be indifferent across a range of applicants $[\underline{q}_s, \bar{q}_s]$ for all $s \in [0, t]$. Thus, the first order condition (8) must hold with equality on $[\underline{q}_s, \bar{q}_s]$ for all $s \in [0, t]$ and the atom quality $r_s$.

To see that such distributional strategies cannot be optimal, consider a firm that deviates by always attracting the best applicants in the atom $\bar{q}_s$, rather than mixing over good and bad applicants. At time $s = 0$, the choice $\bar{q}_0$ is optimal. Moreover, over time the firm's quality rises above $r_s$ since it attracts better applicants. Since the marginal benefit of attracting better applicants, the RHS of (8), strictly increases in $r$, this deviation strictly improves on the posited distributional strategy. This proves that initially identical firms diverge immediately.

[25]When using a distributional strategy, a firm posts an entire distribution of wages $\nu_t = \nu_t(w)$ of wages at any time $t$; we then interpret $r_t(x)$ as the weighted-average talent of firms posting the $x$-ranked wage, and solve for the firm's evolution of talent by taking expectations over the RHS of (5).

### B.2 Proof of Theorem 3

Here we show that worker quality converges. First rewrite firm $x$'s talent evolution (5) as $\dot{r}_t(x) = \alpha(\lambda(q;L) - (1-\Delta(q_t(x)))r_t(x))$. For constant applicants $q$, quality thus drifts towards $\xi(q) = \lambda(q;L)/(1-\Delta(q))$ at rate $\alpha(1-\Delta(q))$. Of course, for any firm $x$, the applicant quality $q_t(x)$ also changes over time, and $q_t'(x)$ is given by (4). We now show that these two nested ODEs, converge to the steady state $q_*(x), r_*(x)$.

First, we derive a contraction property. Define the limits $\underline{q}(x) := \liminf_t q_t(x)$, $\bar{q}(x) := \limsup_t q_t(x)$, $\underline{r}(x) := \liminf_t r_t(x)$, and $\bar{r}(x) := \limsup_t r_t(x)$. Next, interpret (4) as an operator $Q$, mapping firm quality functions $r_t(x)$ into applicant quality functions $q_t(x) = Q[r_t(\cdot)](x)$. We claim that:

$$Q[\xi(\bar{q}(\cdot))](x) \leq \underline{q}(x) \leq \bar{q}(x) \leq Q[\xi(\underline{q}(\cdot))](x). \tag{27}$$

To understand (27), first observe that that $Q$ is antitone: if $r(x) \geq \hat{r}(x)$ for all $x$ then $q(x) = Q(r(\cdot))(x) \leq Q(\hat{r}(\cdot))(x) = \hat{q}(x)$, since $q(x) = \hat{q}(x) = \bar{q}$ and the RHS of (4) increases in $r$. Intuitively, better recruiters introduce more adverse selection. Inequalities (27) then state that if applicant quality was equal to one of its limits, $\underline{q}$ and $\bar{q}$, and quality $r$ was in steady state $r = \xi(q)$, then the induced difference in applicant pools is larger than the original difference.

We prove (27) in two steps. First, since $r_t(x)$ drifts towards $\xi(q_t(x))$, which is asymptotically bounded by $\xi(\underline{q}(x))$ and $\xi(\bar{q}(x))$, we have

$$\xi(\underline{q}(x)) \leq \underline{r}(x) \leq \bar{r}(x) \leq \xi(\bar{q}(x)) \tag{28}$$

for all $x$. Second,

$$\underline{q}(x) = \lim_{t\to\infty} \inf_{t'>t} q_{t'}(x) = \lim_{t\to\infty} \inf_{t'>t} \{Q[r_{t'}(\cdot)](x)\} \geq \lim_{t\to\infty} Q[\sup_{t'>t}\{r_{t'}(\cdot)\}](x) = Q[\bar{r}(\cdot)](x)$$

where the first inequality is the definition of the lim inf, the second the definition of the operator $Q$, the inequality uses the antitonicity of $Q$ – since $r_{t'}(\hat{x}) \leq \sup_{t'>t}\{r_{t'}(\hat{x})\}$ for all $t'$ and $\hat{x}$, we have $Q[r_{t'}(\cdot)](x)$ exceeds $Q[\sup_{t'>t}\{r_{t'}(\cdot)\}](x)$ for all $t'$ and $x$, and hence so does $\inf_{t'>t} Q[r_{t'}(\cdot)](x)$– and the last inequality uses the dominated convergence theorem to exchange the limit $t \to \infty$ and the operator $Q$, as well as the definition of the limsup, $\bar{r}(x) = \lim_{t\to\infty} \sup_{t'>t} r_{t'}(x)$. Together with the analogue argument for $\bar{q}(x)$, we get

$$Q[\bar{r}(\cdot)](x) \leq \underline{q}(x) \leq \bar{q}(x) \leq Q[\underline{r}(\cdot)](x). \tag{29}$$

Since $Q$ is antitone, (28) and (29) imply (27).

To complete the proof of convergence, suppose "inductively" that applicant and firm quality converge above some $\hat{x} \in (0,1]$, i.e. $\underline{q}(x) = \bar{q}(x)$, and hence $\underline{r}(x) = \bar{r}(x)$, for all $x \in (\hat{x},1]$. Fix $\epsilon$, and let $\delta(\epsilon) := \max_{x\in[\hat{x}-\epsilon,\hat{x}]} \left|\bar{q}(x) - \underline{q}(x)\right|$ be the maximum distance the liminf and limsup drift apart over $[\hat{x}-\epsilon,\hat{x}]$. Since $\xi(q)$ is Lipschitz in $q$, and the RHS of (4) is locally Lipschitz in $q$ and $r$,

$$\max_{x\in[\hat{x}-\epsilon,\hat{x}]} \left|Q[\xi(\underline{q}(\cdot))](x) - Q[\xi(\bar{q}(\cdot))](x)\right| \leq K\epsilon\delta(\epsilon)$$

for some constant $K$. Letting $\epsilon < 1/K$, this equation states the exact opposite of (27). Hence we must have $\underline{q}(x) = \bar{q}(x)$ and hence $\underline{r}(x) = \bar{r}(x)$, for all $x \in [\hat{x}-\epsilon,1]$, and thus for all $x \in [0,1]$.

# C Proofs from Section 4

## C.1 Proof that Wages Rise in Productivity

To verify that $V_t'(r;\mu)$ increases in $\mu$, fix an initial workforce $r_t$ and consider the optimal wage policies $\{\hat{q}_s^{H,L}\}$ and the associated worker trajectories $\{r_s^{H,L}\}$ of a high and low productivity firm. If $\hat{q}_t^H = \hat{q}_t^L$, the statement is true. Thus we can assume $q_t^H \neq q_t^L$ and define $T$ as the first time (possibly $\infty$) after $t$ with $r_T^H = r_T^L$. Thus, the trajectories $\{r_s^H\}_{s\in[t,T]}$ and $\{r_s^L\}_{s\in[t,T]}$ are strictly ordered, in that we either have $r_s^H < r_s^L$ for all $s \in (t,T)$, or $r_s^H > r_s^L$ for all $s \in (t,T)$. Write $R_{H,L} = \int_t^T e^{-\rho(s-t)} r_s^{H,L} ds$ and $W_{H,L} = \alpha \int_t^T e^{-\rho(s-t)} W(\hat{q}_s^{H,L}) ds$ for the aggregate workforce and wage bill of firm $\mu$. Since the trajectories $r_s^H$ and $r^L$ reunite at $s = T$, optimality by firms $H$ and $L$ imply $\mu_H R_H - W_H \geq \mu_H R_L - W_L$ and $\mu_L R_H - W_H \leq \mu_L R_L - W_L$. Thus, $(\mu_H - \mu_L)(R_H - R_L) \geq 0$, and hence $r_s^H > r_s^L$ for all $s \in (t,T)$.

## C.2 Proof of Theorem 4

Fix an equilibrium. For any time $t$, let $y_t$ be a firm's rank in the distribution of talent $r$, and $x_t$ its rank in the distribution of wages $w$. Since wages rise in talent and technology, high-technology firms post higher wages than their talent rank, $x_t \geq y_t$, and so $y_t$ (weakly) increases; conversely, low productivity firms have $x_t \leq y_t$, and $y_t$ (weakly) decreases. In either case, $y_t$ converges to some $y$. Since $y_t$ converges uniformly for any countable set of firms - which can be chosen to be dense in the set of all firms' initial ranks - the convergence is indeed uniform for all firms. Clearly, the distribution of limit ranks $y$ is uniform on $[0,1]$, and we can identify firms with their productivity and their limit rank $y$.

Let $F_H, F_L$ be the distribution of limit ranks for high and low productivity firms. To establish stratified technology-talent matching, we will show that the support of $F_H$ lies above the support of $F_L$. Let $\bar{y} := \max \operatorname{supp} F_L$ for the highest limit rank of a low productivity firm, we need to show that there is no "mismatched" high productivity firm, i.e. with $y < \bar{y}$. Assume otherwise, and let $\tilde{y} = \sup\{y \in \operatorname{supp} F_H : y < \bar{y}\}$ be the highest mismatched firm; note that $\tilde{y} \leq \bar{y} \leq 1$, but neither inequality need be strict. We will derive a contradiction by showing that high-technology just below $\tilde{y}$ outbid low-technology firms just above. However, in order to do this, we first have to show that for wage ranks converge to talent ranks on $[\tilde{y}, 1]$.

First, consider firms $[\bar{y}, 1]$. Almost all of these firms are high-technology, so we are in the setting of Section 3.3, and Theorem 3 implies that applicant and worker quality $(q, r)$ converge to their steady state levels $(q_*(y), r_*(y))$ for all $y \in [y, 1]$.

Next, consider firms $[\tilde{y}, \bar{y}]$. Since $\tilde{y}$ is in the support of both $F_H$ and $F_L$, there are both high- and low-technology firms with this limit rank. Write $\tilde{x}_t^H, \tilde{y}_t^H$ and $\tilde{x}_t^L, \tilde{y}_t^L$ for the wage rank and talent rank of the high and low technology firms that converge to $\tilde{y}$. Since both the high wage ranks of the high-technology firm $\tilde{x}_t^H \geq \tilde{y}_t^H$ and the low wage ranks of the low-technology firm $\tilde{x}_t^L \leq \tilde{y}_t^L$ must sustain the same limit quality $\lim \tilde{y}_t^H = \lim \tilde{y}_t^L = \tilde{y}$, the wage ranks, too, must converge to the same limit $\lim \tilde{x}_t^H = \lim \tilde{x}_t^L = \tilde{y}$. Then, since almost all firms with $y \in [\tilde{y}, \bar{y}]$ are low-technology and the order of their wage ranks is time-invariant with the lowest converging to $\tilde{y}$ and the highest being bounded above by $\bar{y}$, the wage rank of any such firm $y$ must converge to $y$. Thus, applicant and worker quality converge to the steady-state quantities, $q_*(y)$ and $r_*(y)$, also for these firms.

To obtain the contradiction, consider a low-technology firm with talent $r_t^L$ limit talent rank $\tilde{y}$, and a high-technology firm with talent $r_t^H$ and limit talent rank $\tilde{y} - \epsilon$, and compare their marginal benefit of better applicants $V_t'(r;\mu)\lambda'(q;r)$. Since talent ranks converge to close-by limits and talent is continuous in rank, the talent gap $r_t^L - r_t^H$ vanishes for small $\epsilon$ and large $t$. Moreover, wage quantiles and hence applicants qualities $\hat{q}_t^L$ and $\hat{q}_t^H$ are $\epsilon$-close for large $t$. Recalling

$$V_t'(r;\mu) = \mu \int_t^\infty e^{-\int_t^s \rho+\alpha(1-\Delta(\hat{q}_u))du} ds$$

from (10), the ratio $V_t'(r_t^H;\mu_H)\lambda'(q;r_t^H)/V_t'(r_t^L;\mu_L)\lambda'(q;r_t^L)$ converges to $\mu_H/\mu_L > 1$ for small $\epsilon$ and large $t$. This contradicts the assumption that the low-technology firm pays higher wages.

### C.3 Wage Dispersion and Peer Effects

To study wage dispersion, observe that (17) give us,[26]

$$\frac{d}{dx} W_t(q_t(x), r_t(x)) = \lambda'(q_t(x); r_t(x)) V_t'(r_t(x)) q_t'(x) - \gamma r_t'(x)$$

Now, suppose that all firms start off with identical talent. We claim that initial wage dispersion is larger with peer effects. At $t = 0$, $r_0(x) \equiv \bar{q}$, and

$$\frac{d}{dx} W_0(q_0(x), \bar{q}) = \lambda'(q_0(x); \bar{q}) V_0'(\bar{q}) q_0'(x)$$

The Envelope Theorem implies that the marginal value of talent is

$$V_t'(r_t) = (\mu + \alpha\gamma) \int_t^\infty e^{-\int_t^s \rho+\alpha(1-\Delta(\hat{q}_u))du} ds$$

is increasing in $\gamma$. Hence wages become more dispersed as $\gamma$ grows.

In steady state, the marginal value is given by (18), while the applicant and talent pool are related by $\lambda(q_*(x), r_*(x)) = r_*(x)$. Substituting these in yields,

$$\frac{d}{dx} W_*(q_*(x), r_*(x)) = \left[ \frac{\mu(1 - \Delta(q_*(x))) - \gamma\rho}{\rho + \alpha(1 - \Delta(q_*(x)))} \right] r_*'(x)$$

Hence if the wage schedule is increasing, as occurs when peer effects are sufficiently small, an increase in $\gamma$ lowers dispersion.

## D Proofs from Section 5

### D.1 Proof that $\zeta(x)$ is Decreasing in Theorem 6

Here we show that equation (20) is strictly negative, and hence that $\zeta(x)$ is decreasing. Using $q'(x) = [\lambda(q(x); r(x)) - q(x)]/x$, multiplying by $x$, dropping the argument $x$, and using linearity in

[26] To determine the level of wages, observe that the lowest firm will choose wages so that utility is zero, i.e. $W_t(q_t(0), r_t(0)) = -\gamma r_t(0)$. In steady state, $r_*(0) = 0$.

$r$ to get $\lambda(q;r) - q = \lambda(q;L) - q + r\Delta(q)$ and $\lambda'(q;r) = \lambda'(q;L) + r\Delta'(q)$, (20) simplifies to

$$\zeta'(x) \overset{\text{sgn}}{=} \Delta'(q)(\lambda(q;r) - q) - \lambda'(q;r)\Delta(q) = \Delta'(q)(\lambda(q;L) - q) - \lambda'(q;L)\Delta(q) \tag{30}$$

since the $r\Delta(q)\Delta'(q)$ terms cancel. Next, substituting

$$\begin{aligned}
\lambda(q;L) - q &= \frac{q}{1-(1-q)p_L} - q = \frac{q(1-q)p_L}{1-(1-q)p_L} \\
\lambda'(q;L) &= \frac{1-(1-q)p_L - qp_L}{(1-(1-q)p_L)^2} = \frac{1-p_L}{(1-(1-q)p_L)^2} \\
\Delta(q) &= \frac{q}{1-(1-q)p_H} - \frac{q}{1-(1-q)p_L} \\
\Delta'(q) &= \frac{1-p_H}{(1-(1-q)p_H)^2} - \frac{1-p_L}{(1-(1-q)p_L)^2}
\end{aligned}$$

into the RHS of (30) and dividing by $\frac{q}{1-(1-q)p_L} > 0$, we get

$$\zeta'(x) \overset{\text{sgn}}{=} (1-q)p_L\left(\frac{1-p_H}{(1-(1-q)p_H)^2} - \frac{1-p_L}{(1-(1-q)p_L)^2}\right) - \frac{1-p_L}{1-(1-q)p_L}\left(\frac{1}{1-(1-q)p_H} - \frac{1}{1-(1-q)p_L}\right).$$

Multiplying with $(1-(1-q)p_L)^2(1-(1-q)p_H)^2 > 0$, this becomes

$$\zeta'(x) \overset{\text{sgn}}{=} (1-q)p_L[\underbrace{(1-p_H)(1-(1-q)p_L)^2 - (1-p_L)(1-(1-q)p_H)^2}_{:=T}] - (1-p_L)(1-(1-q)p_H)(1-q)(p_H - p_L). \tag{31}$$

Substituting

$$\begin{aligned}
T &= 1 - 2(1-q)p_L + ((1-q)p_L)^2 - p_H(1 - 2(1-q)p_L + ((1-q)p_L)^2) \\
&\quad - (1 - 2(1-q)p_H + ((1-q)p_H)^2) + p_L(1 - 2(1-q)p_H + ((1-q)p_H)^2) \\
&= 2(1-q)(p_H - p_L) - (1-q)^2(p_H^2 - p_L^2) - (p_H - p_L)(1 - p_H p_L(1-q)^2) \\
&= (p_H - p_L)\big(2(1-q) - (1-q)^2(p_H + p_L - p_H p_L) - 1\big)
\end{aligned}$$

back into (31) and dividing by $(1-q)(p_H - p_L) > 0$, (31) becomes

$$\begin{aligned}
\zeta'(x) &\overset{\text{sgn}}{=} p_L\big(2(1-q) - (1-q)^2(p_H + p_L - p_H p_L) - 1\big) - (1-p_L)(1-(1-q)p_H) \\
&= p_L\big(-(1-q)p_H + 2(1-q) - (1-q)^2(p_H + p_L - p_H p_L)\big) - (1-(1-q)p_H) \\
&= (1-q)\left[p_H + p_L\left(2 - p_H - (1-q)(p_H + p_L - p_H p_L)\right)\right] - 1.
\end{aligned}$$

To see that this is negative, note that it is linear in $p_H$, and negative at $p_H = 0$ and $p_H = 1$: at $p_H = 0$ we have $(1-q)p_L(2-(1-q)p_L) - 1 < 0$ since $x(2-x) < 1$ for all $x < 1$; at $p_H = 1$ we have $(1-q)(1+p_L q) - 1 < (1-q)(1+q) - 1 = -q^2 < 0$.

# References


Abowd, J., F. Kramarz, and D. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–333.

Acemoglu, D., and J. A. Robinson (2000): "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *Quarterly Journal of Economics*, 115(4), 1167–1199.

Albrecht, J. W., and S. B. Vroman (1992): "Non-Existence of Single-Wage Equilibria in Search Models with Adverse Selection," *Review of Economic Studies*, 59(3), 617–624.

Anderson, A. (2016): "A Dynamic Generalization of Becker's Assortative Matching Result," *Journal of Economic Theory*, (forthcoming).

Anderson, A., and L. Smith (2010): "Dynamic Matching and Evolving Reputations," *Review of Economic Studies*, 77(1), 3–29.

Arai, M. (2003): "Wages, Profits, and Capital Intensity: Evidence from Matched Worker-Firm Data," *Journal of Labor Economics*, 21(3), 593–618.

Atkeson, A., C. Hellwig, and G. Ordoñez (2015): "Optimal Regulation in the Presence of Reputation Concerns," *Quarterly Journal of Economics*, 130(1), 415–464.

Barberà, S., M. Maschler, and J. Shalev (2001): "Voting for Voters: A Model of Electoral Evolution," *Games and Economic Behavior*, 37(1), 40–78.

Beaman, L., and J. Magruder (2012): "Who gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, 102(7), 3574–3593.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813–846.

Bertrand, M., and A. Schoar (2003): "Managing With Style: The Effect Of Managers On Firm Policies," *Quarterly Journal of Economics*, 118(4), 1169–1208.

Board, S., and M. Meyer-ter Vehn (2015): "Relational Contracts in Competitive Labour Markets," *The Review of Economic Studies*, 82(2), 490–534.

Board, S., and M. Meyer-ter-Vehn (2013): "Reputation for Quality," *Econometrica*, 81(6), 2381–2462.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–273.

Burks, S. V., B. Cowgill, M. Hoffman, and M. Housman (2015): "The Value of Hiring through Employee Referrals," *Quarterly Journal of Economics*, 130(2), 805–839.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128(3), 967–1015.

Chakraborty, A., A. Citanna, and M. Ostrovsky (2010): "Two-Sided Matching with Interdependent Values," *Journal of Economic Theory*, 145(1), 85–105.

Chiappori, P.-A., R. J. McCann, and L. P. Nesheim (2010): "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," *Economic Theory*, 42(2), 317–354.

Cole, H. L., G. J. Mailath, and A. Postlewaite (2001): "Efficient Non-contractible Investments in Large Economies," *Journal of Economic Theory*, 101(2), 333–373.

Coverdill, J. E., and W. Finlay (1998): "Fit and Skill in Employee Selection: Insights from a Study of Headhunters," *Qualitative Sociology*, 21(2), 105–127.

DEWAN, T., AND D. P. MYATT (2010): “The Declining Talent Pool of Government,” *American Journal of Political Science*, pp. 267–286.

——— (2014): “Playing for the Winning Team: Selection, Performance, and the Longevity of Organizations,” Working paper, LSE.

FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): “Reallocation, Firm Turnover and Efficiency: Selection on Productiviy or Profitability?,” *American Economic Review*, 98(1), 394–425.

FUCHS, W., B. GREEN, AND D. PAPANIKOLAOU (2016): “Adverse Selection, Slow-moving Capital, and Misallocation,” *Journal of Financial Economics*, 120(2), 286–308.

GARICANO, L. (2000): “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 108(5), 874–904.

GARICANO, L., AND E. ROSSI-HANSBERG (2006): “Organization and Inequality in a Knowledge Economy,” *Quarterly Journal of Economics*, 121(4), 1383–1435.

GREENWALD, B. C. (1986): “Adverse Selection in the Labour Market,” *Review of Economic Studies*, pp. 325–347.

GRETSKY, N. E., J. M. OSTROY, AND W. R. ZAME (1992): “The Nonatomic Assignment Model,” *Economic Theory*, 2(1), 103–127.

GUASCH, J. L., AND A. WEISS (1980): “Wages as Sorting Mechanisms in Competitive Markets with Asymmetric Information: A Theory of Testing,” *Review of Economic Studies*, pp. 653–664.

HARTMAN, P. (2002): *Ordinary Differential Equations*. siam.

HASTINGS, R., AND P. MCCORD (2009): “Netflix Culture: Freedom and Responsibility,” Presentation, http://www.slideshare.net/reed2001/culture-1798664.

HOPENHAYN, H., AND R. ROGERSON (1993): “Job Turnover and Policy Evaluation: A General Equilibrium Analysis,” *Journal of Political Economy*, 101(5), 915–938.

HOPENHAYN, H. A. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60(5), 1127–1150.

HOPENHAYN, H. A. (2014): “Firms, Misallocation, and Aggregate Productivity: A Review,” *Annual Review of Economics*, 6(1), 735–770.

JACK, W., AND R. LAGUNOFF (2006): “Dynamic enfranchisement,” *Journal of Public Economics*, 90(4), 551–572.

JOVANOVIC, B. (1982): “Selection and the Evolution of Industry,” *Econometrica*, 50(3), 649–670.

——— (2014): “Misallocation and Growth,” *American Economic Review*, 104(4), 1149–1171.

KREMER, M., AND E. MASKIN (1996): “Wage Inequality and Segregation by Skill,” Working paper, NBER No. 5718.

KURLAT, P. (2016): “Asset Markets With Heterogeneous Information,” *Econometrica*, 84(1), 33–85.

LAZEAR, E. (2000): “Performance Pay and Productivity,” *The American Economic Review*, 90(5), 1346–1361.

LEVIN, J., AND S. TADELIS (2005): “Profit Sharing and the Role of Professional Partnerships,” *Quarterly Journal of Economics*, pp. 131–171.

Liu, Q., G. J. Mailath, A. Postlewaite, and L. Samuelson (2014): “Stable Matching with Incomplete Information,” *Econometrica*, 82(2), 541–587.

Lizzeri, A., and N. Persico (2004): “Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain’s ‘Age of Reform’,” *Quarterly Journal of Economics*, 119(1), 707–765.

Lockwood, B. (1991): “Information Externalities in the Labour Market and the Duration of Unemployment,” *Review of Economic Studies*, 58(4), 733–753.

Lucas, R. E. (1978): “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, 9(2), 508–523.

Lucas, R. E., and E. C. Prescott (1971): “Investment under Uncertainty,” *Econometrica*, 39(5), 659–681.

McCord, P. (2014): “How Netflix Reinvented HR,” *Harvard Business Review*, 2014(1), 71–76.

Oyer, P., and S. Schaefer (2011): “Personnel Economics: Hiring and Incentives,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4b, pp. 1769–1823. North Holland.

Roberts, K. (2015): “Dynamic Voting in Clubs,” *Research in Economics*, 69(3), 320–335.

Shapley, L. S., and M. Shubik (1971): “The Assignment Game I: The Core,” *International Journal of Game Theory*, 1(1), 111–130.

Sobel, J. (2001): “On the Dynamics of Standards,” *RAND Journal of Economics*, pp. 606–623.

Stiglitz, J. E. (1975): “The Theory of ‘Screening,’ Education, and the Distribution of Income,” *American Economic Review*, 65(3), 283–300.

Waldinger, F. (2012): “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany,” *Review of Economic Studies*, 79(2), 838–861.

——— (2016): “Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge,” *Review of Economics and Statistics*, (forthcoming).