Recruiting Talent*

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

This Version: August 2019

Abstract

We propose a new model of recruiting in which firms compete in wages to attract a high-quality applicant pool; managers then screen the applicants to identify the best talent. The equilibrium exhibits dispersion in wages and productivity: firms with superior screening skills post higher wages, attract better applicants, and recruit more talented workers. This equilibrium leads to an inefficient selection of talent into the industry, and can be improved by policies that reduce wage dispersion. We apply the model to understand the impact of public information (e.g. criminal convictions) and private screening (e.g. job tests) on employment, mismatch and wage dispersion. We also show that firms endogenously develop heterogeneous screening skills when talented workers are better at screening (e.g. via superior referrals).

1 Introduction

The success of most firms is built upon hundreds of individuals who take thousands of decisions, making it critical to identify and recruit the best talent. For example, the Netflix human resource manual 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). Similarly, Google’s head of human resources writes “Hiring is the single most important people activity in any organization. [...] Our greatest single constraint on growth has always, always been our ability to find great people” (Bock, 2015). Within economics, there are many papers that measure the variance of performance, ranging from top executives (Bertrand and Schoar (2003), Bennedsen et al.

*We are particularly grateful to Steve Tadelis for a conversation that inspired this project. We have received useful comments from Jim Albricht, Stepan Aleksenko, Alp Atakan, Roland Benabou, Dirk Bergemann, V Bhaskar, Bob Gibbons, Marina Halac, Johannes Hörner, Peter Kondor, Alessandro Lizzeri, Guido Menzio, Benny Moldovanu, Stephen Morris, Volker Nocke, Kevin Roberts, Larry Samuelson, Susan Vroman, Leeat Yariv, and Diego Zúñiga. We thank seminar audiences at Berlin, Bonn, Cambridge, Edinburgh, ESWM 2017, Games 2016, Gerzensee 2016, ITAM, Leicester, Mannheim, Miami, MIT, NBER Firms 2016, Northwestern/Duke/MIT IO conference 2016, Oxford, Penn, Princeton, Queen Mary, UBC, UCL, UCLA, and Yale. Keywords: adverse selection, assortative matching, firm dynamics, labor market competition, productivity dispersion, recruitment, referrals, wage dispersion. JEL codes: D21, D82, M51.

†UCLA, http://www.econ.ucla.edu/sboard/
‡UCLA, http://www.econ.ucla.edu/mtv/
§UCLA, http://economics.ucla.edu/person/tomasz-sadzik/
to mid-level managers (Lazear et al. (2015), Fenizia (2019)) to blue-collar workers (Lazear, 2000, Benson et al. (2019)). For example, Benson et al. (2020) study 200 business-to-business sales firms and show that, on average, the 75th percentile manager is 58% more productive than the 25th percentile manager, while the 75th percentile salesperson is over 100% percent more productive than the 25th percentile salesperson.

The market for talent is plagued by imperfect information. Unobservable measures of quality, like the AFQT test score, take time to filter into wages (Farber and Gibbons (1996), Altonji and Pierret (2001)). And changes in the employer’s information, in the form of criminal convictions or credit checks, have an observable impact on employment and wages (e.g. Agan and Starr (2017), Bartik and Nelson (2018)). As a result, firms carefully screen applicants, conducting interviews and obtaining referrals. For example, in Behrenz’s (2001) survey, employers report that their most important source of information when hiring are methods of “private screening” in the form of interviews (41%) and personal contacts (25%), as compared to “public information” like references from past employers (21%), references from schools (5%) and the application (3%).

This paper proposes a parsimonious labor market model in which firms screen workers using heterogeneous, private signals (e.g. interviews and referrals). The paper has three main contributions. First, we show that competitive equilibrium exhibits endogenous dispersion in wages and productivity across firms, and derive novel predictions for how public and private information affect wages, employment and firm-worker mismatch. Second, when firms differ in their screening ability, we show that firms with superior recruiting skills post higher wages, and prove that this positive assortative matching is socially inefficient. Third, we show that heterogeneous recruiting skills arise endogenously in a dynamic version of the model in which talented workers are also skilled at recruiting. Thus, talent provides firms a sustainable competitive advantage.

In Section 2, we introduce a static model of labor market competition that is consistent with many features of labor markets (e.g. private information, use of referrals, high expenditure on screening). A continuum of identical firms competes for a continuum of workers who have high or low ability. Talented workers have positive value added, while untalented workers would be better employed outside the industry. Firms attract applicants by posting wages, and then receive independent noisy signals about each applicant. We suppose the market is frictionless and clears from the top: The highest-paying firm attracts all applicants; other firms hire from the remaining, adversely selected pool of workers. The applicant pool quality thus endogenously declines with the wage rank, giving rise to a continuous equilibrium distribution of wages and productivity. We use this benchmark model to show that improved public information raises employment if the information is “conclusive” (e.g. felony records), but lowers employment if the information is “inconclusive” (e.g. credit scores); the latter result contrasts with the standard workhorse model of screening discrimination in which all information is public (Phelps (1972), Simon and Warner (1992)). We also show that an increase in private information (e.g. job tests, employee referral networks) allows high-wage employers to extract more talent from the labor pool, raising the dispersion of wages and productivity, and the segregation of talented workers across firms, helping to explain trends documented by Barth et al. (2016).

In Section 3, we suppose that firms differ in their skill at screening applicants because of differences in their referral networks or decision making ability. We show that firm-worker matching is positive assortative. That is, firms with more skilled recruiters post higher wages, attract better
applicants, and hire more talented recruits. Intuitively, skilled firms have a comparative advantage in hiring from a high-wage applicant pool with a balance of talented workers, rather than hiring from a low-wage pool in which few talented applicants remain. Unlike classic matching models (e.g. Shapley and Shubik (1971), Becker (1973)), the positive assortative matching seen in equilibrium is inefficient. Intuitively, high-wage firms screen applicants first and exert a negative *compositional externality* on low-wage firms by extracting talent from the applicant pool. Positive assortative matching maximizes this externality since the high-wage firms are skilled at extracting talent; this outweighs the private gain from positive assortative matching. Indeed, we show that negative assortative matching, whereby low-skill firms offer high wages and screen first, minimizes the externality and optimally selects talent into the industry. Thus, welfare is increased by policies that reduce wage inequality, lowering the dispersion of talent and the segregation of workers across firms.

In Section 4 we introduce a dynamic version of our model and show that differences in firms' recruiting skills endogenously arise if talented workers are better at recruiting. In the model, the talent of firms evolves as workers retire and today’s recruits become tomorrow’s recruiters. In the unique equilibrium, talented firms post high wages, attract the best applicants and hire talented recruits, reinforcing their initial advantage. If all firms start off with similar talent, the better endowed accumulate talent over time, while the worse endowed hire from poor, deteriorating applicant pools and lose talent. Hence small initial differences are magnified and dispersion arises endogenously. Steady-state dispersion balances two countervailing forces: positive assortative matching which amplifies differences across firms, and imperfect screening which leads to mean reversion and equalizes firms. Talent 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.

1.1 Literature

The static model of sequential screening is closely related to Montgomery (1991) and Kurlat (2016). In Montgomery’s classic model of referrals, talented employees may refer talented applicants, giving rise to heterogeneous recruiting skills. In contrast to our firms who employ a single firm-wide wage policy, Montgomery’s firms compete for referred and non-referred workers in separate markets with separate wages. Thus, our model generates across-firm wage dispersion, whereas his symmetric mixed strategy equilibrium generates within-firm wage dispersion. This difference matters empirically. For example, Kramarz and Skans (2014) find referred workers are paid below-average wages within a firm, but work at firms that pay above-average wages.¹ Our paper also addresses theoretical questions (the role of information, sorting and efficiency) that Montgomery does not address.

Kurlat (2016) studies financial markets with adverse selection in which buyers receive heterogeneous signals about sellers’ assets. Our model assumes that firms’ signals are independent, and shows that firms post different wages, matching is positive assortative and inefficient. In comparison, Kurlat assumes that buyers’ signals are nested, meaning that a more informed buyer knows everything that a less informed buyer knows. He shows that buyers post the same price with ties

¹More generally, the evidence of the within-firm effect of referrals on wages is mixed (Åslund et al. (2014), Burks et al. (2015), Brown et al. (2016), Dustmann et al. (2016), Henskiv and Skans (2016)), while the overall effect of peers on earnings seems strong (Marmaros and Sacerdote (2002), Edin et al. (2003), Zimmerman (2019)).
broken in favor of the less-informed buyers, and equilibrium is efficient. Intuitively, when signals are nested, high-skill firms can screen out all applicants who failed tests of low-skill firms and so do not mind hiring last. In Appendix A we provide a heuristic explanation of how these differences arise.\(^2\)

Our homogeneous-firm model complements the wider literature on wage dispersion. In Albrecht and Vroman (1992) and Burdett and Mortensen (1998), dispersion derives from firms competing for more workers in an economy with search frictions, whereas our dispersion derives from firms competing for better workers in an economy with adverse selection.\(^3\) In many labor markets, the pertinent search friction is in evaluating the quality of applicants, rather than finding them in the first place. Van Ours and Ridder (1992) find that “76% of all vacancies are filled by applicants who arrived during an application period that lasts for about 2 weeks”, leading them to write that “vacancy durations should be interpreted as selection periods and not as search periods for applicants.”\(^4\) This view is consistent with the extensive evidence of adverse selection in the labor market (e.g. Farber and Gibbons (1996)), the widespread use of referrals (e.g. Holzer (1987)), and the significant amount firms spend on screening candidates (e.g. Barron et al. (1985)).

Our heterogeneous-firm model contributes to the literature on firm-worker matching. Becker (1973) observed that if firms and workers are heterogeneous and complementary, then more productive firms hire more talented workers. In a dynamic model, Anderson and Smith (2010) and Anderson (2015) suppose agents match each period and evolve as a function of the match; they show that equilibrium is efficient, and derive sufficient conditions for matching to be positive assortative.\(^5\) In contrast to this literature that focuses on complementary production, our paper focuses on asymmetric information in the hiring process, and shows that complementarities arise endogenously, with more skilled firms posting higher wages. We examine how the resulting wage and productivity distribution changes in response to private screening and public information. We also prove that, unlike Becker, the equilibrium is inefficient even if we allow for production complementarities (see Section 3.3).

While productive complementarities are certainly important, there are settings in which selection plays a bigger role. Gupta (2017) finds that productive managers in insurance companies hire more productive recruits, who remain highly productive even after they switch teams. And Waldinger (2012, 2016) finds that the loss of star faculty at universities in Nazi Germany led to a permanent reduction in the quality of hires, but did not affect the productivity of current faculty.

We also provide a theory of firm dynamics in which a firm’s stock of talent is its key strategic

\(^2\)Other labor market models that feature sequential screening include Lockwood (1991), Ely and Siegel (2013) and Kurlat and Scheuer (2019).


\(^4\)Similarly, Manning (2000) finds that “Employers are able to fill the vast majority of their vacancies suggesting that the number of applicants plays a role only in widening the choice (and hence quality) of applicants.” Relatedly, Kroft and Pope (2014) find that Craigslist lowers housing vacancy rates but not unemployment, concluding that “search frictions might be much more important in the apartment and house rental market than in the labor market.”

\(^5\)See also Prescott and Boyd (1987) and Jovanovic (2014). There is also a literature on matching with incomplete information and interdependent values, e.g. Chakraborty, Citanna, and Ostrovsky (2010) and Liu et al. (2014).
asset. This is most closely related to Montgomery (1991) who argued that today’s talent can be used to acquire talent tomorrow via referrals. However, as discussed in Section 4.6, the models have an important difference: our model generates persistent heterogeneity in talent, whereas in the fully dynamic version of Montgomery’s symmetric equilibrium, talent regresses to the mean and is thus not a sustainable competitive advantage.

There is a broader set of papers on firm dynamics in which competitive advantages stem from technology (Lucas and Prescott (1971), Hopenhayn (1992)), reputation (Jovanovic, 1982) or the stock of labor (Hopenhayn and Rogerson, 1993). By focusing on talent and recruiting, our paper provides a new channel through which firms can sustain a competitive advantage that is particularly relevant for industries ranging from technology to sales. It also gives rise to predictions concerning the dispersion and persistence of productivity, wages and employee quality.

Finally, our paper echoes some themes 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 a government ages, its talent pool depletes and 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 positive feedback loop.6

## 2 A Static Model of Sequential Screening

We first describe our benchmark model of sequential screening. A unit mass of identical firms, each with one vacancy, competes for a unit mass of workers. Workers differ in their talent \( \theta \), with proportion \( q \in (0, 1) \) talented, \( \theta = H = 1 \), and the remainder untalented, \( \theta = L = 0 \). Firms select among applicants by administering a pass/fail test to each applicant. Talented workers always pass the test, whereas untalented workers are screened out with probability \( p \in (0, 1) \).

The labor market is anonymous and perfectly competitive. Firms simultaneously post wages and offer their job to any worker who passes their test. Workers only care about wages, and so accept the highest wage offer they receive. To operationalize this, order firms in terms of the wages they post, and suppose workers then apply to firms from highest to lowest wage. The highest firm screens the anonymous workers in a random order and hires the first who passes their test; the adversely selected remainder then apply to the “second” firm, and so on until all firms and workers are matched. 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.

When a recruiter screens an applicant pool with expected talent \( q \), proportion \( 1 - (1 - q)p \) of the applicants pass the test. Bayes’ rule implies that the fraction of recruits who are talented equals

\[
\lambda(q, p) := \frac{q}{1 - (1 - q)p}.
\]

Clearly, expected talent \( \lambda(q, p) \) increases in both the applicant quality \( q \) and the recruiting/screening

6The key assumption of our dynamic model – that hiring a worker means hiring her judgment – also relates to the literature on dynamic clubs, where today’s members must decide who will make decisions tomorrow. Acemoglu and Robinson (2000), Lizzieri and Persico (2004) and Jack and Lagunoff (2006) consider the extension of the voting franchise, while Sobel (2001), Barberà et al. (2001) and Roberts (2015) consider general club goods.
skills $p$.

Payoffs are as follows. Workers only care about wages, and so accept any job paying more than their outside option, $w \geq 0$. Productivity is 1 for talented workers and 0 for untalented workers. Thus, when a firm posts wage $w$ and attracts applicants $q$, its expected profits are

$$
\pi := \lambda(q, p) - w.
$$

We solve for Nash equilibrium in wages.

**Remarks.** The assumption of frictionless market-clearing “from the top” is standard in the literature on markets with adverse selection (Kurlat (2016), Kurlat and Scheuer (2019)) and goes back at least to Wilson (1980). There are many equivalent ways to model matching in the labor market. For example, one could have all firms evaluate all workers and then have firms pick workers (who passed their test) in order of decreasing wages.

The assumption that firms use binary, pass-fail tests and that talented workers pass the test with certainty is without loss. To see this, assume a more general information structure with finitely many signals $s$ that arise with probability $p^S(s)$. Firms then hire the first applicant with the signal $\tilde{s}$ that maximizes the odds-ratio $\tilde{\ell} = p^H(\tilde{s})/p^L(\tilde{s})$. Recruit quality is then $q = \frac{\tilde{q}}{q^H(\tilde{s})/q^L(\tilde{s})}$ which collapses to (1) when $\tilde{\ell} = p^H(\tilde{s})/p^L(\tilde{s}) = 1/(1 - p)$ for $\tilde{s} = \text{“pass”}$.

We can re-interpret firms’ screening as referrals. Assume that each firm is connected to each talented (resp. untalented) worker via a referral with probability $1 - p^H$ (resp. $1 - p^L$). Firms extend wage offers to their referrals, and referred workers accept their best offer. Re-interpreted in this way, our model differs from Montgomery (1991)’s model of referrals in that (i) wages are set at the firm level rather than conditioned on the talent of the referring employee, and (ii) each firm receives a continuum of referrals and is hence guaranteed to hire a referred worker.

If $w = 0$, then all workers are employed and the game is constant sum. Assuming $w > 0$ introduces a welfare margin, in that talented workers should be employed in the industry, and untalented workers should take the outside option. One can also interpret $w$ as an operating cost for the firm; under this formulation, allocations and payoffs are the same while wages are shifted down by $w$.

The model assumes that it is free to screen a worker. In reality, reading applications, conducting interviews and checking references takes time. Barron et al. (1985) report that the average firm spends about 8 hours screening applicants, with more educated roles requiring more screening. Indeed, (Bock, 2015, p. 76) writes that Google spends 150-500 hours interviewing for each role. Our analysis can easily accommodate such screening costs. In particular, if there is a cost $\kappa$ to screen each applicant and firms screen applicants until one passes their test profit (2) becomes

$$
\pi := \lambda(q, p) - \kappa/(1 - p(1 - q)) - w.
$$

We study this extension in Section 3.3.

### 2.1 Equilibrium

To characterize equilibrium, we first study how the quality of the applicant pool depends on the firm’s rank in the wage distribution. Suppose all firms post different wages, write $x \in [0, 1]$ for the resulting wage quantiles, and $Q(x)$ for applicant quality at wage quantile $x$.

The highest ranked firm
faces applicant pool $Q(1) = \bar{q}$; thus, proportion $\lambda(Q(1), p)$ of its recruits are talented. Since firms select talented workers disproportionately, lower-ranked firms face an adversely selected applicant pool; formally, $Q(x)$ 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 $\lambda(Q(x), p)dx$; hence $d[xQ(x)] = \lambda(Q(x), p)dx$.

Rearranging, the talent pool evolves according to the sequential screening equation

$$Q_x(x) = \frac{\lambda(Q(x), p) - Q(x)}{x}. \quad (3)$$

Since screening is imperfect, some talent remains, $Q(x) > 0$, for all $x > 0$. However, at the bottom, firms pick over the applicants so many times that in the limit no talent remains, $Q(0) = 0.8$

In equilibrium, wages are distributed continuously. To see this, observe that an atom of firms offered the same wage, then – recalling our tie-break rule – a firm could attract discretely better job applicants with a marginal wage raise. Similarly, if there was a gap $[w, w']$ in the wage distribution, then the firm offering $w'$ could attract the same applicants at the lower wage $w$. As a result, the lowest wage in the wage distribution equals the outside option $\underline{w}$.

Turning to equilibrium payoffs, the identical firms compete away all profits. Wages thus coincide with expected productivity, $w(x) = \lambda(Q(x), p)$, where applicant quality $Q(x)$ evolves according to (3). The worst applicant quality $q$ that firms are willing to consider is given by $\lambda(q, p) = \underline{w}$. The fraction of unemployed workers $\overline{x}$ is then implicitly given by $Q(\overline{x}) = q$. Hence, there is unemployment, $\overline{x} > 0$, if and only if $q > 0$, which is the case if and only if $\underline{w} > 0$. To summarize:

**Theorem 1.** Equilibrium wages and productivity are continuously distributed with a minimum of $\underline{w}$. The distribution $w(x)$ is uniquely determined by Bayes’ rule (1) and sequential screening (3).9

### 2.2 Implications for Labor Market Outcomes

Our model is useful for evaluating the impact of information on labor market outcomes such as mismatch, unemployment, and wage and productivity dispersion. Several recent empirical papers study the impact of public information pertaining to felony convictions (Agan and Starr (2017), Doleac and Hansen (2019)) or credit scores (Bartik and Nelson (2018), Ballance et al. (2017)). Such public information segments the labor market and corresponds to a mean-preserving spread of the prior $\tilde{q}$ in our model. Other papers consider the impact of private information such as improved job tests (Autor and Scarborough (2008), Hoffman et al. (2017)) and social network data (Munshi (2003), Beaman (2011)). In our model, this corresponds to a comparative static in the screening skill $p$.

First, consider employment $1 - \overline{x}$. The effect of public information depends on its precision. To see this, observe that expected quality of unemployed and employed workers averages out to ex-ante does not arise in equilibrium.

8To see $Q(x) > 0$ for $x > 0$, note that $(\log Q(x))_x = (\lambda(Q(x), p)/Q(x) - 1)/x = (1/(1 - (1 - Q(x))p) - 1)/x$ is bounded for any fixed $x > 0$ since $p < 1$. On the other hand, $(\log Q(x))_x$ is of order $1/x$, and so the integral $\log Q(x) = \int_x^1 (\log Q)_x$ diverges as $x \to 0$, implying $\log Q(0) = -\infty$, or $Q(0) = 0$.

9Clearly, equilibrium is not unique since the wage distribution is consistent both with firms mixing symmetrically and with firms playing pure strategies; with heterogeneous firms, in Theorem 2, only the latter is possible.
quality
\[ xE[\theta | \text{unemployed}] + (1 - x)E[\theta | \text{employed}] = \bar{q}. \] (4)

Public information introduces a mean-preserving spread in the belief about applicant quality, say, \( \bar{q} \rightarrow \{q', q''\} \). First, consider imprecise information that can be overcome by a private signal, \( \lambda(q', p) > w; \) this might be the case for credit checks or grades at elite business schools. Proposition 1 in Appendix B.2 shows that such information decreases employment. Intuitively, the quality of unemployed workers is fixed at \( \bar{q} \) by free-entry, while the dispersion in \( \bar{q} \) raises the average quality of the employed, so equation (4) implies that employment must fall. In contrast, precise information, such as felony convictions or drug tests, can raise employment. For example, information that spreads the prior such that \( \bar{q} \rightarrow \{0, q''\} \) raises employment as long as \( \lambda(\bar{q}, p) \) is close to \( w \), so there is little employment without the test. Thus, well-intentioned “ban-the-box” legislation might have the perverse effect that certain demographics are entirely shut out from the labor market. These results can help reconcile the fact that “conclusive information” like drug testing (Wozniak, 2015), criminal record checks (Agan and Starr (2017), Doleac and Hansen (2019)), and employee certification (Pallais (2014), Stanton and Thomas (2015)) seem to raise employment and wages, while the effect of “inconclusive information” like credit checks is less clear cut (Bartik and Nelson (2018), Ballance et al. (2017)). This rich set of results has no counter-part in the classic screening models with public and normal signals (e.g. Phelps (1972), Pallais (2014)), where information raises employment as long as \( \bar{q} < w \).10

Second, consider mismatch. Since screening is imperfect, some talented workers remain unemployed, and some untalented workers are employed. Formally, the loss of welfare relative to the first-best allocation is given by:

\[
M = \frac{qx(1 - w)}{\text{Unemployed talented}} + \frac{(1 - \bar{q} - (1 - q)x)w}{\text{Employed untalented}}
\]

This mismatch is empirically important: At a micro level, Fredriksson et al. (2018) show that mismatched workers, as judged by a pre-employment test, have lower earnings growth and higher separation rate. At a macro level, Hsieh et al. (2019) estimate that 20-40% of US economic growth over the last 50 years has come from a better allocation of talent to jobs.

Proposition 2 in Appendix B.3 shows that both public and private information decrease mismatch, as firms are able to make better decisions, raising welfare and wages. In order to compare the relative effect of public and private information, compare (i) our screening model in which each firm independently screens out untalented workers with probability \( p \), and (ii) the “Phelps model” in which all firms receive the same public signal that screens out bad workers with probability \( p \).11 Proposition 3 in Appendix B.3 shows that mismatch in the screening model exceeds the mismatch in the Phelps model. Intuitively, failing a public test black-lists an untalented job applicant, while failing a private test allows him to re-enter the pool and worsens the adverse selection for other firms. More formally, we show that firms’ posteriors in the Phelps model are a mean-preserving

10The effect of private information \( p \) on employment is generally ambiguous. On the upside, better screening at the marginal firm \( x \) raises employment; but on the downside, better screening by inframarginal firms \( x > x \) exacerbate adverse selection and lower employment. See Appendix B.2 for details.
11Phelps (1972)’s original model differs by assuming normal quality and signals.
spread of their posteriors in our private screening model, and thus more informative in the Blackwell order.

Proposition 4 in Appendix B.3 reinforces this finding by showing that, in the screening model, the marginal social return to screening skills at $p = 0$ is zero, $M_p(0) = 0$. This result contrasts with the classic “wisdom of crowds” results (e.g. Pesendorfer and Swinkels (1997)): A continuum of imprecise, independent signals is perfectly informative when aggregated by Bayes’ rule, but is perfectly uninformative when aggregated by our firms in equilibrium.

Third, Theorem 1 can help explain productivity and wage dispersion. On the productivity side, Foster et al. (2008) find that the standard deviation of productivity is 20 log-points within homogeneous 7-digit industries, implying that a 90th percentile firm is 67% more productive than a 10th percentile firm. On the wage side, Katz and Murphy (1992) show that residual wage inequality (after controlling for observables) is even larger, with the 90th percentile worker paid more than 100% more than a 10th percentile worker. The rent sharing literature finds positive correlation between the two, in that high productivity firms also pay high wages. Card et al. (2018) survey this literature, arguing that much of the “rent sharing” comes from the fact that productive firms tend to have productive workers, as seen in our model.

Proposition 5(a) in Appendix B.4 shows that raising screening skills $p$ increases wage and productivity dispersion (in the log-dispersive order). Intuitively, the superior screening skills of the top-wage firms raise their productivity relative to the mean. These top firms extract more talented workers from the applicant pool, lowering the productivity and wages of lower-wage firms (even though their screening skill $p$ increased by the same amount). This finding may help explain the increased dispersion of productivity and wages over the course of the information revolution of the last forty years (Barth et al., (2016)). Consistent with our model, the growth in dispersion seems to be mostly driven by young workers, about whom there is most uncertainty (Smith, 2018).

Finally, our model also speaks to the “AKM-decomposition” of wage dispersion into worker- and firm-effects (Abowd et al., 1999). To capture such a decomposition in our model, consider a dynamic extension wherein (i) workers randomly quit and re-enter the market, and (ii) obtain proportion $b$ of their output via a bonus payment, so the total compensation for worker $\theta \in \{0, 1\}$ equals $w(x) + b\theta$. An AKM wage regression would interpret $w(x)$ as the firm-effect (even though all firms have identical fundamentals) and $b\theta$ as the worker-effect. The model also generates correlation between the two, consistent with Card et al. (2013, 2018). Furthermore, an increase in screening skills increases segregation, thereby increasing the dispersion of firm-effects, worker-effects and their correlation, as seen in Card et al. (2013) and Song et al. (2019).

Proposition 5(b) in Appendix B.4 shows that a reduction in the number of talented worker, $\tilde{q}$, perhaps driven by skill-biased technical change, also leads to an increase in the dispersion of productivity and wages. Intuitively, if the number of talented workers halves, the top firm’s talent drops by less than half, as they are able to screen out many of the untalented workers, leaving proportionally less talent for lower firms, thereby raising inequality.
3 Heterogeneous Screening Skills

We now enrich the model by introducing heterogeneous screening skills $p$ across firms. Indeed, Bloom and Van Reenen (2007) show that firms differ in their ability to manage human capital, and that the skill is positively correlated with firm productivity.

Superior screening skills may come from having better managers. Bender et al. (2018) show that controlling for human capital of workers (and especially, managers) reduces the association between productivity and management policies by 30-50%. Similarly, Bianchi and Giorcelli (2019) show that an HR training program had a large impact on productivity, and was more effective when targeted at middle-managers rather than top-managers. Alternatively, screening skills may result from better institutions and technology. For example, Hoffman et al. (2017) show that a new screening test raised workers’ tenure by 15%. Screening skills can also arise from better referral networks. For example, Google obtains more than half its workers via referrals and, for other applicants, obtains “back door” references via current employees (Bock, 2015, p. 80).

Formally, suppose firms’ recruiting skills $p$ are distributed according to some continuous, strictly increasing cdf $F$ on $[p, \bar{p}]$, where $0 < p < \bar{p} < 1$, allowing us to identify each firm with its screening skill $p$. For a given wage profile, denote the skill of the firm with wage-rank $x$ by $P(x)$. Applicant quality $Q(x)$ evolves according to the sequential screening equation (3), where we replace $p$ by $P(x)$. As in Theorem 1, equilibrium wages are distributed continuously with a minimum of $w$. We then ask: Which firms post higher wages? How are profits and wages determined in equilibrium?

3.1 Positive Assortative Matching

We say there is positive assortative matching (PAM) between firms and applicants if $P(x)$ is increasing, meaning high-$p$ firms are matched with high-$q$ applicant pools. From Becker (1973), we know that equilibrium features PAM if skilled recruiters have a comparative advantage in screening applicants with higher expected talent, i.e. if expected recruit quality $\lambda(q,p)$ is supermodular.

The rest of the paper focuses on the case of scarce talent. This assumption is motivated by industries with relatively few highly productive individuals, such as technology companies or sales. For example, Bock (2015) writes that “Only 10% of your applicants (at best) will be top performers.”

Assumption. Talent is scarce,

$$\lambda(\bar{q}, \bar{p}) \leq 1/2$$

This is a joint condition on the talent distribution and screening skills, and states that even a worker from the unselected pool who passes the test of the most skilled firm is more likely untalented than talented.

Theorem 2. Equilibrium exists and is unique. Firms with skill exceeding a cutoff $p$ enter; matching is then positive assortative.

---

13 This is evaluated by asking questions such as “Do senior managers discuss attracting and developing talented people?” and “Do senior managers get any rewards for bringing in and keeping talented people in the company?”

14 In a previous version of this paper, Board et al. (2017), we characterize equilibrium matching when talent is abundant using a slightly different parametrization of firms’ screening skills. In contrast to the increasing wage function $w(p)$ seen in Theorem 2, equilibrium wages $w(p)$ are hump-shaped.
Proof. Note the partial derivatives
\[
\lambda_p(q, p) = \frac{q(1 - q)}{(1 - p(1 - q))^2} \quad \text{and} \quad \lambda_{qp}(q, p) = \frac{(1 - p(1 - q)) - 2q}{(1 - p(1 - q))^3}.
\] (6)

Since \( \lambda_p \geq 0 \), there exists a \( p \in [\bar{p}, \bar{p}] \) such that firms \( p \geq \bar{p} \) enter, and firms \( p < \bar{p} \) do not. Condition (5) then implies that \( 1 - p(1 - q) \geq 2q \), so \( \lambda_{qp} > 0 \) for all \( q \leq \bar{q}, p \leq \bar{p} \), and matching is positive assortative. That is, higher skill firms post higher wages, and so \( P(x) \) increases.

Intuitively, recruiting skills do not matter if all applicants are either talented or untalented, but they do matter when applicant quality is intermediate. Since \( Q(x) \) is bounded above by \( \bar{q} \), our scarce-talent assumption (5) implies that skilled firms have a comparative advantage at screening better applicants.

We can now construct the equilibrium. Given PAM, a firm’s rank in the skill distribution equals its equilibrium wage rank, \( F(p) = x \), and we can identify a firm by this rank \( x \). The skills of the firm with wage-rank \( x \) is then given by \( P(x) = F^{-1}(x) \), its applicant quality \( Q(x) \) is determined by the sequential screening equation (3), and the recruit quality \( \lambda(Q(x), P(x)) \) by Bayes’ rule, (1).

From this we can derive the entry threshold \( p \), wages, and profits. Denote the equilibrium wage required to attract applicants of quality \( q \) by \( W_q(q) \). The marginal firm pays the outside option, so employment \( x \) is given by \( \lambda(Q(x), P(x)) = w \), which determines the entry threshold via \( F(p) = x \).

Wages are determined by firm \( x \)'s first-order condition,
\[
W_q(Q(x)) = \lambda_q(Q(x), P(x)).
\] (7)

Finally, profits \( \Pi(p) \) follow by the envelope condition
\[
\Pi_p(P(x)) = \lambda_p(Q(x), P(x)).
\] (8)

Intuitively, productivity \( \lambda(q, p) \) depends on both the applicant pool quality and the screening ability; workers capture the marginal benefit of the former and firms capture the marginal benefit of the latter.

The so constructed wage profile is the only possible candidate for an equilibrium. To verify that the wages are indeed optimal, it suffices to note that marginal profits \( \lambda_q(q, p) - W_q(q) \) single-cross in \( p \) and matching is positive assortative; hence the FOC (7) implies global optimality.

Theorem 2 captures a natural complementarity in the recruiting function \( \lambda \), whereby recruiting skills \( p \) are relatively more valuable at intermediate applicant pools \( q \) with high uncertainty about quality, than at low values of \( q \) with low uncertainty. Thus, skilled firms pay high wages, attract high-quality applicants, recruit talented workers, and achieve high productivity and profits. It also shows that heterogeneous screening skills raise both the dispersion of productivity and the level of profits in the economy.

3.2 Equilibrium Inefficiency

Equilibria in matching models with transferable utility are typically efficient (e.g. Shapley and Shubik (1971), Becker (1973)). Surprisingly, this welfare theorem fails in our model. In particular, Theorem 3 shows that mismatch is maximized by positive assortative matching and minimized by
negative assortative matching (NAM). The key difference to standard matching models is the compositional externality discussed in the introduction. The quality of applicant pools is endogenous, and depends on the screening skill of higher-paying firms. Intuitively, high-skilled firms pick out more talented workers than low-skilled firms and introduce more adverse selection. This externality is maximized by PAM and minimized by NAM.\(^\text{15}\)

For a simple example, suppose firms are either skilled, \( p > 0 \), or unskilled, \( p = 0 \); to abstract from entry, assume firms are on the short side of the market and the minimum wage \( w \) does not bind, so all firms enter. Under PAM, first the skilled firms hire from the unselected pool \( q \); unskilled firms then hire from an adversely selected pool with quality \( q < \bar{q} \). In contrast, under NAM, unskilled firms hire from the unselected pool \( q \). Crucially, they do not change the quality of the talent pool, and so the skilled firms also hire from a pool of quality \( \bar{q} \). Since the unskilled firm impose no compositional externality on the skilled firms, more talented workers are hired under NAM than under PAM, raising social surplus.\(^\text{16}\)

To argue the inefficiency of the competitive equilibrium more generally, consider a planner who can direct all firms’ entry and wage decisions but is subject to the same informational frictions.\(^\text{17}\) Her problem is to choose employment \( 1 - x \) and a matching function, or equivalently screening order, \( P : [x, 1] \rightarrow [p, \bar{p}] \) to minimize mismatch, or equivalently, maximize surplus

\[
\int_{x}^{1} [\lambda(Q(x), P(x)) - w] dx. \tag{9}
\]

For example, in equilibrium, matching is positive assortative \( P^{\text{PAM}}(x) = F^{-1}(x) \), with entry cutoff \( x \) satisfying \( \lambda(Q(x), P^{\text{PAM}}(x)) = w \); assuming the same entry behavior, NAM is characterized by \( P^{\text{NAM}}(x) = F^{-1}(x + 1 - x) \). In contrast to the standard assignment model, the surplus of the \( x \)-ranked firm in the integrand of (9) depends on \( P(x') \) for \( x' > x \) via the applicant quality \( Q(x) \).

We first argue that the planner wants the highest skilled firms to enter, and hence the matching function \( P \) is measure-preserving with range \( [p, \bar{p}] \), where \( F(p) = x \). Increasing skills \( P(x) \) at wage rank \( x \) increases the surplus at that rank, \( \lambda(Q(x), P(x)) \), but reduces applicant quality \( Q(x') \) at lower ranks \( \bar{x} \in [x, x] \). This negative indirect effect diminishes but does not overturn the positive direct effect. Formally, we show in the proof of Theorem 3 that the indirect effect scales down the

\(^{15}\)Compositional externalities are seen in other contexts. In models of directed search, they arise when wage offers affect the distribution of workers entering the labor market (Albrecht et al., 2010) or when workers’ job search creates “phantom vacancies” because employers fail to remove their filled vacancies from the market (Albrecht et al., 2017). In status games, consumption of a status good lowers the population rank of others and lowers their utility (Hopkins and Kornienko, 2004). In position auctions, the probability a customer looks at a low-ranked advertiser depends on the quality of higher-ranked advertisers (Athey and Ellison (2011)). And in school matching, the order in which seats are filled affects the students that other schools attract (Dur et al., 2018).

\(^{16}\)This argument is incomplete. In fact, skilled firms are worse off under NAM. While unskilled firms do not reduce pool quality, they do reduce the pool size \( x \). And while \( x \) does not directly affect skilled firms’ recruiting, it magnifies the adverse selection induced by high-wage skilled firms on low-wage skilled firms since the former extract the same amount of talent from a smaller pool, as captured by the denominator \( x \) in (3). Despite this complication, Theorem 3 implies that NAM minimizes mismatch.

\(^{17}\)In particular, the planner cannot communicate firms’ test results to each other. With a continuum of firms and independent tests, allowing such communication would trivially solve any mismatch.
direct effect by a factor
\[ \exp \left( - \int_{\hat{x}}^{x} \lambda_{q}(Q(\hat{x}), P(\hat{x})) \, d\hat{x} \right). \] (10)

Turning to the screening order \( P(\cdot) \), we claim that surplus (9) increases whenever a low-skill firm \( p \) is promoted from wage rank \( x \) to \( x' > x \) past firms with higher skills \( P(\hat{x}) > p \) for \( \hat{x} \in [x, x'] \). Thus PAM minimizes surplus, and NAM maximizes it. In a discrete analogue of our model, consider two firms with screening skills \( p \) and \( p' = p + dp \) at adjacent wage ranks \( x \) and \( x' = x + dx \) under PAM. The corresponding applicant quality equals \( q = q' - \phi(q', p', x')dx \) and \( q' \), where

\[ \phi(q, p, x) = (\lambda(q, p) - q)/x \]

is the compositional externality from the RHS of the sequential screening equation (3). How does switching their screening order affect their joint recruited talent and hence surplus? First, the low-skill firm \( p \) now screens the better pool; since \( \lambda(q, p) \) is supermodular, this lowers total surplus of the two firms by \( \lambda_{p}\phi_{p}dpdx \). In Figure 1 this is represented by the shift from white circles representing PAM to the black circles representing NAM. If the pool quality was exogenous as in Becker, this would be the end of the story. In our model, there is a second effect: firm \( p \) extracts \( \phi_{p}dpdx \) less talent, increasing surplus by \( \lambda_{q}\phi_{p}dpdx \). In Figure 1, this is represented by the applicant quality of firm \( p' \) rising from \( q' \) to \( q'' \). The next inequality shows that the second term outweighs the first, and so moving the low-skill firm \( p \) ahead of the high-skill firm \( p' \) raises surplus

\[ \frac{\lambda_{q}p}{\lambda_{q}} = 2 \frac{1 - q}{1 - p(1 - q)} - \frac{1}{1 - p} < \frac{1 - q}{1 - p(1 - q)} = \frac{\lambda_{p}}{\lambda} < \frac{\lambda_{p}}{\lambda - q} = \frac{\phi_{p}}{\phi}. \] (11)

In words, marginal recruiting success \( \lambda_{q} \) is less sensitive to \( p \) than absolute recruiting success \( \lambda \), which in turn is less sensitive to \( p \) than the compositional externality \( \phi \). Thus, total surplus of the two firms \( p, p' \) is maximized by NAM; the effect on aggregate surplus (9) must be discounted by (10), but the sign remains unchanged. The compositional externality overturns one of the most fundamental insights of the assignment model, namely the first welfare theorem.

**Theorem 3.** For any level of unemployment \( x > 0 \), PAM minimizes surplus (9) and NAM maximizes surplus.

*Proof.* See Appendix C.2.

Theorem 3 has practical implications. In some applications, the planner might be able to directly implement negative assortative matching. For example, in the NFL “reverse” draft the lowest-ranked football teams pick first. This lowers productivity dispersion and makes games more competitive; Theorem 3 suggests that the NFL draft also maximizes sorting of talent into the league.

If the planner cannot observe the skill of different firms, she can still improve welfare by restricting the set of admissible wages. In equilibrium, high-skill firms offer weakly higher wages than low-skill firms by Theorem 2. Thus, the planner’s optimal policy is a single wage, inducing firms to select in a random order. We argue in Appendix C.3 that the same outcome can be achieved by a wage cap. This argument suggests that the NCAA’s ban on paying athletes may raise talent in competitive college sports by preventing colleges with the best scouts bidding away the best athletes.
and lowering the quality of the marginal programs. Such a policy improves surplus in our model, but may also lower the (monetary) utility of college athletes.

3.3 Discussion of the Compositional Externality

In the competitive equilibrium, all firms face the same equilibrium wage schedule \( W(Q(x)) \) and hence ignore the compositional externality they impose on others, which depends on their screening skills \( p \). In contrast, Pigouvian wages that support the efficient, negative assortative matching would charge firm \( p = P(x) \) for its externality on all lower-wage firms. Proposition 6 in Appendix C.4 shows that in a two-firm version of our model, such Pigouvian taxes are implemented by a second-price auction. In a second-price auction both firms post wages, the high-wage firm screens unselected applicants and pays the low firm’s wage, and the low firm faces adverse selection and pays the workers’ outside option \( w \). As always in a second-price auction, equilibrium bids reflect bidders’ values. Specifically, the wage difference \( w - \hat{w} \) equals the benefit from hiring first and thereby avoiding adverse selection. Since the low-skilled firm faces more adverse selection, it actually bids more that the high-skilled firm, leading to negative assortative matching in equilibrium.\(^{18}\)

The compositional externality and the equilibrium inefficiency seen in Theorem 3, is robust to productive complementarities and screening costs. First, consider a Becker-style model in which a firm’s revenue is multiplicative in firm type and (expected) worker type \( p \cdot \lambda(q, p) \), so high-type firms also have a higher marginal product. As one would expect, Proposition 7(a) in Appendix C.5 shows equilibrium sorting remains positive assortative (Theorem 2). More surprisingly, Proposition 7(b)\(^{18}\)This result is analogous to other efficiency results in auctions with externalities (e.g. Gilbert and Newbery (1982)).
shows that the compositional externality overcomes both the exogenous Becker-complementarity and the endogenous screening-complementarity, and PAM remains inefficient. Second, consider the screening cost model mentioned in Section 2, in which profit (2) becomes $\pi := \lambda(q,p) - \kappa/(1 - p(1 - q)) - w$. Proposition 8(a) in Appendix C.6 shows that the cost term reinforces PAM since skilled firms are more selective and hence more sensitive to applicant quality. Thus, equilibrium is consistent with the finding that high-wage firms have more applications (Belot et al. (2018), Banfi and Villena-Roldán (2019)) and interview more applicants (Barron et al. (1985)). Moreover, Proposition 8(b) shows that the equilibrium inefficiency continues to hold.

Finally, we note that the compositional externality also leads firms to overinvest in their recruiting skills if firms can choose their recruiting skill $p$ at cost $c(p)$. This is most clearly seen when the game is constant sum, $w = 0$, so all investment is wasteful. More generally, firms that invest highly into screening skills also post high wages in equilibrium, and the social return of investment falls short of the private return $\lambda_p(Q(x),p)$ by the factor (10). This result contrasts with the finding of efficient investment found in classic matching models, e.g. Cole et al. (2001). Overinvestment is important in practice. Bock (2015, p. 60) urges companies to spend more on screening recruits, writing that “The presence of a huge training budget is not evidence that you’re investing in your people. It’s evidence that you failed to hire the right people to begin with. […] At Google, we front-load our people investment. This means the majority of our time and money spent on people is invested in attracting, assessing, and cultivating new hires.” Alas, not every firm can hire the best.

4 Talent as a Sustainable Competitive Advantage

We now embed our static labor market into a model of firm dynamics to study the evolution of talent over time. The key premise is that firms with more talented employees are more skilled at recruiting. For example, Gupta (2017) shows that more productive sales managers hire more productive salespeople, while Waldinger (2012, 2016) finds that the loss of star professors in Nazi Germany led to a permanent reduction in the quality of hires. This advantage may be because good employees provide better referrals or have better underlying cognitive skills.

Referrals are a crucial source of information for firms. More than a third of US jobs are filled through referrals (Holzer, 1987). Referrals are also higher quality than average applicants: In field experiments they have better performance, even after controlling for observables (Beaman and Magruder (2012), Pallais and Sands (2016)); in observational studies, they have a higher probability of being hired and better unobservable skills (i.e. AFQT scores), despite worse schooling (Burks et al. (2015), Brown et al. (2016), Hensvik and Skans (2016)).

Productive workers provide better referrals. Beaman and Magruder (2012)’s field experiment finds that productive workers refer more people and high-performing referrals, whereas unproductive workers’ referrals are no better than non-referrals. Similarly, Pallais and Sands (2016) write that “a referrer’s performance is a strong predictor of her referral’s performance”. In observational studies, Burks et al. (2015) find that in trucking, “there are large differences in profits between referrals from high-productivity referrers compared to low-productivity referrers”, while Hensvik and Skans (2016) write that “the cognitive abilities of linked entrants are positively associated with the abilities of
their linked incumbents” and that “firms with more productive employees rely more on social ties.”¹⁹

The second rationale for positive correlation of productivity and recruiting skills is the idea that it takes competence to assess competence, which underlies the famous Dunning-Krueger effect: “The skills you need to produce a right answer are exactly the skills you need to recognize what a right answer is.”²⁰ This is seen in Beaman and Magruder (2012), where “high-ability participants are able to predict their referrals’ ability [...] low-ability participants, on the other hand, are not systematically able to predict their referrals’ performance.” At a more systematic level, more talented employees may be able to design better technologies and institutions to better screen applicants.

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, amplifying their talent advantage. We characterize firm values, wages, and profits over time, and show that the economy converges to a steady state which exhibits persistent dispersion in talent, wages and productivity. In equilibrium, the positive assortative firm-applicant matching offsets the regression to mediocrity that results from imperfect screening. Importantly, steady-state talent differences are sustained by firms’ endogenous wage choices and thus arise for any degree of correlation between productivity and recruiting skills, however weak.

Our results are of particular interest because such persistent dispersion does not arise in the classic referrals model of Montgomery (1991). The model has been influential since it “allows for positive profits from referrals, [...] presents a natural rationale for endogenous skill segregation across firms, [and] shows how firms’ inability to observe worker ability ex-ante can generate wage inequality” (Hensvik and Skans, 2016). Ironically, the symmetric equilibrium in Montgomery’s model does not generate long-run differences in profits, skills or wage levels across firms, as we show in Section 4.6.

### 4.1 Model

Time \( t \geq 0 \) is continuous. There is a unit mass of firms, each with a unit mass 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 \( adt \) workers retire, leaving firms with vacancies. In the job market, there are then \( adt \) open jobs and \( adt \) applicants, of whom fraction \( \bar{q} \) are talented. Analogous to the static model, firms compete for these applicants by posting life-time wages \( w(t) \).

We assume that talented workers have an advantage in recruiting, either because of superior referral networks, or because of better judgment. We describe this relationship by a skill function \( \psi(r) \) satisfying \( \psi_r > 0 \) and \( \psi_{rr} \geq 0 \). For example, if an employee with talent \( \theta \in \{L, H\} \) has screening skill \( p^\theta \) and firms ask random employees to act as recruiters, possibly by providing a referral, then the skill function is linear, \( \psi(r) = p^L + r(p^H - p^L) \). To give a sense for the size of these differences, Pallais and Sands’s (2016) oDesk experiment found that average referred workers submitted work 70% of the time, while referrals by the worst-performing 20% of referrers submitted

---

¹⁹This positive association is a natural consequence of homophily along traits like intelligence (e.g. McPherson et al. (2001)). Indeed, Brown et al. (2016) find that “most referrals take place between a provider and a recipient with similar characteristics in terms of age, gender, ethnicity, education, and division and staff level within the corporations.”

work less than 57% of the time. Firms thus desire talented workers both for the immediate increase in productivity, and for the benefit of having skilled recruiters in the future. Indeed, Bock (2015, p. 85) writes “the first step to building a recruiting machine is to turn every employee into a recruiter”.

If a firm with talent $r$ posts wage $w$ with rank $x$ at time $t$, it attracts applicants with quality $Q(x, t)$ and hires recruits of quality $\lambda(Q(x, t), \psi(r))$. Writing $R(x, t)$ for the talent of the firm with wage rank $x$ at time $t$, $Q(x, t)$ is determined by the sequential screening equation,

$$Q_x(x, t) = \frac{\lambda(Q(x, t), \psi(R(x, t))) - Q(x, t)}{x}$$

and $Q(1, t) = \bar{q}$, (12)

as in (3). In turn, the evolution of a given firm’s talent $r(t)$ with wage rank $x(t)$ is given by the difference between its inflow $\lambda$ and outflow $r$,

$$r_t(t) = \alpha \left( \lambda(Q(x(t), t), \psi(r(t))) - r(t) \right).$$

(13)

Turning to payoffs, workers maximize lifetime wages $w(t)$, while firms’ revenue equals $r(t)$. To abstract from entry and exit, suppose that there is no outside option, $w = 0$. A firm’s problem is to choose wages to maximize total discounted profits. Denoting the discount rate by $\beta > 0$, its value function is

$$V(r, s) = \max_{\{w(t)\}_{t \geq s}} \int_s^\infty e^{-\beta(t-s)}(r(t) - \alpha w(t)) dt,$$

(14)

where $r(t)$ evolves according to (13) with initial condition $r(s) = r$.

An equilibrium is given by a wage path $\{w(t)\}_{t \geq 0}$ for every firm, so that given the induced wage ranks $x(w, t)$ and applicant qualities $Q(x, t)$, 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.\footnote{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.}

#### 4.2 Firm’s Problem

First, we study a firm’s optimal wage path $\{w(t)\}_{t \geq 0}$ for any given applicant function $Q(x(w, t), t)$ without imposing equilibrium restrictions on other firms. As in Section 3, it is convenient to write $W(q, t)$ for the wage required to attract applicants $q$ at time $t$, and let the firm optimize directly

21With $w = 0$, the game is constant-sum and there is no scope for inefficiency. In an earlier version of this paper, Board et al. (2017), we showed that a version of our inefficiency result, Theorem 3 extends to the dynamic model.

22Note that the “wages” $w(t)$ are really one-time payments to each of the $\alpha$ newly hired employees at time $t$. More realistically, one could model worker compensation as constant flow wage $(\beta + \alpha)w(s)$ until retirement, but it simplifies the accounting to have firms incur these costs up-front.

23As 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 4.3 that equilibrium strategies are almost always pure.

24This 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.
over the applicant pools \( \{q_t\}_{t \geq 0} \). After this change of variable, the firm’s Bellman equation becomes

\[
\beta V(r, t) = \max_q \{r - \alpha W(q, t) + \alpha (\lambda(q, \psi(r)) - r)V_r(r, t) + V_t(r, t)\}. \tag{15}
\]

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

\[
W_q(q, t) = \lambda_q(q, \psi(r))V_r(r, t). \tag{16}
\]

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

The RHS of (16) is increasing in \( r \), so firms with more talent have a higher marginal benefit from attracting better applicants, yielding positive assortative matching. Intuitively, firms with more talent have higher marginal benefit from better applicants \( \lambda_q(q, \psi(r)) \) because of the supermodularity of \( \lambda \) as in Section 3, and higher marginal value of talent \( V_r(r, t) \) because \( V(r, t) \) is convex in \( r \) (see Appendix D.1). Hence, wages are dynamic complements: an increase in today’s wage raises tomorrow’s talent, and thereby tomorrow’s optimal wage.

To compute equilibrium wages from the first-order condition (16), we write \( r(u) \) and \( q(u) \) for the equilibrium trajectory of talent and applicant quality, and apply the envelope theorem and the law of motion of firm talent (13) to compute

\[
V_r(r(t), t) = \int_t^\infty e^{-\int_t^s \beta + \alpha (1-\lambda_p(q(u),\psi(r(u))))\psi_r(r(u))} du ds \tag{17}
\]

as we argue formally in Appendix D.1. Intuitively, the future benefit of better employees is discounted both at the interest rate \( \beta \) and the retirement rate \( \alpha \). But selective recruiting raises the persistence of firm talent or, equivalently, reduces the talent decay rate by a factor \( 1 - \lambda_p\psi_r \).

### 4.3 Equilibrium

Given the single-firm analysis, it is straightforward to characterize equilibrium. 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 D.2). 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 steps

1. **Allocations.** At time \( t \), applicant quality \( Q(x, t) \) is determined by sequential screening (12).

   The evolution of firm \( x \)’s talent \( R(x, t) \) is then given by the firm dynamics equation (13) with \( x(t) \equiv x \).

2. **Payoffs.** Firm \( x \)’s marginal value of talent is determined by (17), with \( q(u) = Q(x, u) \). Using this, wages \( W(q, t) \) are given by the first-order condition (16), with \( r = R(x, t), q = Q(x, t) \), and \( W(0, t) = w = 0 \).

Given these wages, the FOC (16) implies global optimality of the HJB (15) because the net benefit
of attracting marginally better applicants \( \lambda_q(q, \psi(r))V_r(r, t) - W_q(q, t) \) single-crosses in \( r \). Standard verification theorems then imply that the policy functions are indeed optimal. To summarize:

**Theorem 4.** Equilibrium exists and is essentially unique. Firm-applicant matching is positive assortative and the distribution of talent has no atoms at any \( t > 0 \).

*Proof.* Only the last claim, about no atoms at \( t > 0 \), remains to be shown. See Appendix D.2.

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

To understand the evolution of talent \( R(x, t) \) and applicant quality \( Q(x, t) \), consider Figure 2. At \( t = 0 \), all firms employ average workers, with quality \( \bar{q} = 0 \). 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 \( Q(1, t) = \bar{q} \), and so (13) implies that its talent grows monotonically and converges to a steady state. For lower-ranked firms, the dynamics are more subtle. For example, firm \( x = 0.5 \) initially improves as its recruits are more talented than retirees. However, as higher-ranked firms become better at identifying talent, its applicant pool deteriorates and its quality eventually falls back.

The top firms initially lose money as they post high wages and invest in talent. This ultimately raises both their productivity and their screening skill, giving them an advantage in the labor market, and delivering a steady stream of profits. Over time, rents shift from workers to firms: Firms earn zero lifetime value, with early workers paid more than their productivity and later workers paid less, as differentiated firms compete less intensely over workers. The way small initial differences are amplified over time resembles the dynamics in Giorcelli (2019), where a management training program in post-war Italy generated productivity gains that grew from 15% in the first year to 49% after 15 years. One channel was that “better managed firms paid higher average wages to their workers, which may indicate that trained managers were able to hire/retain better workers”.

### 4.4 Steady State

Steady-state talent among recruits and applicants \( \{R^*(x), Q^*(x)\} \) is easily characterized. First, the talent of each firm’s recruits and retirees balance,

\[
\lambda(q, \psi(r)) = r. \tag{18}
\]

This has a unique fixed point, \( r = \rho(q) \), since \( \lambda(q, \psi(r)) - r \) is convex (see Appendix D.1), positive at \( r = 0 \), negative at \( r = 1 \), and hence crosses zero exactly once. Naturally, firms with better applicants have higher talent; formally, \( \rho(q) \) increases since \( \lambda(q, \psi(r)) - r \) rises in \( q \) and single-crosses from above in \( r \).

Second, substituting \( R^*(x) = \rho(Q^*(x)) \) into the sequential screening equation (12), steady-state applicant quality \( Q^*(x) \) is given by

\[
Q^*_x(x) = \frac{\lambda(Q^*(x), \psi(\rho(Q^*(x)))) - Q^*(x)}{x}. \tag{19}
\]
Figure 2: Equilibrium Dynamics of Applicant and Recruit Quality. All firms start with talent $\bar{q} = 0.25$, and choose wages optimally as characterized in the text. The recruiting function is $\psi(r) = 0.4 + 0.9r$ and turnover is $\alpha = 0.2$.

Together (18) and (19) pin down firm $x$’s talent $R^*(x)$ and applicant quality $Q^*(x)$ in steady state, independent of turnover $\alpha$ and the discount rate $\beta$. Differentiating, steady-state talent dispersion is given by:

$$R^*_x(x) = \rho_q(Q^*(x))Q^*_x(x) = \frac{\lambda_q(Q^*(x), \psi(R^*(x)))}{1 - \lambda_p(Q^*(x), \psi(R^*(x)))\psi_r(R^*(x))} \cdot Q^*_x(x).$$  \hfill (20)

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

**Theorem 5.** 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 $R(x, t)$ converges to $R^*(x)$.

**Proof.** See Appendix D.3

To show convergence, Figure 2 suggests a “proof by induction”. The top firm recruits from a pool of constant quality, so equation (13) implies that its talent converges exponentially to steady state. Then consider firm $x$ close to 1. As the talent of higher firms converges, firm $x$’s applicant pool converges, and then firm $x$’s talent converges, too, by equation (13). The formal proof is more complicated because $x$ is continuous; it proceeds by showing 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-Lindelöf theorem.

Footnotes:

25The turnover rate $\alpha$ determines the rate at which talent converges to steady state, while the interest rate $r$ determines firm and worker shares in steady state.

26Note that the denominator in this expression is positive around the steady state: While $1 - \lambda_p(q, \psi(r))\psi_r(r)$ may be negative for arbitrary values of $r$, this cannot happen in steady state where $r = \rho(q)$, since we know that $\rho(q)$ increases in $q$ and, by the implicit function theorem, $R_q(q) = \lambda_q(q, \psi(q))\psi_r(\rho(q))/\left(1 - \lambda_p(q, \psi(q))\psi_r(\rho(q))\right)$. 
In steady state, talent/productivity $R(x)$ are dispersed. If all firms were hiring from the same applicant pool $q$, talent differences from the steady state level $r - \rho(q)$ would decay exponentially. The link between talent and recruiting skills slows this decay but cannot stop it. Rather, what stops the decay 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. Equation (20) highlights that any degree of correlation between productivity and recruiting skills $\psi_r > 0$ gives rise to non-vanishing talent dispersion. In particular, if we take the direct talent-skill relationship $\psi_r$ to zero, (20) shows that the strategic effect in itself gives rise to dispersion $R^*_x = \lambda_q Q^*_x$, while the direct link merely amplifies this dispersion by a factor $1 - \lambda_p \psi_r$.

As discussed in Section 2, improved screening (e.g. referral quality) may help explain the increasing differences in productivity, wages and talent between firms. These forces are magnified in the dynamic model. In the short-run, if we fix firms’ talent and increase signal accuracy $\psi(\cdot)$, productivity dispersion rises as high-paying firms fish out more of the talented workers from the applicant pool; over time, dispersion is further amplified as talent accumulates at the top firms, further raising their screening ability and the talent of their recruits.

Talent thus generates a sustainable competitive advantage. One might wonder why a firm with untalented workers doesn’t compete more aggressively in wages to build its talent over time. While this is a feasible strategy, it is simply too expensive. High-talent firms have a higher marginal benefit from raising wages because recruiting skills and applicant quality are complements in the job market. The model captures a comparative advantage of Google which Paul Otellini, the CEO of Intel, called their “self-replicating talent machine” (Bock, 2015, p. 67).

### 4.5 Rent Sharing

The model generates predictions about how value is shared between workers and firms. In steady state the marginal value of talent (17) simplifies to

$$V^*_r(\rho(q)) = \frac{1}{\beta + \alpha (1 - \lambda_p (q, \psi(\rho(q))) \psi_r(\rho(q)))}.$$  

(21)

As in (17), the marginal product of talent is annuitized at the interest rate $\beta$ and the local talent decay rate $\alpha(1 - \lambda_p \psi_r)$. Substituting (21) into the first-order condition (16), flow wages in steady-state equilibrium $(\beta + \alpha)W^*(q)$ are given by

$$(\beta + \alpha)W^*_q = \frac{(\beta + \alpha) \lambda_q}{\beta + \alpha (1 - \lambda_p \psi_r)}.$$  

(22)

where we omit arguments for legibility. Steady-state flow profits $\Pi^*_q$ in turn are given by

$$\Pi^*_q = \rho_q - (\beta + \alpha)W^*_q = \frac{\beta \lambda_p \psi_r}{(1 - \lambda_p \psi_r) (\beta + \alpha (1 - \lambda_p \psi_r))}.$$  

(23)

\[^{27}\text{Indeed, since the equilibrium correspondence is upper hemi-continuous, even in the limit case where } \psi(r) \equiv p \text{ positive assortative matching remains an equilibrium and gives rise to non-trivial dispersion. We do not formally allow for this limit case since equilibrium, and hence its steady-state, are no longer unique; indeed all firms are indifferent between all wages in the equilibrium support at all times.}\]
Equations (22) and (23) show an increase in turnover raises both the level of flow wages and their dispersion. To see this, observe that an increase in turnover, $\alpha$, has no impact on the steady-state distribution of talent $R^\ast(x)$, but raises the rate at which the economy converges to the steady state. Equations (22) and (23) then imply that flow wages $(\beta+\alpha)W^\ast(Q^\ast(x))$ rise and flow profits $\Pi^\ast(Q^\ast(x))$ fall for all firms $x$. Intuitively, when turnover is high, a firm’s stock of talent quickly depletes and a low-talent firm can achieve almost the same profits as a high-talent firm by mimicking its wage policy; this intensifies competition and drives up wages. Higher turnover also increases wage dispersion in both the dispersive order and the log-dispersive order (see Appendix D.4). Intuitively, shifting surplus from firms to workers disproportionally benefits high-wage workers, while $W^\ast(Q^\ast(0)) = 0$, irrespective of $\alpha$.

4.6 Is Talent Always a Sustainable Competitive Advantage?

Talent differences emerge and persist in Theorem 5 since positive assortative matching overcomes the mean-regression induced by noisy recruiting. Here we discuss the assumptions that underlie this argument, helping us understand in which industries talent may be a sustainable competitive advantage.

The first key assumption is that wages are set at the firm level, rather than being individually negotiated. In Montgomery (1991), talented workers refer talented applicants to their current employer. The firm then makes applicant-specific wage offers. Multiple firms may share a referred worker, so wage competition gives rise to a mixed strategy equilibrium along the lines of Burdett and Judd (1983) in which referred and non-referred workers are offered different wages by the same firm. We now argue that Montgomery’s symmetric equilibrium does not generate persistent dispersion. For an apples-to-apples comparison with our model, assume an infinite time horizon, continuous time, and that firms employ a unit mass of workers with aggregate talent $r(t)$ who retire at rate $\nu$. When a firm has a vacancy, a random current employee is asked for a referral. If the employee is talented (resp. untalented), the expected talent of the recruit in equilibrium equals $\bar{\lambda}$ (resp. $\lambda$), where $0 < \lambda < \bar{\lambda} < 1$. Since wage offers are individual, the expected talent of a referred worker does not depend on the number of other talented workers that firm employs. Talent evolution is then governed by $r_t = \alpha(\lambda + r(\bar{\lambda} - \lambda) - r)$ and converges to $r^\ast = \lambda/(1 - \lambda + \lambda)$, irrespective of initial talent. In contrast, our firms have firm-specific wage policies. If we interpret the recruiting skill function $\psi(r)$ as the firm sampling its current employees for referrals, our model predicts that a talented worker recruits better referrals if their colleagues are also talented. This is because high-talent firms pay high wages, and help convert the recommendation of the talented worker into a talented recruit. This breaks the separability seen in Montgomery, implying that better firms retain their talent advantage in the long-run.

Second, we assume non-decreasing returns to talent, in that revenue is linear in talent $r$, while screening skills $\psi(r)$ are weakly convex in $r$; together, these assumptions generate positive assortative matching. The model provides strict incentives for positive assortative matching, so a little concavity does not affect equilibrium sorting and the distribution of steady-state talent. But sufficient concavity induces negative assortative matching and gives rise to convergence; for an extreme example, if flow revenue equaled $\max\{r_t, \bar{q}\}$, then all firms randomize over wages to maintain talent at $\bar{q}$. Similarly, if it suffices to have 10% of talented employees to guarantee good screening
outcomes, then firms below the 10% threshold would bid aggressively to improve their screening.\footnote{As discussed in Section 3, positive assortative matching relies on a number of other assumptions. If talent is not scarce (5), the most skilled firms may pay lower wages than some firms with average recruiting skills (see Board et al. (2017) Appendix A.2 for details). And if there is imperfect competition, low-skilled firms may bid more to avoid the externalities imposed on them by high-skilled firms (see Section 3.3).}

Third, we assume that firms only differ in their talent. In Board et al. (2017) we allow firms to have different production technology in addition to different talent, and suppose the two are complements. The idea is illustrated by Netflix’s HR manual: “In procedural work, the best are two times better than average. In creative/inventive work, the best are ten times better than average” (Hastings and McCord, 2009). Whether the firms converge then depends on the relative persistence of these two state variables. If technology is immobile then talent ultimately lines up with technology, so high-tech firms eventually become high-talent. However, if the industry is fast moving and talented people make new inventions, then technology ultimately lines up with talent, so high-talent firms eventually become high-tech. Indeed, universities seem to fall into the latter category: Waldinger (2016) argues the loss of human capital at German universities had a large, persistent effect on output, whereas the loss of physical capital had a small, temporary effect.

Finally, our firms have a continuum of workers. This means that the firms’ stock of talent moves slowly and firms do not leap-frog each other. If instead we modeled firms in terms of discrete individuals, a lucky hire would allow firm $r$ to jump over firm $r + \epsilon$; it would then raise its wages, helping it to stay ahead. Formally, equilibrium would still exhibit dispersion in steady-state talent, but the randomness in the hiring process renders a firm’s talent an irreducible Markov chain, so any firm cycles through all positions in the distribution eventually. Of course, if firms are large, the regression to the mean would be slow, and we can think of the perfect persistence in Theorem 5 as an idealization of this economic force.

5 Conclusion

In their survey on personnel economics, Oyer and Schaefer (2011) write that “This literature has been very successful in generating models and empirical work about incentive systems […] The literature has been less successful at explaining how firms can find the right employees in the first place.” By studying the equilibrium interaction of firms’ recruiting strategies, we hope this paper takes a step in this direction.

The paper has three major contributions. First it provides a simple framework to understand how screening generates predictions about wages, employment and mismatch. It thus enriches Phelps’s classic (1972) model of statistical discrimination to distinguish between public information (e.g. felony convictions, credit scores) and private information (e.g. job tests, employee referral networks). It also provides a theory of wage and productivity dispersion, providing a different interpretation of firm-effects in matched employee-employer data.

Second, the paper shows that skilled firms post higher wages, attract better applicants, hire better recruits and obtain higher productivity and profits. It also shows that this positive assortative matching is inefficient, in contrast to the classic result of Becker (1973) in which complementarities are exogenous. This compositional externality can be mitigated by policies that reduce wage dispersion.
Finally, we propose a new model of firm dynamics based on the tenet that talented workers are better at identifying talented applicants. Initially similar firms diverge over time and the economy converges to a steady state featuring persistent dispersion in talent, wages and productivity. We thus complement the classic model of Montgomery (1991) by showing that referrals can generate across-firm inequality, even in the very long run.

The model relates to a number of fields and can be taken in a number of directions. Industrial economists may wish to allow quality to be multi-dimensional, so the ability to acquire talent affects the direction the firm takes in both the product and labor market. Organizational economists may ask how a firm can improve the quality of its employees by placing authority in the hands of more recent hires or agents with a better track record of hiring. Labor economists might wish to add search frictions, so higher wages both increase the quality of applicants and the probability of finding a suitable worker. Network economists may wish understand how job-to-job transitions affect firms’ ability to aggregate information and sort workers into the industry. And macroeconomists may be interested in how the talent is reallocated between firms or sectors in response to individual or aggregate shocks to productivity.
Appendix

A Comparison with Kurlat (2016)

In this section we clarify the relationship between our model and that of Kurlat (2016) with “false positives”, and two types of firms: skilled and unskilled. The most significant difference is that our model assumes that firms’ tests are conditionally independent, whereas Kurlat assumes “nested” signals in that an untalented worker who passes the test of a skilled recruiter also passes the test of an unskilled recruiter.\(^{29}\) Workers thus fall into one of four categories: A) talented workers, who automatically pass both tests, B) untalented workers who pass both tests, C) untalented workers who pass the test of the unskilled recruiter but fail the test of the skilled recruiter, and D) untalented workers who fail both tests.\(^{30}\)

This seemingly small difference from our model results in very different equilibria. To see why, assume that \(w\) is small, so all skilled firms and some unskilled firms enter. Since unskilled firms hire proportionally from categories A, B, and C, they do not impose a compositional externality on other firms. But skilled firms, who hire proportionally from categories A and B, but screen out C workers, impose a negative externality on unskilled firms. Thus, unskilled firms have an incentive to outbid skilled firms but not vice versa (and neither type of firm has an incentive to outbid its own type). The equilibrium wage distribution is thus degenerate, with all entering firms offering some wage \(w^*\) and workers endogenously breaking ties in favor of unskilled firms. The wage \(w^*\) is determined by the entry condition of unskilled firms, and the proportion of unskilled firms that enter is determined by the condition that all workers in categories A and B are employed. This equilibrium has negative assortative matching and efficient aggregate sorting, in sharp contrast to the positive assortative matching and inefficient aggregate sorting of our model with independent signals.

B Proofs from Section 2

In this appendix we study the effect of private information (i.e. an increase in screening skills \(p\)) and public information on applicant quality, unemployment, mismatch, and wage and productivity dispersion. Additional public information is captured by a mean-preserving spread of \(\bar{q}\). To interpret this, recall that public information segments the labor market, so rather than hiring from one pool of workers with average quality \(q\), firms hire from various pools of quality \(q'\), where \(E[q'] = q\). Thus, public information increases/decreases an aggregate labor market outcome, such as unemployment or mismatch, if this outcome is convex/concave in \(\bar{q}\).

\(^{29}\)Other differences between our model and Kurlat’s are superficial. His focus on asset markets and ours on labor markets is purely semantic; indeed, the follow-up paper Kurlat and Scheuer (2019) is phrased in terms of labor markets. The impatience of his distressed sellers would correspond to a reservation wage of talented workers in our model. Finally, his model allows for richer interaction across many markets, but the unique equilibrium described in his Proposition 1 features a single market and is equivalent to the equilibrium in an auction where buyers bid for assets, and then sellers apply from top to bottom bid with buyers screening applying sellers and accepting the first seller who passes their test, exactly as in our model.

\(^{30}\)Category A corresponds to the “green assets” in Kurlat’s table I, category C to “red assets”, and category D to the “black assets”.

25
B.1 Applicant Quality

Write the applicant quality for a firm with rank $x$ as $Q(x, \bar{q}, p)$ to make explicit the effect of aggregate talent $\bar{q}$ and screening skills $p$. Recall that

$$\lambda(q, p) := \frac{q}{1 - p(1 - q)} \quad \text{and} \quad \phi(x, q, p) := \frac{\lambda(q, p) - q}{x}.$$ 

**Lemma 1.** Applicant quality $Q(x, \bar{q}, p)$ is:

(a) Increasing in wage rank $x$ with derivative $Q_x(x, \bar{q}, p) = \phi(x, Q(x, \bar{q}, p), p)$.

(b) Increasing in aggregate talent $\bar{q}$ with derivative

$$Q_{\bar{q}}(x, \bar{q}, p) = \exp \left( - \int_x^1 \phi_q(\bar{x}, Q(\bar{x}, \bar{q}, p), p) d\bar{x} \right). \quad (24)$$

(c) Decreasing in screening skills $p$ with derivative

$$Q_p(x, \bar{q}, p) = - \int_x^1 \exp \left( - \int_x^{\bar{x}} \phi_q(\bar{x}, Q(\bar{x}, \bar{q}, p), p) d\bar{x} \right) \phi_p(\bar{x}, Q(\bar{x}, \bar{q}, p), p) d\bar{x}.$$ 

(d) Log-submodular in $(x, \bar{q})$.

(e) Log-supermodular in $(x, p)$.

*Proof.* (a) is the sequential screening equation (3).

(b) and (c) follow from the theory of ordinary differential equations, e.g. Hartman (2002, Theorem 3.1), whereby the solution of the ODE $Q_x(x, \bar{q}, p) = \phi(x, Q(x, \bar{q}, p), p)$ with boundary condition $Q(1, \bar{q}, p) = \bar{q}$ satisfy $Q_{\bar{q}} = \phi_q Q_{\bar{q}}$ and $Q_p = \phi_q Q_p + \phi_p$ with boundary conditions $Q_q(1, \bar{q}, p) = 1$ and $Q_p(1, \bar{q}, p) = 0$.

(d) and (e) follow because

$$\left( \log Q(x, \bar{q}, p) \right)_x = \frac{\lambda(Q(x, \bar{q}, p), p) - Q(x, \bar{q}, p)}{Q(x, \bar{q}, p)_x} = \frac{1}{x} \left( \frac{1}{1 - p(1 - Q(x, \bar{q}, p))} - 1 \right)$$

falls in $\bar{q}$ by (b) and rises in $p$ by (c). 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. A rise in screening skills $p$ does not affect applicant quality at high-ranking firms but depresses it at lower-ranking firms, raising dispersion. 

B.2 Unemployment

To analyze equilibrium unemployment, it is useful to first establish a closed-form solution for the inverse of applicant quality $Q(x, \bar{q}, p)$. We first show that this inverse is given by

$$X(q, \bar{q}, p) = \left( \frac{1 - \bar{q}}{1 - q} \right)^{\frac{1}{p}} \left( \frac{q}{\bar{q}} \right)^{\frac{1-p}{p}}.$$ 


Indeed, differentiating,

\[ X_q = \left( \frac{1}{p} \frac{1}{1-q} + \left( \frac{1}{p} - 1 \right) \frac{1}{q} \right) X = \left( \frac{q + (1-p)(1-q)}{pq(1-q)} \right) X = \frac{X}{\lambda(q,p) - q} \]

which is the inverse of (3).

Turning to unemployment, applicant quality at the marginal firm solves \( \lambda(q,p) = \frac{q}{1-p(1-q)} = \frac{q}{1-w/p} \) and hence the cutoff applicant quality equals \( q = Q(p,w) := \frac{w(1-p)}{1-wp} \). Thus, as long as \( Q(p,w) \leq q \), equilibrium unemployment is given by

\[ X(q,p,w) = X(Q(p,w),\bar{q},p) = \left( \frac{1 - \bar{q}}{1 - Q(p,w)} \right)^{\frac{1}{p}} \left( \frac{Q(p,w)}{\bar{q}} \right)^{\frac{1-w}{p}} \]

otherwise, if \( Q(p,w) \geq \bar{q} \), we set \( X(q,p,w) \equiv 1 \). Analyzing (25) as a function of \( q \) is straightforward:

**Proposition 1.** Equilibrium unemployment \( X(q,p,w) \) is strictly decreasing and convex in \( q \) for \( \bar{q} \in [Q(p,w),1] \), and constant, equal to 1 for \( q \leq Q(p,w) \). Hence, an “imprecise” public signal about the unselected applicant pool \( \bar{q} \), i.e. for which the support of the posterior is included in \([Q(p,w),1]\), raises unemployment.

**Proof.** Differentiating (25)

\[ X_q = -\left( \frac{1}{p} \frac{1}{1-q} + \left( \frac{1-p}{p} \right) \frac{1}{q} \right) X = -\left( \frac{1-p(1-q)}{p(1-q)\bar{q}} \right) X < 0 \]

\[ X_{qq} = X \left( \left( \frac{1-p(1-q)}{p(1-q)\bar{q}} \right)^2 - \left( \frac{1-p(1-q)}{p(1-q)\bar{q}} \right) \right) = X \frac{1-p}{(p(1-q)\bar{q})^2} > 0 \]

as required. \( \square \)

Proposition 1 characterizes unemployment as a function of aggregate talent \( \bar{q} \). As a function of screening skills \( p \), the picture is less clear. Rewriting (25) as

\[ X = \exp \left( \frac{1}{p} \log \left( \frac{1 - \bar{q}}{1 - Q(p,w)} \right) + \left( \frac{1}{p} - 1 \right) \log \left( \frac{Q(p,w)}{\bar{q}} \right) \right) \]

we get

\[ X_p = \left( \frac{1}{p} \cdot \frac{1}{1 - Q(p,w)} + \frac{1-p}{p} \cdot \frac{1}{Q(p,w)} \right) Q_p(p,w) + \frac{1}{p^2} \log \left( \frac{\bar{q}}{Q(p,w)} \cdot \frac{1 - Q(p,w)}{1 - q} \right) X \]

(26)

The first (negative) term captures increased screening skills of the marginal firm, which reduces the break-even applicant quality, \( Q_p(p,w) = -\frac{w(1-w)}{(1-wp)^2} < 0 \), and hence lowers unemployment. The second (positive) term captures increased screening skills at infra-marginal firms, which lowers applicant quality at the marginal firm, hence raising unemployment. Which effect dominates is in general ambiguous.

To illustrate, consider small values of \( p \). When \( \bar{q} < w \), everyone is unemployed absent private signals, so private information lowers unemployment. When \( \bar{q} > w \), no-one is unemployed absent
private signals, so private information raises unemployment. However, this negative effect vanishes for small $p$.

**Lemma 2.** Assume $\bar{w} \neq \bar{q}$.

(a) If $\bar{q} < \bar{w}$, then $X(\bar{q}, p, w) \equiv 1$ for $p \in [0, p]$, where $p$ solves $\lambda(\bar{q}, p) = w$.

(b) If $\bar{q} > \bar{w}$ then $X(\bar{q}, 0, w) = X_p(\bar{q}, 0, w) = 0$.

**Proof.** Part (a) is obvious. If $\bar{q} < \bar{w}$, and so firms would not hire without screening, employment also vanishes for small $p \in [0, p]$.

The first equation in part (b), $X(\bar{q}, 0, w) = 0$, is obvious, too. If $\bar{w} < \bar{q}$, firms are willing to hire without screening, so when $p = 0$ and there is no selection, all workers are hired. The second equation in part (b) $X_p(\bar{q}, 0, w) = 0$ follows from (26) by l’Hopital’s rule, since $X \propto \exp(-1/p)$ term goes to zero much faster than the $1/p$-terms go to $\infty$. In fact, all higher-order $p$-derivatives of $X$ vanish at $p = 0$, too. Thus, while unemployment is strictly positive for any $\bar{w}, p > 0$, it is very small for $p$ close to 0; alternatively, applicant quality $Q(x, p)$, which converges to $Q(0, p) = 0$ for any $p > 0$, remains pretty large for $x > 0$ and $p$ close to 0.

### B.3 Mismatch

We first show that, unsurprisingly, both public information (a mean-preserving spread in $\bar{q}$) and private information (and increase in $p$) reduce mismatch.

**Proposition 2.** We have the following:

(a) Mismatch is concave in $\bar{q}$, and strictly concave when $\bar{q} \in [q, 1]$. Thus, public information decreases mismatch.

(b) Mismatch decreases in screening skills $p$.

**Proof.** For part (a) we note that mismatch equals

$$M = qX(\bar{q}, p, w)(1 - w) + (1 - \bar{q} - (1 - q)X(\bar{q}, p, w))w = (1 - \bar{q})w - (w - \bar{q})X(\bar{q}, p, w)$$

(27)

for $\bar{q} \geq q$; below it is simply $\bar{q}(1 - w)$. The result then follows by Proposition 1.

Part (b) follows by a rather different argument, noting that mismatch is inversely related to aggregate surplus. In the proof of Theorem 3 in Appendix C, we show that the marginal effect of screening skills at any firm rank $x'$ on aggregate surplus equals

$$\lambda_p(Q(x', \bar{q}, p), p) \exp \left( - \int_{x}^{x'} \frac{\lambda_q(Q(x, \bar{q}, p), p)}{x} dx \right) > 0.$$ 

Since mismatch is the inverse of surplus, the effect of rank-$x'$ screening skills on mismatch is negative for any $x'$, and hence also in aggregate when all firms’ screening skills $p$ rise simultaneously, as we assume here.

The next result compares our screening model with the “Phelps model”, where all firms use the same public signal that screens out untalented workers with probability $p$. Alternatively, one might contrast these two models by stating that screening tests are conditionally independent in our model, but perfectly correlated in the Phelps model.
Proposition 3. Equilibrium mismatch is smaller with public information (the Phelps model) than with private information (the screening model).

Proof. With public information, workers’ public posteriors equal 0 (after failing the test) or \( \lambda(\bar{q},p) \) (after passing the test). This is a mean-preserving spread of the distribution of posteriors at the time of hiring in equilibrium, which has an atom at \( \lambda(\bar{q},p) = w \) and then a continuous part with support \([\lambda(\bar{q},p), \lambda(\bar{q},p)]\). Thus, public signals are Blackwell-sufficient for the private signals (at the point of hiring), and thus reduce mismatch by Proposition 2(a).

Finally, we consider the value of imprecise signals in the screening model.

Proposition 4. Assume \( \bar{q} \neq w \). In the screening model, the marginal value of screening skills vanishes at the bottom, \( M_p(0) = 0 \).

Proof. For \( \bar{q} < w \), the result is obvious since \( \lambda(\bar{q},p) < w \) for small \( p \in [0,p] \). Thus no-one is hired and \( M(p) \) is flat on \([0,p]\). For \( \bar{q} > w \), the result is surprising. Differentiating (25), we get \( M_p = -X_p(w - q) + X(w - Q_p) \), and the result follows by Lemma 2(b). Intuitively, with no information everyone is employed and the point of information is to weed out some untalented applicants. While every single test achieves that, workers get so many attempts at passing each test, that asymptotically all of them do. Good for employment, but bad for mismatch. More strongly, all higher-order derivatives of \( M \) in \( p \) also vanish at \( p = 0 \).

As a comparison, consider mismatch in the “Phelps model”. If \( \bar{q} > w \), then all applicants who pass the test are employed, so mismatch is solely due to the opportunity cost of untalented workers who slip through the screening, \( M = -w(1 - \bar{q})p \). Thus mismatch is linear in screening skills. If \( \bar{q} < w \), the same result obtains if screening is sufficiently precise, \( p > p \); otherwise, if \( p < p \), no one is employed, and so \( M \equiv \bar{q}(1 - w) \).

B.4 Wage and Productivity Dispersion

To study dispersion we use the demanding log-dispersive order (Shaked, 1982). That is, wage and productivity dispersion increases in \( \bar{q}, p \) if \( \lambda(Q(x', \bar{q},p),p)/\lambda(Q(x, \bar{q},p),p) \) increases in \( \bar{q}, p \) for all \( x' > x \).

\footnote{In this terminology, Proposition 1 states that applicant quality dispersion \( Q(x', \bar{q},p)/Q(x, \bar{q},p) \) decreases in \( \bar{q} \) and increases in \( p \).}

Proposition 5. We have the following:
(a) The dispersion of wages and productivity falls in \( \bar{q} \).
(b) The dispersion of wages and productivity increases in \( p \) if talent is scarce (5).

Proof. We wish to show that \( \lambda(Q(x, \bar{q},p),p) = \frac{Q(x, \bar{q},p)}{1 - p + pQ(x, \bar{q},p)} \) is log-submodular in \( (x, \bar{q}) \) and log-supermodular in \( (x, p) \). For part (a) we compute

\[
(\log \lambda(Q(x, \bar{q},p),p))_{\bar{q}} = (\log Q(x, \bar{q},p))_{\bar{q}} - \frac{pQ(x, \bar{q},p)}{1 - p + pQ(x, \bar{q},p)} = (\log Q(x, \bar{q},p))_{\bar{q}} \left( \frac{1 - p}{1 - p + pQ(x, \bar{q},p)} \right)
\]
which falls in $x$ since $Q(x, \tilde{q}, p)$ rises in $x$ and is log-submodular in $x$ and $\tilde{q}$ by Proposition 1(a) and (d). Intuitively, the larger proportional decline of applicant quality at lower-ranked firms is aggravated by the concavity of recruit quality $\lambda(q, p)$ in applicant quality $q$.

For part (b) we compute
\[
(\log \lambda(Q(x, \tilde{q}, p), p))_p = \frac{\lambda_q(Q(x, \tilde{q}, p), p)Q_p(x, \tilde{q}, p) + \lambda_p(Q(x, \tilde{q}, p), p)}{\lambda(Q(x, \tilde{q}, p), p)}
\]
and, omitting arguments for legibility,
\[
(\log \lambda(Q(x, \tilde{q}, p), p))_{px} = \frac{1}{x^2} \left[ \lambda\lambda_{qq}Q_xQ_p + \lambda\lambda_{pq}Q_x - \lambda_qQ_x\lambda_qQ_p + (\lambda\lambda_{qpx} - \lambda_q\lambda_pQ_x) \right]
\]
To see that this is positive, recall $Q_x = \phi > 0, Q_\tilde{q} > 0, Q_p < 0$ from Lemma 1, and $\lambda_{qq} < 0$ and $\lambda_{qp} > 0$ from (6), using (5). Then the first three terms are positive. Using $Q_x = \phi$, the term in brackets then equals $\lambda_q(\lambda_\phi \rho - \lambda_p \phi) = \lambda_q(\lambda \lambda_p - \lambda_p(\lambda - q))/x > 0.

\section{Proofs from Section 3}

In this section we prove our main inefficiency result, Theorem 3, and extend it to two model variants with production complementarities and screening costs, introduced in Section 3.3. To accommodate these variants, we first formulate a more general model where the surplus per worker when a firm with skills $p$ hires from an applicant pool with quality $q$ is given by a general function $\omega(q, p)$. The baseline model corresponds to $\omega(q, p) = \lambda(q, p) - \omega$, production complementarities correspond to $\omega(q, p) = h(p) \cdot \lambda(q, p) - \omega$, and screening costs to $\omega(q, p) = \lambda(q, p) - \kappa/(1 - p(1 - q)) - \omega$.

We develop the apparatus to analyze aggregate surplus for general surplus functions $\omega(q, p)$ in C.1, specialize to the baseline model and prove Theorem 3 in C.2, prove some auxiliary claims from Section 3 in C.3 and C.4, and then extend the inefficiency result to models with production complementarities and screening costs in C.5 and C.6.

\subsection{Aggregate Surplus for General Surplus Functions $\omega(q, p)$}

Fix the number of entering firms, and thereby the cutoff $x$. Aggregate surplus under matching $P(\cdot)$ is given by
\[
S(P(\cdot)) = \int_{\hat{x}}^{1} \omega(Q(x), P(x))dx
\]
(28)

Consider adding a new firm $p$ at rank $x$. This has three effects on surplus. First, firm $p$ hires from an applicant pool with quality $Q(x)$ and so generates surplus $\omega(Q(x), p)$. Second, firm $p$ pushes out the marginal firm, $P(x)$. Third, it lowers the ranking of intermediate firms, $\hat{x} < x$, thereby lowering their applicant quality, $Q(\hat{x})$. To quantify the latter externality, note that firm $p$’s hiring reduces applicant quality just below $x$ by $dQ(x) = \phi(Q(x), p, x)dx$. For lower ranking firms $\hat{x}$, this effect is mitigated since firm $p$ pushes intermediate firms $\hat{x} \in [\hat{x}, x]$ to lower wage ranks and thereby reduces their externality by $-\phi_{x}(Q(\hat{x}), p(\hat{x}), \hat{x})$. By standard results on ODEs in the proof
of Lemma 1, firm $p$’s total effect on applicant quality $Q(\hat{x})$ is given by
\[
\chi(\hat{x}; p, x, P(\cdot)) = -\phi(Q(x), p, x) \exp \left( - \int_{\hat{x}}^{x} \phi_q \right) - \int_{\hat{x}}^{x} \left[ \exp \left( - \int_{\hat{x}}^{\hat{x}} \phi_q \right) \left( -\phi_x(Q(\hat{x}), P(\hat{x}), \hat{x}) \right) \right] d\hat{x} 
\]
(29)
where we dropped the arguments in the integrand $\phi_q = \phi_q(Q(\hat{x}), P(\hat{x}), \hat{x})$ to enhance legibility. Putting these three effects together, firm $p$’s (infinitesimal) net-contribution to surplus when assigned to wage rank $x$ in matching $P(\cdot)$ is thus given by
\[
s(p, x, P(\cdot)) = \omega(Q(x), p) + \int_{\hat{x}}^{x} \chi(\hat{x}; p, x)\omega_q(Q(\hat{x}), P(\hat{x}))d\hat{x} - \omega(Q(x), P(\hat{x})). \tag{30}
\]

We wish to evaluate the effect of changing a matching function $P(\cdot)$ into a second matching function $P'(\cdot)$. To do this, suppose we move one firm $p$ from rank $\bar{x}(p)$ to $\bar{x}(p)$ given matching $P(\cdot)$. The (infinitesimal) change in surplus equals
\[
\int_{\bar{x}(p)}^{\bar{x}(p)} s_x(p, x, P(\cdot))dx.
\]

More generally, let us sequentially move all firms $p \in [p, \bar{p}]$ in order of increasing $p$. The matching function changes over the course of this transformation, and we write $P^p(\cdot)$ for the matching function after firms $p' < p$ have been shifted; thus $P^p_2(\cdot) = P(\cdot)$ and $P^p_1(\cdot) = P'(\cdot)$.\footnote{There are multiple ways to transform $P(\cdot)$ into $P'(\cdot)$ in increasing order of $p$. In a discrete analogue, if $p_1 < p_2 < p_3$, we can transform $\{p_1, p_3, p_2\}$ into $\{p_3, p_1, p_2\}$ by either moving $p_1$ to second position, or by moving $p_1$ to the top and then moving $p_2$ above it. Formally, a transformation is fully specified by the original matching $P(\cdot)$ and the intermediate target positions $\bar{x}(\cdot)$; the intermediate matchings $P^p(\cdot)$ as well as $\bar{x}(\cdot)$ are generated endogenously. In the two applications used in the proofs of Lemma 4 and Theorem 3 we move each firm immediately into its final position, $P'(\bar{x}(p)) = p$, but this need not be the case in general.}

The aggregate change in surplus is given by
\[
S(P'(\cdot)) - S(P(\cdot)) = \int_{p}^{\bar{p}} \left[ \int_{\bar{x}(p)}^{\bar{x}(p)} s_x(p, x, P^p(\cdot))dx \right] dF(p). \tag{31}
\]

We first establish a general formula for the integrand in (31).

**Lemma 3.** The marginal surplus of moving firm $p$ past wage rank $x$ given matching $P(\cdot)$ with $P(x) = \hat{p}$ equals
\[
s_x(p, x, P(\cdot)) = \left[ \omega_q(Q(x), p)\phi(Q(x), \hat{p}, x) - \omega_q(Q(x), \hat{p})\phi(Q(x), p, x) \right]
\]
\[\quad - \left[ \phi_q(Q(x), p, x)\phi(Q(x), \hat{p}, x) + \phi_x(Q(x), p, x) - \phi_q(Q(x), \hat{p}, x)\phi(Q(x), p, x) - \phi_x(Q(x), \hat{p}, x) \right] \gamma \]
where
\[
\gamma = \gamma(x, P(\cdot)) = \int_{x}^{\hat{x}} \exp \left( - \int_{x}^{\hat{x}} \phi_q(Q(\hat{x}), P(\hat{x}), \hat{x})d\hat{x} \right) \omega_q(Q(\hat{x}), P(\hat{x}))d\hat{x}. \tag{33}
\]
To understand these equations, consider swapping \( p \) and \( \hat{p} \), and let us drop the arguments \( Q(x) \) and \( x \) to enhance legibility. The first term in the first line of (32) \( \omega_q(p)\phi(\hat{p}) \) is the increased surplus contribution of firm \( p \) when selecting before \( \hat{p} \) from a pool with \( dQ(x) = \phi(\hat{p})dx \) higher quality. Conversely, the second term \( \omega_q(\hat{p})\phi(p) \) is \( \hat{p} \)'s decreased surplus when selecting after \( p \).\(^{33}\)

The term in the second line of (32) is the effect of switching \( p \) and \( \hat{p} \)'s screening order on residual pool quality \( Q(x - dx) \). By screening applicants \( q + \phi(\hat{p})dx \) from a pool of size \( x + dx \) (rather than applicants \( q \) from pool size \( x \)), firm \( p \)'s selection increases by \([\phi_q(p)\phi(\hat{p}) + \phi_x(p)]dx\); conversely firm \( \hat{p} \)'s selection decreases by \([\phi_q(\hat{p})\phi(p) + \phi_x(\hat{p})]dx\). This change in applicant quality \( dQ(x) \) affects revenue \( \omega(Q(\hat{x}), P(\hat{x})) \) of all lower-ranking firms \( \hat{x} \in [x, \bar{x}] \), where the exponential term in (33) equals \( dQ(\hat{x})/dQ(x) \) and captures differential selection by intermediate firms \( \hat{x} \in [x, \bar{x}] \).

Note that
\[
s_x(P(x), x, P(\cdot)) = 0; \quad (34)
\]
intuitively, switching the screening order of two identical firms \( p = \hat{p} = P(x) \) makes no difference.

**Proof.** Differentiating (30)
\[
s_x(p, x, P(\cdot)) = \omega_q(Q(x), p)Q_x(x) + \chi(x; p, x)\omega_q(Q(x), P(x)) + \int_{\underline{x}}^{\bar{x}} \frac{\chi_x(\hat{x}; p, x)\omega_q(Q(x), P(\hat{x}))d\hat{x}}{\phi_q(x; p, x)}.
\]
Since \( Q_x(x) = \phi(Q(x), P(x), x) \) and \( \chi(x; p, x) = -\phi(Q(x), p, x) \), and recalling \( P(x) = \hat{p} \), the first two terms correspond to the first line of (32).

As for the last term, \( \chi_x(\hat{x}; p, x) \) equals
\[
- [\phi_q(Q(x), p, x)Q_x(x) + \phi_x(Q(x), p, x) - \phi(Q(x), p, x)\phi_q(Q(x), P(x), x) - \phi_x(Q(x), P(x), x)] \exp \left(-\int_{\underline{x}}^{\bar{x}} \phi_q \right)
\]
where the term is square-brackets equals the square-bracket term in the second line of (32), while integration over \( \hat{x} \in [x, \bar{x}] \) yields \( \int_{\underline{x}}^{\bar{x}} \exp \left(-\int_{\underline{x}}^{\bar{x}} \phi_q \right) \omega_q(Q(\hat{x}), P(\hat{x})))d\hat{x} = \gamma \).

We cannot determine the sign of (32) in general. But at the lowest wage rank \( x = \underline{x} \), the analysis simplifies because the integral domain in (33) collapses and so \( \gamma = 0 \), yielding a necessary condition for optimality of PAM that is easy to check. We will check that this condition is violated for the model variants in Propositions 7 and 8, below.

**Lemma 4.** Assume that \( 1 - x \) firms \( p \in [p, \bar{p}] \) enter. If
\[
\frac{\omega_q(\underline{x}, p)}{\omega_q(\underline{x}, \underline{p})} < \frac{\phi_p(\underline{x}, p)}{\phi_p(\underline{x}, \underline{p})}
\]
then PAM does not maximize surplus.

**Proof.** Let us transform \( P(\cdot) = P^{PAM}(\cdot) \) into another matching \( P'(\cdot) \) with NAM for \( x \in [\underline{x}, \bar{x} + \epsilon] \) and PAM for \( x \in (\bar{x} + \epsilon, 1] \). Intuitively, if there were 10 firms with skill \( p_1 < \ldots < p_{10} \) then this

\(^{33}\)To relate these terms to the proof sketch in the body of the paper, set \( \hat{p} = p + dp \) and \( \omega = \lambda - w \). Then \( \omega_q(p)\phi(\hat{p}) - \omega_q(p)\phi(p) = (\lambda_q\phi_p - \lambda_p\phi)dp \), where the first (positive term) captures the decreased adverse selection when the low-skill firm \( p \) is shifted ahead of the high-skill firm \( \hat{p} \), and the second term the loss of surplus due the supermodularity of \( \lambda \).
would mean swapping \( p_1 \) and \( p_2 \), so firms are ranked \( \{p_2, p_1, \ldots, p_{10}\} \), from lowest to highest. Formally \( P'(x) = F^{-1}(x + \epsilon - (x - x)) \) for \( x \in [x, x + \epsilon] \), and \( P'(x) = F^{-1}(x) \) for \( x > x + \epsilon \), where \( \epsilon > 0 \) is small. We transform \( P(\cdot) \) into \( P'(\cdot) \) by shifting firms \( p \in [p, F^{-1}(x + \epsilon)] \) (in rising order of \( p \)) to their \( P'(\cdot) \)-wage rank \( \hat{x}(p) = x + \epsilon - [F(p) - F(p')] \). Since lower firms \( p' < p \) have already been shifted to \( \hat{x}(p') > \hat{x}(p) \) at firm \( p \)'s “turn”, firm \( p \) starts at rank \( \hat{x}(p) = F(p') < \hat{x}(p) \) and is shifted exclusively past firms with higher screening skills \( P_p(x) = F^{-1}[F(p) + x - \hat{x}] \geq p \), recalling the definition of the matching function \( P_p(\cdot) \) at \( p \)'s “turn”.

We now argue that given (35) the net value of this transformation (31) is positive. Using (34), the integrand in (31) equals

\[
s_x(p, x, P_p(\cdot)) = s_x(P_p(x), x, P_p(\cdot)) - \int_p^{P_p(x)} s_{xp}(\hat{p}, x, P_p(\cdot))d\hat{p} = -\int_p^{P_p(x)} s_{xp}(\hat{p}, x, P_p(\cdot))d\hat{p}.
\]

To see that this is positive, differentiate (32) with respect to \( p \), and evaluate for firm \( p \) at the marginal wage rank \( \bar{x} \). The derivative of the second line is zero because \( \gamma = 0 \) for \( x = \bar{x} \). Thus,

\[
s_{xp}(\bar{x}, x, P_{PAM}(\cdot)) = \omega_{xp}(Q(x), p)\phi(Q(x), p, x) - \omega_q(Q(x), p)\phi_p(Q(x), p, x)
\]

which is negative by (35). Now observe that \( (\hat{p}, x, P_p(\cdot)) \) converges to \( (p, x, P_{PAM}(\cdot)) \) for all \( \hat{p}, p \in [p, F^{-1}(x + \epsilon)] \) and \( x \in [x, x + \epsilon] \) as \( \epsilon \to 0 \), using for instance the topology of uniform convergence on matching functions \( P(\cdot) \).

\[\square\]

C.2 Proof of Theorem 3

We now return to the baseline model where surplus is given by \( \omega(q, p) = \lambda(q, p) - w \). The comparison of the elasticities (11) implies (35), and so Lemma 4 already implies that PAM is inefficient. In particular, the proof of Lemma 4 shows that locally near the cutoff \( x \) shifting low-skill firms ahead of high-skill firms increases surplus.

Theorem 3 claims more strongly than that surplus is minimized by PAM and maximized by NAM. We show this by arguing that for \( \omega(q, p) = \lambda(q, p) - w \) our local argument at the bottom of the wage distribution in fact holds globally.

Indeed, (32) simplifies considerably:

**Lemma 5.** When \( \omega(q, p) = \lambda(q, p) - w \), the marginal surplus of moving firm \( p \) past wage rank \( x \) given matching \( P(\cdot) \) with \( P(x) = \hat{p} \) equals

\[
s_x(p, x, P(\cdot)) = [\lambda_q(p)\phi(\hat{p}) - \lambda_q(\hat{p})\phi(p)] (1 - \gamma/x)
\]

where we dropped arguments \( Q(x) \) and \( x \) for legibility, and \( 1 - \gamma/x = \exp\left( -\int_{\bar{x}}^{x} \frac{\lambda_q(Q(\bar{x}).P(\bar{x}))}{x} d\bar{x} \right) \in (0, 1) \).

The term in square-brackets corresponds to the incremental talent hired by firms \( p \) and \( \hat{p} \) when the former is shifted ahead of the latter. The effect on aggregate surplus is scaled down by the factor \( 1 - \gamma/x \) as in (10) since incremental talent hired by the marginal firms \( p \) and \( \hat{p} \) reduces the talent hired by lower-ranking firms.
Proof. We will show that the marginal surplus (32) equals

\[ s_x(p, x, P(\cdot)) = \left[ \lambda_q(p)\phi(\hat{p}) - \lambda_q(\hat{p})\phi(p) \right] - \left[ \lambda_q(p)\phi(\hat{p}) - \lambda_q(\hat{p})\phi(p) \right] \gamma_x \]  

(37)

and thus collapses to (36). Using the definition of \( \omega(q, p) \), the first line in (32) equals the first square-bracket term in (37). Turning to the second line in (32) and recalling \( \phi = \frac{\lambda - q}{x} \), elementary algebra implies

\[ \phi_q(p)\phi(\hat{p}) + \phi_x(p) - \phi_q(\hat{p})\phi(p) - \phi_x(\hat{p}) = \frac{1}{x} \left[ \lambda_q(p)\phi(\hat{p}) - \lambda_q(\hat{p})\phi(p) \right]. \]  

(38)

Multiplying by \( \gamma \), the second line in (32) equals the second square-bracket term in (37), as required.

To evaluate \( \gamma \) observe that

\[ \exp \left( - \int_{\hat{x}}^x \phi_q \right) = \exp \left( - \int_{\hat{x}}^x \frac{\lambda_q(Q(\hat{x}), P(\hat{x})) - 1}{\hat{x}} d\hat{x} \right) = \frac{x}{\hat{x}} \exp \left( - \int_{\hat{x}}^x \frac{\lambda_q(Q(\hat{x}), P(\hat{x}))}{\hat{x}} d\hat{x} \right). \]

Equation (33) thus simplifies to

\[
\gamma = x \int_{\hat{x}}^x \exp \left( - \int_{\hat{x}}^x \frac{\lambda_q(Q(\hat{x}), P(\hat{x}))}{\hat{x}} d\hat{x} \right) \frac{\lambda_q(Q(\hat{x}), P(\hat{x}))}{\hat{x}} d\hat{x} \\
= x \int_{\hat{x}}^x \frac{d}{d\hat{x}} \exp \left( - \int_{\hat{x}}^x \frac{\lambda_q(Q(\hat{x}), P(\hat{x}))}{\hat{x}} d\hat{x} \right) d\hat{x} = x \left[ 1 - \exp \left( - \int_{\hat{x}}^x \frac{\lambda_q(Q(\hat{x}), P(\hat{x}))}{\hat{x}} d\hat{x} \right) \right].
\]

(39)

as required.

We now show that transforming an arbitrary matching \( P(\cdot) \) into \( P'(\cdot) = P^{\text{NAM}}(\cdot) \) raises surplus. Intuitively, if there were 10 firms with skill \( p_1 < \ldots < p_{10} \), we would shift firm \( p_1 \) to the highest position, then shift firm \( p_2 \) to the second-highest position, and so on. Formally, we shift type-\( p \) firms in rising order of \( p \) to their NAM-rank \( \bar{x}(p) = 1 - [F(p) - F(p_0)] \). At firm \( p \)'s turn, i.e. in matching \( P^p(\cdot) \), lower-skill firms \( p' < p \) have already been shifted to their NAM rank in \( \bar{x}(p') > \bar{x}(p) \), and so firm \( p \) starts at \( \bar{x}(p') \leq \bar{x}(p) \) and is shifted past higher-skill firms \( P^p(x) \geq p \) for all \( x \in [\bar{x}(p), \bar{x}(p)] \).

Lemma 5 implies that at each stage,

\[ s_{xp}(p, x, P^p(\cdot)) = [\lambda_{xp}(p)\phi(\hat{p}) - \phi_x(p)\lambda_q(\hat{p})] (1 - \gamma / x) < 0 \]

by (11). Using (34), the marginal effect of increasing a firm’s rank is always positive,

\[ s_x(p, x, P^p(\cdot)) = s_x(P^p(x), x, P^p(\cdot)) - \int_{P^p(x)} P^p(p) s_{xp}(\bar{p}, x, P^p(\cdot)) d\bar{p} > 0 \]

(40)

since \( p < P^p(x) \). Thus, the integrand in (31) is positive for all \( p \in [p, \bar{p}] \) and \( x \in [\bar{x}(p), \bar{x}(p)] \) and NAM maximizes aggregate surplus. The analogue argument implies that PAM minimizes aggregate surplus.

Finally we establish that the planner indeed wants the firms with the highest screening skills to
enter the market. To see this, differentiate firm $p$’s contribution to surplus (30) with respect to $p$

$$s_p(p, x, P(\cdot)) = \lambda_p(Q(x), p) + \int_{\mathbb{R}} \chi_p(\bar{x}; p, x)\lambda_q(Q(\bar{x}), P(\bar{x}))d\bar{x}$$

$$= \lambda_p(Q(x), p) - \phi_p(Q(x), p, x)\int_{\mathbb{R}} \exp \left( - \int_{\mathbb{R}} \phi_q \right) \lambda_q(Q(\bar{x}), P(\bar{x}))d\bar{x}$$

$$= \lambda_p(Q(x), p) - \left( \frac{\lambda_p(Q(x), p)}{x} \right) x \left[ 1 - \exp \left( - \int_{\mathbb{R}} \frac{\lambda_q(Q(\bar{x}), P(\bar{x}))}{\bar{x}} d\bar{x} \right) \right]$$

$$= \lambda_p(Q(x), p) \exp \left( - \int_{\mathbb{R}} \frac{\lambda_q(Q(\bar{x}), P(\bar{x}))}{\bar{x}} d\bar{x} \right) > 0.$$  

where the second equality uses the definition of the externality $\chi$ in (29), and the third equality uses the definition of $\gamma$ (33) and its evaluation in (39).

### C.3 Wage Caps

Restricting firms to a single wage $\bar{w}$ implements a random screening order, which is constrained efficient by the proof of Theorem 3. The wage level $\bar{w}$ does not affect the screening order, but rather pins down the marginal entering firm $\bar{p} = \bar{p}(\bar{w})$ via the indifference condition $\int_{F(\bar{p})}^1 \lambda(Q(x), \bar{p})dx = \bar{w}$. The optimal wage level $\bar{w}$ is the one that maximizes surplus $\int_{\mathbb{R}}^1 \int_{F(p)}^1 \lambda(Q(x), p)dx - w\exp dF(p)$. Since the marginal firm $\bar{p}$ exerts a negative externality on firms $p > \bar{p}$, we know that firms pay workers more than their outside option, $\bar{w} > w$.

Here we argue that firms have no incentives to underbid, offering $w < \bar{w}$; hence, making $\bar{w}$ a wage cap that firms are free to underbid also implements the second-best outcome. To see this, we need to check that the marginal firm $\bar{p}$ (which is most tempted to underbid) does not want to cut its wage to $w$ (which is the most profitable deviation). Indeed, note that at $w = w$ firm $\bar{p}$ exerts no externality on other firms or workers, so captures its full contribution to social surplus. By definition of $\bar{p}$, this social surplus is zero when the firm offers $\bar{w}$ and screens at a random rank, as instructed by the planner. When the firm disobeys the planner and posts $w = \underline{w}$, the contribution to social surplus is thus negative, meaning the firm’s profits are also negative.

### C.4 Second-Price Auction

Two firms $i$ have mass $n_i$ vacancies and screening skills $p_i^i$. They compete for workers in a second-price auction with outside option $w$. Suppose each firm consists of many independent divisions, who share screening skills and are bound by a common wage policy, but do not share information about applicants.

**Proposition 6.** The second-price auction with two firms has an equilibrium in weakly dominant strategies; this equilibrium exhibits negative assortative matching.

**Proof.** When firm $i$ screens first, applicant quality follows (3), with $P(x) = p_i^i$ for $x \geq 1 - n_i$ and $P(x) = p_j^j$ for $x \leq 1 - n_i$. Firm $i$’s productivity is then given by $\bar{\lambda}_i := \int_{1-n_i}^1 \lambda(Q(x), p_i^i)dx$ and firm $j$’s by $\Lambda_j := \int_{1-n_i}^{1-n_j} \lambda(Q(x), p_j^j)dx$. As usual in second-price auctions, bidding one’s value $w_i^i = \underline{w} + \bar{\lambda}_i - \Lambda_i$ is a weakly dominant strategy. When both firms bid their values, firm $i$ wins
(and screens first) if and only if $\bar{\lambda}^i - \lambda^i \geq \bar{\lambda}^j - \lambda^j$; that is, if and only if having firm $i$ screen first maximizes aggregate surplus. Since negative assortative matching maximizes surplus by Theorem 3, the low-skilled firm bids the higher wage in equilibrium to avoid the externality imposed by the high-skilled firm.

C.5 Production Complementarities

Theorem 3 is in stark contrast to the efficiency of equilibrium in the standard assignment model, where equilibrium also features PAM when the production function is supermodular. We now show that equilibrium continues to be inefficient even with exogenous complementarities. For simplicity, we posit a multiplicative production function, $\omega(q,p) = h(p)\lambda(q,p) - w$ with $h, h_p > 0$.

**Proposition 7.** Suppose firm $p$’s output equals $h(p)\lambda(q,p) - w$.

(a) Equilibrium exists, and is unique. All firms above some threshold $p \geq p^*$ enter and sort according to PAM.

(b) If $h(p) = p$, PAM is inefficient.

**Proof.** (a) Surplus $\omega$ is increasing in $p$ and supermodular, $\omega_{qp} = \lambda_{qp}h + \lambda_qh_p > 0$, so in equilibrium firms $p \geq p^*$ enter and matching is PAM. Equilibrium existence then follow as in Theorem 2. Indeed, applicant quality $Q(x)$ and recruit quality $\lambda(Q(x),p)$ are identical to the baseline model, and the complementary production function only affects equilibrium wages and profits.

(b) To check the necessary condition for efficiency (35), note first that

$$\frac{\lambda_{qp}}{\lambda_q} = -\frac{1}{1-p} + 2\frac{1-q}{1-p(1-q)} = \frac{2(1-p)(1-q) - (1-p(1-q))}{(1-p)(1-p(1-q))}$$

and

$$\frac{\phi_p}{\phi} = \lambda_p \cdot \frac{1}{\lambda - q} = \frac{q(1-q)}{(1-p(1-q))^2} \cdot \frac{1-p(1-q)}{q(1-q)} = \frac{1}{1-p(1-q)} \cdot \frac{1}{p}.$$ 

Thus, equilibrium is inefficient (and could be improved by re-ordering low-wage firms) if

$$\frac{\phi_p}{\phi} - \frac{\omega_{qp}}{\omega_q} = \left(\frac{\phi_p}{\phi} - \frac{\lambda_{qp}}{\lambda_q}\right) - \frac{h'}{h} = \frac{q}{(1-p)(1-p(1-q))} + \frac{1}{p} - \frac{h'}{h}$$

is positive. This depends on the degree of supermodularity. For the standard specification with $h(p) = p$, it is positive. That is, the effect of reducing the compositional externality outweighs two sources of supermodularity: one in the recruiting function $\lambda(q,p)$, and another one in the production function $\omega = \lambda p - w$.

C.6 Screening Costs

Here we show that Theorems 2 and 3 extend to a model with screening costs, where net surplus is given by $\omega(q,p) = \lambda(q,p) - \kappa/(1-p(1-q)) - w$. Observe that screening costs $\kappa/(1-p(1-q))$ fall in applicant quality $q$ but increase in screening skills $p$ since more skillful firms interview more candidates.\(^{34}\)

\(^{34}\)By adopting this surplus function, we implicitly assume that firms prefer to screen candidates rather than hiring a random, unscreened candidate. This mild condition is satisfied if (i) the minimum wage $w$ is
Proposition 8. Suppose there is a cost $\kappa \geq 0$ to screen each applicant.

(a) Equilibrium exists, and is unique. All firms above some threshold $p \geq p^*$ enter and sort according to PAM.
(b) PAM is inefficient.

Proof. (a) As in Theorem 2, skilled firms post higher wages since

$$\omega_{qp} = \lambda_{qp} - \left( \frac{\kappa}{1 - p(1 - q)} \right)_{pq} = 1 - p(1 - q) - 2q + \frac{1 + p(1 - q)}{(1 - p(1 - q))^{\kappa}} \kappa > 0.$$ 

As before, the first term is positive since $\lambda = q/(1 - p(1 - q)) < 1/2$. Additionally, the second term is always positive. Thus $\omega(q, p)$ is supermodular on $q \in [0, \hat{q}]$, as required. As for entry, note that $\omega_p = \omega(1 - q)/(1 - p(1 - q)) \geq 0$. Equilibrium existence and uniqueness then follow as in Theorem 2.

(b) We apply Lemma 4 to show that PAM is inefficient. Indeed

$$\frac{\omega_{qp}}{\omega_q} = 2 \frac{1 - \kappa}{1 - p(1 - q)} - \frac{1 - \kappa}{1 - p(1 - \kappa)} < \frac{1 - \kappa}{1 - p(1 - q)} = \frac{\lambda_p}{\lambda} < \frac{\lambda_p}{\lambda - q} = \frac{\phi_p}{\phi}$$

where the first inequality follows because $\kappa < q$, which in turn follows from $\omega(q, p) = (q - \kappa)/(1 - p(1 - q)) - w \geq 0$.

Proposition 8 shows that the main insights of our paper carry over to a model with screening costs. This extension also creates a role for different information structures because our “perfect bad news” screening is no longer without loss. We can show that Proposition 8 extends to “symmetric signals” where high (resp. low) types pass the screening tests with probability $p$ (resp. $1 - p$), and firms differ in their screening skill $p \in (0.5, 1)$.

However, if firms receive “perfect good news” information whereby untalented applicants fail all tests, while firms differ in their probability $p$ of identifying talented workers, then equilibrium matching is NAM and equilibrium is efficient. Intuitively, all firms are equally effective at screening and hire only talented workers, $\lambda = 1$, but skilled firms are more efficient and need not search as long to find a talented worker. We dislike this signal structure because of its counter-factual predictions that all firms hire workers of the same quality, and high-quality firms have lower search expenditure.

D Proofs from Section 4

D.1 Derivation of Equation (17) and Proof that $V(r)$ is Convex

The envelope theorem applied to (14) implies

$$V_r(r(t), t) = \int_t^\infty e^{-\beta(s-t)} \frac{\partial r(s)}{\partial r(t)} ds.$$ 

sufficiently high such that $\omega(q, p) > q$ for any firm $p$ willing to hire a worker, $\lambda(q, p) \geq w$, or (ii) screening additionally screens out unmodelled “terrible” types of workers.
To compute the integrand, we write the solution of the talent evolution (13) as a function of its initial condition \( r_s(s, r) = \xi(s, r), s \) where \( \xi(s, r) = \alpha(\lambda(Q(s), \psi(r)) - r) \) and \( r(t, r) = r \). As in the proof of Lemma 1, Hartman (2002, Theorem 3.1) implies \( r_{st} = \xi_r(r) \) with boundary condition \( r_s(t, r) = 1 \). Hence

\[
\frac{\partial r(s)}{\partial r(t)} = r_L(s, r(t)) = \exp \left( \int_s^t \xi_u(r(u), u)du \right) = \exp \left( -\alpha \int_s^t [1 - \lambda_p(Q(u), \psi(r(u)))\psi_r(r(u))]du \right)
\]

implying (17).

To see that \( V(r, t) \) is convex in \( r \), we differentiate again to obtain

\[
V_{rr}(r(t), t) = \int_t^\infty e^{-\beta(s-t)} \frac{\partial^2 r(s)}{\partial r(t)^2} ds =
\]

\[
= \int_t^\infty e^{-\beta(s-t)} \frac{\partial r(s)}{\partial r(t)} \alpha \int_s^t \left[ \lambda_{pp}(Q(u), \psi(r(u)))\psi_r(r(u))^2 + \lambda_p(Q(u), \psi(r(u)))\psi_{rr}(r(u)) \right] \frac{\partial r(u)}{\partial r(t)} du ds
\]

which is positive since all four derivatives \( \lambda_p = \frac{q(1-q)}{(1-p(1-q))^2}, \lambda_{pp} = \frac{2q(1-q)^2}{(1-p(1-q))^2}, \psi_r, \psi_{rr} \) are positive, the first three strictly. For an intuition consider the random recruiter example where \( \psi(r) = p^L + r(p^H - p^L) \). Intermediate levels of recruiting skills \( r \) have the drawback that the firm’s wage must strike a compromise between the firm’s low-skill and high-skill recruiters, while with homogeneous recruiters, \( r = 0 \) or 1, can choose the optimal wage for all.

### D.2 Proof of Theorem 4

Here we complete the proof of Theorem 4 by arguing 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 \( r(t) \) and write \( [x^0, x^1] \) for the talent-ranks of these firms. Since optimal wages rise in talent and hence talent differences never vanish, firms in the atom must have identical talent \( r(s) \) for all \( s < t \). At any time \( s < t \) 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, posting both high and low wages to attract good and bad applicants; they must thus be indifferent across a range of applicants \( [q^0(s), q^1(s)] \) for all \( s < t \). Thus, the first order condition (16) must hold with equality on \( [q^0(s), q^1(s)] \) for all \( s < t \) 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 \( q^1(s) \), rather than mixing over good and bad applicants. At time \( s = 0 \), the choice \( q^1(0) \) 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 (16), strictly increases in \( r \), this deviation strictly improves on the posted distributional strategy. This proves that initially identical firms diverge immediately.

---

35When using a distributional strategy, a firm posts an entire distribution of wages \( \nu = \nu(w, t) \) of wages at any time \( t \); we then interpret \( \psi(R(x, t)) \) as the weighted-average skill of firms posting the \( x \)-ranked wage, and solve for the firm’s evolution of talent by taking expectations over the RHS of (13).
D.3 Proof of Theorem 5

Here we show that firm $x$’s talent $R(x, t)$ and applicant quality $Q(x, t)$ converge to their steady state levels $R^*(x)$ and $Q^*(x)$. For constant applicants $Q(x, t) \equiv q$, talent drifts towards $\rho(q)$. The complication is that firm $x$’s applicant quality also changes over time, with $Q_x(x, t)$ given by (12) and $r_t(x, t)$ by (13).

First, we establish a contraction property. Define the limits $\underline{Q}(x) := \liminf_t Q(x, t)$, $\bar{Q}(x) := \limsup_t Q(x, t)$, $\underline{R}(x) := \liminf_t R(x, t)$, and $\bar{R}(x) := \limsup_t R(x, t)$. Next, interpret (12) as an operator $\mathcal{Q}$, mapping firm quality functions $R(\cdot, t)$ into applicant quality functions $Q(\cdot, t) = \mathcal{Q}[R(\cdot, t)](\cdot)$. We claim that:

$$\mathcal{Q}[\rho(Q(\cdot))](x) \leq \underline{Q}(x) \leq Q(x) \leq \bar{Q}(x) \leq \mathcal{Q}[\rho(\bar{Q}(\cdot))](x). \quad (42)$$

To understand (42), first observe that that $\mathcal{Q}$ is antitone: if $R(x) \geq \hat{R}(x)$ for all $x$ then $Q(x) = \mathcal{Q}(R(\cdot))(x) \leq \mathcal{Q}(\hat{R}(\cdot))(x) = \bar{Q}(x)$, since $Q(1) = \bar{Q}(1) = \bar{q}$ and $\phi(q, \psi(r), x)$ increases in $r$. Intuitively, better recruiters introduce more adverse selection. Inequalities (42) then state that if applicant quality was equal to one of its limits, $Q$ and $\bar{Q}$, and talent $r$ was in steady state, then the induced difference in applicant pools is larger than the original difference.

We prove (42) in two steps. First, since $R(x, t)$ drifts towards $\rho(Q(x, t))$, which is asymptotically bounded by $\rho(\underline{Q}(x))$ and $\rho(\bar{Q}(x))$, we have

$$\rho(\underline{Q}(x)) \leq R(x) \leq \bar{R}(x) \leq \rho(\bar{Q}(x)) \quad (43)$$

for all $x$. Second,

$$\underline{Q}(x) = \lim_{t \to \infty} \inf_{t' > t} Q(x, t') = \lim_{t \to \infty} \inf_{t' > t} \{\mathcal{Q}[R(\cdot, t')](x)\} \geq \lim_{t \to \infty} \mathcal{Q}[\sup_{t' > t} \{R(\cdot, t')\}](x) = \mathcal{Q}[\bar{R}(\cdot)](x)$$

where the first equality is the definition of the lim inf, and the second the definition of the operator $\mathcal{Q}$. The inequality uses the antitonicity of $\mathcal{Q}$: since $\bar{R}(\hat{x}, t') \leq \sup_{t' > t} \{R(\hat{x}, t')\}$ for all $t' > t$ and $\hat{x}$, we know that $\mathcal{Q}[\bar{R}(\cdot, t')](x)$ (weakly) exceeds $\mathcal{Q}[\sup_{t' > t} \{R(\cdot, t')\}](x)$ for all $t' > t$ and $x$, and hence so does $\inf_{t' > t} \mathcal{Q}[R(\cdot, t')](x)$. The last inequality uses the dominated convergence theorem to exchange the limit $t \to \infty$ and the operator $\mathcal{Q}$, as well as the definition of the lim sup, $\bar{R}(x) = \lim_{t \to \infty} \sup_{t' > t} R(x, t')$. Together with the analogue argument that $\bar{Q}(x) \leq \mathcal{Q}[R(\cdot)](x)$, and applying the antitone operator $\mathcal{Q}$ to (43) we get

$$\mathcal{Q}[\rho(\bar{Q}(\cdot))](x) \leq \mathcal{Q}[\bar{R}(\cdot)](x) \leq \bar{Q}(x) \leq \mathcal{Q}[R(\cdot)](x) \leq \mathcal{Q}[\rho(\bar{Q}(\cdot))](x)$$

establishing (42).

To complete the proof of convergence, suppose “inductively” that applicant and firm quality converge above some $\hat{x} \in (0, 1]$, i.e. $Q(x) = \bar{Q}(x)$, and hence $R(x) = \bar{R}(x)$, for all $x \in (\hat{x}, 1]$. Fix $\epsilon$, and let $\delta(\epsilon) := \max_{x \in [\hat{x} - \epsilon, \hat{x}]} |\bar{Q}(x) - \bar{Q}(x)|$ be the maximum distance the liminf and limsup drift
apart over \([\bar{x} - \epsilon, \bar{x}]\). Since \(\rho(q)\) is locally Lipschitz in \(q\) with constant \(K\),\(^{36}\) we have

\[
\max_{x \in [\bar{x} - \epsilon, \bar{x}]} \left| \rho(Q(\cdot))(x) - \rho(\hat{Q}(\cdot))(x) \right| \leq K \delta(\epsilon).
\]

Next, since \(Q[R(\cdot)](x)\) solves (12), the RHS of which is locally Lipschitz in \(q\) and \(r\) with constant \(K'\), and choosing \(\epsilon < 1/K'(1 + K)\) we get

\[
\max_{x \in [\bar{x} - \epsilon, \bar{x}]} \left| Q[\rho(Q(\cdot))](x) - Q[\rho(\hat{Q}(\cdot))](x) \right| \leq K' \epsilon (1 + K) \delta(\epsilon) < \max_{x \in [\bar{x} - \epsilon, \bar{x}]} \left| \hat{Q}(x) - Q(x) \right|.
\]

contradicting (42). Hence we must have \(\hat{Q}(x) = Q(x)\) and hence \(R(x) = \hat{R}(x)\), for all \(x \in [\bar{x} - \epsilon, 1]\), and thus for all \(x \in [0, 1]\).

### D.4 How Turnover affects Wage and Productivity Dispersion

Equation (22) implies that \((\beta + \alpha)(W^*(Q^*(x')) - W^*(Q^*(x)))\) rises in \(\alpha\) for all \(x' > x\), and hence steady-state wage variation grows in the dispersive order. Here we argue that \(W^*(Q^*(x'))/W^*(Q^*(x))\) and \(\Pi^*(Q^*(x'))/\Pi^*(Q^*(x))\) increase in \(\alpha\) for all \(x' > x\), and hence wage and profit variation also grows in the more demanding log-dispersive order.

Formally, we argue that steady-state wages and profits as functions of applicant quality and turnover \((\beta + \alpha)W^*(q, \alpha)\) and \(\Pi^*(q, \alpha)\) are log-supermodular in \(q\) and \(\alpha\). This implies that they are also log-supermodular as functions of rank \(x\) and turnover \(\alpha\) since applicant quality \(q = Q(x)\) is a monotone transformation of \(x\) that does not depend on turnover \(\alpha\).

Recall that a function \(\xi(q, \alpha)\) with \(\xi(0, \alpha) = 0\) is log-supermodular in \(q, \alpha\) if its partial derivative \(\xi_q(q, \alpha)\) is log-supermodular: For then

\[
(\log \xi(q, \alpha))_q = (\log \int_0^q \xi_q(\hat{q}, \alpha) d\hat{q})_q = \frac{\xi_q(q, \alpha)}{\int_0^q \xi_q(\hat{q}, \alpha) d\hat{q}}
\]

rises in \(\alpha\), since \(\xi_q(q, \alpha)/\xi_q(\hat{q}, \alpha)\) rises in \(\alpha\) for all \(\hat{q} < q\).

Thus, it suffices to show that marginal wages and profits (22) and (23) are log-supermodular in \(q\) and \(\alpha\). The only factor that depends on both \(q\) and \(\alpha\) is \(1/\eta(q, \alpha) := 1/ (\beta + \alpha (1 - \lambda_p(q, \psi(\rho(q)))) \psi_r(\rho(q)))\), and indeed

\[
(\log \eta(q, \alpha))^{-1} q, \alpha = \eta_p \eta_q - \eta_q \eta \eta = \frac{-(1 - \lambda_p \psi_r)(\alpha \lambda_p \psi_r)_q + (\lambda_p \psi_r)_q (\beta + \alpha (1 - \lambda_p \psi_r))}{\eta^2} = \frac{(\lambda_p \psi_r)_q \eta^3}{\eta^2} > 0
\]

since \((\lambda_p(q, \psi(\rho(q)))) \psi_r(\rho(q))_q = (\lambda_p + \lambda_q \psi_r \rho_q) \psi_r + \lambda_p \psi_r \rho_q > 0\).

\(^{36}\)Indeed, recall that \(\rho(q) = r \in [0, 1)\) solves \(\lambda(q, \psi(r)) - r = 0\). Thus, by convexity of the LHS we have \(\lambda_p(q, \psi(\rho(q))) \psi_r(\rho(q)) - 1 < (\lambda(q, \psi(1)) - 1)/(1 - \rho(q)) < 0\) and so \(R_q(q) = -\frac{\lambda_p(q, \psi(\rho(q))) \psi_r(\rho(q))}{1 - \lambda(q, \psi(1)))} < 0\).
References


