Signaling to Experts[*]

Pablo Kurlat **Florian Scheuer**

Stanford University University of Zurich

February 2018

Abstract

We study competitive equilibria in a signaling economy with heterogeneously informed buyers. In terms of the classic Spence (1973) model of job market signaling, firms have access to direct but imperfect information about worker types, in addition to observing their education. Firms can be ranked according to the quality of their information, i.e. their expertise. In equilibrium, some high type workers forgo signaling and are hired by better informed firms, which make positive profits. Workers' education decisions and firms' use of their expertise are strategic complements, allowing for multiple equilibria. We characterize wage dispersion and the extent of signaling as a function of the distribution of expertise among firms. The market can create insufficient or excessive incentives for firms to acquire information, and we provide a formula to measure this inefficiency. Our model can also be applied to a variety of other signaling problems, including securitization, corporate financial structure, insurance markets, or dividend policy.

1 Introduction

We study competitive markets with the following features: Sellers are privately informed about their own type; they can take a publicly observable action that is differentially costly for different types; buyers can directly observe imperfect information about sellers' types; and the quality of this information is heterogeneous across buyers. The first two features define a standard signaling environment.[1] Our objective is to move beyond the special case, studied extensively, where buyers are completely uninformed and rely exclusively on the public signal to form beliefs about sellers' types. Instead, we investigate the effect of adding the third and fourth features, buyers' heterogeneous direct information, on equilibrium prices and allocations.

---

[1]Throughout, we refer to a *signaling* rather than a *screening* problem. Traditionally, which term is used depends on which party proposes contract terms. Since in our setup there are markets for all possible contracts, the distinction vanishes.

Our running example is an extension of the canonical Spence (1973) model of job market signaling: workers have private information about their own productivity; education is purely wasteful but is more costly for less productive workers so it can be used to signal; and firms have heterogeneous expertise in directly assessing workers' productivities, in addition to verifying their education. For instance, firms have access to such direct information through tests, interviews, referrals or trial periods, and differ in their ability to extract accurate predictions from them. We ask how differences in recruiting expertise across firms affect the equilibrium: What wages do more- versus less-expert firms offer, which workers do they hire, how much profit do they make, what education levels do they require, and what are the implications for social welfare?

While we exposit our setup and results in terms of this application, our model is general and can be used to answer these basic questions for many signaling and screening problems. How do investors' abilities to directly assess a company's profitability affect IPO prices, incentives for insiders to retain undiversified shareholdings, and the payment of positive dividends? What are firms' incentives to engage in costly brand-building or to offer warranties if consumers have heterogeneous ability to find out about product quality directly, e.g. by studying product reviews? How does the use of different risk assessment models across insurance companies affect equilibrium deductibles and premiums? What are the effects of buyers' heterogeneous pricing technologies on the design and tranching of asset-backed securities?

Returning to labor markets, we focus on the most parsimonious setting with two worker types and consider configurations for firms' direct information that allow us to rank firms by their expertise, i.e. their probability of making mistakes: the "false positives" case where firms may observe good signals from low-productivity workers and the opposite case, with "false negatives." We assume that each firm hires a single worker; such capacity constraints are crucial to rule out trivial solutions where the most-expert firms hire all workers.

Our first task is to define a notion of competitive equilibrium that applies to this environment. We assume that each combination of a wage and an education level defines a separate market. Any worker is allowed to apply for a job in any market (provided he acquires the level of education prescribed by that market) and any firm can recruit in any market. For workers, markets are partially exclusive: naturally, they commit to a single education level but can apply for jobs at many different wages, effectively setting a reservation wage. Markets do not necessarily clear: in any given market, workers can apply for jobs and not get them and firms can try to hire workers and not find them. When hiring, firms need not hire randomly from the pool of applicants: they can reject some applicants and only hire from among those they find acceptable, but only to the extent that their own direct information allows them to tell workers apart. Equilibrium requires that workers' expectations of their chances of finding work in each market and firms' beliefs about what workers they will encounter in each market

be consistent with each other and with firm recruiting and worker education decisions.

As is common in signaling models, the set of equilibria depends on what beliefs agents can entertain regarding markets where in equilibrium there is no trade. A crucial technical contribution of this paper is to construct restrictions on these out-of-equilibrium beliefs that deliver a unique and plausible equilibrium in the familiar uninformed-buyers benchmark, yet still guarantee equilibrium existence and tractability in the general case of heterogeneous expertise. We propose the following conditions: First, for any market where a firm has well-defined beliefs about what acceptable workers it would encounter, these beliefs can only place weight on workers who would find it (weakly) optimal to apply to that market. Second, if a firm does not have well-defined beliefs about acceptable workers it would encounter, we impose that any workers that would be acceptable to the firm must expect that, if they were to apply for a job in that market, they would get one for sure.

For the benchmark where firms have no direct information, our definition ensures that the least-cost separating allocation is the unique equilibrium. Our refinement implies that pooling is inconsistent with equilibrium: at slightly higher education levels than a putative pooling allocation, firms must believe that they will only encounter high type workers because they are the ones most willing to choose higher education, and therefore firms could profitably deviate.

For the false positives case, the following pattern emerges. Low worker types get no education and high types get either no education or enough education to fully separate. Firms with sufficiently accurate information recruit zero-education workers at a wage w^* that leaves high productivity workers indifferent between signaling and not signaling, and make positive profits. These firms face both high and low productivity applicants, so they can only profit if they are able to reject a sufficient proportion of low types. Firms with less accurate information recruit either educated workers at a wage equal to the high types' productivity, or zero-education workers at a wage equal to low types' productivity, and make zero profits in either case. Two simple conditions summarize any equilibrium: an indifference condition that requires the marginal firm to make zero profits by hiring zero-education workers at wage w^*, and a market clearing condition requiring high type workers who forgo education to indeed find jobs at wage w^*.

This tractable structure allows us to study comparative statics. We find that more high types forgo education if their signaling cost is higher. This intuitive property does *not* hold in the uninformed-firms benchmark, where the total amount of education does not respond to high types' costs. With heterogeneity among firms, a higher cost of education raises the cost of fully separating. This lowers the wage w^* that high types are willing to accept in order to avoid signaling, which in turn makes it worthwhile for more lower-expertise firms to hire them. We also find that an increase in demand for workers results in more high types forgoing education, another empirically appealing property that cannot be generated in the uninformed-firms benchmark. Finally, an improvement in the distribution of expertise leads to

less signaling. Hence, our model can explain, for instance, the share of firms paying dividends, the retention of equity by entrepreneurs, or the degree of under-insurance, as a function of the quality of investors' (or insurers') direct information.

Our model features strategic complementarities between high-quality workers' signaling decisions and firms' recruiting decisions. If enough high productivity workers forgo education, the pool of applicants in zero-education markets will improve. This induces less expert firms to recruit zero-educated workers, which in turn allows more high type workers to forgo education. As a result, the model may feature multiple equilibria, each with different proportions of high types choosing to forgo education. The least cost-separating allocation, where all high types get enough education to separate, is always one of these equilibria: if all high types signal, there is no hope to hire them without requiring the signal, and therefore firms' expertise is useless. When there are multiple equilibria, they can be Pareto-ranked. The signal is a pure deadweight cost, and the equilibrium with less signaling is preferred by everyone.

One feature of the classic signaling and screening model that has been criticized is a discontinuity as the buyers' prior becomes degenerate. The symmetric information case involves no signaling, but in the presence of even a minimal mass of low types, the high types must emit a non-trivial signal to separate. Our model offers a natural way to smooth out this stark property: there always exists an equilibrium that continuously approaches the full information limit, both as the share of low types vanishes and as buyers' direct information becomes perfect. A similar discontinuity arises in the standard signaling model when the signaling costs of the two types converge: whenever the costs differ, there is a discrete amount of signaling, but no signaling when they are equal. We show that our model overcomes this discontinuity as well.

We then ask whether the equilibrium provides the right incentives for firms to improve their expertise. Do firms invest the socially optimal amount in developing better evaluation methods for job applicants, pricing tools for financial securities, valuation models for company shares, or risk scoring for insurance applicants? Following the approach in Kurlat (2016b), we compute the ratio of the marginal social value to the marginal private value of improved expertise; a ratio greater than one implies that the market doesn't provide sufficient incentives to invest in expertise, and vice versa. We find that this ratio is equal to the product of the relative cost of education for high types (which captures the deadweight cost of signaling) and the elasticity of the fraction of high types who forgo signaling with respect to increased demand. The first term is, by assumption, lower than one; we show that the second is above one so the product could be higher or lower than one. Thus, the market could under- or over-provide incentives to invest in expertise, and our formula allows to determine the resulting inefficiency as a function of these two "sufficient statistics," both of which can in principle be measured.

Finally, we characterize equilibrium in the false negatives case, which we show to be essentially unique. Productive workers now make different choices depending on how transparent

they are, that is, how many firms are able to identify them as high types. Those most easily identified forgo education and are paid their productivity. Less transparent workers also forgo education but now earn a range of lower wages. They are hired in part by non-selective firms in markets where low types also apply, so wages must be low enough to allow these non-selective firms to break even on whatever pool of applicants they face. The least transparent productive workers instead resort to education in order to separate from low types. Therefore, our model provides a novel theory of wage dispersion among equally productive (and educated) workers based on how easy it is to evaluate their productivities.

The equilibrium is again pinned down by two conditions: an indifference condition for the marginal high type worker choosing whether to signal or not, and a market clearing condition that requires all high type workers who do not signal to find a job. In terms of comparative statics, we now find that the level of education needed to fully separate from low types is endogenous. For instance, unlike in the no-information benchmark and perhaps surprisingly, it is increasing in the demand for workers. The reason is that more demand means more selective hiring at every level of expertise, and thus less non-selective hiring. This makes low types worse off because they have a lower chance of being hired above their reservation wage. In turn, this increases how much education is necessary to separate from them. Hence, interestingly, our model with false negatives predicts that higher demand leads to polarization in signaling: fewer high types signal, but those who do ought to do so more intensely. We also characterize wage dispersion as a function of demand, signaling costs, and the productivity distribution.

Related Literature. This paper introduces heterogeneous expertise among buyers into the canonical competitive signaling and screening environments due to Spence (1973) and Rothschild and Stiglitz (1976). To this purpose, we develop a notion of equilibrium that builds on concepts proposed by Gale (1996), Guerrieri et al. (2010), Guerrieri and Shimer (2014) and Kurlat (2016a), all of which are based on the idea that different contracts define different markets and the probability of trade, rather than the price, is the market-clearing variable. This allows us to naturally incorporate capacity constraints among buyers and to study the extensive margin of trade, which is crucial in many relevant settings, such as labor or financial markets.

The way in which we model heterogeneity of information on the buyers' side—and hence their ability to distinguish between sellers based on their own direct assessment rather than just the publicly observable signal or screening device—is borrowed directly from Kurlat (2016a). However, Kurlat (2016a) studies a single-dimensional environment, where the set of contracts is just the set of prices, so public signaling is ruled out. Our paper instead incorporates a second dimension, allowing us to capture signaling or screening through, for example, education, underinsurance, equity retention, dividends or advertising. On a technical side, incorporating signaling requires us to model buyers' beliefs associated with off-equilibrium actions, a challenge

that we tackle here but that is not present in Kurlat (2016a). Similar to our paper, Gale (1996) and Guerrieri et al. (2010) also allow for general, multidimensional contracts. Relative to them, however, our contribution is to relax the assumption that buyers are completely and uniformly uninformed, by introducing heterogeneous information for buyers.

The refinement on beliefs that we impose is closely related to the D1 criterion proposed by Cho and Kreps (1987), the condition for a refined equilibrium proposed by Gale (1996), and the conditions on beliefs imposed by Guerrieri et al. (2010) for contracts that are not traded in equilibrium. It is based on the idea that, in markets with zero supply in equilibrium, buyers anticipate that, if they were to place demand there, they would only attract the sellers (among those they do not reject based on their direct assessment) who are willing to accept the lowest probability of trade. Indeed, our requirement that if a firm's beliefs about a market are not well defined, the workers that the firm would accept must be confident that they could get a job for sure is the natural generalization of the infinite-tightness condition imposed by Guerrieri et al. (2010) to our framework with heterogeneous information. The refinement eliminates the traditional reasons for multiplicity that emerge in signaling games when out-of-equilibrium beliefs are left unrestricted. By contrast, the multiplicity we find in the false positives case is due to an entirely orthogonal force, namely the strategic complementary between signaling and the use of expertise, which vanishes in the classic no-information benchmark.

More broadly, our work relates to the literature that followed Rothschild and Stiglitz (1976) on competition in multidimensional contracts with asymmetric information (see e.g. Miyazaki, 1977; Wilson, 1977; Dubey and Geanakoplos, 2002; Bisin and Gottardi, 2006; Netzer and Scheuer, 2014; and Azevedo and Gottlieb, 2017). Similarly, there is an extensive literature that has applied the Spence (1973) signaling model to various settings, including corporate finance (Leland and Pyle, 1977; Ross, 1977), dividend policy (Bhattacharya, 1979; Bernheim, 1991), security design (DeMarzo and Duffie, 1999; DeMarzo, 2005), and brand-building (Nelson, 1974; Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1986), to name a few. None of these two strands of literature, however, have attempted to move beyond the polar case where sellers are informed and buyers are uninformed. Our paper provides a general analysis of how heterogeneous information affects equilibrium in all these situations.

Daley and Green (2014) also study an environment where the possibility of signaling coexists with direct information ("grades"), and find conditions such that the equilibrium features either partial or complete pooling. They assume that grades are equally observable by all firms, so they have no role for expertise on the firm side. Feltovich et al. (2002) also consider an environment with (homogeneous) direct information in addition to signaling, and find that—in a model with three types—the highest types may refrain from signaling to distinguish themselves from the medium types, a behavior they refer to as "countersignaling." A similar feature emerges in our model in the false negatives case, where some high types separate through signaling while

others pool with low types in terms of the signal they emit, relying instead on expert buyers to identify them.

Board et al. (2016) share our interest in the idea that firms differ in their ability to tell apart high- and low-quality job applicants. In their setup, however, workers do not make any decisions, so whether or not they know their own productivity does not matter. This rules out any way in which workers may signal their private information, or be screened other than through firms' direct assessment of them. Instead, in our model, workers can emit a publicly observable signal, such as education, that can be used to convey information about their productivity. In addition, Board et al. (2016) assume that firms' direct information is independent across firms, whereas we work with a nested information structure where more-expert buyers know strictly more than less-expert ones.

The rest of this paper is organized as follows. Section 2 introduces the model and briefly illustrates a number of well-known applications. Section 3 provides our equilibrium definition and Section 4 shows that it gives rise to a unique equilibrium in the standard signaling environment where firms are uninformed. In Section 5, we characterize the set of equilibria with false positives, including its comparative-static and welfare properties, and Section 6 considers the case of false negatives. Finally, Section 7 concludes. All proofs are relegated to the Appendix.

2   The Economy

Our model is intended to capture a generic signaling setting. For clarity, we present our results in terms of Spence's original job market signaling model. However, the only critical assumptions are perfect competition, heterogeneous information, and the existence of some action (the signal) that is inefficient from a first-best point of view but involves different costs for different sellers. Our results therefore apply to any setting with these features, and we provide some alternative interpretations of the model below.

2.1   Job Market Signaling

There is a unit measure of workers indexed by i, uniformly distributed in the interval $[0, 1]$. Each worker is endowed with a single unit of labor. Worker i's productivity is

$$q(i) = \begin{cases} q_L & \text{if } i < \lambda \\ q_H & \text{if } i \geq \lambda \end{cases} \tag{1}$$

with $q_L < q_H$. We refer to workers with $i < \lambda$ and $i \geq \lambda$ as low and high types, respectively. A worker's index i is private information. Workers of the same type but different indices i all have the same productivity; they differ only in terms of how easy it is for firms to identify them, as

specified below.

Workers can choose a publicly observable level of education e, which has no effect on their productivity. If worker i chooses a level of education e and gets a job at a wage w, his utility is $w - c(i)e$, where

$$c(i) = \begin{cases} c_L & \text{if } i < \lambda \\ c_H & \text{if } i \geq \lambda. \end{cases}$$

We assume $c_L > c_H$, so low types experience a higher utility cost of obtaining education.

Up to here, the model coincides with the basic Spence (1973) signaling model. Our innovation is to introduce a notion of expertise, i.e. that firms possess heterogeneous information about the workers they encounter. Formally, there is a continuum of firms of measure greater than one, indexed by $\theta \in [0, 1]$. The measure of firms over $[0, 1]$ is denoted by F and assumed to be continuous. When firm θ analyzes worker i, it observes a direct signal

$$x(i, \theta) = \begin{cases} 0 & \text{if } i < \theta \\ 1 & \text{if } i \geq \theta. \end{cases} \tag{2}$$

If $\theta = \lambda$, this signal allows the firm to perfectly infer the worker's productivity. If $\theta < \lambda$, the firm makes "false positive" mistakes: it observes positive signals from a subset of the low type workers. If $\theta > \lambda$, the firm makes "false negative" mistakes. We assume that firms can be perfectly ranked by their expertise, so one of two cases applies: either F has support in $[0, \lambda]$ or it has support in $[\lambda, 1]$. We consider both configurations subsequently. In either case, θ measures the firm's expertise: the closer θ is to λ, the better the firm is at correctly identifying a worker's productivity.

Each firm can hire at most one worker. Equivalently, we could assume that buyers have limited funds (and are unable to borrow) to leverage their expertise, which may be more natural in some of our financial market applications sketched below. Either way, some form of capacity constraints are needed to keep the problem interesting by preventing the best-informed buyers from implementing all trades. If a firm hires worker i at wage w its profits are $q(i) - w$.

Thus, our key innovation compared to the canonical signaling model is that buyers have access to direct, even though imperfect, information about sellers, rather than relying exclusively on self-selection. Moreover, the quality of this information is heterogeneous.[2] For example, some managers have better judgement in assessing the talent of job applicants, as in Board et al. (2016), or recruiters may run tests or interviews (see e.g. Guasch and Weiss, 1980, and Lockwood, 1991). Another channel of direct information about workers is through referrals. For

---

[2]Whenever there is no heterogeneity across firms (so the support of F is concentrated at a single value of θ), our model collapses back to the standard signaling problem. If $\theta < \lambda$, then all workers $i \in [0, \theta)$ are fully identified as low types, and all $i \in [\theta, 1]$ look indistinguishable to all firms. Hence, the former group of workers get their first-best outcome, and a standard signaling model without expertise applies to the latter population, with a share of low types equal to $(\lambda - \theta)/(1 - \theta)$. Similarly, if $\theta > \lambda$, we obtain a standard signaling environment where the share of low types is λ/θ.

example, Beaman and Magruder (2012) and Burks et al. (2015) show empirically that better employees make more and better referrals, and that firms differ in the degree to which their employees can predict the performance of their referrals.

2.2   Other Interpretations

As is common to signaling models, the crucial feature is that the signal e is costly and satisfies a single-crossing property. For the job market signaling application, single crossing can be verified by letting $u(w, e) = w - c(i)e$ and computing the marginal rate of substitution:

$$-\frac{\partial u(w, e)/\partial e}{\partial u(w, e)/\partial w} = c(i),$$

which is higher for low types.

There are many other signaling settings that are formally isomorphic to our baseline model. We briefly describe four of them in the following.

Securitization.   Consider first the security design problem of DeMarzo and Duffie (1999). A continuum $i \in [0, 1]$ of originators each own a pool of assets that generate future cash flow y. The distribution of these cash flows is privately known to the originators, and given by $G_L(y)$ if $i < \lambda$ and $G_H(y)$ if $i \geq \lambda$, where G_H first-order stochastically dominates G_L, and they have common support. The originators prefer receiving cash over holding their risky assets, for instance because they have access to other profitable investment opportunities, or because they have superior ability in valuing assets and therefore want to raise cash to fund new asset purchases. Formally, they value future cash flows from their unissued assets at discount factor $\alpha < 1$. They face a pool of small, heterogeneously informed, buyers who do not discount, so the efficient allocation calls for selling all assets. Of course, due to their private information, the originators face a lemons problem when selling their assets. To raise cash, they therefore issue a limited-liability security backed by their assets. DeMarzo and Duffie (1999) show that, under general conditions, it is optimal to sell a high-quality, senior claim to the assets (i.e. debt) and retain the remaining, risky equity tranche as "skin in the game," i.e. a signal of asset quality. Let Y denote the upper bound of G_k, $k = L, H$; let $Y - e$ denote the face value of the debt tranche, and w denote its price per unit of face value. Then issuer i's payoff is

$$u(w, e) = (Y - e)w + \alpha \int \max\{y + e - Y, 0\} \, dG_k(y)$$

with $k = L$ if $i < \lambda$ and $k = H$ if $i \geq \lambda$. The marginal rate of substitution is

$$-\frac{\partial u\left(w,e\right)/\partial e}{\partial u\left(w,e\right)/\partial w} = \frac{w - \alpha\left[1 - G_k\left(Y - e\right)\right]}{Y - e}.$$

By first-order stochastic dominance (FOSD), this is higher for low types and therefore satisfies single crossing. Finally, suppose each buyer demands one unit of face value of the asset-backed security. Then the buyer's payoff is $q_k(e) - w$ just like in our baseline model, where

$$q_k(e) \equiv \int \min\left\{\frac{y}{Y - e}, 1\right\} dG_k\left(y\right)$$

because each unit of the security has face value $Y - e$, so buying one unit of face value means buying $1/(Y - e)$ securities.

Our model thus captures the equilibrium in this classic tranching problem with the additional feature that buyers are heterogeneously informed about the quality of the asset-backed security. This may involve differential knowledge of aspects of the underlying asset pool or, more importantly, special expertise in the pricing of these securities (such as proprietary pricing models). For instance, Bernardo and Cornell (1997) provide empirical evidence for significant variation in valuations of mortgage-backed securities (with the winning bid exceeding the median bid by over 17% on average) even though all buyers were sophisticated investors or intermediaries. They conclude that this variability is due to differences in pricing technology. Mattey and Wallace (2001) document heterogeneity of this variability across different mortgage-backed securities, suggesting that some securities are easier to price than others.

Financial Structure of Firms.   Our next example is a variant of the corporate finance problem studied by Leland and Pyle (1977). Each entrepreneur i owns a project whose future payoff, privately known, is given by (1). As in the previous example, entrepreneurs are impatient, so their own valuation for their project's return is $\alpha q\left(i\right)$, and they wish to sell their project to heterogeneously informed investors. To signal the quality of their project, entrepreneurs can publicly announce that they will retain a fraction e of the equity of their firm. If an entrepreneur sells a fraction $1 - e$ of his firm at a price per unit of w then his utility will be $w\left(1 - e\right) + \alpha q\left(i\right)e$. The marginal rate of substitution is

$$-\frac{\partial u\left(w,e\right)/\partial e}{\partial u\left(w,e\right)/\partial w} = \frac{w - \alpha q\left(i\right)}{1 - e},$$

which, again, is higher for low types. If an investor buys one unit of firm i at a unit price w his profits are $q\left(i\right) - w$. Heterogeneous information among investors could be the result of differential experience in this particular industry, differential contacts with company insid-

ers, or differential access to analyst reports, which make some investors better than others at distinguishing good from bad projects.[3]

Insurance. Our model can also be mapped into the Rothschild and Stiglitz (1976) insurance problem. A continuum $i \in [0, 1]$ of risk-averse households each have wealth X and will suffer a loss of d with probability $1 - q(i)$. $q(i)$ is given by (1) and is privately known to the household. They face a pool of small, risk-neutral, heterogeneously informed insurance companies, so the efficient allocation calls for households to be fully insured. Insurance companies offer policies that cover the loss d minus a deductible e, in exchange for an up-front premium $(1 - w)(d - e)$, so that $1 - w$ is the implicit probability of loss that makes the insurance contract actuarially fair. If a household gets contract (w, e), its utility is

$$u(w, e) = qv(X - (1 - w)(d - e)) + (1 - q)v(X - (1 - w)(d - e) - e),$$

where $v(\cdot)$ is the household's von Neumann-Morgenstern utility function. The marginal rate of substitution is

$$-\frac{\partial u(w, e)/\partial e}{\partial u(w, e)/\partial w} = \frac{1}{d - e}\left(w - \frac{q}{q + (1 - q)\frac{v'(X - (1-w)(d-e)-e)}{v'(X - (1-w)(d-e))}}\right).$$

It is straightforward to show that this is decreasing in q and therefore satisfies single crossing. If an insurance company covers one unit of losses from household i at an implicit probability $1 - w$, then its profits are $1 - w - (1 - q(i)) = q(i) - w$.[4] Heterogeneous information among insurance companies could be the result of some of them having larger actuarial databases or more sophisticated predictive models that allow them to tell apart riskier from safer types.

Dividend Policy. Finally, consider the dividend puzzle, which observes that firms pay dividends even though their tax treatment is less favorable than that of share repurchases. The dividend signaling hypothesis (going back to Bhattacharya, 1979) explains this corporate payout policy by viewing dividends as a costly signal to convey private information about profitability (see e.g. Bernheim and Wantz, 1995, for empirical evidence). Formally, suppose a continuum $i \in [0, 1]$ of firms will each produce a random, i.i.d. stream of cash flows $\{y_t\}_{t=1}^{\infty}$. The distribution of y is privately known to the incumbent shareholder and given by $G_L(y)$ if $i < \lambda$ and

---

[3]Leland and Pyle (1977) model the cost of retention as risk-bearing by a risk-averse entrepreneur, rather than reduced investment by an entrepreneur who can reinvest his proceeds from selling the project at an above-market rate of return $r = 1/\alpha - 1 > 0$, as we do here following DeMarzo and Duffie (1999). Though the interpretation is similar, the mechanics in Leland and Pyle's model are therefore closer to the Rothschild and Stiglitz (1976) insurance application we sketch below.

[4]Since each contract covers $d - e$ losses, covering one unit of losses means that the insurance company enters into $1/(d - e)$ contracts.

$G_H(y)$ if $i \geq \lambda$, where G_H first-order stochastically dominates G_L. The conditional means are $\mathbb{E}_i(y) = rq(i)$, where r is the interest rate and $q(i)$ is given by (1). The incumbent shareholder announces a dividend e to be paid at $t = 1$ and then sells all its shares (cum-dividend) to heterogeneously informed outside investors. Dividends are taxed at a rate τ. Furthermore, following Bhattacharya (1979), if the cash flow y_1 is less than the announced dividend e, the incumbent agrees to provide the firm with a loan to finance the shortfall, at a cost $\beta(e - y_1)$. Letting $w - \tau e$ denote the price paid by investors, the payoff for the incumbent shareholder is

$$u(w, e) = w - \tau e - \beta \int_0^e (e - y) \, dG_k(y)$$

with $k = L$ if $i < \lambda$ and $k = H$ if $i \geq \lambda$. The marginal rate of substitution is

$$-\frac{\partial u(w, e)/\partial e}{\partial u(w, e)/\partial w} = \tau + \beta G_k(e).$$

By FOSD, this is higher for low types and thus satisfies single crossing. An outside investor's profit is given by the net present value of the firm's cash flows $q(i)$ minus the dividend tax τe minus the price paid $w - \tau e$, for a total of $q(i) - w$, just like in the benchmark model.

3   Equilibrium

There are many markets $m \in M$ open simultaneously. Each market is defined by a wage $w(m)$ and an education level $e(m)$ required for any worker who is hired. Each worker decides in which markets to apply for a job, and each firm decides in which markets to solicit applications from workers and how to use its information to screen applicants.

There is no guarantee for either workers or firms of finding a counterparty in market m; both firms and workers may face rationing in any given market. Markets are assumed to be non-exclusive: both workers and firms may visit many markets simultaneously. However, workers are committed to a single education level, so they can only apply for jobs in markets that correspond to their chosen education. This aligns well with the natural timing, where education is determined before entering the labor market. Similarly, in the corporate finance applications, it corresponds to situations where the design of the security (the size of the junior tranche), the financial structure of the firm (the retained equity) or the amount of dividends to be paid out are determined first, and then the securities or firm shares are offered in potentially multiple markets with different unit prices.

3.1 Worker's Problem

Each worker is endowed with one unit of labor. We assume that workers, having picked a signal e, choose what markets to apply to by setting a reservation wage $\underline{w}$ and applying for jobs in every market m with $e(m) = e$ and $w(m) \geq \underline{w}$.[5]

In order to decide how much education to obtain and what reservation wage to set, a worker needs to know what are the chances of finding a job in each market. This is summarized by a *rationing function* $\mu : \mathbb{R}^+ \times \mathbb{R}^+ \times [0,1] \to \mathbb{R}^+$, which is taken as given by workers but determined in equilibrium. We assume that labor is divisible and the law of large numbers applies, so the quantity $\mu(w; e, i)$ denotes how many units of labor of worker i will remain unsold if he obtains education e and sets a reservation wage w. Thus, μ is weakly increasing in w, and $d\mu(w; e, i)$ measures how many units of labor worker i with education e sells at wage w. If the function $\mu(\cdot; e, i)$ has more than one single step, the worker ends up supplying fractional amounts of labor in more than one market.[6]

Worker i's problem is

$$\max_{e, \underline{w}} \int_{\underline{w}}^{\infty} w \, d\mu(w; e, i) - c(i) e \qquad \text{s.t.} \quad \mu(\underline{w}; e, i) \geq 0. \tag{3}$$

The worker's problem can be simplified by noticing that since there is no disutility of labor, the worker will always try to sell his entire endowment of labor, so the constraint in (3) binds. Therefore if worker i chooses education e, his reservation wage will be

$$\underline{w}(e, i) = \sup \left\{ w \in \mathbb{R}^+ : \mu(w; e, i) \leq 0 \right\}$$

and his expected wage will be

$$\bar{w}(e, i) = \int_{w \geq \underline{w}(e,i)} w \, d\mu(w; e, i),$$

so his problem reduces to just choosing an education level:

$$\max_{e} \bar{w}(e, i) - c(i) e. \tag{4}$$

---

[5]Since workers prefer higher to lower wages, this is always an optimal strategy. However, since there may be markets where worker i knows it is impossible to find a job, it does not matter whether he applies to them or not, so the worker may be indifferent between several strategies. Kurlat (2016a) shows that a reservation wage is the only optimal strategy that is robust to a small chance of finding a job in every market.

[6]Instead of supplying deterministic fractional amounts of labor in different markets, equivalently workers could be hired probabilistically across different markets, as long as they apply to all markets above a reservation wage before knowing in which one they end up being hired. In what follows, we sometimes allude to this alternative, stochastic interpretation of rationing.

We will allow for the possibility that workers randomize over education levels (or that identical workers choose different signals), and denote by $\pi_i(e)$ the probability that a worker i chooses education level e.

3.2 Firm's Problem

When a firm observes applicants, it may use its information to select which ones to hire, to the extent that it can tell them apart. A feasible hiring rule for firm θ is a function $\chi : [0,1] \to \{0,1\}$ that is measurable with respect to its information set, that is:

$$\chi(i) = \chi(i') \quad \text{whenever} \quad x(i,\theta) = x(i',\theta).$$

A firm will reject applicants with $\chi(i) = 0$ and hire from among those i for whom $\chi(i) = 1$, which we describe as χ-acceptable. Let X denote the set of possible hiring rules.

In order to decide what market to hire from and what hiring rule to apply, a firm needs to form beliefs $G(\cdot; m, \chi)$ about what workers it will be drawing from should it choose to hire in market m with hiring rule χ. If the firm thinks it will find χ-acceptable workers in market m, then $G(\cdot; m, \chi)$ is a well-defined probability distribution on $\{i \in [0,1] : \chi(i) = 1\}$; otherwise $G(\cdot; m, \chi) = \emptyset$. Let g denote the density or pmf of G, which we assume is well-defined.

Firm θ's problem is

$$\max_{m,\chi} \int [q(i) - w(m)] \, dG(i; m, \chi) \quad \text{s.t.} \quad \chi \text{ feasible for } \theta. \tag{5}$$

As with workers, we allow for firms to randomize over hiring decisions (or identical firms to make different choices). We denote by the measure δ_θ on $M \times X$ the distribution of choices by firms of type θ across markets and hiring rules. Each firm can hire at most one worker, so $\delta_\theta(M \times X) \leq 1$.

3.3 Aggregation

In equilibrium, the rationing function faced by workers needs to be consistent with supply and demand in each market. Moreover, firms' beliefs must line up with the supply of workers. To formulate these conditions, we next compute aggregate supply and demand as implied by the worker and firm problems.

First, the supply of workers of index i that apply to market m is

$$S(i,m) = \pi_i(e(m)) \, \mathbb{I}(\underline{w}(e(m), i) \leq w(m)). \tag{6}$$

A total of $\pi_i(e(m))$ workers of index i choose the level of education that enables them to apply

14

to market m. Of these, only those whose optimal reservation wage $\underline{w}\left(e\left(m\right),i\right)$ is below the wage $w\left(m\right)$ that defines market m actually apply to market m.

Second, demand D is

$$D\left(M_0, X_0\right) = \int \delta_\theta \left(M_0, X_0\right) dF\left(\theta\right) \qquad (7)$$

for any $M_0 \times X_0 \subseteq M \times X$. D is a measure on $M \times X$ that describes how many total workers firms demand at every combination of markets and hiring rules. Since the rationing function faced by workers is defined over reservation wages, it is useful to let

$$M\left(w, e\right) = \{m \in M : e\left(m\right) = e, w\left(m\right) \geq w\}$$

be the set of markets that specify an education level of exactly e and a wage of at least w. This allows us to calculate hires H as

$$H\left(w; e, i\right) = \int_{M(w,e) \times X} g\left(i; m, \chi\right) dD\left(m, \chi\right). \qquad (8)$$

$H\left(w; e, i\right)$ is the number of workers of index i with education level e who are hired in markets where the wage is at least w. For a firm that imposes hiring rule χ in market m, the number of workers of type i that it hires is consistent with its beliefs. Hence equation (8) embodies a rational expectations assumption. Total hires are obtained by adding across different hiring rules and markets using the demand measure D. Thus, H summarizes the total demand for workers with reservation wage w (for a given index i and signal e), which in equilibrium will need to be in line with rationing μ and the supply given by (6).

3.4   Equilibrium Definition

These concepts allow us to define an equilibrium.

Definition 1. An equilibrium consists of:

i. An education decision π_i for each worker i

ii. A demand decision δ_θ for each firm θ

iii. A rationing function μ

iv. Beliefs $G\left(\cdot; m, \chi\right)$

such that

i. *Worker optimization.* For every i, e such that $\pi_i\left(e\right) > 0$, e solves worker i's problem (4), taking μ as given.

ii. *Firm optimization.* Every (m, χ) in the support of $\delta_\theta (\cdot)$ solves firm θ's problem (5), and $\delta_\theta (M, X) \leq 1$.

iii. *Rationing consistency.* $\mu (w; e, i)$ satisfies

$$d\mu (w (m) ; e (m) , i) S (i, m) = dH (w (m) ; e (m) , i) \tag{9}$$

and is weakly decreasing in i for every (w, e).

iv. *Belief consistency.* Beliefs satisfy:

(a) If $\int_i \chi (i) S (i, m) \, di > 0$ then

$$\int_i q (i) \, dG (i; m, \chi) = \frac{\int_i q(i)\chi(i)S(i,m)di}{\int_i \chi(i)S(i,m)di}. \tag{10}$$

(b) If $G (\cdot; m, \chi) \neq \emptyset$ then for any i in the support of $G (\cdot; m, \chi)$

   i. $\chi (i) = 1$

   ii. $e (m)$ solves worker i's problem

   iii. $\underline{w} (e, i) \leq w (m) .$

(c) If $G (\cdot; m, \chi) = \emptyset$, then $\mu (w (m) , e (m) , i) = 0$ for all i such that $\chi (i) = 1$.

The worker and firm optimization requirements are standard: workers choose education and firms choose markets and acceptance rules to maximize their payoffs. The rationing consistency requirement (9) requires that for markets where $S (i, m) > 0$, the fraction of labor worker i sells in market m equals the ratio of i-hires to i-applicants. For markets where $S (i, m) = 0$, (9) imposes no constraints on $d\mu$. We do impose that μ must be weakly decreasing in i. This rules out low types being more optimistic than high types about the wages they would obtain at some off-equilibrium levels of education, which would be inconsistent with any rational conjecture about what firms might be doing in off-equilibrium markets.

The belief consistency requirement specifies what beliefs firms can hold in equilibrium. Condition (iva) just imposes Bayes' Rule for the expected productivity of workers for positive-supply markets. It says that in any market with a positive supply of χ-acceptable workers, firms expect the average productivity to correspond to a representative sample of them.[7]

---

[7]If firms expect to draw from the entire supply of χ-acceptable workers who apply to market m, then Bayes' Rule implies beliefs over χ-acceptable workers will be:

$$g (i; m, \chi) = \frac{\chi (i) S (i, m)}{\int_i \chi (i) S (i, m) \, di},$$

which implies that (10) holds. If firms with different hiring rules hire in the same market, the allocation will depend on the order in which they choose workers, as early firms may skew the sample that later firms draw

The key challenge in constructing a tractable equilibrium notion in our model with signaling is how to discipline firms' beliefs in zero-supply markets. We propose the following refinement, which will guarantee equilibrium uniqueness in the no-information benchmark (and at the same time preserve equilibrium existence throughout).[8] Condition (ivb) (which is automatically satisfied in positive-supply markets) says that beliefs can only place weight on χ-acceptable workers that would be willing to look for a job in market m. In other words, a firm can never expect to find in market m a worker who could obtain higher utility by choosing a different level of education, or who can find a job for sure at the same level of education but a higher wage. Condition (ivc) says that if a firm is certain that it cannot find χ-acceptable workers in market m, then it must be that χ-acceptable workers are certain to find a job at a wage at least $w\,(m)$ if they choose education $e\,(m)$. In other words, firms can only believe that a market is empty of acceptable workers if being guaranteed a job at wage $w\,(m)$ is not sufficient to persuade them to choose this market. Notice that a firm's beliefs about market m do not depend on whether the firm decides to recruit workers in market m. Firms perceive themselves to be small and therefore take the decisions of workers as given.

Condition (ivc) is closely related to the infinite-tightness condition in Guerrieri et al. (2010) and Guerrieri and Shimer (2014). In their setup, for every market there either is at least one worker type who finds that market optimal or the market tightness is infinite. In the first case, this allows firms to have well-defined beliefs about which workers they would encounter; in the second, workers would match for sure. Condition (ivc) generalizes this idea. There is no unique tightness measure for each market because the same market may be tight for some workers but not others, depending on whether firms who hire in this market find them acceptable. The requirement that workers must expect to match for sure if firms think they cannot find them must therefore be imposed separately for the workers in each χ-acceptance group.

4   The No-Information Benchmark

We now characterize equilibrium for the case where F is a point mass at $\theta = 0$ (or equivalently at $\theta = 1$), i.e. when all firms are completely uninformed. This corresponds to the classic signaling environment. For this case, the least-cost separating allocation emerges as the unique

---

from. Kurlat (2016a) assumes that there are separate markets for each possible ordering over hiring rules and shows that under a "false positives" information structure, trade will take place in markets where the less selective firms hire first. This means that no one's sample is skewed by earlier firms, so all firms indeed draw from the entire pool of applicants. In the equilibrium under "false negatives", more selective firms may hire first, so not all firms draw from the entire pool, and beliefs do not correspond to simply applying Bayes' Rule to the entire pool of χ-acceptable workers. However, only firms who only accept high types end up hiring after more selective firms, so even though the distribution is skewed, its mean is unaffected and condition (10) holds.

[8]Modeling beliefs is not an issue in Kurlat (2016a) as he rules out signaling, so there are no markets corresponding to off-equilibrium signals.
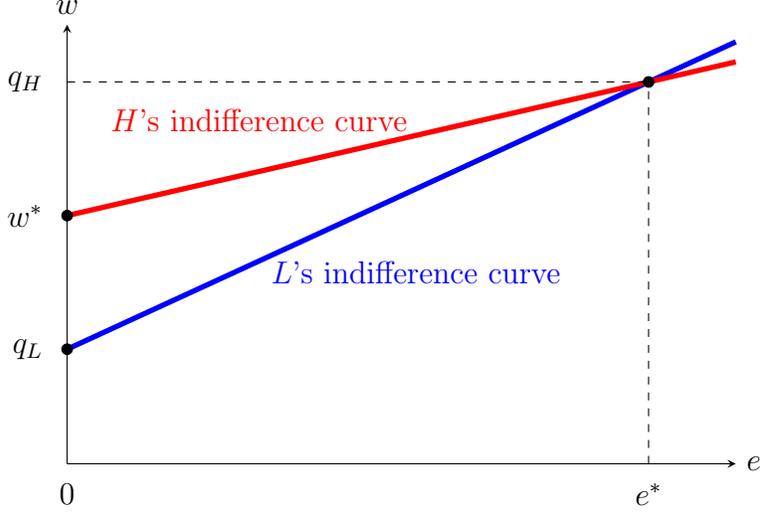
Figure 1: The least-cost separating allocation.

equilibrium. In this allocation, low types get no education, high types get just enough education to separate with

$$e^* = \frac{q_H - q_L}{c_L}, \tag{11}$$

and each type is paid their own productivity, as illustrated in Figure 1.

Proposition 1. *If F is a point mass at $\theta = 0$, there is a unique equilibrium, given by:*

i. *Education decisions:*

$$\pi_i(e) = \begin{cases} 1 & \text{if } i < \lambda \text{ and } e = 0 \\ 1 & \text{if } i \geq \lambda \text{ and } e = e^* \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

ii. *Demand decisions:*

$$\delta(m) = \begin{cases} \lambda/F(0) & \text{if } w(m) = q_L \text{ and } e(m) = 0 \\ (1-\lambda)/F(0) & \text{if } w(m) = q_H \text{ and } e(m) = e^* \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

iii. *Rationing function:*

$$\mu(w, e, i) = \mathbb{I}(w \geq \min\{q_L + c_L e, q_H + c_H(e - e^*)\}) \tag{14}$$

iv. *Beliefs:*

$$g(i; m) = \begin{cases} \frac{1}{\lambda}\mathbb{I}(i < \lambda) & \text{if } e(m) < e^*, w \geq q_L + c_L e(m) \\ \frac{1}{1-\lambda}\mathbb{I}(i \geq \lambda) & \text{if } e(m) \geq e^*, w \geq q_H + c_H(e(m) - e^*) \\ \emptyset & \text{otherwise} \end{cases} \tag{15}$$

18

The equilibrium is constructed by setting μ as a step function at the lower envelope of the indifference curves of both types, which makes low types indifferent between any $e \in [0, e^*]$ and high types indifferent between any $e \geq e^*$. Therefore, $e = 0$ for low types and $e = e^*$ for high types is indeed optimal. This is then sustained by firms' belief that in the range $[0, e^*]$ they will only encounter low types above the lower envelope and no one at all below, and similarly for high types above e^*. Hence there are no profits in any market, so firms are trivially optimizing.

The key step in establishing uniqueness is to rule out pooling, i.e. markets with positive supply of both high and low types. This follows the standard logic based on single-crossing. If there was pooling at a level of education e', then high types would require a lower wage than low types to be willing to choose $e = e' + \epsilon$. Hence firms that consider hiring in a market with $e = e' + \epsilon$ and a wage that leaves high types indifferent must believe that they will only encounter high types, which for small ϵ must be more profitable than hiring at e'.

The types of deviations to pooling contracts that may lead to non-existence in Rothschild and Stiglitz (1976) are not profitable because each firm perceives itself to be small. A job with $e = 0$ and $w = q_H - c_H e^* + \epsilon$ is strictly preferred to the equilibrium by all workers and if a firm was large and could hire the entire population it could break the equilibrium by offering to hire everyone in this market, which would be profitable for low values of λ. Here, if a small firm tries to hire in this market, it will not attract any high types, because they know that they will be competing with all the low types for an infinitesimal chance to be hired and will have to settle for $w = q_L$ if they are not. Formally, this is captured by the assumption that beliefs do not depend on whether a firm decides to recruit in a particular market. This is the same logic that leads to existence and uniqueness in Guerrieri et al. (2010).

5   False Positives

In this section, we turn to the case where F has full support on the interval $[0, \lambda]$, so firm θ misperceives all applicants $i \in [\theta, \lambda]$ as high types even though they are in fact low types. We construct a class of candidate equilibria in which firms with sufficiently accurate signals hire high types at a wage below q_H. We then show that any equilibrium must lie within this class. However, in order to actually be an equilibrium, a candidate equilibrium must satisfy an additional condition. Using this result, we establish existence and discuss the possibility of multiple equilibria. We then study comparative statics and analyze welfare implications.

5.1   Candidate Equilibria

Define w^* as

$$w^* = q_H - c_H e^*. \tag{16}$$

This is the wage that gives high types who get no education the same utility as under the least-cost separating allocation, as shown in Figure 1. Let m^* be the market with $e(m^*) = 0$ and $w(m^*) = w^*$.

Suppose that all the low types and a fraction π^* of the high types were to choose $e = 0$ and apply for jobs in market m^*, while the rest of the high types choose $e = e^*$ and get jobs at $w = q_H$. Any firm that hires in market m^* will face a mix of $\pi^*(1-\lambda)$ high type and λ low type applicants. It is profitable to hire workers from this pool as long as the firm is sufficiently good at screening out low types. If firm θ decides to hire in market m^* it will impose hiring rule $\chi_\theta(i) = \mathbb{I}[i \geq \theta]$ in order to screen out as many low types as possible, leaving only $\lambda - \theta$ χ-acceptable ones. Therefore it ends up hiring $\frac{\lambda - \theta}{\lambda - \theta + \pi^*(1-\lambda)}$ low types and $\frac{\pi^*(1-\lambda)}{\lambda - \theta + \pi^*(1-\lambda)}$ high types. Using (10), the firm's profits will be

$$\Pi(\theta) = \frac{(\lambda - \theta)q_L + \pi^*(1-\lambda)q_H}{\lambda - \theta + \pi^*(1-\lambda)} - w^*. \tag{17}$$

Equation (17) implies that profits are increasing in θ. Firms whose information enables them to screen out a higher proportion of low types will be hiring a better pool of workers. This defines a cutoff θ^* such that firms can make profits in market m^* if and only if $\theta > \theta^*$:

$$\theta^* = \max\left\{0, \lambda - \pi^*(1-\lambda)\frac{q_H - w^*}{w^* - q_L}\right\}. \tag{18}$$

The calculations above assume that some high types are indeed willing to apply to market m^*. For this to be true, it must be the case that they are sure they will find a job, since they can always guarantee themselves the same utility by choosing $e = e^*$ and getting a job that pays $w = q_H$. This means that there must be enough firms above θ^* to hire all $\pi^*(1-\lambda)$ high types who forgo education and apply to market m^*. By the arguments above, each firm $\theta \geq \theta^*$ hires $\frac{\pi^*(1-\lambda)}{\lambda - \theta + \pi^*(1-\lambda)}$ high types, so in equilibrium we need

$$\int_{\theta^*}^{\lambda} \frac{1}{\lambda - \theta + \pi^*(1-\lambda)} dF(\theta) = 1. \tag{19}$$

Note that low-type workers do *not* guarantee themselves a job at w^* since some of the firms hiring in market m^* will reject them, so they will sell a fraction of their labor at $w = q_L$.

The indifference condition (18) and the market clearing condition (19) define two relationships between the cutoff firm θ^* and the fraction of high types π^* that forgo signaling. Both of these relationships are downward sloping, as shown in Figure 2.

The indifference condition (18) is downward-sloping because if more high types decide to forgo education, they improve the pool of workers available for hire in market m^*, allowing
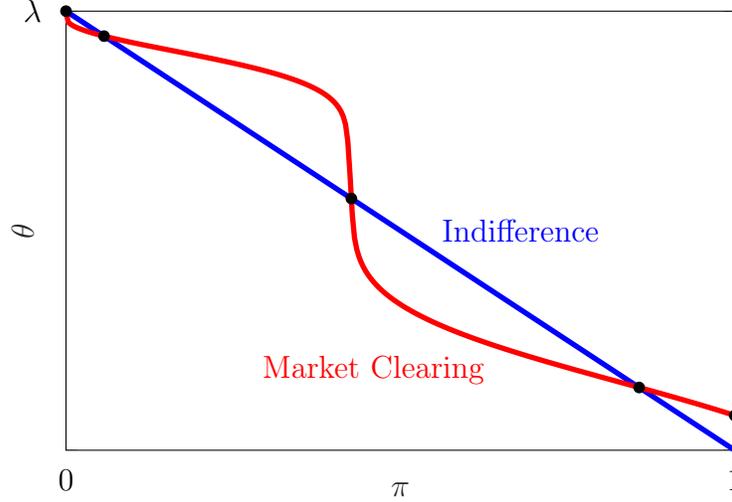
Figure 2: Indifference and market clearing conditions for the false positives case.

less-informed firms to earn profits. The market clearing condition (19) is downward sloping because if more high types decide to forgo education, they can only find jobs in market m^* if additional firms decide to hire there. In other words, high type workers' decision to forgo education and firms' decision to hire in market m^* are *strategic complements*. The more high types forgo education, the more profitable it is for any given firm to hire in m^*; the more firms hire in m^*, the more attractive it is for high type workers stop signaling.

Our candidate equilibria will be of two possible kinds. The first is interior equilibria as described above, where the market clearing condition intersects the indifference condition. In these candidate equilibria, a fraction π^* of the high types choose not to signal. High-expertise firms can tell them apart from the low types relatively well. They are therefore willing to hire them at a wage w^*, knowing that it is unlikely that they mistakenly hire a low type. However, some low types with $i \geq \theta^*$ do manage to mislead firms into hiring them at w^*, while the rest of them get jobs at wage q_L. Low-expertise firms hire either at wage q_L without requiring a costly signal, or at wage q_H requiring education e^*; in either case they do not need to be selective because only one type of worker is to be found in each of these markets.

The second type of equilibrium, a corner equilibrium where the market clearing condition is above the indifference condition at $\pi = 1$, is similar except that nobody signals. The fact that the curve corresponding to condition (19) lies above the curve corresponding to (18) means that there are more firms willing to hire at w^* than the total mass of workers they would accept. As a result, high-θ firms bid up the wage higher than w^*, leading some firms to drop out until the number of firms willing to pay this wage equals the number of workers they end up hiring.

The equilibrium wage w^{**} is such that the marginal buyer θ^{**} makes zero profits:

$$w^{**} = \frac{(\lambda - \theta^{**}) \, q_L + (1 - \lambda) \, q_H}{1 - \theta^{**}} \geq w^* \tag{20}$$

where θ^{**} is such that all high type workers find a job at w^{**}:

$$\int_{\theta^{**}}^{\lambda} \frac{1}{1 - \theta} dF(\theta) = 1 \tag{21}$$

For example, in Figure 2, there is a total of five candidate equilibria.

The corner equilibrium is equivalent to the one described by Kurlat (2016a) in an economy that does not allow for signaling, which is also summarized by a marginal buyer and a price such that the marginal buyer is indifferent and all high type sellers trade.[9] It is clear that his indifference and market clearing conditions, which are equivalent to (20) and (21), always have a unique intersection. In other words, there is no multiplicity among corner equilibria. The potential for multiple equilibria in our model is confined to interior equilibria, which involve signaling. The driving force is the strategic complementarity described above, which is different from the forces that may lead to multiplicity in Akerlof (1970) (where adverse selection depends on the price) or in canonical signaling models (where different off-equilibrium beliefs can be self-sustaining). Indeed, with our refinement on beliefs, the uninformed-firm benchmark has a unique equilibrium. The multiplicity we identify here relies on both the presence of signaling and heterogeneous information among firms.

Proposition 2. *In any equilibrium, all low types choose $e = 0$. Furthermore, either*

i. High types' decisions are:

$$\pi_i(e) = \begin{cases} \pi^* & \text{if } e = 0 \\ 1 - \pi^* & \text{if } e = e^* \end{cases} \tag{22}$$

and firms' decisions are:

$$\delta_\theta(m^*, \chi) = 1 \quad \text{if and only if } \theta \geq \theta^* \text{ and } \chi(i) = \mathbb{I}(i \geq \theta)$$

where π^* and θ^* satisfy equations (18) and (19); or:

---

[9]There are three minor differences: first, Kurlat (2016a) allows sellers to have a positive value for retaining the good, while we assume it to be zero so all units of labor are sold in equilibrium; second, he assumes $q_L = 0$ so even though there is a two-price equilibrium like the one here where a fraction of low types trade at a price of q_L, there is also a one-price equilibrium where they do not; finally, Kurlat (2016a) models buyers' capacity constraints in terms of dollars rather than in terms of quantities, so the price appears in the market-clearing condition. None of these differences affects the basic idea.

ii. All high types choose $e = 0$ and firms' decisions are:

$$\delta_\theta \left(m^{**}, \chi \right) = 1 \quad \text{if and only if } \theta \geq \theta^{**} \text{ and } \chi(i) = \mathbb{I}(i \geq \theta)$$

where θ^{**} satisfies (21) *and m^{**} is such that $e(m^{**}) = 0$ and $w(m^{**}) = w^{**}$ given by* (20).

Proposition 2 establishes that the only possible equilibria are those described above. The key is to establish that high and low types cannot coexist at any level of education other than $e = 0$, that is there is no pooling at positive signaling levels. The logic is similar, though somewhat subtler, than in the uninformed-firms benchmark.

Suppose that low and high types coexisted at some market m with $e(m) > 0$, as illustrated in Figure 3. In the uninformed-firms benchmark, this is ruled out by specifying beliefs about market m': if they are well defined, they can only put weight on high types, since they find m' more attractive than m while low types do not; and they must be well defined because otherwise $\mu \left(w(m'), e(m'), i \right) = 0$, so high types would prefer $e(m')$. With differentially informed firms, this logic does not go through: the firms hiring in market m may be screening out low types, so low types' expected wage at $e = e(m)$ could be lower than that of high types. Thus, it is possible that both high and low types find $e(m')$ more attractive than $e(m)$.

The argument for ruling out pooling at m is slightly different. Suppose first that the highest firm that hires in market m has $\bar{\theta} = \lambda$ (i.e. it can tell workers apart perfectly). In this case, we can rule out pooling by considering this firm's beliefs about market m''. Since this firm's beliefs can only include high types, it only cares about the wage it pays and therefore finds market m'' preferable over market m. If instead the highest firm that hires in market m has $\bar{\theta} < \lambda$, then there are low types $i \in \left(\bar{\theta}, \lambda \right)$ whose chance of getting a job in market m is just as good as the high types, and firm $\bar{\theta}$ can rule them out in the usual way by hiring in market m'.

5.2 Beliefs, Rationing and Possible Deviations

In order for a candidate equilibrium to actually be an equilibrium, both firms and workers must be attaining their maximum possible payoffs. To verify whether this is the case, we need to specify firm beliefs and rationing functions for all markets. A candidate equilibrium will be an equilibrium if firms' beliefs about the workers in off-equilibrium markets are sufficiently pessimistic and the rationing functions faced by workers at off-equilibrium levels of education are sufficiently high so that the choices specified by the candidate equilibrium are indeed optimal. We will next compute the most pessimistic beliefs for firms that are consistent with a candidate equilibrium allocation and worker optimization, and check whether firms can increase their profits in some off-equilibrium market.

Consider a candidate equilibrium of the first type, given by an intersection of (18) and (19).
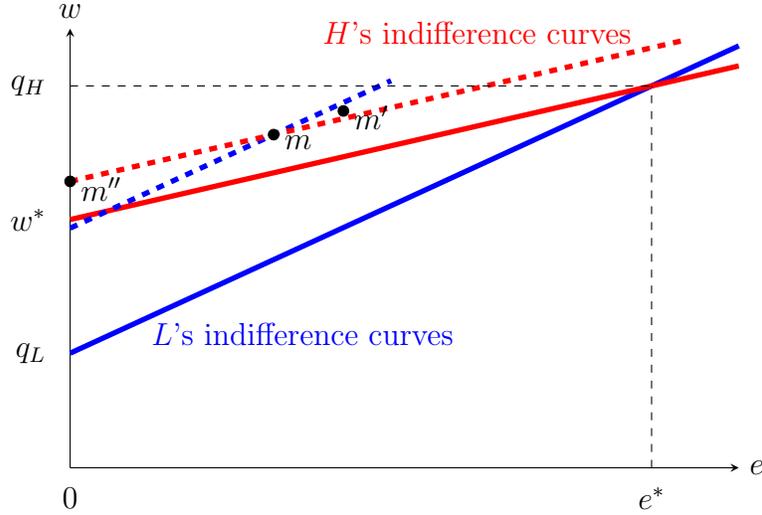
Figure 3: Ruling out pooling at $e > 0$.

Workers' equilibrium utility is given by

$$u(i) = \begin{cases} q_L & \text{if } i < \theta^* \\ q_L + \int_{\theta^*}^{i} \frac{1}{\pi^*(1-\lambda)+\lambda-\theta} dF(\theta)(w^* - q_L) & \text{if } i \in [\theta^*, \lambda) \\ w^* & \text{if } i \geq \lambda. \end{cases} \quad (23)$$

Low types with $i < \theta^*$ do not fool any firm into hiring them in market m^* and therefore end up obtaining q_L. High types are indifferent between market m^* or choosing $e = e^*$ to ensure themselves $w = q_H$, which results in a payoff of w^*. Low types with $i \in [\theta^*, \lambda)$ get hired partly in market m^* and partly at wage q_L; the respective proportions depend on what fraction of firms they are able to mislead, so their payoff is increasing in i.

Let $\tilde{w}(e, i)$ be the wage that would make worker i indifferent between their equilibrium payoff and choosing education e:

$$\tilde{w}(e, i) = \begin{cases} u(i) + c_L e & \text{if } i < \lambda \\ u(i) + c_H e & \text{if } i \geq \lambda. \end{cases} \quad (24)$$

Now suppose firm θ considers hiring in market m. What beliefs could it consistently hold about the workers it will encounter there? Is it possible that it believes it will find χ_θ-acceptable low types, i.e. workers with $i \in [\theta, \lambda)$? For this to be the case, the following must hold:

$$\tilde{w}(e(m), \theta) \leq \tilde{w}(e(m), i) = \bar{w}(e(m), i) \leq \bar{w}(e(m), \lambda) \leq \tilde{w}(e(m), \lambda). \quad (25)$$

The first inequality follows from the fact that $u(i)$ and therefore $\tilde{w}(e, i)$ is increasing in i. The second step follows from the requirement that if beliefs place weight on type i, then i must be indifferent between $e(m)$ and his equilibrium choice. The third follows from monotonicity of μ
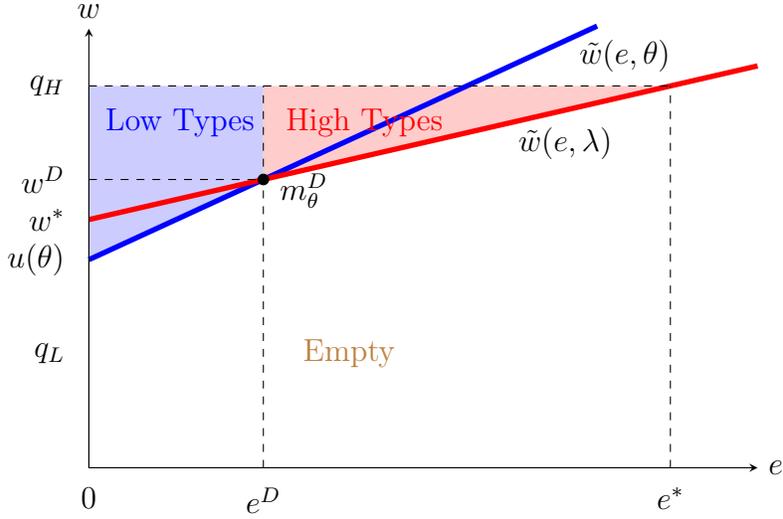
24

Figure 4: Beliefs for firm θ

in i, which implies that $\bar{w}$ is also monotonic in i. The last inequality follows from the fact that otherwise worker λ could exceed his equilibrium payoff by choosing $e(m)$. Therefore, the only markets where firm θ can place beliefs on χ_θ-acceptable low types are those with education levels that worker $i = \theta$ is willing to choose for weakly lower wages than high types.

Could it be, instead, that firm θ's beliefs are not well-defined? Only if

$$w(m) \le \tilde{w}(e(m), \theta), \tag{26}$$

since otherwise the requirement that $\mu(w(m), e(m), \theta) = 0$ implies that some χ_θ-acceptable worker could exceed his equilibrium payoff by choosing $e(m)$.

Together, conditions (25) and (26) imply that for any market m such that $\tilde{w}(e(m), \lambda) < \tilde{w}(e(m), \theta)$ and $w(m) > \tilde{w}(e(m), \lambda)$, firm θ's beliefs $G(\cdot; m, \chi_\theta)$ can only place weight on high types. Figure 4 illustrates the regions of different possible beliefs. High type workers obtain utility w^* in equilibrium while the lowest-i low type that firm θ would accept obtains $u(\theta) < w^*$. Their indifference curves intersect at education level e^D. Worker $i = \theta$ is willing to choose education levels $e \in (0, e^D)$ for lower wages than high types, and therefore firm θ can consistently believe that in any market with $e \le e^D$ and $w \ge \tilde{w}(e, \theta)$ it will only face low types. However, for education levels $e > e^D$, high types would be willing to choose them for lower wages than low types. Therefore, firm θ must believe that in any market with $e > e^D$ and $w > \tilde{w}(e, \lambda)$ it will only face high types.

The fact that firm θ can identify markets where it is sure to find only high types does not necessarily mean that it can obtain higher profits than in equilibrium by hiring in these markets. Wages are higher, so profits could be higher or lower. Denote the lowest-wage market where

firm θ's beliefs are guaranteed to only include high types by m_θ^D, which satisfies

$$w\left(m_\theta^D\right) = \tilde{w}\left(e\left(m_\theta^D\right), \lambda\right) = \tilde{w}\left(e\left(m_\theta^D\right), \theta\right).$$

Using (11), (16), (23) and (24) and rearranging, the maximal profits that firm θ can obtain by hiring in market m_θ^D are

$$\Pi^D(\theta) = q_H - w\left(m_\theta^D\right) = \frac{c_H}{c_L}\int_{\theta^*}^{\theta} \frac{1}{\pi^*(1-\lambda) + \lambda - \theta} dF(\theta)(q_H - q_L).$$

Hence, by (17), profits in market m_θ^D exceed those that firm θ obtains in equilibrium if and only if

$$\frac{\lambda - \theta}{\pi^*(1-\lambda) + \lambda - \theta} > \frac{c_H}{c_L}\int_{\theta}^{\lambda} \frac{1}{\pi^*(1-\lambda) + \lambda - t} dF(t). \tag{27}$$

If this condition is satisfied for some θ, then the candidate equilibrium is not actually an equilibrium because there is no set of beliefs consistent with equilibrium that does not permit profitable deviations. It is easy to construct examples where condition (27) is or is not satisfied. Figure 5 shows an economy with multiple candidate equilibria. For candidate equilibrium (π_1^*, θ_1^*), condition (27) does not hold for any $\theta > \theta_1^*$, so it is indeed an equilibrium. Instead, for candidate equilibrium (π_2^*, θ_2^*), condition (27) holds for some $\theta > \theta_2^*$, so it is not an equilibrium.[10]

Hiring in a market like m_θ^D is similar to the cream-skimming deviations that are used to break putative pooling equilibria in Rothschild and Stiglitz (1976) and related models, including the uninformed-firms benchmark of Section 4. In candidate equilibria where some high types choose $e = 0$, they end up being hired in market m^*, where they are pooled with low types. Just like in the benchmark, the possible deviation involves peeling off high types from market m^* by requiring an action that is more costly for low types than for them. However, there are two important differences. First, unlike in the Rothschild and Stiglitz (1976) model, purely local deviations do not work. A firm cannot cream-skim the high types off a pooling contract by requiring a small amount of extra education. Since low types are hired from the m^* pool at lower rates than high types, they obtain lower utility. Therefore, they find deviations *more* attractive than high types as long as they involve only a small amount of extra signaling. In order to repel the low types, the deviating firm must require a sufficiently larger signal.

Second, in order to profit, the deviating firm must use both sources of information in combination: direct assessment and signaling. A purely uninformed firm cannot profitably deviate because in order to repel the lowest-i low types (who cannot get jobs at m^* at all) it must require $e = e^*$ and therefore pay at least q_H to attract high types, at which point the deviation is no longer profitable. In order to profitably deviate, a firm must possess sufficient expertise

---

[10] The example uses $q_H = 1, q_L = 0.4, c_H = 0.9, c_L = 0.15, \lambda = 0.6, f(\theta) = 0.5[\sin(13.3(\theta - \lambda)^{0.4} + 2.2\pi) + 1]^{2.8}$.
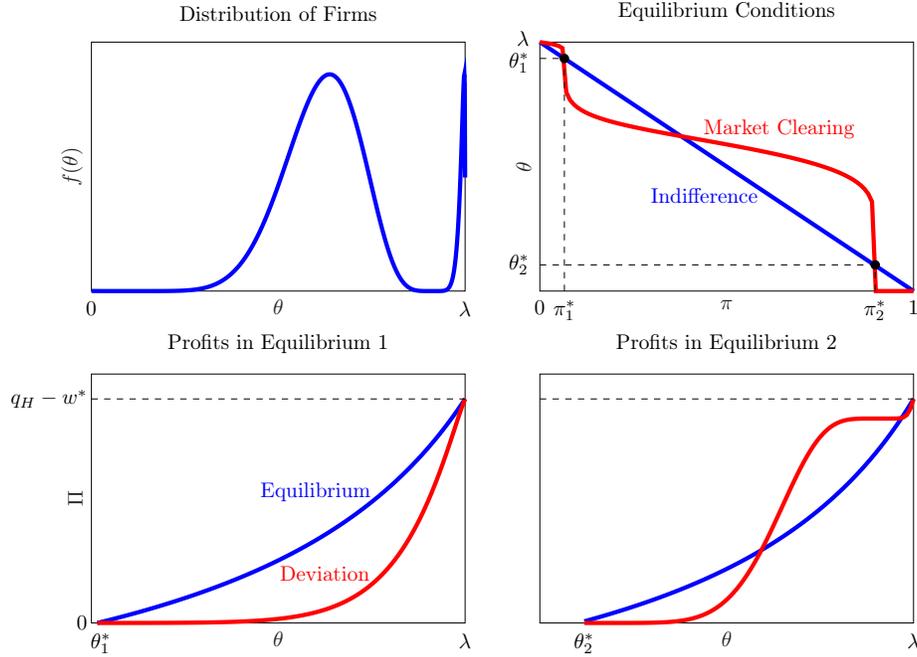
Figure 5: Candidate equilibria that do or do not satisfy condition (27).

to be able to reject the lowest-i low types directly and then rely on the signal to screen out the higher-i low types.

Using a similar reasoning, if the candidate equilibrium is of the second type, with the market clearing curve above the indifference curve at $\pi = 1$, we find that firm θ can find a profitable deviation whenever

$$\frac{\lambda - \theta}{1 - \theta} > \frac{c_H}{c_L - c_H} \frac{1 - \lambda}{1 - \theta^{**}} \int\limits_{\theta}^{\lambda} \frac{1}{1-t} dF(t). \tag{28}$$

Instead, if conditions (27) or (28) respectively are not satisfied, then the candidate equilibria are indeed equilibria. This result is formally stated in the following proposition:

Proposition 3.   *i. Suppose π^* and θ^* satisfy equations (18) and (19) and condition (27) does not hold for any $\theta \in [\theta^*, \lambda]$. Then the worker and firm decisions described in Proposition 2 part i are part of an equilibrium.*

   ii. Suppose θ^{**} and w^{**} satisfy equations (20) and (21) with $w^{**} \geq w^*$ and condition (28) does not hold for any $\theta \in [\theta^{**}, \lambda]$. Then the worker and firm decisions described in Proposition 2 part ii are part of an equilibrium.

All equilibrium objects as specified in Definition 1 are provided in the Appendix. Proposition 3 immediately implies that $\pi^* = 0$, $\theta^* = \lambda$ is always an equilibrium, since it satisfies equations (18) and (19) and condition (27) does not hold for $\theta = \lambda$. Hence, there is always an equilibrium
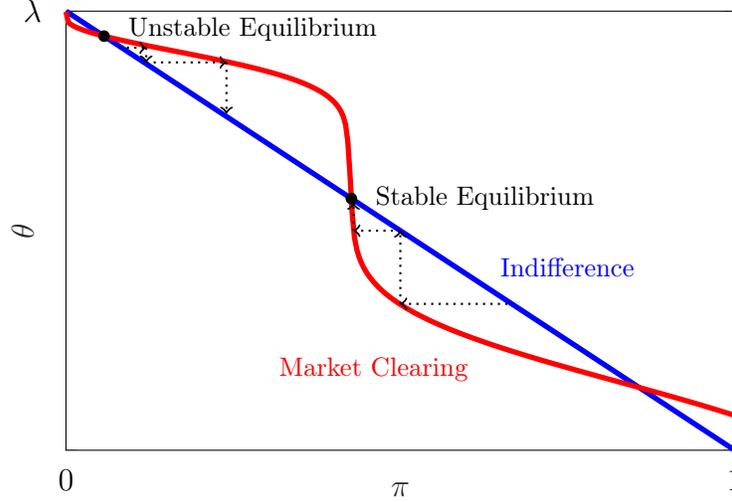
Figure 6: Stable and unstable equilibria

that coincides with the benchmark, in which signaling choices perfectly separate workers and firms' expertise is irrelevant. Depending on parameters, there may exist others in addition.

5.3   Comparative Statics

To characterize how the equilibria in our model move with parameters, we focus on interior equilibria where the locus of the market clearing condition (19) is steeper than of the indifference condition (18). In an informal sense, these equilibria are stable, while equilibria where the indifference condition is steeper are unstable. To see this, consider the following iterative reasoning. Suppose firms mistakenly conjectured that a fraction $\pi^* + \epsilon$ of high type workers choose $e = 0$. Under this conjecture, the indifference condition (18) defines a cutoff firm that makes zero profits by hiring in market m^*. Using this cutoff, the market clearing condition defines a new fraction of high type workers that can choose $e = 0$ and be sure to find a job, and so on. As Figure 6 illustrates, this iteration converges back to (π^*, θ^*) if (19) is steeper than (18), and diverges otherwise. To be clear, there is nothing in our definition of equilibrium that requires this type of stability. But since these equilibria are more robust to small miscalculations, we focus on them. As usual, most comparative statics would be reversed in the unstable equilibria.

In order to make the comparative statics clearer, we reparametrize the model as follows. First, let each firm hire Δ workers instead of just one. Changes in Δ can then be interpreted as changes in the demand for workers, keeping the screening abilities of firms constant. Second, assume that the mass of low type workers is $\hat{\lambda}$, distributed uniformly in the interval $[0, \lambda]$, with a density $\frac{\hat{\lambda}}{\lambda}$; correspondingly, the mass of high types is $1 - \hat{\lambda}$, distributed uniformly in the interval $[\lambda, 1]$ with a density $\frac{1-\hat{\lambda}}{1-\lambda}$. Changes in $\hat{\lambda}$ have the interpretation of changes in the fraction of low
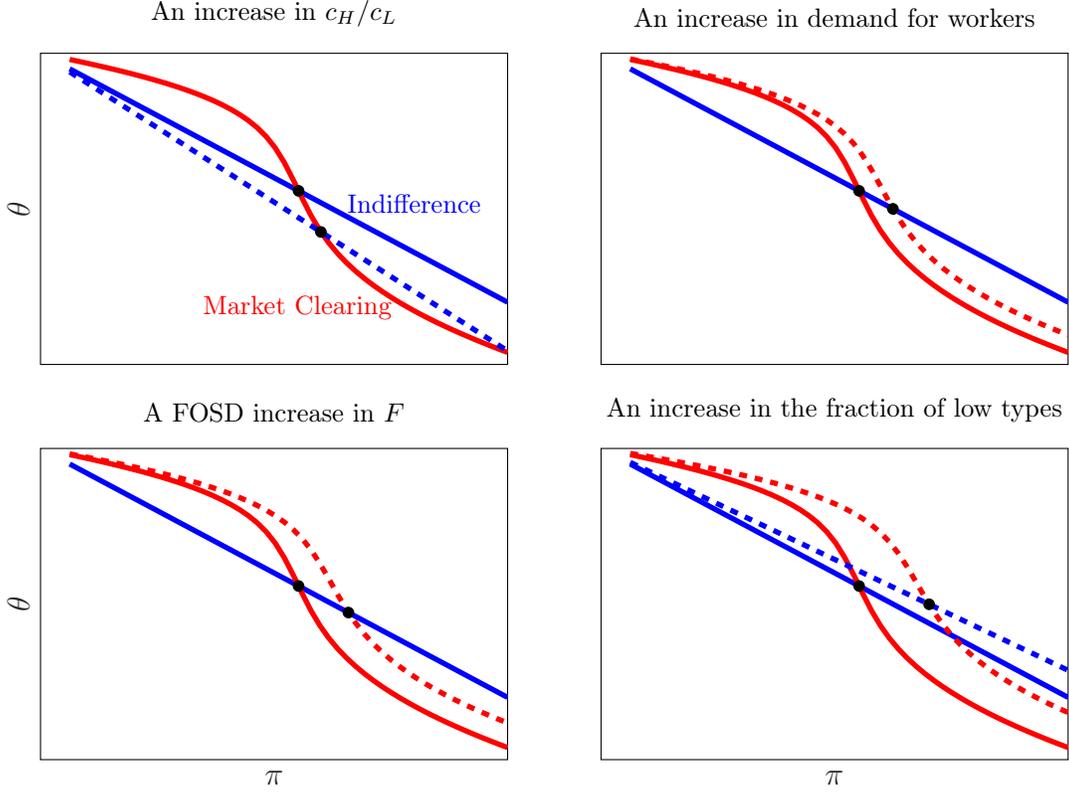
<div align="center">Figure 7: Comparative statics</div>

types, leaving their relative detectability constant. The following proposition summarizes how π^* and θ^* depend (locally) on parameters around a stable equilibrium:

Proposition 4.   *i. π^* is increasing and θ^* decreasing in the cost ratio c_H/c_L*

   ii. π^* is increasing and θ^* decreasing in the demand for workers Δ

   iii. π^* is increasing and θ^* decreasing with a FOSD increase in the expertise distribution F

   iv. π^* and θ^* are invariant with respect to productivities q_H and q_L

   v. π^* and θ^* could both be increasing or decreasing with the fraction of low type workers $\hat{\lambda}$.

These comparative statics are illustrated in Figure 7. The logic of part i is as follows. By equation (11), c_L governs how much education the high types need to obtain in order to separate from the low types. c_H, in turn, governs how much it costs them to obtain this education, so c_H/c_L governs how much utility they obtain if they choose e^*. Since w^* is the wage that makes them indifferent between $e = e^*$ and $e = 0$, it also depends on this cost ratio. Using (16),

$$w^* = \left(1 - \frac{c_H}{c_L}\right) q_H + \frac{c_H}{c_L} q_L, \tag{29}$$

so wages in the m^*-market are lower when high types' relative cost of signaling is higher. Having to pay lower wages means that, other things being equal, less informed firms are willing to hire in market m^*, so the cutoff θ^* falls. In turn, this means that firms hiring in market m^* will hire more total workers, so the fraction of high types that can choose $e = 0$ and still make sure to find a job rises. Graphically, the indifference condition shifts downwards. Note that the no-information benchmark, somewhat unappealingly, does not have this property: all high types choose $e = e^*$ and e^* does not depend on c_H, so high types do *not* respond to a higher cost of signaling by signaling less. Allowing for heterogeneously informed firms overturns this counterintuitive feature: equilibrium forces do lead workers to respond on the extensive margin, with fewer of them signaling as it gets more expensive.

An increase in demand means that, for any given cutoff θ^*, firms hire more workers in market m^* so, other things being equal, the number of high types who can forgo education and find a job at m^* rises. In turn, this justifies lower-θ firms hiring in m^*. Graphically, there is an outward shift in the market clearing condition. Again, the intuitive property that the amount of wasteful signaling falls when demand is high is absent when buyers are uninformed: in that case, all high types signal no matter what demand is. By contrast, our model generates the plausible prediction, for instance, that more high types forgo signaling through education (or that the amount of retained equity falls) in boom times.

A FOSD increase in the distribution of θ means that firms are able to screen out more low types, and therefore hire more high types (the market clearing condition shifts outwards). Therefore, more high types are able to forgo education and still find a job, which means π^* rises. In turn, this lowers the cutoff level of expertise for firms to be willing to hire in market m^*. Better tools for directly evaluating job applicants, firm shares or asset-backed securities reduce the need to signal through education, dividends or retained equity tranches, respectively; better tools for scoring insurance applicants reduce the need for them to be screened through high deductibles, etc. In this way, direct information substitutes for traditional signaling.

Finally, productivities have no effect on the equilibrium levels of π^* and θ^*. As equation (29) shows, the wage in market m^* is a weighted average of q_H and q_L. Therefore, no matter what these productivities are, the indifferent firm θ^* will be the one whose pool of acceptable workers includes a proportion of exactly $\frac{c_H}{c_L}$ low types. If, say, the productivity of low types was lower, the wage that defines market m^* adjusts exactly so as to leave firm θ^* indifferent. Moreover, the fraction of low types can have different effects on the equilibrium depending on the rest of the parameters. From equation (58), higher $\hat{\lambda}$ induces an upward shift in the indifference condition, but the market clearing condition can shift in either direction.

5.4 Continuity in the Symmetric Information, No-Signaling and No-Information Limits

One counterintuitive feature of the uninformed-firms benchmark is that it is discontinuous in the buyers' prior. If all workers have the same productivity there is no information asymmetry and no signaling in equilibrium. However, as soon as there is even an infinitesimal mass of low types, high types will signal enough to separate. The following result shows that this unappealing property vanishes in our model.

Proposition 5.

i. For $\hat{\lambda}$ low enough, there exists a corner equilibrium with $\pi^* = 1$, and $\lim_{\hat{\lambda} \to 0} w^{**} = q_H$.

ii. Let F^* be a mass point at $\theta = \lambda$ and zero otherwise. For any continuous F sufficiently close to F^* (under the total variation distance), there exists a corner equilibrium with $\pi^* = 1$, and $\lim_{F \to F^*} w^{**} = q_H$.

One way to approach the symmetric information limit is by taking $\hat{\lambda} \to 0$, since $\hat{\lambda} = 0$ implies symmetric information. As $\hat{\lambda} \to 0$, there is always a corner equilibrium where nobody signals, and $w^{**} \to q_H$. Hence, this equilibrium smoothly approaches the first best outcome. Pure signaling with $\pi^* = 0$ is also an equilibrium for any positive $\hat{\lambda}$, so the discontinuity does not go away entirely, but the set of equilibria is upper hemi-continuous in $\hat{\lambda}$. A second direction to approach the symmetric information limit is making the distribution F approach a mass point at $\theta = \lambda$, since that limit also implies symmetric information. Again, a corner equilibrium without signaling always exists sufficiently close to the limit, so the set of equilibria is upper hemi-continuous in this dimension as well.[11]

A second form of discontinuity in the uninformed-firms benchmark arises with respect to the cost of signaling. For any $c_H/c_L < 1$, high types will signal enough to fully separate, whereas when $c_H/c_L = 1$ the signal does not allow high types to separate and pooling allocations result. The current model, instead, is upper hemi-continuous as $c_H/c_L \to 1$. In the opposite limit, as signaling becomes cheap, the model reduces to the uninformed-firms benchmark.

Proposition 6.   *i. For c_H/c_L sufficiently close to 1, there exists a corner equilibrium with $\pi^* = 1$.*

ii. For c_H/c_L sufficiently close to 0, the only equilibrium has $\pi^* = 0$

---

[11]By contrast, in the degenerate case where F has full mass at some $\theta < \lambda$, a corner equilibrium never exists (the right-hand side of (28) is zero at θ in this case, so there is always a profitable deviation.) Intuitively, when all firms are equally well informed, our model collapses to standard signaling model and only the full separation equilibrium exists. Hence, heterogeneity of information is crucial to obtain the continuity results in this section.

Part i of Proposition 6 establishes that if signaling is sufficiently expensive, there is an equilibrium with no signaling, where all workers pool at $e = 0$. If within this limiting case one takes the limit as F becomes degenerate at 0 (meaning firms have no information), then this reduces to the pooling allocation in Akerlof (1970). Conversely, part ii establishes that if signaling is sufficiently cheap, then the only equilibrium allocation is the benchmark separating allocation and firms' expertise is not used.

5.5 Welfare

Pareto-Ranking Equilibria. The only reason why allocations in the model are not first-best efficient is that education is socially wasteful; since workers' productivity is the same in any firm and they never end up unemployed, their allocation across firms is irrelevant. Total payoffs are just output minus the cost of education:

$$W = \lambda q_L + (1 - \lambda) q_H - (1 - \lambda) (1 - \pi^*) c_H e^*. \tag{30}$$

Therefore, from a total surplus perspective, equilibria with higher π^* are better. This does not immediately imply that they are Pareto superior; expected wages for different workers are different across equilibria so it is possible that there could be winners and losers from shifting from one equilibrium to another. The following result establishes that interior equilibria can indeed be Pareto-ranked against each other, but cannot be Pareto-ranked against a corner equilibrium with $\pi^* = 1$ if such an equilibrium exists:

Proposition 7.   *i. Let (π_1^*, θ_1^*) and (π_2^*, θ_2^*) be part of two equilibria, each satisfying (18) and (19), with $\pi_2^* > \pi_1^*$. The second equilibrium Pareto dominates the first.*

   ii. Let $(\pi_1^*, \theta_1^*) = (0, \lambda)$ be part of an equilibrium and let θ^{**} be part of an equilibrium satisfying (21). The second equilibrium Pareto dominates the first.

   iii. Let (π_1^*, θ_1^*) with $\pi_1^* > 0$ be part of an equilibrium, satisfying (18) and (19) and let θ^{**} be part of an equilibrium satisfying (21). Neither equilibrium Pareto dominates the other.

In comparing interior equilibria, it is straightforward to show that firms are better off in the higher-π^* equilibrium, since wages are the same and there is a better pool of workers at m^*. High type workers are indifferent because their payoff is w^*. They critical step is to show that low type workers are also better off. They gain from the fact that more firms are hiring in market m^*, which (other things being equal) increases their chances of earning w^* but lose from the fact that there are more high type workers looking for work at m^*, which lowers their chance of being hired by any given firm. However, using the fact that in both equilibria high

types must be hired for sure it is possible to show that the first effect dominates, so low types also prefer the higher-π^* equilibrium.

The corner equilibrium (if it exists) cannot be Pareto ranked against interior equilibria. Since the wage is higher and the cutoff firm is lower, workers are better off in the corner equilibrium. However, the best firms are worse off since they have to pay higher wages and their accurate signals mean they benefit little from the improved pool of workers. Intermediate firms with $\theta \in (\theta^{**}, \theta_1^*)$ are better off in the corner equilibrium since they make positive profits while they would make zero profits in the interior equilibrium.

Information Acquisition. We next ask whether the equilibrium gives firms the right incentives to acquire expertise, following the same approach as Kurlat (2016b). Consider an individual firm and suppose it could invest in becoming better at screening workers. This will affect its profits and also, by affecting the equilibrium, the economy's total deadweight cost of education. Let $\Pi(\theta)$, given by equation (17), denote the individual firm's profits and $W(\theta)$ the equilibrium total payoffs given by equation (30) (ignoring their distribution across workers and firms), both expressed as a function of the individual firm's screening technology θ. Assume the firm's cost of acquiring its screening technology is $c(\theta)$, where $c(\cdot)$ is increasing and convex. Taking the equilibrium as given, firms will invest until the marginal cost of better screening equals the marginal benefit: $c'(\theta) = \Pi'(\theta)$. A social planner interested in minimizing deadweight costs would instead want the firm to invest up to the point where $c'(\theta) = W'(\theta)$. Using the model, we can compute the ratio

$$r(\theta) = \frac{W'(\theta)}{\Pi'(\theta)}. \tag{31}$$

If $r(\theta) > 1$, the marginal social value of better screening is greater than the marginal cost. This would provide a rationale for subsidizing investments in such things as assessment models for job applicants, risk scoring models in insurance markets or pricing models for stocks and financial derivatives. Conversely, if $r(\theta) < 1$, there would be a case for taxing those investments.

The following proposition provides a formula for $r(\theta)$ that relates it to equilibrium objects which, in principle, could be measured, and places a lower bound on it. Denote by

$$\eta \equiv \frac{1}{\pi^*} \left. \frac{\partial \pi^*}{\partial \Delta} \right|_{\Delta=1} \tag{32}$$

the elasticity of the share of high types who do not signal with respect to an increase in demand.

Proposition 8.   *i. The ratio of social to private marginal value of expertise is $r(\theta) = \frac{c_H}{c_L}\eta$.*

   ii. The elasticity η is greater than 1.

33

First, Proposition 8 establishes, perhaps surprisingly, that $r(\theta)$ does not depend on θ. One might have conjectured that the misalignment of incentives would be different for firms that, e.g. due to different cost functions $c(\cdot)$, choose different θ in equilibrium. Yet, it turns out that, if the market under- or over-provides incentives to improve direct screening, it does so uniformly for all firms.

Second, Proposition 8 shows that r can be written as the product of the signaling cost ratio and the demand elasticity of π^*. The ratio c_H/c_L enters the formula because, by equation (30), it governs the deadweight cost of signaling for a high type that chooses e^*.

To understand the role of the elasticity of π^* with respect to demand, observe that, again by (30), $W'(\theta)$ crucially depends on how the equilibrium π^* changes in response to an individual firm's screening technology θ. If a firm improves its screening technology, it will reject more low type applicants and therefore hire more high types, so the market clearing condition shifts outwards. Recall from Section 5.3 that demand affects the equilibrium through exactly the same channel: by producing an outward shift in the market clearing condition. Hence, η precisely summarizes the effect of a firm's expertise on π^*. In particular, we show in the proof of Proposition 8 that the overall effect on π^* depends on the size of the shift to the market clearing condition and on the difference between the slopes of the indifference and market clearing conditions. For example, when these slopes are very similar, π^* will respond strongly to a firm's expertise and η will be large.

Overall, the result implies that it is desirable to encourage investments in direct screening if the cost of signaling is relatively similar for high and low types (which makes the deadweight cost of signaling high) and if the signaling decisions of high types are highly sensitive to demand (which would make them highly sensitive to improved screening as well). For example, higher dividend taxes make the signaling costs of different types more similar, thereby making an underinvestment in expertise more likely. Moreover, the cost ratio and the demand elasticity of π^* are *sufficient* to determine the magnitude of r. Conditional on these two statistics, knowledge of other parameters, such as the shape of the cost function $c(\cdot)$, are not required. As usual with sufficient statistics though, η is of course endogenous to the equilibrium.

The second part of Proposition 8 establishes a lower bound of 1 on the elasticity η, which in turn implies a lower bound of c_H/c_L on r. To understand this, suppose there is an increase in demand of $\Delta\%$. If the mix of workers in market m^* remained constant, each firm in $[\theta^*, \lambda]$ would hire $\Delta\%$ more high types, implying an elasticity of 1. However, precisely because π^* increases, the mix of workers available in market m^* improves, so each firm increases its hiring of high types by more than $\Delta\%$. Furthermore, higher π^* means that marginal firms enter market m^*, further increasing demand. The strength of this last effect depends on the density $f(\theta^*)$ of firms near the cutoff θ^*. Since this density could be very high (to the point where the slopes of the indifference and market clearing conditions are the same, leading to an unbounded response

of π^* to Δ), there is no upper bound on r.

The magnitude of r depends on the relative importance of the various externalities from a firm choosing its screening technology. First, in an interior equilibrium, improved screening always helps other firms, since it leads more high types to forgo education and improves the mix of workers available at m^*. Second, it is neutral for high type workers since they get a payoff of w^* regardless. Third, the effect on low types with $i > \theta$ is also positive. In principle, there are offsetting effects: these workers benefit from having more firms hiring in market m^* and lose from having more high type workers looking for work in m^*. However, just like when one compares across equilibria, the market clearing condition implies that the first effect dominates. Lastly, for low types with $i < \theta$ the effect is ambiguous, because better screening increases their chances of being rejected. If this last effect is negative and strong enough, the sum of the externalities could be negative, which would lead to $r < 1$.

If instead of being in an interior equilibrium the economy is at a corner equilibrium with $\pi^* = 1$, it is immediate that improved screening has no marginal social value, since no worker is signaling. It would still have a positive marginal *private* value, so $r = 0$. In this region, better screening by one firm has a *negative* effect on other firms, since it does not improve the pool of workers in market m^{**} but drives up the wage w^{**} .

6   False Negatives

We finally turn to the other case where firms can be ranked by their expertise, with "false negative" mistakes, so the support of F is $[\lambda, 1]$. All low types are now fully detectable for all firms. High types, in turn, can be ranked in terms of how easy they are to detect. Higher-i workers are relatively transparent, since most firms can tell (with certainty) that they are high types, while lower-i high types are relatively obscure, since they can only be identified as high types by the smarter, lower-θ firms. Unlike the false positives case, firms now face a nontrivial decision as to what hiring rule to use. There may be markets where a firm θ observes $x(i, \theta) = 0$ for all the workers that apply (so if it insisted on hiring only workers with a positive signal it would not be able to hire at all) but the firm knows that in equilibrium some high type workers with $i \in [\lambda, \theta)$ do apply there, so it may want to hire from the pool of *all* applicants. We refer to this as non-selective hiring.

6.1   Equilibrium without Bunching

For expositional purposes, assume first that the density of firms $f(\theta)$ is strictly increasing, meaning that there is a higher density of less-informed firms. In this subsection, we describe the equilibrium that emerges in this case, which is illustrated in Figure 8. A formal characterization
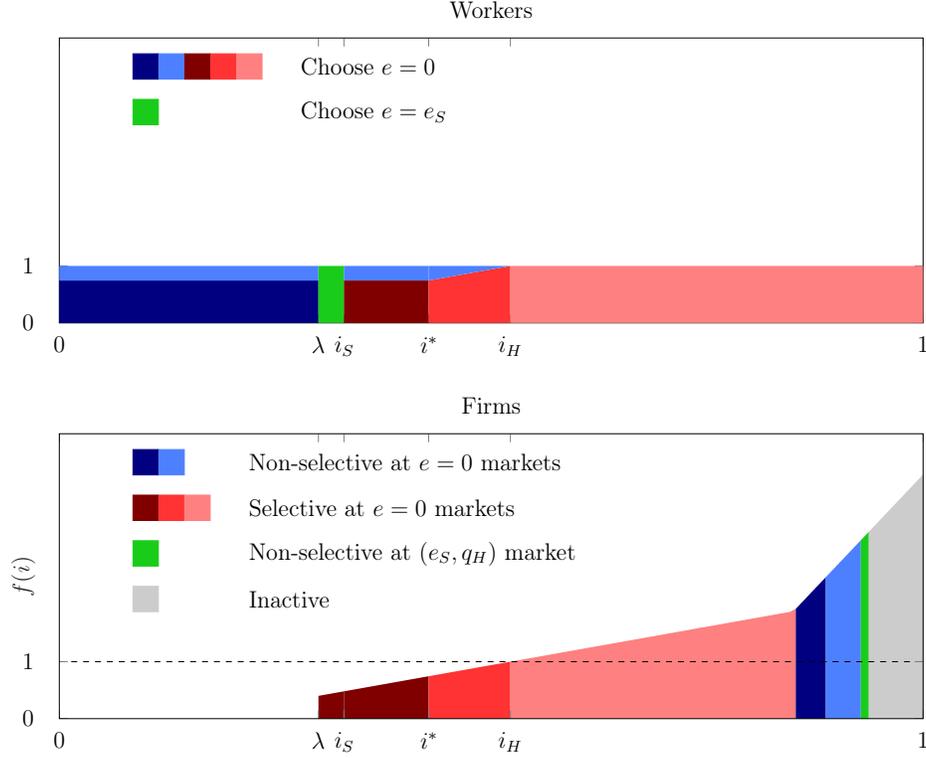
35

Figure 8: Equilibrium matching of workers and firms with false negatives

of the fully general case is provided in the Appendix.

Define i_H by

$$f\left(i_H\right) = 1. \tag{33}$$

Workers with $i \geq i_H$ are the most transparent. Take any $i' > i_H$ and consider the interval $[i_H, i']$. Firms with $\theta \in [i_H, i']$ can identify good workers in the interval $[i_H, i']$ but cannot identify any high type worker with $i < i_H$. Equation (33), together with a strictly increasing density imply that $F\left(i'\right) - F\left(i_H\right) > i' - i_H$, so there are more firms than workers in the interval $[i_H, i']$. Firms will compete for these workers, bidding their wage up to q_H. In equilibrium, these workers choose $e = 0$, set a reservation wage of q_H, and are hired (selectively) by firms with $\theta \geq i_H$ in market $(0, q_H)$.

For $i < i_H$, there are fewer firms that can identify workers as high types than there are workers. The most obscure ones, in the interval $[\lambda, i_S]$, will signal by choosing enough education that the low types do not want to imitate them. Letting u_L be the utility achieved by low types, the education level needed to separate is

$$e_S = \frac{q_H - u_L}{c_L}. \tag{34}$$

Since $u_L \geq q_L$, this satisfies $e_S \leq e^*$. These workers will be hired in market (e_S, q_H) by firms with $\theta > i_H$, who hire non-selectively knowing that only high types apply to this market.

36

Workers in the intermediate range (i_S, i_H) will choose $e = 0$ and be partially hired non-selectively by firms with $\theta > i_H$ and partially hired selectively by the highest-θ firm that can still identify them as good. Let $\underline{w}(0, i)$ denote the reservation wage of one such worker i. This will be increasing in i: more transparent workers do not need to apply to lower-wage markets in order to find jobs. Therefore, a firm that hires non-selectively in market $(0, \underline{w}(0, i))$ will be hiring from a pool that includes high types in the interval $[i_S, i]$ (but not those above that), plus all the low types. The zero-profit condition for such a firm is

$$\underline{w}(0, i) = \frac{(i - i_S) q_H + \lambda q_L}{i - i_S + \lambda}. \tag{35}$$

Firms with $\theta < i_H$ will hire selectively in the cheapest market where they can identify high type workers. This means that they will hire $f(i)$ high types in market $(0, \underline{w}(0, i))$. In turn, this implies worker i must be able to sell his other $1 - f(i)$ units of labor at higher wages, so selling $f(i)$ units in market $(0, \underline{w}(0, i))$ exhausts his endowment. Since at those higher wages only firms with $\theta > i$ hire selectively, and they all reject worker i, all $1 - f(i)$ units must be sold to non-selective firms. This implies that *any* worker below i, including the low types, can sell $1 - f(i)$ units of labor across markets with wages at least $\underline{w}(0, i)$. The payoff for a high type worker in the interval (i_S, i_H) is therefore

$$u(i) = f(i) \underline{w}(0, i) + \int_i^{i_H} \underline{w}(0, i') \, df(i'). \tag{36}$$

Since $u(i)$ is increasing, this defines a cutoff worker i^* that is indifferent between $e = 0$ and signaling, which delivers a payoff of $q_H - c_H e_S$:

$$f(i^*) \underline{w}(0, i^*) + \int_{i^*}^{i_H} \underline{w}(0, i') \, df(i') = q_H - c_H e_S. \tag{37}$$

Market $(0, \underline{w}(0, i^*))$ is the lowest-wage market at which both high and low types apply, so all $f(i^*)$ units of labor that low types cannot sell at wages above $\underline{w}(0, i^*)$ are sold in market $(0, q_L)$. Hence, the utility of low types is

$$u_L = f(i^*) q_L + \int_{i^*}^{i_H} \underline{w}(0, i') \, df(i'). \tag{38}$$

Replacing (34), (35) and (38) into (37) and simplifying gives the following indifference condition for the marginal worker i^*:

$$\Gamma(i^*, i_S) = f(i^*) \left(\frac{c_H}{c_L} - \frac{\lambda}{i^* - i_S + \lambda} \right) - \lambda \left(1 - \frac{c_H}{c_L} \right) \int_{i^*}^{i_H} \frac{1}{i - i_S + \lambda} df(i) = 0. \tag{39}$$

Equation (39) defines a positive relationship between the measure $i_S - \lambda$ of high type workers who signal and the worker i^* who is indifferent between signaling and not signaling. It is straightforward to show that i^* is increasing in i_S. Intuitively, workers who signal do not apply for jobs in $e = 0$ markets. Higher i_S (more signaling) means the pool of applicants for non-selective firms worsens, so in order to maintain zero-profits the wage must fall (equation (35)). In turn, this means that the utility of both high and low type workers falls (equations (36) and (38)). It falls more for high types because low types sell a fraction of their labor in market $(0, q_L)$, where the wage is unaffected by higher i_S. Hence, other things equal, higher i_S makes signaling to separate more attractive, so the indifferent type i^* rises.

The final equilibrium condition requires that all workers below i^* who choose $e = 0$ must obtain the same utility as worker i^*, since otherwise they would prefer to signal. This implies that their reservation wage must also be $\underline{w}(0, i^*)$, and they must be hired for sure. There is a measure $i^* - i_S$ of workers in this category, and a fraction $1 - f(i^*)$ are hired at wages above $\underline{w}(0, i^*)$, so $f(i^*)(i^* - i_S)$ must be hired in market $(0, \underline{w}(0, i^*))$. The measure of firms who are capable of detecting them is $F(i^*)$, and therefore in equilibrium we must have the following market clearing condition:

$$F(i^*) = f(i^*)(i^* - i_S). \tag{40}$$

Equation (40) defines another positive relationship between i_S and i^*. If more of the obscure workers decide to signal, then the most informed firms will work their way up to hire slightly-less-obscure workers.

Equations (39) and (40) hold for an interior equilibrium where some range of workers are indeed hired by non-selective firms. However, it is possible that all workers below i_H prefer to signal rather than being hired at a wage low enough to attract non-selective firms, so there is a corner equilibrium with $i^* = i_H$. For this corner equilibrium, the market clearing condition (40) and definition (33) imply $i_S = i_H - F(i_H)$. Also, in this corner equilibrium, there is no non-selective hiring, so $u_L = q_L$ and $e_S = e^*$. This will be an equilibrium if workers just below i_H indeed prefer to signal:

$$\underbrace{q_H - c_H e^*}_{\text{utility of signaling}} > \underbrace{\frac{F(i_H) q_H + \lambda q_L}{F(i_H) + \lambda}}_{\text{wage for nonselective firms to break even}}$$

so, using (11),

$$1 - \frac{c_H}{c_L} > \frac{F(i_H)}{F(i_H) + \lambda}, \tag{41}$$

which is equivalent to $\Gamma(i_H, i_H - F(i_H)) < 0$. Figure 9 plots equations (39) and (40) for two examples, one with an interior solution and one with a corner solution.

Even though both the indifference and the market clearing condition are upward sloping, the following result establishes that they intersect at most once, leading to a unique equilibrium.
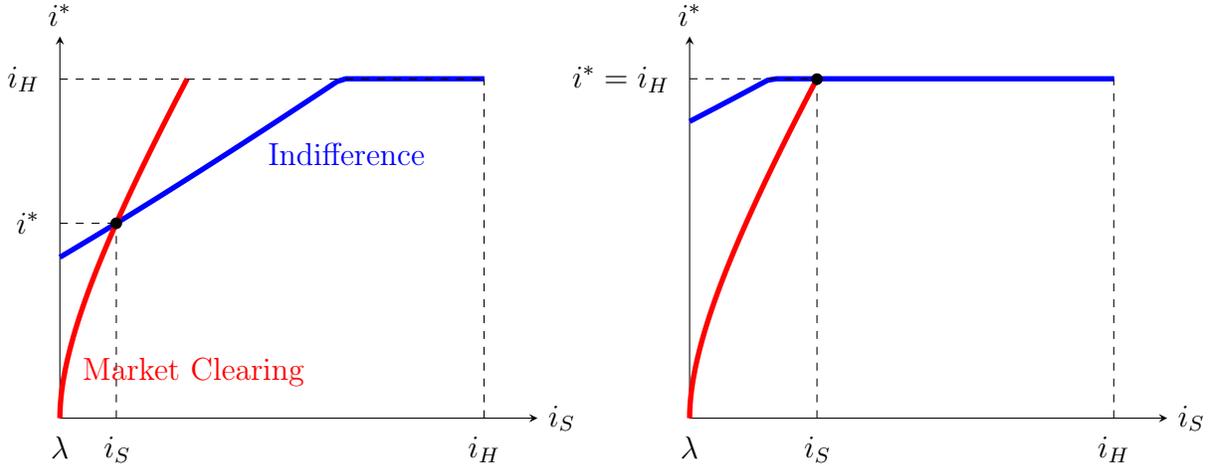
Figure 9: Indifference and market clearing with false negatives; interior and corner equilibrium

Proposition 9. *If condition* (41) *holds, the system of equations* (39), (40) *has no solution. Otherwise, it has a unique solution.*

Taking stock, in contrast to the case with false positives, this model is able to generate dispersion in reservation wages among workers who are equally productive *and educated*, depending on how transparent they are. In particular, high types $i \in [i^*, i_H)$, who all select $e = 0$, set reservation wages $\underline{w}(0, i)$ given by (35), which are increasing in i. Similarly, the model can explain, for instance, different prices for asset-backed securities for which both the structure of tranches and the underlying cash flows are similar, but which differ in how many buyers have access to accurate pricing models to evaluate them. Interestingly, this dispersion is driven by break-even conditions of firms that are *not* making use of expertise. Moreover, the false negatives case predicts that the level of signaling e_S will in general depend on the distribution of expertise (again unlike the false positives case).

The structure of equilibrium is similar to the frequently observed pattern of signaling and "countersignaling" (Feltovich et al., 2002): It is the hard-to-identify high types who must use the costly signal in order to differentiate themselves from low types. By contrast, the most obvious high types can be confident that expert buyers are able to tell them apart, thus eliminating the need for signaling. The setup in Feltovich et al. (2002) features three different levels of worker productivity; in our two-type model, countersignaling instead emerges because high types differ in their transparency. Moreover, our model generates the intuitive prediction that expected wages of those high types who "countersignal" increase in their transparency.

Proposition 12 in the Appendix deals with the general case (allowing for f not to be monotone) and shows that the same characterization of equilibrium goes through when using the "ironed" density of firms. Hence, when f is decreasing, there will be "bunching" of different workers at the same reservation wage. We show that an equilibrium always exists and is gener-
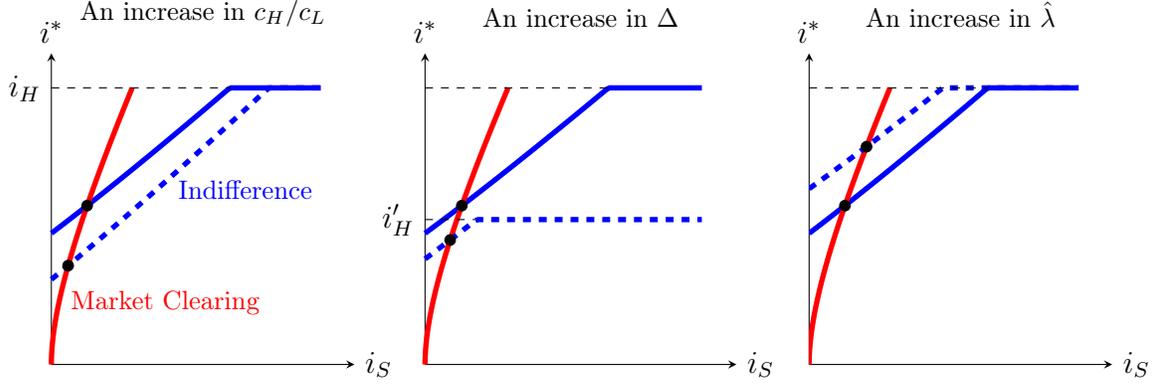
Figure 10: Comparative statics with false negatives

ically unique.

6.2   Comparative Statics

Finally, the following proposition summarizes how the number of workers who signal, the level of signaling needed to separate from low types, and the identity of the marginal worker i^* depend (locally) on parameters around an interior equilibrium.[12]

Proposition 10.   *i. i_S, i^* and e_S are decreasing in the ratio of signaling costs c_H/c_L*

ii. i_S and i^* are decreasing but e_S is increasing in the demand for workers Δ

iii. i_S and i^* are invariant with respect to productivities q_H and q_L; e_S is increasing in $q_H - q_L$

iv. i_S, i^* and e_S are increasing in the fraction of low type workers $\hat{\lambda}$.

These comparative statics are illustrated in Figure 10. The first result shows that an increase in the ratio c_H/c_L reduces the amount of signaling e_S. This is in contrast to the no-information benchmark (and the false positives case), where those who signal choose e^*, which only depends on c_L. Recall that c_H/c_L governs the cost for high types of signaling enough to separate from low types. If this rises, then fewer high types will choose to signal. Graphically, the indifference condition shifts downwards. Since this improves the pool of workers in $e = 0$ markets, the wage that nonselective firms can pay rises. Since low types sell a fraction of their labor to those firms, their utility rises. This, in turn, implies that it takes less education to separate from them because their net gain from earning the wage q_H is lower, so e_S falls. For example, this result predicts that an increase in dividend taxes, raising c_H/c_L, leads to both a smaller fraction of firms paying dividends and a lower amount of dividends paid per firm.

---

[12]Changes in the demand for workers and in the fraction of low types are parametrized as in Section 5.3.

An increase in demand means that at every level of expertise there are more selective hires, and therefore fewer non-selective hires, so it is harder for low types and obscure high types to get hired above their reservation wage, and more likely that they will have to take their reservation wage. This makes low types worse off, and therefore it requires more signaling to successfully separate from them. In turn, this implies that fewer high type workers end up signaling. Hence, unlike the no-information benchmark and the false positives case, higher demand leads to polarization: fewer workers choose positive education but those who do, interestingly, choose a higher quantity.

As in the false-positives case, q_H and q_L drop out of equations (39) and (40). Equation (36) implies that high types who choose $e = 0$ get a wage that is a weighted average of q_H and q_L. In turn, equations (34) and (38) imply that the payoff of high types who signal is also a weighted average of q_H and q_L. Therefore, the indifferent type i^* is such that the weights on q_H and q_L are the same under both alternatives, regardless of q_H and q_L. As is true in the no-information and the false positives case, a greater gap between q_H and q_L makes it more attractive for low types to mimic high types, so separation requires a more intense signal.

Finally, an increase in the fraction of low types means that the wage that nonselective firms are willing to pay falls. This lowers the utility of obscure high types, so that more of them choose to signal, and also the utility of low types, which means that more signaling is needed to separate from them. Graphically, the indifference condition shifts upwards.

6.3 Continuity in the Symmetric Information and No-Signaling Limits

As in the false positives case, the model behaves continuously in the symmetric information limit and in the expensive signaling limit.

Proposition 11.

  i. $\lim_{\hat{\lambda} \to 0} \underline{w}(0, i) = q_H$, $\lim_{\hat{\lambda} \to 0} i^* = \lambda$, and $\lim_{\hat{\lambda} \to 0} i_S = \lambda$, for all $i \in [i^*, i_H)$

  ii. Let F^* be a mass point at $\theta = \lambda$. $\lim_{F \to F^*} i_H = \lambda$.

  iii. $\lim_{\frac{c_H}{c_L} \to 1} i_S = \lambda$.

Part (i) says that as the fraction of low types goes to zero, nobody signals and everyone is paid q_H. Part ii says that as firms become fully informed, again nobody signals and all high types are paid q_H. Part iii says that if signaling is sufficiently expensive, no workers are going to signal in equilibrium. In all cases the equilibrium allocations are continuous in the limit, unlike in the uninformed-firms benchmark.

7 Conclusion

We have developed a general theory to analyze competitive equilibria in economies where buyers possess heterogeneous information about sellers and contracts are multidimensional, specifying both a price and a signal. These information and contracting patterns are the feature of many markets, including labor, asset and insurance markets, as we have illustrated through a series of examples. Our notion of equilibrium allows us to obtain strong results: we prove that an equilibrium always exists, it may not be unique in the false-positives case but is generically unique in the false-negatives case, and it may not be efficient. Moreover, we uncover a tractable structure to characterize it in both cases, based on the intersection of an indifference and a market clearing condition. This allows us to provide results on comparative statics. Our model predicts intuitive and continuous equilibrium responses to, for instance, changes in the prior, demand, signaling costs or expertise that cannot be generated in the canonical model with uninformed buyers.

We expect that our framework can be extended to study other structures of buyers' direct information, including ones where firms cannot be perfectly ranked by their expertise, such as when both false positive and negative errors occur. In this case, we conjecture the equilibrium to feature a combination of the two pure cases we have analyzed: high types are hired in a similar way as in the false negatives-case, except that those in $[i_S, i_H)$ are partly hired by selective false-positive firms, because those firms have an advantage over non-selective firms by being able to screen out some low types.

Our model may also be a useful starting point to study a number of richer environments. First, a market for information may arise, where better informed firms sell their information to less informed ones (e.g. in the form of analyst reports), instead of just trading on it themselves. To prevent the price of information from dropping to zero, some form of capacity constraints would again be required, which would effectively change the distribution of expertise in our model. Second, many of our applications have a dynamic aspect, where the costly signal involves a delay in trading. Our approach could be used to consider settings where some direct information is revealed to buyers gradually at heterogeneous rates, and one could explore how this affects the timing pattern of trades. These issues are left for future research.

References

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Azevedo, Eduardo M. and Daniel Gottlieb, "Perfect Competition in Markets With Adverse Selection," *Econometrica*, 2017, *85* (1), 67–105.

Beaman, Lori and Jeremy Magruder, "Who gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, 2012, *102*, 3574–3593.

Bernardo, Antonio and Bradford Cornell, "The Valuation of Complex Derivatives by Major Investment Firms," *Journal of Finance*, 1997, *52*, 785–798.

Bernheim, B. Douglas, "Tax Policy and the Dividend Puzzle," *RAND Journal of Economics*, 1991, *22*, 455–476.

_ **and Adam Wantz**, "A Tax-Based Test of the Dividend Signaling Hypothesis," *American Economic Review*, 1995, *85*, 532–551.

Bhattacharya, Sudipto, "Imperfect information, Dividend Policy, and "The Bird in the Hand" Fallacy," *Bell Journal of Economics*, 1979, *10*, 259–270.

Bisin, Alberto and Piero Gottardi, "Efficient Competitive Equilibria with Adverse Selection," *Journal of Political Economy*, 2006, *114* (3), 485–516.

Board, Simon, Moritz Meyer-ter-Vehn, and Tomasz Sadzik, "Recruiting Talent," 2016. UCLA Working paper.

Burks, Stephen, Bo Cowgill, Mitchell Hoffman, and Michael Housman, "The Value of Hiring through Employee Referrals," *Quarterly Journal of Economics*, 2015, *130*, 805–839.

Cho, In-Koo and David M. Kreps, "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 1987, *102* (2), 179–221.

Daley, Brendan and Brett Green, "Market Signaling with Grades," *Journal of Economic Theory*, 2014, *151*, 114–145.

DeMarzo, Peter, "The Pooling and Tranching of Securities: A Model of Informed Intermediation," *Review of Financial Studies*, 2005, *18*, 1–35.

_ **and Darrell Duffie**, "A Liquidity-Based Model of Security Design," *Econometrica*, 1999, *67* (1), 65–99.

Dubey, Pradeep and John Geanakoplos, "Competitive Pooling: Rothschild-Stiglitz Reconsidered," *The Quarterly Journal of Economics*, 2002, *117* (4), 1529–1570.

Feltovich, Nick, Richmond Harbaugh, and Ted To, "Too cool for school? Signalling and countersignalling," *RAND Journal of Economics*, 2002, pp. 630–649.

Gale, Douglas, "Equilibria and Pareto Optima of Markets with Adverse Selection," *Economic Theory*, 1996, *7* (2), 207–235.

Guasch, J. Luis and Andrew Weiss, "Wages as Sorting Mechanisms in Competitive Markets with Asymmetric Information: A Theory of Testing," *Review of Economic Studies*, 1980, *47*, 653–664.

Guerrieri, Veronica and Robert Shimer, "Dynamic Adverse Selection: A Theory of Illiquidity, Fire Sales, and Flight to Quality," *American Economic Review*, 2014, *104* (7), 1875–1908.

_ , _ , **and Randall Wright**, "Adverse Selection in Competitive Search Equilibrium," *Econometrica*, 2010, *78* (6), 1823–1862.

Kihlstrom, Richard E. and Michael Riordan, "Advertizing as a Signal," *Journal of Political Economy*, 1984, *92*, 427–450.

Kurlat, Pablo, "Asset Markets with Heterogeneous Information," *Econometrica*, 2016, *84*, 33–85.

_ , "The Social Value of Financial Expertise," 2016. Stanford University Working Paper.

Leland, Hayne E. and David H. Pyle, "Informational Asymmetries, Financial Structure, and Financial Intermediation," *The Journal of Finance*, 1977, *32* (2), 371–387.

Lockwood, Ben, "Information Externalities in the Labour Market and the Duration of Unemployment," *Review of Economic Studies*, 1991, *58*, 733–753.

Mattey, Joe and Nancy Wallace, "Housing-Price Cycles and Prepayment Rates of U.S. Mortgage Pools," *Journal of Real Estate Finance and Economics*, 2001, *23*, 161–184.

Milgrom, Paul R. and John Roberts, "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 1986, *94*, 796–821.

Miyazaki, Hajime, "The Rat Race and Internal Labor Markets," *The Bell Journal of Economics*, 1977, *8* (2), 394–418.

Nelson, Philip, "Advertising as Information," *Journal of Political Economy*, 1974, *82*, 729–754.

Netzer, Nick and Florian Scheuer, "A Game Theoretic Foundation of Competitive Equilibria with Adverse Selection," *International Economic Review*, 2014, *55* (2), 399–422.

Ross, Stephen A., "The Determination of Financial Structure: The Incentive-Signalling Approach," *Bell Journal of Economics*, 1977, *8*, 23–40.

Rothschild, Michael and Joseph E. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 1976, *90* (4), 630–49.

Spence, Michael, "Job Market Signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Wilson, Charles, "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 1977, *16* (2), 167 – 207.

Appendix

Proof of Proposition 1

i. The proposed $\{\pi, \delta, \mu, G\}$ is an equilibrium.

(14) implies that low types are indifferent between any $e \in [0, e^*]$ and high types are indifferent between any $e \geq e^*$, so education decisions (12) solve the workers' problem. (15) implies that firms can make zero profits by hiring in market $\{e = 0, w = q_L\}$ (where there are only low types) or $\{e = e^*, w = q_H\}$ (where there are only high types), and any other market has either $G(\cdot; m) = \emptyset$ or results in losses. Therefore (13), which places demand only in markets $\{e = 0, w = q_L\}$ and $\{e = e^*, w = q_H\}$ and yields zero profits, is an optimal choice. Furthermore, (13) implies $\delta_\theta(M, X) = \frac{1}{F(0)} < 1$ so no firm hires more than one worker. (14) implies that the reservation wages are:

$$\underline{w}(e, i) = \min\{q_L + c_L e, q_H + c_H(e - e^*)\} \tag{42}$$

Replacing (42) and (12) in (6) implies that supply is:

$$S(i, m) = \begin{cases} \mathbb{I}(i < \lambda) & \text{if } e(m) = 0, w(m) \geq q_L \\ \mathbb{I}(i \geq \lambda) & \text{if } e(m) = e^*, w(m) \geq q_H \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

and replacing (13) in (7) implies that demand is:

$$D(m) = \begin{cases} \lambda & \text{if } e(m) = 0, w(m) = q_L \\ 1 - \lambda & \text{if } e(m) = e^*, w(m) = q_H \\ 0 & \text{otherwise} \end{cases}$$

and using (8), hires are:

$$H(w; e, i) = \begin{cases} \mathbb{I}(i < \lambda) & \text{if } e = 0, w \leq q_L \\ \mathbb{I}(i \geq \lambda) & \text{if } e = e^*, w \leq q_H \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

Equations (43) and (44) imply that (9) holds. Finally, (43) implies that beliefs (15) satisfy the consistency condition (10) from Definition 1 in nonempty markets. Since low types find $e \in [0, e^*]$ optimal and high types find $e \geq e^*$ optimal, and reservation wages are given by (42), then beliefs

satisfy condition (ivb) when they are well defined, and they are not well defined only at wages where $\mu(w; e, i) = 1$ for all i, so condition (ivc) is satisfied as well.

ii. The above equilibrium is unique

(a) In any equilibrium, each firm makes zero profits in each market m where it places demand $\delta(m) > 0$. If there was a firm that made strictly negative profits in some market m with $\delta(m) > 0$, it could increase profits by setting $\delta(m) = 0$. On the other hand, suppose there is a firm that makes strictly positive profits in some market m. Recall that $F(1) > 1$, so there must exist a strictly positive measure of firms that each hire strictly less than one worker. Any firm who hires less than one worker could increase its profits by raising $\delta(m)$, so it cannot be optimizing.

(b) In any equilibrium, there does not exist a market m such that $S(i, m) > 0$ both for some $i < \lambda$ and some $i' \geq \lambda$. Otherwise, pick such a market m with a positive supply of both high and low types. Consider a market m' with $e(m') = e(m) + \epsilon$ and $w(m') \in (\bar{w}(e(m), i') + c_H \epsilon, q_H)$. Suppose type $i < \lambda$ is in the support of $G(\cdot; m')$. This requires

$$\bar{w}(e(m'), i) - c_L e(m') \geq \bar{w}(e(m), i) - c_L e(m)$$

Rearranging gives $\bar{w}(e(m'), i) - \bar{w}(e(m), i) \geq c_L \epsilon$. Since firms cannot discriminate, it follows that $\bar{w}(e(m), i)$ is the same for all i. Also, since $\mu(w(m'), i, e(m'))$ is constrained to be increasing in i, $\bar{w}(e(m'), i)$ is also weakly increasing in i. Therefore:

$$\bar{w}(e(m'), i') - \bar{w}(e(m), i') \geq c_L \epsilon$$

which implies $\bar{w}(e(m'), i') - \bar{w}(e(m), i') > c_H \epsilon$. This contradicts the premise that type i' finds $e(m)$ optimal. Hence, no $i < \lambda$ can be in the support of $G(\cdot; m)$. If the support of $G(\cdot,; m)$ only includes $i \geq \lambda$, then firms could make profits by hiring in market m', which contradicts part (iia). Therefore it must be that $G(\cdot; m') = \emptyset$. This implies that $\mu(w(m'), e(m'), i') = 1$, which in turn implies $\bar{w}(e(m'), i') > \bar{w}(e(m), i') + c_H \epsilon$, which contradicts the premise that type i' finds $e(m)$ optimal.

(c) In any equilibrium, all low types obtain a payoff of q_L. Suppose that they obtained a payoff $q'_L > q_L$. This implies that they are hired with positive probability in a market with $w(m) > q_L$. By part (iib), the supply in this market only includes low types, which implies negative profits for firms, contradicting part (iia). Suppose that they obtained a payoff $q'_L < q_L$ and consider a market m with $e(m) = 0$ and $w(m) \in (q'_L, q_L)$. If $G(\cdot; m) \neq \emptyset$, then firms can make profits by hiring in this market; otherwise, $\mu(w(m), e(m), i) = 0$, which means low type workers can obtain a payoff $w(m) > q'_L$ by choosing $e = 0$.

(d) In any equilibrium, all high types obtain payoff $q_H - c_H e^*$. Suppose first that they obtained a payoff $u_H > q_H - c_H e^*$. If they did so by selecting $e' < e^*$, then this implies $\bar{w}(e', i) - c_H e' > q_H - c_H e^*$, which implies $\bar{w}(e', i) - c_L e' > q_L$ and since $\bar{w}(e', i)$ is the same for all i, this implies that low types can obtain a payoff higher than q_L. If instead $e' \geq e^*$, this implies they are hired with positive probability at a wage $w > q_H$ and hence strictly negative profits for firms, contradicting (iia). Second, suppose they obtained a payoff $u_H < q_H - c_H e^*$. This means that for any $i \geq \lambda$ it must be that $\bar{w}(e^*, i) < q_H$, and therefore $\bar{w}(e^*, i) < q_H$ for $i < \lambda$ as well. Consider a market m with $e(m) = e^*$ and $w(m) \in (u_H + c_H e^*, q_H)$. $G(\cdot; m)$ must be well defined because otherwise $\mu(w(m), e^*, i) = 1$ so high types can obtain a payoff of at least $w(w) - c_H e^* > u_H$ by choosing education e^*. But the support of $G(\cdot; m)$ cannot include low types because choosing e^* implies a payoff of $\bar{w}(e^*, i) - c_L e^* < q_H - c_L e^* < q_L$;

46

and the support of $G(\cdot; m)$ cannot include only high types because then firms could make profits by hiring in market m.

(e) Step (iic) implies that all low types select $e = 0$ and reservation wage q_L and face no rationing in market $\{e = 0, w = q_L\}$. Step (iid) implies that all high types must select $e = e^*$ and reservation wage q_H and face no rationing in market $\{e = e^*, w = q_H\}$. This determines (12) as well as (14) and (15) in these markets. It also requires that there is total demand λ in market $\{e = 0, w = q_L\}$ and demand $1 - \lambda$ in $\{e = e^*, w = q_H\}$, thus (13) must hold. For all other markets, (14) and (15) then follow from the equilibrium definition.

Proof of Proposition 2

We first show that, in any equilibrium, all low types choose $e = 0$ and reservation wage $w = q_L$. Some fraction $\pi \in [0, 1]$ of the high types choose $e = 0$ and reservation wage $w = w^*$ (if $\pi < 1$) or $w \geq w^*$ (if $\pi = 1$). The rest of the high types choose $e = e^*$ and reservation wage $w = q_H$. We prove this claim based on the following sequence of steps:

i. By the same argument as in the proof of Proposition 1, all firms make non-negative profits in all markets m where they place positive demand $\delta_\theta(m) > 0$.

ii. Firms' profits must be weakly increasing in θ. To see this, suppose that $\theta' > \theta$ but firm θ makes strictly higher profits than θ', and let m be such that $\delta_\theta(m) > 0$. But by hiring in market m, firm θ' must make profits at least as high as firm θ since both firms accept all the high types but firm θ' rejects more low types.

iii. There exists some $\bar{\theta}$ such that all firms $\theta \leq \bar{\theta}$ make zero profits, and $F(\bar{\theta}) > 0$. To see this, recall that at least a measure $F(1) - 1 > 0$ of firms must hire less than one worker. If any of these firms make profits in some market, then as in the proof of Proposition 1 they could increase their profits by increasing their hiring. Hence, there is a strictly positive measure of firms that make zero profits. The claim then follows from the monotonicity of profits in θ.

iv. In any equilibrium, low types obtain a payoff of at least q_L. Suppose that they obtained a payoff $q_L' < q_L$ and consider a market m with $e(m) = 0$ and $w(m) \in (q_L', q_L)$. If $G(\cdot; m) \neq \emptyset$, then firms $\theta < \bar{\theta}$ can make profits by hiring in this market; otherwise, $\mu(w(m), e(m), i) = 0$, which means low type workers can obtain a payoff $w(m) > q_L'$ by choosing $e = 0$.

v. In any equilibrium, high types obtain a payoff of at least $w^* = q_H - c_H e^*$. Suppose they obtained a payoff $u_H < q_H - c_H e^*$. This means that for any $i \geq \lambda$ it must be that $\bar{w}(e^*, i) < q_H$, and therefore $\bar{w}(e^*, i) < q_H$ for $i < \lambda$ as well. Consider a market m with $e(m) = e^*$ and $w(m) \in (u_H + c_H e^*, q_H)$. $G(\cdot; m, \chi_\theta)$ must be well defined for all θ because otherwise $\mu(w(m), e^*, \lambda) = 0$ so high types can obtain a payoff of at least $w(m) - c_H e^* > u_H$ by choosing education e^*. But the support of $G(\cdot; m, \chi_\theta)$ cannot include low types for any θ because choosing e^* implies a payoff of $\bar{w}(e^*, i) - c_L e^* < q_H - c_L e^* = q_L$; and the support of $G(\cdot; m, \chi_\theta)$ cannot include only high types for $\theta < \bar{\theta}$ because then firms $\theta < \bar{\theta}$ could make profits by hiring in market m.

vi. For any $i \geq \lambda$ and any e, $\mu(w, e, i)$ is a step function that only takes values 0 and 1. To see this, consider two wage levels $w' > w$. For any i such that $\mu(w, e, i) > 0$, worker optimization implies that $S(i, \{e, w\}) = S(i, \{e, w'\})$. Since no firm discriminates between different high types, if $\mu(w, e, i) > 0$ for some $i \geq \lambda$, then this is true for any $i \geq \lambda$. Furthermore since μ is decreasing in i, this is true for all $i < \lambda$ as well. Therefore if $\mu(w, e, i) > 0$ for some $i \geq \lambda$ then $S(i, \{e, w\}) = S(i, \{e, w'\})$ for all i. Therefore no firm will find it optimal to hire at wages between w and w', which implies $\mu(w, e, i)$ is constant in the interval $[w, w')$.

47

vii. In any equilibrium, all low types get education $e = 0$. To see this, assume to the contrary that some $i < \lambda$ chooses $e' > 0$.

(a) By (iv), we have that $\bar{w}(e', i) \geq q_L + c_L e' > q_L$. Together with step (i), this requires that some high types also choose e', because otherwise firms would make losses by paying more than q_L. Denote by u_H be the payoff obtained by high types. Step (vi) implies that any high type that chooses e' gets a job in market m' where $e(m') = e'$ and $w(m') = u_H + c_H e'$. Therefore there are no high types in any market with $e = e'$ and $w < w(m')$, which means that low types who choose e' can only find jobs at $w = w(m')$ or $w = q_L$. Let $\underline{\theta}$ be the lowest firm type that hires in a market m'.

(b) Let

$$S_H = \int_\lambda^1 S(i, m') \, di$$

be the measure of high types that choose e'. The payoff for type $i < \lambda$ of choosing e' is bounded above by

$$\bar{u}_i(e') = q_L + \int_{\underline{\theta}}^i \frac{1}{S_H} dF(\theta)(w(m') - q_L) - c_L e'$$

which is lower than q_L for i sufficiently close to $\underline{\theta}$. Let $\underline{i}$ be the lowest worker type such that there is a $\delta_1 > 0$ such that all workers $i \in [\underline{i}, \underline{i} + \delta_1]$ choose $e = e'$. We know that $\underline{i} > \underline{\theta}$.

(c) Assume that all firms in the interval $\theta \in [\underline{\theta}, \underline{i}]$ choose to hire in market m'. Since there are no workers with $i < \underline{i}$ in market m', then upon hiring in market m', any firm $\theta \leq \underline{i}$ hires from the entire pool of applicants, without rejecting any. Since this hiring rule is available to all firms, part (iii) implies that all $\theta \leq \underline{i}$ firms must make zero profits by hiring in market m'. For this to be true, it must mean that they cannot make profits in any other market, including any markets with $e = 0$. But any firm with $\theta > \underline{i}$ will be able to reject some workers in the interval $[\underline{i}, \underline{i} + \delta]$, which implies it can make strictly positive profits by hiring in market m'. Therefore, all firms in the interval $(\underline{i}, \underline{i} + \delta_1]$ hire in market m'. This in turn implies that if worker $\underline{i}$ is willing to choose e', then worker $\underline{i} + \delta_1$ *strictly* prefers e', since, compared to worker $\underline{i}$, he has a higher chance of being hired in market m' and the same chance of being hired in any other market. By continuity, this implies that there is a number $\delta_2 > \delta_1$ such that all workers in $[\underline{i}, \underline{i} + \delta_2]$ choose $e = e'$. Repeating the same reasoning, this implies that there is a strictly increasing sequence $\{\delta_n\}$ such that all workers in $[\underline{i}, \delta_n]$ choose e'. Therefore all workers $i \in [\underline{i}, \lambda]$ choose $e = e'$.

   i. Let $\hat{\theta}$ be the highest firm type that hires in market m', and assume $\hat{\theta} = \lambda$. Consider market m'' with $e(m'') = 0$ and $w(m'') = u_H + \epsilon$. If conditional beliefs $G(\cdot; m'', \chi_{\hat{\theta}})$ are well-defined, then they can only include high types because firm $\hat{\theta}$ only accepts high types. For sufficiently small ϵ, this implies that firm $\hat{\theta}$ could make higher profits in market m'' than in market m', a contradiction. Instead, if $G(\cdot; m'', \chi_{\hat{\theta}}) = \emptyset$, this requires $\mu(u_H + \epsilon, 0, i) = 0$ for all $i \geq \lambda$, which implies that $e = 0$ is a better choice than e' for high types, again a contradiction.

   ii. Assume instead that $\hat{\theta} < \lambda$. All workers in $[\hat{\theta}, \lambda]$ are accepted by all firms who hire in market m', which implies that they get wage $w(m')$ for sure and a payoff of $u_H - (c_L - c_H) e'$. Now consider a market $\tilde{m}$ with $e(\tilde{m}) = e' + \epsilon$ and $w(\tilde{m}) \in$

48

$(w(m') + c_H \epsilon, w(m') + c_L \epsilon)$. Suppose type $i \in \left(\hat{\theta}, \lambda \right)$ is in the support of $G\left(\cdot; \tilde{m}, \chi_{\hat{\theta}} \right)$. This requires

$$\bar{w}\left(e\left(\tilde{m} \right), i \right) - c_L e\left(\tilde{m} \right) \geq u_H - (c_L - c_H) e' \quad \Rightarrow \quad \bar{w}\left(e\left(\tilde{m} \right), i \right) - \left[u_H + c_H e' \right] \geq c_L \epsilon$$

Since $\bar{w}\left(e\left(\tilde{m} \right), i \right)$ must be increasing in i, this implies that for any high type worker i':

$$\bar{w}\left(e\left(\tilde{m} \right), i' \right) - \left[u_H + c_H e' \right] \geq c_L \epsilon \quad \Rightarrow \quad \bar{w}\left(e\left(\tilde{m} \right), i' \right) - \left[u_H + c_H e' \right] \geq c_H \epsilon$$

which contradicts the premise that i' finds e' optimal. Hence, no $i \in \left(\hat{\theta}, \lambda \right)$ can be in the support of $G\left(\cdot; \tilde{m}, \chi_{\hat{\theta}} \right)$. If the support of $G\left(\cdot; \tilde{m}, \chi_{\hat{\theta}} \right)$ only includes $i \geq \lambda$, then for small enough ϵ, firm $\hat{\theta}$ would find it more profitable to hire in market $\tilde{m}$ than in market m'. Therefore it must be that $G\left(\cdot; m' \right) = \emptyset$. This implies that $\mu\left(w\left(\tilde{m} \right), e\left(\tilde{m} \right), i' \right) = 0$, which in turn implies that high types prefer $e\left(\tilde{m} \right)$ to $e\left(m' \right)$, again a contradiction.

(d) Assume instead that some firm $\tilde{\theta} \in (\underline{\theta}, \underline{i})$ prefers to hire in some market $m'' \neq m'$. This implies firms $\theta \in \left(\tilde{\theta}, \underline{i} \right)$ also prefer m'' over m', since they hire from the same pool of workers as firm $\tilde{\theta}$ in market m' but from a more selected pool in other markets. Therefore if worker $i = \tilde{\theta}$ prefers e'' over e', so will worker $\underline{i}$, since he obtains the same payoff as worker $i = \bar{\theta}$ upon choosing e' but weakly higher upon choosing e''. This contradicts the assumption that worker $\underline{i}$ chooses e'.

viii. In any equilibrium, the high types select either $e = 0$ or $e = e^*$. If some types selected $e > e^*$, then by step (v) this would require paying them $w > q_H$ and therefore involve negative profits for firms. On the other hand, suppose some high types set $e' \in (0, e^*)$ and let $\underline{w}$ be their reservation wage. For any $w \geq \underline{w}$, beliefs can only place weight on high types since by part (vii), no low types choose e'. This implies that if $\underline{w} < q_H$, any firm, including those with $\theta < \bar{\theta}$, could make profits by hiring in market $(e', \underline{w})$, which contradicts part (i). Therefore we must have $\underline{w} = q_H$. Note that this implies that there can only be a single $e \in (0, e^*]$ such that $\pi_i(e) > 0$ for some $i \geq \lambda$ since otherwise, the high types would only select the lowest such e. Let π be the fraction of high types who choose $e = 0$ and $1 - \pi$ the fraction who choose $e = e'$. Since they must be indifferent, it follows that high types who choose $e = 0$ get a wage of $w' = q_H - c_H e'$. Since all low types choose $e = 0$, firms will find it profitable to hire in market $(e = 0, w = w')$ iff

$$\frac{\pi(1 - \lambda) q_H + (\lambda - \theta) q_L}{\pi(1 - \lambda) + (\lambda - \theta)} > q_H - c_H e'$$

This defines the cutoff firm θ' such that firms with $\theta < \theta'$ make zero profits. Furthermore, this implies that all workers with $i < \theta'$ do not get hired in market $(e = 0, w = w')$ and therefore obtain a payoff of q_L. Let $\Theta \subseteq [0, \theta']$ be the set of firms who hire workers in market m' defined by $e = e'$ and $w = q_H$. Since all high types who choose e' get a job at $w = q_H$ it follows that $F(\Theta) = 1 - \pi$. Suppose worker $i' < \theta'$ chooses $e = e'$. His chance of finding a job at wage q_H will be given by:

$$1 - \mu\left(q_H, e', i' \right) = \frac{F\left(\Theta \cap [0, i'] \right)}{1 - \pi} = \frac{F\left(\Theta \cap [0, i'] \right)}{F(\Theta)}$$

Since F is continuous, then for i' sufficiently close to θ', $\mu\left(q_H, e', i' \right)$ will be arbitrarily close to 0, and therefore (since $e' < e^*$), $(1 - \mu(q_H, e', i'))q_H - c_L e' > q_L$. Thus, there is a low type who would prefer $e = e'$ to $e = 0$, which contradicts step (vii).

ix. Let u_H be the equilibrium payoff of high types.

(a) If $u_H > w^*$, then it must be that all high types choose $e = 0$ and get hired at a wage $w^{**} = u_H$. Firms will find it profitable to hire in this market if

$$\frac{(1 - \lambda) q_H + (\lambda - \theta) q_L}{(1 - \lambda) + \lambda - \theta} > u_H$$

This defines a cutoff $\bar{\theta}$, so (20) holds. Furthermore, since all high types must be hired at this wage, (21) must hold.

(b) If $u_H = w^*$, then high types are indifferent between choosing $e = 0$ and getting hired at wage w^* and choosing $e = e^*$ and getting hired at a wage q_H. Let π be the fraction that choose $e = 0$. Firms will find it profitable to hire in market $(e = 0, w = w^*)$ iff

$$\frac{\pi (1 - \lambda) q_H + (\lambda - \theta) q_L}{\pi (1 - \lambda) + (\lambda - \theta)} > w^*$$

This defines the cutoff $\bar{\theta}$, so (18) holds. Furthermore, since all high types must be hired at w^*, (19) must hold too.

Proof of Proposition 3

We first state the equilibrium objects $\{\pi, \delta, \mu, G\}$.

Interior Equilibrium.

(a) Education decisions:

$$\pi_i(e) = \begin{cases} 1 & \text{if } i < \lambda \text{ and } e = 0 \\ \pi^* & \text{if } i \geq \lambda \text{ and } e = 0 \\ 1 - \pi^* & \text{if } i \geq \lambda \text{ and } e = e^* \\ 0 & \text{otherwise} \end{cases} \tag{45}$$

(b) Demand decisions:

$$\delta_\theta(m, \chi) = \begin{cases} 1 & \text{if } \theta \geq \theta^*, w(m) = w^*, e(m) = 0 \text{ and } \chi(i) = \mathbb{I}(i \geq \theta) \\ \frac{\lambda - \varphi^*}{F(\theta^*)} & \text{if } \theta < \theta^*, w(m) = q_L, e(m) = 0 \text{ and } \chi(i) = 1 \forall i \\ \frac{(1 - \lambda)(1 - \pi^*)}{F(\theta^*)} & \text{if } \theta < \theta^*, w(m) = q_H, e(m) = e^* \text{ and } \chi(i) = 1 \forall i \\ 0 & \text{otherwise} \end{cases} \tag{46}$$

where

$$\varphi^* = \int_{\theta^*}^{\lambda} \frac{\lambda - \theta}{\pi^*(1 - \lambda) + \lambda - \theta} dF(\theta)$$

(c) Rationing function:

$$\mu(w, e, i) = \begin{cases} 1 & \text{if } e = 0, w > w^* \\ 1 - \int_{\theta^*}^{\min\{i, \lambda\}} \frac{1}{\lambda - \theta + \pi^*(1 - \lambda)} dF(\theta) & \text{if } e = 0, w^* \geq w > q_L \\ \mathbb{I}(w \geq \min\{\tilde{w}(e, i), \tilde{w}(e, \lambda)\}) & \text{otherwise} \end{cases} \tag{47}$$

50

(d) Beliefs: for selection rule $\chi(i) = \mathbb{I}(i \geq \theta)$,

$$g(i; m, \chi) = \begin{cases} \frac{\mathbb{I}(i \in [\theta, \lambda)) + \pi^* \mathbb{I}(i \geq \lambda)}{\pi^*(1-\lambda) + \lambda - \theta} & \text{if } e(m) = 0, w(m) \geq w^* \\ \frac{\mathbb{I}(i \in [\theta, \lambda])}{\lambda - \theta} & \text{if } e(m) = 0, w^* > w(m) \geq q_L \\ \text{unit measure at } i = \theta & \text{if } e(m) \in \left(0, e\left(m_\theta^D\right)\right), w(m) \geq \tilde{w}(e(m), \theta) \\ \frac{\mathbb{I}(i \geq \lambda)}{1-\lambda} & \text{if } e(m) \geq e\left(m_\theta^D\right), w(m) \geq \tilde{w}(e(m), \lambda) \\ \emptyset & \text{for any other } m \end{cases} \quad (48)$$

and for selection rule $\chi(i) = 1 \forall i$,

$$g(i; m, \chi) = \begin{cases} \frac{\mathbb{I}(i < \lambda) + \pi^* \mathbb{I}(i \geq \lambda)}{\pi^*(1-\lambda) + \lambda} & \text{if } e(m) = 0, w(m) \geq w^* \\ \frac{\mathbb{I}(i < \lambda)}{\lambda} & \text{if } e(m) = 0, w^* > w(m) \geq q_L \\ \frac{\mathbb{I}(i < \theta^*)}{\theta^*} & \text{if } e(m) \in (0, e^*), w(m) \geq \tilde{w}(e(m), 0) \\ \frac{\mathbb{I}(i \geq \lambda)}{1-\lambda} & \text{if } e(m) \geq e^*, w(m) \geq \tilde{w}(e(m), \lambda) \\ \emptyset & \text{for any other } m \end{cases} \quad (49)$$

We now verify that $\{\pi, \delta, \mu, G\}$ satisfies all the equilibrium conditions from Section 3.4. (47) implies that low types $i \in [0, \lambda)$ are indifferent between any $e \in \left[0, e(m_i^D)\right]$ and high types are indifferent between any $e \geq 0$, so the education decisions (45) solve the workers' problem. The beliefs (48) and (49) together with the fact that condition (27) does not hold imply that firms $\theta \geq \theta^*$ maximize profits by hiring selectively in market $\{e = 0, w = w^*\}$. All other firms make zero profits by hiring non-selectively either in market $\{e = 0, w = q_L\}$ or $\{e = e^*, w = q_H\}$, and any other market has either $G(\cdot; m, \chi_\theta) = \emptyset$ or results in losses. Therefore the demands (46) are an optimal choice. Furthermore, (46) is constructed such that $\delta_\theta(M, X) = 1$ for all $\theta \geq \theta^*$ and $\delta_\theta(M, X) \leq 1$ for all $\theta < \theta^*$ as required by equilibrium. The rationing function (47) implies that the reservation wages are:

$$\underline{w}(e, i) = \min\{\tilde{w}(e, i), \tilde{w}(e, \lambda)\} \quad (50)$$

Replacing (50) and (45) in (6) implies that supply is:

$$S(i, m) = \begin{cases} \mathbb{I}(i < \lambda) + \pi^* \mathbb{I}(i \geq \lambda) & \text{if } e(m) = 0, w(m) \geq w^* \\ \mathbb{I}(i < \lambda) & \text{if } e(m) = 0, w(m) \in [q_L, w^*) \\ (1 - \pi^*) \mathbb{I}(i \geq \lambda) & \text{if } e(m) = e^*, w(m) \geq q_H \\ 0 & \text{otherwise} \end{cases} \quad (51)$$

and replacing (46) in (7) implies that demand in market m for a set of selection rules $X_0(\theta') = \{\chi(i) = \mathbb{I}(i \geq \theta) : \theta \in [0, \theta']\}$ is:

$$D(m, X_0(\theta')) = \begin{cases} \lambda - \varphi^* & \text{if } e(m) = 0, w(m) = q_L \\ \max\{F(\theta') - F(\theta^*), 0\} & \text{if } e(m) = 0, w(m) = w^* \\ (1 - \pi^*)(1 - \lambda) & \text{if } e(m) = e^*, w(m) = q_H \\ 0 & \text{otherwise} \end{cases}$$

and using (8), hires are:

$$H(w; e, i) = \begin{cases} \mathbb{I}(i < \lambda) + \pi^* \mathbb{I}(i \geq \lambda) & \text{if } e = 0, w \leq q_L \\ \int_{\theta^*}^{\min\{i, \lambda\}} \frac{\mathbb{I}(i \in [\theta^*, \lambda)) + \pi^* \mathbb{I}(i \geq \lambda)}{\pi^*(1-\lambda) + \lambda - \theta} dF(\theta) & \text{if } e = 0, w \in (q_L, w^*], i \geq \theta^* \\ (1 - \pi^*) \mathbb{I}(i \geq \lambda) & \text{if } e = e^*, w \leq q_H \\ 0 & \text{otherwise} \end{cases} \quad (52)$$

51

Equations (51) and (52) imply that (9) holds. Finally, (51) implies that beliefs (48) and (49) satisfy the consistency condition (10) from Definition 1 in nonempty markets. Since low types i find $e \in \left[0, e(m_i^D)\right]$ optimal and high types find any $e \geq 0$ optimal, and reservation wages are given by (50), then beliefs satisfy condition (ivb) when they are well defined, and they are not well defined only at wages where $\mu(w; e, i) = 0$ for all i such that $\chi(i) = 1$, so condition (ivc) is satisfied as well.

  Corner Equilibrium.

(a) Education decisions:

$$\pi_i(e) = \begin{cases} 1 & e = 0 \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

(b) Demand decisions:

$$\delta_\theta(m, \chi) = \begin{cases} 1 & \text{if } \theta \geq \theta^{**},\ w(m) = w^{**},\ e(m) = 0 \text{ and } \chi(i) = \mathbb{I}(i \geq \theta) \\ \frac{1 - F(\lambda) + F(\theta^{**})}{F(\theta^{**})} & \text{if } \theta < \theta^{**},\ w(m) = q_L,\ e(m) = 0 \text{ and } \chi(i) = 1 \forall i \\ 0 & \text{otherwise} \end{cases} \tag{54}$$

(c) Rationing function:

$$\mu(w, e, i) = \begin{cases} 1 & \text{if } e = 0,\ w > w^{**} \\ 1 - \int_{\theta^{**}}^{\min\{i, \lambda\}} \frac{1}{1 - \theta} dF(\theta) & \text{if } e = 0,\ w^{**} \geq w > q_L \\ \mathbb{I}\left(w \geq \min\{\tilde{w}(e, i), \tilde{w}(e, \lambda)\}\right) & \text{otherwise} \end{cases} \tag{55}$$

(d) Beliefs: for selection rule $\chi(i) = \mathbb{I}(i \geq \theta)$,

$$g(i; m, \chi) = \begin{cases} \frac{\mathbb{I}(i \geq \theta)}{1 - \theta} & \text{if } e(m) = 0, w(m) \geq w^{**} \\ \frac{\mathbb{I}(i \in [\theta, \lambda))}{\lambda - \theta} & \text{if } e(m) = 0, w^{**} > w(m) \geq q_L \\ \text{unit measure at } i = \theta & \text{if } e(m) \in \left(0, e\left(m_\theta^D\right)\right), w(m) \geq \tilde{w}(e(m), \theta) \\ \frac{\mathbb{I}(i \geq \lambda)}{1 - \lambda} & \text{if } e(m) \geq e\left(m_\theta^D\right), w(m) \geq \tilde{w}(e(m), \lambda) \\ \emptyset & \text{for any other } m \end{cases} \tag{56}$$

and for selection rule $\chi(i) = 1 \forall i$,

$$g(i; m, \chi) = \begin{cases} 1 \forall i & \text{if } e(m) = 0, w(m) \geq w^{**} \\ \frac{\mathbb{I}(i < \lambda)}{\lambda} & \text{if } e(m) = 0, w^{**} > w(m) \geq q_L \\ \frac{\mathbb{I}(i < \theta^{**})}{\theta^{**}} & \text{if } e(m) \in (0, e^{**}), w(m) \geq \tilde{w}(e(m), 0) \\ \frac{\mathbb{I}(i \geq \lambda)}{1 - \lambda} & \text{if } e(m) \geq e^{**}, w(m) \geq \tilde{w}(e(m), \lambda) \\ \emptyset & \text{for any other } m \end{cases} \tag{57}$$

with $e^{**} = (w^{**} - q_L)/(c_L - c_H)$

Checking that this satisfies all the equilibrium requirements is analogous to the interior equilibrium case.

Proof of Proposition 4

Using the reparametrization of the model in terms of Δ and $\hat{\lambda}$, it is straightforward to show that equations (18) and (19) become

$$\theta^* = \lambda - \pi^* \frac{\lambda}{\hat{\lambda}} \left(1 - \hat{\lambda}\right) \frac{q_H - w^*}{w^* - q_L} \tag{58}$$

$$\int_{\theta^*}^{\lambda} \frac{\Delta}{(\lambda - \theta)\frac{\hat{\lambda}}{\lambda} + \pi^* \left(1 - \hat{\lambda}\right)} dF(\theta) = 1. \tag{59}$$

Replacing (11) and (16) into (58), the indifference condition reduces to:

$$\theta^* = \lambda - \pi^* \frac{\lambda}{\hat{\lambda}} \left(1 - \hat{\lambda}\right) \frac{1}{\frac{c_L}{c_H} - 1} \tag{60}$$

Let $\theta^I (\pi^*, p)$ and $\theta^M (\pi^*, p)$ represent the solutions to (60) and (59) respectively, where p is a parameter. The equilibrium value of π^* is given by a solution to the equation $\theta^I (\pi^*, p) - \theta^M (\pi^*, p) = 0$. Using the implicit function theorem, the derivatives of π^* and θ^* with respect to parameter p are given by:

$$\frac{d\pi^*}{dp} = \frac{\frac{\partial \theta^M}{\partial p} - \frac{\partial \theta^I}{\partial p}}{\frac{\partial \theta^I}{\partial \pi^*} - \frac{\partial \theta^M}{\partial \pi^*}} \tag{61}$$

$$\frac{d\theta^*}{dp} = \frac{\partial \theta^I}{\partial p} + \frac{\partial \theta^I}{\partial \pi^*} \frac{\partial \pi^*}{\partial p} \tag{62}$$

By assumption, the denominator of (61) is positive, and equation (60) implies that $\frac{\partial \theta^I}{\partial \pi^*}$ is negative.

i. (60) implies that θ^I is decreasing in $\frac{c_H}{c_L}$, whereas $\frac{c_H}{c_L}$ does not appear in equation (59). Using this in equations (61) and (62) gives the result.

ii. (59) implies that θ^M is increasing in Δ, and Δ does not appear in equation (60). Using this in equations (61) and (62) gives the result.

iii. The distribution F does not appear in equation (60). Rewrite equation (59) as

$$\int_{0}^{\lambda} \frac{\mathbb{I}(\theta \geq \theta^*)}{(\lambda - \theta)\frac{\hat{\lambda}}{\lambda} + \pi^* \left(1 - \hat{\lambda}\right)} dF(\theta) = 1$$

The term inside the integral is an increasing function of θ. Therefore a FOSD increase in F implies that the left hand side of (59) increases, so θ^* must rise to maintain equality. Using this in equations (61) and (62) gives the result.

iv. This follows because neither q_H nor q_L appear in equations (61) and (62).

v. Taking derivatives in equation (60) yields $\partial \theta^I / \partial \hat{\lambda} = \pi^* \lambda (c_L/c_H - 1)^{-1} \hat{\lambda}^{-2} > 0$. Using the implicit function theorem in equation (59):

$$\frac{\partial \theta^M}{\partial \hat{\lambda}} = - \frac{\int_{\theta^*}^{\lambda} \left[\frac{\lambda - \theta}{\lambda} - \pi^*\right]^{-2} dF(\theta) \left[(\lambda - \theta^*)\frac{\hat{\lambda}}{\lambda} + \pi^* (1 - \lambda)\right]}{f(\theta^*)}$$

which could be positive or negative, so equation (61) and (62) cannot be signed in general.

Proof of Proposition 5

i. Equation (20) generalizes to

$$w^{**} = \frac{(\lambda - \theta^{**}) \frac{\hat{\lambda}}{\lambda} q_L + \left(1 - \hat{\lambda}\right) q_H}{(\lambda - \theta^{**}) \frac{\hat{\lambda}}{\lambda} + \left(1 - \hat{\lambda}\right)} \tag{63}$$

Therefore for $\hat{\lambda}$ low enough, $w^{**} > w^*$ so there is a candidate corner equilibrium. Condition (28) generalizes to

$$(\lambda - \theta) \frac{\hat{\lambda}}{\lambda} \frac{(\lambda - \theta^{**}) \frac{\hat{\lambda}}{\lambda} + \left(1 - \hat{\lambda}\right)}{\left[(\lambda - \theta) \frac{\hat{\lambda}}{\lambda} + \left(1 - \hat{\lambda}\right)\right]} > \frac{c_H}{c_L - c_H} \left(1 - \hat{\lambda}\right) \left(\int_\theta^1 \frac{1}{(\lambda - t) \frac{\hat{\lambda}}{\lambda} + \left(1 - \hat{\lambda}\right)} dF(t) \right)$$

which cannot hold for sufficiently low $\hat{\lambda}$, so the candidate equilibrium is indeed an equilibrium. Furthermore, taking the limit in (63) we obtain $\lim_{\hat{\lambda} \to 0} w^{**} = q_H$.

ii. Equation (21) implies that $\lim_{F \to F^*} \theta^{**} = \lambda$, which implies, using (20), that $\lim_{F \to F^*} w^{**} = q_H$ for F sufficiently close to F^*, so a candidate equilibrium with the desired properties exists. Furthermore, as $\theta^{**} \to \lambda$, condition (28) cannot hold so the candidate equilibrium is indeed an equilibrium.

Proof of Proposition 6

i. Using (29), $\lim_{\frac{c_H}{c_L} \to 1} w^* = q_L$ so using (20) we have $w^{**} > w^*$, so there is a candidate corner equilibrium. Furthermore, as $c_H/c_L \to 1$, condition (28) cannot hold so the candidate equilibrium is indeed an equilibrium.

ii. Using (29), $\lim_{\frac{c_H}{c_L} \to 0} w^* = q_H$. Using (18), this implies $\theta^* \to \lambda$, so condition (19) cannot hold for any $\pi^* > 0$.

Proof of Proposition 7

i. By equation (17), firm profits are increasing in π^*, and since w^* is the same across equilibria, firms are better off in the higher-π^* equilibrium. High type workers obtain a payoff of w^* in both equilibria, so they are indifferent. Using (23), workers with $i \leq \theta_2^*$ get a payoff of q_L in both equilibria, so they are also indifferent. Workers with $i \in (\theta_2^*, \theta_1^*]$ get q_L in the first equilibrium and more than q_L in the second, so they are better off in the second. For workers with $i \in (\theta_1, \lambda)$, their payoff is:

$$u(i) = q_L + \int_{\theta^*}^i \frac{1}{\pi^* (1 - \lambda) + \lambda - \theta} dF(\theta) (w^* - q_L)$$

$$= w^* - \int_i^\lambda \frac{1}{\pi^* (1 - \lambda) + \lambda - \theta} dF(\theta) (w^* - q_L)$$

where we used (19). This is increasing in π^*, so they are also better off in the second equilibrium.

ii. In the first equilibrium all firms make zero profits, so they are better off in the second equilibrium. Low type workers get a payoff of q_L in the first equilibrium, but those with $i > \theta^{**}$ get more in the second equilibrium. High type workers get a payoff of w^* in the first equilibrium but w^{**} in the second, so they are also better off.

iii. By equation (17), for θ sufficiently close to λ, firm θ's profits approach $q_H - w$, so $w^* < w^{**}$ implies they are higher in the first equilibrium. High type workers get a payoff of w^* in the first equilibrium but w^{**} in the second, so they are better off in the second.

Proof of Proposition 8

i. Using (17) yields

$$\Pi'(\theta) = \frac{\pi^*(1-\lambda)}{[\lambda - \theta + \pi^*(1-\lambda)]^2}(q_H - q_L) \tag{64}$$

and using (30)

$$W'(\theta) = (1-\lambda)\, c_H e^* \frac{\partial \pi^*}{\partial \theta}. \tag{65}$$

Around an interior equilibrium, equations (18) and (19) imply

$$\frac{\partial \pi^*}{\partial \theta} = \frac{\overbrace{\dfrac{(\lambda - \theta^*) + \pi^*(1-\lambda)}{[\lambda - \theta + \pi^*(1-\lambda)]^2 f(\theta^*)}}^{\text{Sensitivity of market clearing to expertise of one firm}}}{\underbrace{\dfrac{\int_{\theta^*}^{\lambda}(1-\lambda)[(\lambda-\theta)+\pi^*(1-\lambda)]^{-2}\,dF(\theta)\,[(\lambda-\theta^*)+\pi^*(1-\lambda)]}{f(\theta^*)}}_{\text{Slope of market clearing}} - \underbrace{(1-\lambda)\dfrac{1}{\frac{c_L}{c_H}-1}}_{\text{Slope of indifference}}}. \tag{66}$$

Replacing (64), (65) and (66) into (31), we obtain

$$r(\theta) = \frac{\frac{c_H}{c_L}}{\pi^*(1-\lambda)\left[\int_{\theta^*}^{\lambda}[(\lambda-\theta)+\pi^*(1-\lambda)]^{-2}\,dF(\theta) - \frac{f(\theta^*)}{\pi^*(1-\lambda)+\lambda-\theta^*}\frac{1}{\frac{c_L}{c_H}-1}\right]}. \tag{67}$$

Applying formula (61) and definition (32) yields

$$\eta = \frac{[\lambda - \theta^* + \pi^*(1-\lambda)]/[\pi^* f(\theta^*)]}{\int_{\theta^*}^{\lambda}(1-\lambda)[(\lambda-\theta)+\pi^*(1-\lambda)]^{-2}\,dF(\theta)\,[(\lambda-\theta^*)+\pi^*(1-\lambda)]/f(\theta^*) - (1-\lambda)/\left(\frac{c_L}{c_H}-1\right)}. \tag{68}$$

Replacing (68) into equation (67) and simplifying gives the result.

ii. Rearranging (68) and using that $f(\theta^*) \geq 0$:

$$\eta \geq \frac{1}{\pi^*(1-\lambda)\int_{\theta^*}^{\lambda}[(\lambda-\theta)+\pi^*(1-\lambda)]^{-2}\,dF(\theta)} \tag{69}$$

Now rearrange the market clearing condition (19) as

$$\int_{\theta^*}^{\lambda}[\pi^*(1-\lambda)]^{-1}[(\lambda-\theta)+\pi^*(1-\lambda)]^{-1}\,dF(\theta) = [\pi^*(1-\lambda)]^{-1}$$

Since $[(\lambda - \theta) + \pi^* (1 - \lambda)]^{-1}$ is increasing in θ, this implies

$$\pi^* (1 - \lambda) \int_{\theta^*}^{\lambda} [(\lambda - \theta) + \pi^* (1 - \lambda)]^{-2} dF(\theta) \leq 1$$

Replacing in equation (69) gives the result.

Proof of Proposition 9

Solving (40) for i_S and replacing in (39), a solution requires:

$$\Delta(i^*) \equiv f(i^*) \left[\frac{c_H}{c_L} - \lambda \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right] - \left(1 - \frac{c_H}{c_L} \right) \int_{i^*}^{i^H} \frac{\lambda}{i - i^* + \frac{F(i^*)}{f(i^*)} + \lambda} df(i) = 0 \qquad (70)$$

Taking the derivative and rearranging:

$$\frac{\partial \Delta}{\partial i^*} = f'(i^*) \frac{c_H}{c_L} \left(1 - \lambda \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right) + \lambda \left[f(i^*) \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-2} \left(1 - \frac{F(i^*) f'(i^*)}{f(i^*)^2} \right) \right.$$
$$\left. - \left(1 - \frac{c_H}{c_L} \right) \int_{i^*}^{i^H} \left(i - i^* + \frac{F(i^*)}{f(i^*)} + \lambda \right)^{-2} df(i) \left(\frac{F(i^*) f'(i^*)}{f(i^*)^2} \right) \right]$$
$$\geq f'(i^*) \frac{c_H}{c_L} \left(1 - \lambda \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right) + \lambda \left[f(i^*) \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-2} \left(1 - \frac{F(i^*) f'(i^*)}{f(i^*)^2} \right) \right.$$
$$\left. - \left(1 - \frac{c_H}{c_L} \right) \int_{i^*}^{i^H} \left(i - i^* + \frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} df(i) \left(\frac{F(i^*) f'(i^*)}{f(i^*)^2} \right) \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right]$$

where the inequality follows because $i > i^*$. If i^* satisfies (70), then:

$$\frac{\partial \Delta}{\partial i^*} \geq f'(i^*) \frac{c_H}{c_L} \left(1 - \lambda \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right) + \lambda \left[f(i^*) \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-2} \left(1 - \frac{F(i^*) f'(i^*)}{f(i^*)^2} \right) \right.$$
$$\left. - f(i^*) \left[\frac{c_H}{c_L} - \lambda \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right] \left(\frac{F(i^*) f'(i^*)}{f(i^*)^2} \right) \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-1} \right]$$
$$= \lambda f(i^*) \left(\frac{F(i^*)}{f(i^*)} + \lambda \right)^{-2} > 0$$

so the function $\Delta(i^*)$ is increasing at any i^* such that $\Delta(i^*) = 0$. Condition (41) is equivalent to $\Delta(i_H) < 0$. Furthermore,

$$\Delta(\lambda) = - \left(1 - \frac{c_H}{c_L} \right) \left[f(\lambda) + \int_{i^*}^{i^H} \frac{\lambda}{i} df(i) \right] < 0$$

Therefore, if (41) holds, there can be no $i^* \in [\lambda, i_H]$ that satisfies $\Delta(i^*) = 0$ because $\Delta(\lambda) < 0$ and $\Delta(i_H) < 0$ and Δ must be increasing at any solution. Instead, if condition (41) does not hold, $\Delta(i_H) \geq 0$, so by continuity and using the fact that Δ is increasing at any solution, there is exactly one $i^* \in [\lambda, i_H]$ that satisfies $\Delta(i^*) = 0$.

Proof of Proposition 10

For simplicity, we focus on the case without bunching (extending the proof to allow for bunching is notationally cumbersome but straightforward). Let $i^{*I}(i_S, p)$ and $i^{*M}(i_S, p)$ represent the solutions to (39) and (40) respectively, where p is a vector of parameters. The equilibrium value of i_S is given by a solution to the equation $i^{*I}(i_S, p) - i^{*M}(i_S, p) = 0$. Using the implicit function theorem, the derivatives of i^* and i_S with respect to parameter p are given by:

$$\frac{di_S}{dp} = \frac{\frac{\partial i^{*M}}{\partial p} - \frac{\partial i^{*I}}{\partial p}}{\frac{\partial i^{*I}}{\partial i_S} - \frac{\partial i^{*M}}{\partial i_S}} \tag{71}$$

$$\frac{di^*}{dp} = \frac{\partial i^{*I}}{\partial p} + \frac{\partial i^{*I}}{\partial i_S}\frac{\partial i_S}{\partial p} \tag{72}$$

The denominator of (71) is negative, and equation (39) implies that $\frac{\partial i^{*I}}{\partial i_S}$ is positive. Furthermore, the implicit function theorem implies that

$$\frac{\partial i^{*I}}{\partial p} = -\frac{\frac{\partial \Gamma(i^*, i_S; p)}{\partial p}}{\frac{\partial \Gamma(i^*, i_S; p)}{\partial i^*}} \tag{73}$$

and equation (39) implies that $\frac{\partial \Gamma(i^*, i_S; p)}{\partial i^*}$ is positive.

i. Using (39),

$$\frac{\partial \Gamma(i^*, i_S; \frac{c_H}{c_L})}{\partial \frac{c_H}{c_L}} = f(i^*) + \lambda \int_{i^*}^{i_H} \frac{1}{i - i_S + \lambda} d\bar{f}(i) \geq f(i^*) + \lambda \frac{[f(i_H) - f(i^*)]}{i_H - i_S + \lambda} > 0$$

so using (73) $\frac{\partial i^{*I}}{\partial \frac{c_H}{c_L}} < 0$. Since $\frac{\partial i^{*M}}{\partial \frac{c_H}{c_L}} = 0$, using this in (71) and (72) implies that $\frac{\partial i^*}{\partial \frac{c_H}{c_L}} < 0$ and $\frac{\partial i_S}{\partial \frac{c_H}{c_L}} < 0$. In turn, (38) implies that:

$$\frac{\partial u_L}{\partial \frac{c_H}{c_L}} = -f'(i^*)\frac{\partial i^*}{\partial \frac{c_H}{c_L}}(\underline{w}(0, i^*) - q_L) > 0$$

and (34) then implies that $\frac{\partial e_S}{\partial \frac{c_H}{c_L}} < 0$.

ii. Introducing variable demand Δ leaves equations (39) and (40) unchanged except that equation (33) generalizes to $\Delta\bar{f}(i_H) = 1$, so

$$\left.\frac{\partial i_H}{\partial \Delta}\right|_{\Delta=1} = -\frac{1}{\bar{f}'(i_H)} \tag{74}$$

Therefore

$$\left.\frac{\partial \Gamma(i^*, i_S; \Delta)}{\partial \Delta}\right|_{\Delta=1} = -\frac{\partial i_H}{\partial \Delta}\frac{\lambda}{i_H - i_S + \lambda}\bar{f}'(i_H)\left(1 - \frac{c_H}{c_L}\right) = \frac{\lambda}{i_H - i_S + \lambda}\left(1 - \frac{c_H}{c_L}\right) > 0$$

so using (73) $\frac{\partial i^{*I}}{\partial \Delta} < 0$. Since $\frac{\partial i^{*M}}{\partial \Delta} = 0$, using this in (71) and (72) implies $\frac{\partial i_S}{\partial \Delta} < 0$ and $\frac{\partial i^*}{\partial \Delta} < 0$.

Now assume towards a contradiction that e_S falls. This implies that the utility of the marginal

57

high type i^*, given by $u(i^*) = q_H - c_H e_S$ must rise. Equation (36) generalizes to:

$$u(i) = \Delta \bar{f}(i) \underline{w}(0, i) + \int_{i^*}^{i_H} \Delta \underline{w}(0, i')\, d\bar{f}(i')$$

so evaluating at i^* and taking the total derivative with respect to Δ:

$$\frac{du(i^*)}{d\Delta} = \frac{\partial u(i^*)}{\partial \Delta} + \frac{\partial u(i^*)}{\partial i^*}\frac{\partial i^*}{\partial \Delta} \tag{75}$$

with

$$\frac{\partial u(i^*)}{\partial \Delta} = \bar{f}(i^*)\underline{w}(0, i^*) + \int_{i^*}^{i_H} \underline{w}(0, i')\, d\bar{f}(i') - \frac{\partial i_H}{\partial \Delta}\underline{w}(0, i_H)\bar{f}'(i_H) = -\left[\underline{w}(0, i_H) - u(i^*)\right] < 0$$

(using (74) to replace $\frac{\partial i_H}{\partial \Delta}$), and $\partial u(i^*)/\partial i^* = \bar{f}(i^*)\partial \underline{w}(0, i^*)/\partial i^* > 0$. Replacing in (75) and using the assumption that $\frac{du(i^*)}{d\Delta} \geq 0$

$$-\left[\underline{w}(0, i_H) - u(i^*)\right] + \bar{f}(i^*)\frac{\partial \underline{w}(0, i^*)}{\partial i^*}\frac{\partial i^*}{\partial \Delta} \geq 0 \Rightarrow \frac{\partial i^*}{\partial \Delta} \geq \frac{\underline{w}(0, i_H) - u(i^*)}{\bar{f}(i^*)\frac{\partial \underline{w}(0, i^*)}{\partial i^*}} > 0,$$

which contradicts the first part of the result.

iii. The fact that i^* and i_S do not depend on q_H and q_L follows because neither q_H nor q_L appear in equations (39) and (40). Using (34), (37) and (38):

$$e_S = \frac{q_H - q_L}{c_L}f(i^*)\frac{i^* - i_S}{i^* - i_S + \lambda},$$

which is increasing in $q_H - q_L$.

iv. Letting the fraction of low types be $\hat{\lambda}$, equations (33), (35) and (39) generalize, respectively, to:

$$f(i^H) = \frac{1 - \hat{\lambda}}{1 - \lambda}, \qquad \underline{w}(0, i) = \frac{\frac{1-\hat{\lambda}}{1-\lambda}(i - i_S)q_H + \hat{\lambda}q_L}{\frac{1-\hat{\lambda}}{1-\lambda}(i - i_S) + \hat{\lambda}}, \text{ and} \tag{76}$$

$$\Gamma(i^*, i_S) = \frac{c_H}{c_L}f(i^*) - \hat{\lambda}\left(1 - \frac{c_H}{c_L}\right)\int_{i^*}^{i_H}\left[\frac{1}{\frac{1-\hat{\lambda}}{1-\lambda}(i - i_S) + \hat{\lambda}}\right]df(i') - f(i^*)\hat{\lambda}\frac{1}{\frac{1-\hat{\lambda}}{1-\lambda}(i^* - i_S) + \hat{\lambda}} = 0 \tag{77}$$

while the market clearing condition is unchanged. Therefore

$$\frac{\partial \Gamma(i^*, i_S; \hat{\lambda})}{\partial \hat{\lambda}} = -\left(1 - \frac{c_H}{c_L}\right)\int_{i^*}^{i_H}\frac{\frac{i-i_S}{1-\lambda}}{\left(\frac{1-\hat{\lambda}}{1-\lambda}(i - i_S) + \hat{\lambda}\right)^2}df(i') - f(i^*)\frac{\frac{i^*-i_S}{1-\lambda}}{\left(\frac{1-\hat{\lambda}}{1-\lambda}(i^* - i_S) + \hat{\lambda}\right)^2} < 0$$

so using (73), $\frac{\partial i^{*I}}{\partial \hat{\lambda}} > 0$. Since $\frac{\partial i^{*M}}{\partial \hat{\lambda}} = 0$, using this in (71) and (72) implies that $\frac{\partial i^*}{\partial \hat{\lambda}} > 0$ and

$\frac{\partial i_S}{\partial \hat{\lambda}} > 0$. In turn, (38) implies that:

$$\frac{\partial u_L}{\partial \hat{\lambda}} = -f'(i^*)\frac{\partial i^*}{\partial \hat{\lambda}}(\underline{w}(0, i^*) - q_L) + \frac{\partial i_H}{\partial \hat{\lambda}}\underline{w}(0, i_H)f'(i_H) + \int_{i^*}^{i_H}\frac{\partial \underline{w}(0, i')}{\partial \hat{\lambda}}df(i') < 0$$

and (34) then implies that $\frac{\partial e_S}{\partial \hat{\lambda}} > 0$.

Proof of Proposition 11

i. The first statement follows directly from equation (76), the second from equation (77) and the last from equation (81).

ii. Since the measure of firms is assumed to be greater than 1, for F sufficiently close to F^*, then $\bar{f}(i) > 1$ for all i, which implies $i_H = \lambda$.

iii. Equation 39 implies that $\lim_{\frac{c_H}{c_L} \to 1} i^* - i_S = 0$. Using this, equation 40 implies that $\lim_{\frac{c_H}{c_L} \to 1} F(i^*) = 0$, which implies $\lim_{\frac{c_H}{c_L} \to 1} i^* = \lambda$ and therefore $\lim_{\frac{c_H}{c_L} \to 1} i_S = \lambda$.

Equilibrium with False Negatives in the General Case

For the case where $f(i)$ is not monotone, the argument in Section 6.1 needs to be modified. Consider two workers, i and i' with $i^* < i < i' < i_H$ and assume $f(i) > f(i')$. The argument above, unmodified, implies that any worker will be able to sell a fraction $1 - f(i')$ of his labor to non-selective firms at wages above $\underline{w}(0, i')$. This means that only $f(i')$ of i-workers will be available for hire in market $(0, \underline{w}(0, i))$, which is less than the $f(i)$ workers that firms with $\theta = i$ want to hire. Realizing this, firms would bid up the wage, displacing non-selective firms. To characterize exactly what will happen, it is useful to define $\bar{F}(\theta)$ as the convex hull of $F(\theta)$, i.e. the highest convex function on $[\lambda, 1]$ such that $\bar{F}(\theta) \leq F(\theta)$:

$$\bar{F}(\theta) \equiv \min_{\omega, \theta_1, \theta_2} \{\omega F(\theta_1) + (1 - \omega)F(\theta_2)\}$$
$$\text{s.t. } \omega \in [0, 1], \ \theta_1, \theta_2 \in [\lambda, 1] \text{ and } \omega\theta_1 + (1 - \omega)\theta_2 = \theta.$$

The corresponding density $\bar{f}(\theta)$, which is weakly increasing, is the "ironed" version of the original density $f(\theta)$. We now show how the analysis in Section 6.1 extends to this general case, replacing F with $\bar{F}$. Let i_H be defined as
$$i_H \equiv \min_{i \in [\lambda, 1]} \{i : \bar{f}(i) \geq 1\}.$$

This generalizes the definition of i_H in (33), allowing both for the possibility of ironing and the case where $f(i) > 1$ for all i (in which case trivially $i_H = \lambda$). Let the reservation wage for type $i \in [i^*, i_H]$ be given by
$$\underline{w}(0, i) = \max_{i'} \frac{(i' - i_S)q_H + \lambda q_L}{i' - i_S + \lambda} \quad \text{s.t.} \quad \bar{f}(i') = \bar{f}(i). \tag{78}$$

Hence, when $\bar{f}$ is strictly increasing, this coincides with (35), but in a flat region (due to ironing), $\underline{w}(0, i)$ equals the reservation wage of the type at the top of the ironing range. In other words, in intervals $[i_0, i_1]$ where the ironed density $\bar{f}$ is constant, there will be "bunching:" all workers set the same reservation wage $\bar{w}(0, i_1)$ and firms $\theta \in [i_0, i_1]$ hire them in the same market, which is the cheapest market where they can identify high types. Based on the same steps as underlying (39) but using (78)

instead of (35), we obtain

$$\Gamma(i, i_S) = \bar{f}(i)\left(\frac{c_H}{c_L} - \frac{\lambda}{i^b(i) - i_S + \lambda}\right) - \lambda\left(1 - \frac{c_H}{c_L}\right)\int_i^{i_H}\frac{1}{i' - i_S + \lambda}d\bar{f}(i') \tag{79}$$

where $i^b(i) = \max\left\{i' : \bar{f}(i') = \bar{f}(i)\right\}$. Let i^* and i_S solve

$$i^* = \min_{i \in [\lambda, 1]}\left\{i : \Gamma(i, i_S) \geq 0\right\} \tag{80}$$

and

$$\bar{F}(i^*) = \bar{f}(i^*)(i^* - i_S). \tag{81}$$

Equation (80) generalizes the indifference condition (39) to account for the fact that, with bunching, the reservation wage function (78) and hence $\Gamma(i, i_S)$ can be discontinuous in i. Note that, by (80), whenever i^* falls into a bunching region, it corresponds to the lower end of it.

These definitions allow us to state the following general existence and uniqueness result:

Proposition 12. *There exists a generically unique equilibrium:*

    i. All low types $i \in [0, \lambda)$ choose $e = 0$ and reservation wage $\underline{w}(0, i) = q_L$.

    ii. All high types $i \in [i_H, 1]$ choose $e = 0$ and $\underline{w}(0, i) = q_H$.

    iii. For $i \in [\lambda, i_H)$, the equilibrium takes one of the following two possible forms:

        (a) An interior equilibrium where i_S and i^* solve (80) and (81) and:

            i. A measure $i_S - \lambda$ of high types with $i \in [\lambda, i^*)$ choose $e = e_S$ and $\underline{w}(e_S, i) = q_H$

            ii. All other high types with $i \in [\lambda, i^*)$ choose $e = 0$ and $\underline{w}(0, i) = \underline{w}_S$ with $\underline{w}_S$ such that

$$\bar{f}(i^*)\underline{w}_S + \int_{i^*}^{i_H}\underline{w}(0, i')d\bar{f}(i') = q_H - c_H e_S \tag{82}$$

            iii. All $i \in [i^*, i_H)$ choose $e = 0$ and $\underline{w}(0, i)$ as given by (78)

        (b) A corner equilibrium where $\Gamma(i_H, i_H - F(i_H)) < 0$ and:

            i. A measure $F(i_H)$ of high types with $i \in [\lambda, i_H)$ choose $e = 0$ and $\underline{w}(0, i) = w^*$

            ii. All other high types with $i \in [\lambda, i_H)$ choose $e = e^*$ and $\underline{w}(e^*, i) = q_H$.

The corner equilibrium is of the same form as described in Section 6.1. If the equilibrium is interior, the proposition encompasses two cases. Either there is no bunching at i^*, in which case our previous analysis goes through: the indifference condition $\Gamma(i^*, i_S) = 0$ implies that type i^* is just indifferent between signaling or not, and all high types below i^* who do not signal set reservation wage $\underline{w}_S = \underline{w}(0, i^*)$. The other case allows for i^* to be in a bunching region. Because there is a discontinuity in $u(i)$ at i^* in this case, i^* is given by the smallest i that still prefers choosing $e = 0$ over signaling (so $u(i^*) > q_H - c_H e_S$ and hence $\Gamma(i^*, i_S) > 0$). All high types $i < i^*$ are indifferent between signaling or not, and those who do not signal set a reservation wage $\underline{w}_S < \underline{w}(0, i^*)$ that is so high that there is no non-selective hiring. All firms' decisions are described in the proof below.

When there is bunching at the bottom (i.e. on the interval $[\lambda, i^*]$), the market clearing condition (81) implies $i_S = \lambda$, so there is no signaling whatsoever in equilibrium. This occurs when there is high density of precisely informed buyers relative to less informed ones.

We now provide a proof of Proposition 12, establishing first the uniqueness and then the existence of the stated equilibrium.

Uniqueness.   We prove uniqueness based on the following sequence of steps:

i. By the same arguments as in the proof of Proposition 2:

   (a) all firms make non-negative profits in each market

   (b) profits are decreasing in θ

   (c) all firms $\theta \geq \bar{\theta}$ make zero profits in equilibrium, with $\bar{\theta} \in [\lambda, 1)$ and $F(1) - F(\bar{\theta}) > 0$

   (d) low types obtain a payoff of at least q_L

   (e) high types obtain a payoff of at least $w^* = q_H - c_H e^*$.

ii. Because all low types are indistinguishable for all firms, all low types must obtain the same utility. Denote this by u_L.

iii. Utility for workers is weakly increasing in i. This follows immediately from the fact that $\mu(w; e, i)$ is weakly increasing in i.

iv. In any equilibrium, all low types choose $e = 0$. Suppose some low types choose $e' > 0$. By step (id), we have $\bar{w}(e', i) \geq q_L + c_L e'$ for all $i < \lambda$. Consider all markets m with $e(m) = e'$ and $w(m) > q_L$ where low types are hired with positive probability. For low types to be hired, in any such market m there must be firms that hire non-selectively, setting $\chi(i) = 1$ for all i. By step (ia), $S(i, m) > 0$ for some $i \geq \lambda$ in all these markets m.
Let w' be the highest wage where anyone choosing e' is hired with positive probability and let m' be such that $e(m') = e'$ and $w(m') = w'$. Suppose first that some high types $i' \geq \lambda$ are hired in market m' by selective firms $\theta \leq i'$ setting selection rule $\chi_\theta(i) = \mathbb{I}(i \geq \theta)$. The equilibrium payoff of these high types must be $u'_H \leq w' - c_H e'$. Consider a market $\tilde{m}$ with $e(\tilde{m}) = 0$ and $w(\tilde{m}) \in (w' - c_H e', w')$. Then $G(\cdot; \tilde{m}, \chi_\theta)$ must be well defined since otherwise $\mu(w(\tilde{m}), 0, i') = 0$, so type i' could obtain a payoff of at least $w(\tilde{m}) > u'_H$ by choosing $e = 0$. The support of $G(\cdot; \tilde{m}, \chi_\theta)$ can only include high types by construction of χ_θ. But this would imply that firm θ could increase its profits by hiring high types in market $\tilde{m}$ instead of market m' at wage $w(\tilde{m}) < w'$. Hence, everyone in market m' must be hired by non-selective firms. Since this is feasible for any firm and by step (ia), all firms must make zero profits in market m'. This implies $w' < q_H$.
Because all firms hire non-selectively in market m', $\mu(w', e', i) = \mu'$ is the same for all i. Suppose first that $\mu' < 1$. Consider a market m'' with $e(m'') = e'$ and $w(m') = w' - \epsilon$. Then for sufficiently small $\epsilon > 0$, $S(i, m'') = S(i, m')$ for all i. Since profits are zero in market m', all firms could make positive profits by hiring in market m'', contradicting (i). Hence we must have $\mu' = 0$. This implies that the equilibrium payoff of the low types is $u'_L = w' - c_L e'$ and the equilibrium payoff of those high types who choose $e = e'$ is $u'_H = w' - c_H e'$.
Consider a market m'' such that $e(m'') = e' + \epsilon$ and $w(m'') \in (w' + c_H \epsilon, w' + c_L \epsilon)$. Suppose $\chi(i) = 1$ for all i. Then $G(\cdot; m'', \chi)$ must be well defined since otherwise $\mu(w(m''), e(m''), i) = 0$ for all i, so all high types who choose e' could obtain payoff $w(m'') - c_H e(m'') > w' - c_H e' = u'_H$, a contradiction. The support of $G(\cdot; m'', \chi)$ cannot include low types since $w(m'') - c_L e(m'') < w' - c_L e' = u'_L$. The support of $G(\cdot; m'', \chi)$ cannot include only high types since then any firm $\theta > \bar{\theta}$ could make strictly positive profits in market m'' for $\epsilon \in (0, (q_H - w')/c_L)$. This delivers the final contradiction.

v. Any high type who chooses $e > 0$ must set reservation wage $\underline{w}(e, i) = q_H$. Suppose otherwise, then there exists a market m with $e(m) = e$ and $w(m) < q_H$ such that $S(i, m) > 0$ for some $i > \lambda$. Since there are no low types in market m by step (iv), any firm $\theta > \bar{\theta}$ could then make positive profits by hiring non-selectively in market m, contradicting step (ic).

61

vi. No firm hires high types selectively at any $e > 0$. Suppose there was a high type $i > \lambda$ who is hired in market m with $e(m) > 0$ and $w(m) = q_H$ by a firm $\theta < i$ that sets selection rule $\chi_\theta(i) = \mathbb{I}(i \geq \theta)$. Consider market m' with $e(m') = 0$ and $w(m') \in (q_H - c_H e(m), q_H)$. Then $G(\cdot; m, \chi_\theta) \neq \emptyset$ since otherwise $\mu(w(m), 0, i) = 0$, so type i could obtain a payoff $w(m') > q_H - c_H e(m)$ by choosing $e = 0$. Since by construction the support of $G(\cdot; m, \chi_\theta)$ only includes high types and since $w(m') < q_H$, firm θ can increase its profits by hiring high types in market m' rather than m, a contradiction. Hence, all high types selecting $e > 0$ are hired by firms using selection rule $\chi(i) = 1$ for all i.

vii. If any high types choose some education $e_S > 0$, it must satisfy $q_H - c_L e_S = u_L$. To see this, suppose first that some high types choose $e \in (0, e_S)$. By step (v), they set reservation wage q_H and by step (vi) they are hired by non-selective firms. However, this implies that the low types, by choosing e, could obtain $q_H - c_L e > u_L$, a contradiction. Suppose next that some high types choose $e > e_S$. Consider some market m with $e(m) = e_S$ and $w(m) = q_H - \epsilon$ and selection rule $\chi(i) = 1$ for all i, which is feasible for all firms. For sufficiently small ϵ, the support of $G(\cdot; m, \chi)$ must be well defined, since otherwise $\mu(w(m), e, i) = 0$ and those high types choosing e could do better by choosing education e_S and reservation wage $w(m)$. The support cannot include low types because $q_H - c_L e < u_L$. Since $w(m) < q_H$, firms $\theta \geq \bar{\theta}$ could therefore make strictly positive profits in market m, contradicting step (ic).

viii. Define
$$\bar{w}_S \equiv q_H - c_H e_S = \left(1 - \frac{c_H}{c_L}\right) q_H + \frac{c_H}{c_L} u_L. \tag{83}$$

There exists a cutoff i^* such that: for $i < i^*$, high types' utility is $u(i) = \bar{w}_S$ and for $i \geq i^*$, utility is $u(i) \geq \bar{w}_S$ and $e = 0$. Steps (v) and (vii) imply that high types who choose $e > 0$ must obtain utility equal to $\bar{w}_S$. Therefore the only possible way to obtain higher utility is to choose $e = 0$. The result then follows from step (iii).

ix. The reservation wage $\underline{w}(0, i)$ of all types $i \geq i^*$ (who choose $e = 0$) is weakly increasing in i. This follows from the fact that $\underline{w}(0, i)$ solves $\mu(w, 0, i) = 0$, and $\mu(w, 0, i)$ is weakly increasing in w and weakly decreasing in i.

x. If some type $i \geq 0$ who chooses $e = 0$ is hired by a selective firm, this can only occur at $i's$ reservation wage $\underline{w}(0, i)$. To see this, consider a market m with $e(m) = 0$ and $w(m) = w$ where a high type $i \geq \lambda$ is hired by a selective firm $\theta < i$ setting $\chi_\theta(i) = \mathbb{I}(i \geq \theta)$, and suppose $\mu(w, 0, i) > 0$. This implies $S(i, m') > 0$ for some m' with $e(m') = 0$ and $w(m') = w - \epsilon$. As a result, firm θ could increase its profits by shifting demand to market m' using the same selection rule.

xi. There does not exist a market m with $e(m) = 0$ and $w(m) > q_L$ where all firms hire non-selectively. Suppose there were such a market and let m be the highest-wage market where all firms hire non-selectively. All firms must make zero profits in m and $\mu(w(m), 0, i) = \bar{\mu}$ for all i. Suppose $\bar{\mu} > 0$. Consider a market m' with $e(m') = 0$ and $w(m') = w(m) - \epsilon$. For sufficiently small $\epsilon > 0$, $S(i, m') = S(i, m)$ for all i. Then all firms could make positive profits by hiring in market m', contradicting (i). Hence we must have $\bar{\mu} = 0$. This implies that there can only be a single such market m, and that all workers must obtain utility of at least $w(m)$ in equilibrium. Let $\bar{i}$ denote the highest $i \in [\lambda, 1]$ such that $S(i, m) > 0$ (by zero profits and $w(m) > q_L$, we must have $\bar{i} > \lambda$). To ensure that no firm wants to hire selectively in market m, all firms $\theta \leq \bar{i}$ must at least make profits $q_H - w(m)$ in equilibrium, i.e. they must hire high types in some market $m' \neq m$ with $w(m') \leq w(m)$. However, because all workers obtain utility of at least $w(m)$, there cannot be any supply of workers in market m'.

xii. All types $i > i_H$, who select $e = 0$ by step (viii), must set reservation wage q_H. Their reservation wage cannot be above q_H since no firm would hire at such a wage. Suppose some $\tilde{i} > i_H$ sets a reservation wage below q_H. This implies that all firms $\theta \in (i_H, \tilde{i})$ maximize profits by hiring selectively at the reservation wage of worker $i = \theta$, which is below q_H because reservation wages must be monotonic by step ix. The total number of workers these firms would hire is $F(\tilde{i}) - F(i_H) \geq \bar{F}(\tilde{i}) - \bar{F}(i_H) \geq \tilde{i} - i_H$. The first inequality follows from the fact that $F(\theta) \geq \bar{F}(\theta)$ for all θ by construction of $\bar{F}$, and $F(i_H) = \bar{F}(i_H)$ by definition of i_H. The second follows from the fact that $\bar{f}(\theta) \geq 1$ for all $\theta \geq i_H$. Moreover, generically the second inequality is strict. This implies that $\mu(w, 0, i) < 0$ for some worker i in this interval, which cannot be part of an equilibrium.

xiii. Consider first the case where the equilibrium is interior with $i^* < i_H$ and let $i_S - \lambda$ denote the measure of high type workers who choose $e = e_S$. For all other high type workers in $[\lambda, i_H]$, who choose $e = 0$, the reservation wage must be as described in Proposition 12, with the cutoff i^* defined in step (viii) given by (80).

Suppose first that there exist workers in $[\lambda, i_H]$ with higher reservation wages than those defined in Proposition 12, and let $\tilde{i}$ be the highest worker such that for some $\epsilon > 0$, reservation wages are higher than those defined in Proposition 12 for all $i \in (\tilde{i} - \epsilon, \tilde{i})$.

(a) If $\tilde{i} \in (i^*, i_H]$ is in a region where $\bar{f}(i)$ is strictly increasing, let $\underline{\tilde{w}}(0, i)$ be the reservation wages set by $i \in (\tilde{i} - \epsilon, \tilde{i})$. Define markets $M(\tilde{i}) = \{m : e(m) = 0, w(m) = \underline{\tilde{w}}(0, i), i \in (\tilde{i} - \epsilon, \tilde{i})\}$. Firms $\theta \leq \tilde{i} - \epsilon$ can find high types in markets with wages below $\underline{\tilde{w}}(0, i - \epsilon)$, so they don't want to hire selectively in any market $m \in M(\tilde{i})$. Therefore total selective hiring in markets $m \in M(\tilde{i})$ will be $\int_{\tilde{i}-\epsilon}^{\tilde{i}} d\bar{F}(i)$. By construction of $\tilde{i}$, all workers $\tilde{i} + \varepsilon$, $\varepsilon > 0$, set reservation wage $\underline{w}(0, \tilde{i} + \varepsilon)$ given by (78). By step (xi), a fraction $\bar{f}(\tilde{i} + \varepsilon)$ of them are hired by selective firms, and step (x) implies that the selective hiring occurs at their reservation wage $\underline{w}(0, \tilde{i} + \varepsilon)$. (Since $\tilde{i} \leq i_H$, we have $\bar{f}(\tilde{i} + \varepsilon) \leq 1$.) Taken together, this implies that a share $1 - \bar{f}(\tilde{i} + \varepsilon)$ of workers $\tilde{i} + \varepsilon$ must be hired by nonselective firms at or above $\underline{w}(0, \tilde{i} + \varepsilon)$. Continuity of $\bar{f}$ then implies that a fraction $1 - \bar{f}(\tilde{i})$ of workers of type $\tilde{i}$ will be hired by nonselective firms at wages at or above $\underline{w}(0, \tilde{i})$. Suppose first that $\underline{\tilde{w}}(0, i)$ is strictly increasing in $(\tilde{i} - \epsilon, \tilde{i})$. For each $i \in (\tilde{i} - \epsilon, \tilde{i})$, all workers $i' > i$ set reservation wages above $\underline{\tilde{w}}(0, i)$, so the supply of workers in market $\tilde{m}(i)$ includes $i - i_S$ high types and λ low types. Therefore (78) and the fact that $\underline{\tilde{w}}(0, i) > \underline{w}(0, i)$ imply that no firms want to hire non-selectively in any market $\tilde{m}(i)$. Alternatively, suppose $\underline{\tilde{w}}(0, i)$ is flat in $(\tilde{i} - \epsilon, \tilde{i})$ at level $\underline{w}(0, \tilde{i})$. Since a fraction $1 - \bar{f}(\tilde{i})$ of workers of type $\tilde{i}$ will be hired by nonselective firms at wages at or above $\underline{w}(0, \tilde{i})$, the same must then be true for all $i \in (\tilde{i} - \epsilon, \tilde{i}]$. In both cases, the total measure of workers in $(\tilde{i} - \epsilon, \tilde{i}]$ not hired at wages at or above $\underline{w}(0, \tilde{i})$ by nonselective firms is $\bar{f}(\tilde{i}) \epsilon$. Since $\bar{f}(i)$ is strictly increasing, $\bar{f}(\tilde{i}) \epsilon > \int_{\tilde{i}-\epsilon}^{\tilde{i}} d\bar{F}(i)$, which implies that $\mu(\underline{\tilde{w}}(0, i), 0, i) < 1$ for some workers $i \in (\tilde{i} - \epsilon, \tilde{i})$, and therefore setting reservation wage $\underline{\tilde{w}}(0, i)$ cannot be optimal.

(b) If $\tilde{i} \in (i^*, i_H]$ is in a region where $\bar{f}(i)$ is constant or if $\tilde{i} < i^*$ then this implies that the reservation wage of worker $\tilde{i}$ is higher than that of some worker $i' > \tilde{i}$, which would violate step (ix).

(c) If $\tilde{i} = i^*$, this would imply that some workers $i \in [\lambda, i^*]$ set a reservation wage $\underline{\tilde{w}}(0, i) > \underline{w}_S$. This can only occur without violating step (ix) when i^* corresponds to the lower end of a bunching region and $\underline{w}_S < \underline{w}(0, i^*)$. Let i' be the lowest $i \in [\lambda, i^*]$ such that $\underline{\tilde{w}}(0, i) > \underline{w}_S$ for all $i > i'$. We must have $i' > \lambda$ since otherwise no-one signals by (82). No firm $\theta < i'$ wants

to hire any type $i > i'$ since they maximize profits by hiring in market $(0, \underline{w}_S)$. Hence, total selective hires in (i', i^*) are given by $F(i^*) - F(i') \leq \bar{F}(i^*) - \bar{F}(i') < i^* - i'$. The first inequality follows from the fact that, since i^* is the lower end of a bunching region, we have $F(i^*) = \bar{F}(i^*)$. The second follows from the definition of i_H and the fact that $i^* < i_H$. Since there is no non-selective hiring (if there was, step (a) would apply), this implies that $\mu(\underline{\tilde{w}}(0, i), 0, i) > 0$ for some workers $i \in (i', i^*)$, and therefore setting reservation wage $\underline{\tilde{w}}(0, i)$ cannot be optimal.

Suppose next that there exist workers with lower reservation wages than those defined in Proposition 12, and let $\tilde{i}$ be the highest worker such that for some $\epsilon > 0$, reservation wages are lower than those defined in Proposition 12 for all $i \in (\tilde{i} - \epsilon, \tilde{i})$.

(a) If $\tilde{i} \in [i^*, i_H]$ is in a region where $\bar{f}(i)$ is strictly increasing, take some $i' \in (\tilde{i} - \epsilon, \tilde{i})$ with a reservation wage $w' < [(i' - i_S)q_H + \lambda q_L] / [i' - i_S + \lambda]$ and consider the market m' with $e(m') = 0$ and $w(m') = w'$. The supply of workers in this market includes all low types (a measure λ) and at least the high types $i \in [\lambda, i']$ who do not signal (a measure at least $i' - i_S$). Therefore, a firm that hired non-selectively in market m' would make profits of at least

$$\frac{(i' - i_S)q_H + \lambda q_L}{i' - i_S + \lambda} - w' > 0$$

Since this is feasible for all firms, it contradicts (ic).

(b) If $\tilde{i} \in [i^*, i_H]$ is in a region $[i_0, i_1]$ where $\bar{f}(i)$ is constant, then a fraction $1 - \bar{f}(i_1)$ of all workers $i < i_1$ are hired by nonselective firms at wages at least $\underline{w}(0, i_1)$. The measure of firms in $(i_0, \tilde{i})$ is $F(\tilde{i}) - F(i_0) \geq \bar{F}(\tilde{i}) - \bar{F}(i_0) = \bar{f}(i_1)(\tilde{i} - i_0)$, with strict inequality in the generic case where the original density f is not exactly constant and equal to $\bar{f}(i_1)$. For all these firms, it is profit maximizing to hire selectively at the reservation wage of worker $i = \theta$, which implies that $\mu(w, 0, i) < 0$ for some i in this interval, and which therefore cannot be part of an equilibrium.

(c) If $\tilde{i} < i^*$ then equation (80) implies that the utility of all workers $i \leq \tilde{i}$ is below $\bar{w}_S$. By (82) and (83), they would be better off choosing $e = e_S$ and setting reservation wage $\underline{w}(e_S, i) = q_H$.

xiv. In any interior equilibrium, the cutoff i^* defined in step (viii) must also satisfy (81). To see this, observe first that it cannot be that all workers $i \in [\lambda, i^*]$ signal. If this were the case, consider the beliefs of a firm $\theta \in [\lambda, i^*)$ in a market m with $e(m) = 0$ and $w(m) = \underline{w}_S + \varepsilon$. Then $G(\cdot; m, \chi_\theta) \neq \emptyset$ since otherwise $\mu(w(m), 0, i) = 0$ for $i \in [\lambda, i^*)$, so these workers could get payoff in excess of $\bar{w}_S$ by choosing $e = 0$ instead of signaling. Moreover, the support of $G(\cdot; m, \chi_\theta)$ cannot include high types since otherwise some firms could increase their profits by shifting demand to market m for ε sufficiently small. Hence, for all firms $\theta \in [\lambda, i^*)$, it is profit maximizing to hire selectively in market m with $e(m) = 0$ and $w(m) = \underline{w}_S$, so total selective hires will be $F(i^*) = \bar{F}(i^*)$, where the equality follows from the fact that, by (80), if i^* falls in a bunching region, it corresponds to the lower end of it. On the other hand, the measure of workers who are not hired by nonselective firms at higher wages is $\bar{f}(i^*)(i^* - i_S)$. Hence, if $\bar{F}(i^*) > \bar{f}(i^*)(i^* - i_S)$, then $\mu(w, 0, i) < 0$ for some i in this interval, which cannot be part of an equilibrium. If $\bar{F}(i^*) < \bar{f}(i^*)(i^* - i_S)$, then $\mu(w, 0, i) > 0$ for some i in this interval, so $\underline{w}_S$ cannot be the optimal reservation wage.

xv. For the case of a corner equilibrium, note first that $\Gamma(i_H, i_H - F(i_H)) < 0$ implies $\underline{w}(0, i_H) < w^*$, where we used $\bar{f}(i_H) = 1$ (abstracting from the trivial case $i_H = \lambda$, which is fully characterized

by step (xii)). Together with steps (ie) and (ix), this means that there cannot be any nonselective hiring. Hence, $u_L = q_L$ and $e_S = e^*$. Moreover, since there are only $F(i_H)$ firms with $\theta < i_H$ that can hire selectively at $e = 0$, this immediately implies that at least a measure $i_S - \lambda = i_H - \lambda - F(i_H)$ of workers must signal. All other workers in $[\lambda, i_H]$ must set reservation wage w^*. They cannot set a reservation lower than w^* by step (ie). Suppose some workers set a higher reservation wage and let i' be the lowest $i \in [\lambda, i_H)$ such that $\underline{w}(0, i) > w^*$ for all $i > i'$. We must have $i' > \lambda$ since otherwise no-one would signal. No firm $\theta < i'$ wants to hire any type $i > i'$ since they maximize profits by hiring in market $(0, w^*)$. Hence, total selective hires in (i', i_H) are given by $F(i_H) - F(i') \leq \bar{F}(i_H) - \bar{F}(i') < i_H - i'$. The first inequality follows from $F(i_H) = \bar{F}(i_H)$ and the second from the definition of i_H. Since there is no non-selective hiring, this implies that $\mu(\underline{w}(0, i), 0, i) > 0$ for some workers $i \in (i', i_H)$, and therefore setting reservation wage $\underline{w}(0, i)$ cannot be optimal.

xvi. By the same argument as in step (xiv), in the corner equilibrium not every worker with $i < i_H$ can signal. Moreover, again by the same argument as in step (xiv), it is not possible that the measure $i_S - \lambda$ of workers who signal exceeds $i_H - F(i_H) - \lambda$. Hence, together with the previous step, we must have $i_S = i_H - F(i_H)$.

xvii. Finally, we show that there is a unique solution to equations (80) and (81). The argument in the proof of Proposition (9) applies, except that, with bunching, the function $\Delta(i^*)$ is no longer continuous. From (79) we see that Γ is still continuous in i_S but, as i increases, jumps up at the lower end of each bunching interval. This is because when i enters a bunching region, $i^b(i)$ jumps to the upper end of that region. As a result, $\Delta(i) \equiv \Gamma(i, i - \bar{F}(i)/\bar{f}(i))$ is continuous in i except when i is the lower end of a bunching interval, in which case $\Delta(i)$ discontinuously jumps up at that point as i increases. Recall that the solution to (80) and (81) is $i^* = \min_{i \in [\lambda, i_H]} \{i | \Delta(i) \geq 0\}$. Together with the result from Proposition 9 that $\Delta'(i) > 0$ when i is not in a bunching region and $\Delta = 0$, this implies the following:

   (a) If $\Delta(i_H) < 0$, then $\Delta(i) < 0$ for all $i \in [\lambda, i_H]$, so there cannot be any solution to (80) and (81) and the corner equilibrium is the unique equilibrium.

   (b) If $\Delta(i_H) \geq 0$, then either $\Delta(i) > 0$ for all $i \in [\lambda, i_H]$, in which case $i^* = \lambda$, or there exists a unique solution $i^* \in (\lambda, i_H]$. Hence, if there is an interior equilibrium, it is also unique.

Existence. We have already established the existence of a solution to equations (80) and (81). We now provide the equilibrium beliefs and rationing functions and verify that all demand decisions maximize firms' profits and all workers maximize utility.

   Interior Equilibrium.

   (a) Education decisions:
$$\pi_i(e) = \begin{cases} 1 & \text{if} \begin{cases} i \geq i_S \text{ and } e = 0 \\ i \in [\lambda, i_S) \text{ and } e = e_S \\ i < \lambda \text{ and } e = 0 \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

(b) Rationing function:

$$\mu\left(w,e,i\right)=\begin{cases}\mathbb{I}\left(w\geq q_{H}\right) & \text{if} & e=0,\ i\geq i_{H}\\[4pt]\bar{f}(i^{r}(w)) & \text{if} & \begin{cases}e=0,\ i\in[i^{*},i_{H}),\ w\in(\underline{w}(0,i),\underline{w}(0,i_{H})]\\ e=0,\ i\in[\lambda,i^{*}),\ w\in(\underline{w}_{S},\underline{w}(0,i_{H})]\\ e=0,\ i\in[0,\lambda),\ w\in(q_{L},\underline{w}(0,i_{H})]\end{cases}\\[12pt]0 & \text{if} & \begin{cases}e=0,\ i\in[i^{*},i_{H}],\ w\leq\underline{w}(0,i)\\ e=0,\ i\in[\lambda,i^{*}),\ w\leq\underline{w}_{S}\\ e=0,\ i\in[0,\lambda),\ w\leq q_{L}\end{cases}\\[12pt]1 & \text{if} & e=0,\ i\leq i_{H},\ w>\underline{w}(0,i_{H})\\[4pt]\mathbb{I}(w\geq\tilde{w}(e,i)) & \text{if} & e>0,\ i\geq\lambda\\[4pt]\mathbb{I}\left(w\geq\min\{\tilde{w}(e,i),\tilde{w}(e,\lambda)\}\right) & \text{if} & e>0,\ i<\lambda\end{cases}$$

where

$$i^{r}(w)=\min_{i\in[i^{*},i_{H}]}\left\{i:\underline{w}(0,i)\geq w\right\},$$

$\tilde{w}(e,i)$ is given by (24) and

$$u\left(i\right)=\begin{cases}q_{H} & \text{if} & i>i_{H}\\ \bar{f}(i)\underline{w}(0,i)+\int_{i}^{i_{H}}\underline{w}(0,i')d\bar{f}(i') & \text{if} & i\in[i^{*},i_{H}]\\ \bar{f}(i^{*})\underline{w}_{S}+\int_{i^{*}}^{i_{H}}\underline{w}(0,i)d\bar{f}(i) & \text{if} & i\in[\lambda,i^{*})\\ \bar{f}(i^{*})q_{L}+\int_{i^{*}}^{i_{H}}\underline{w}(0,i)d\bar{f}(i) & \text{if} & i<\lambda\end{cases}$$

(c) Demand decisions:

$$\delta_{\theta}\left(m,\chi\right)=\begin{cases}1 & \text{if} & \begin{cases}\theta\in[\lambda,i^{*}),\ w(m)=\underline{w}_{S},\ e(m)=0\ \text{and}\ \chi(i)=\mathbb{I}(i\geq\theta)\\ \theta\in[i^{*},i_{H}),\ w(m)=\underline{w}(0,\theta),\ e(m)=0\ \text{and}\ \chi(i)=\mathbb{I}(i\geq\theta)\\ \theta\in[i_{H},\theta^{*}),\ w(m)=q_{H},\ e(m)=0\ \text{and}\ \chi(i)=\mathbb{I}(i\geq\theta)\end{cases}\\[12pt]\frac{\lambda(1-\bar{f}(i^{*}))}{F(1)-F(\theta^{*})} & \text{if }\theta\geq\theta^{*},\ w(m)=q_{L},\ e(m)=0\ \text{and}\ \chi(i)=1\forall i\\[6pt]\frac{i_{S}-\lambda}{F(1)-F(\theta^{*})} & \text{if }\theta\geq\theta^{*},\ w(m)=q_{H},\ e(m)=e_{S}\ \text{and}\ \chi(i)=1\forall i\end{cases}$$

where θ^{*} is such that $F(\theta^{*})-F(i_{H})=1-i_{H}$.

The non-selective demand in markets m with $e(m)=0$ and $w(m)=\underline{w}(0,i)$ with $i\in[i^{*},i_{H})$ and $\bar{f}'(i)>0$ remains to be specified. For an small interval of types $[i_{0},i_{0}+\Delta]$ the change in the probability of being hired non-selectively is:

$$\bar{f}\left(i_{0}+\Delta\right)-\bar{f}\left(i_{0}\right)\approx\bar{f}'\left(i_{0}\right)\Delta$$

Using that in a non-bunching region, $i^{r}\left(w\right)=\frac{i_{S}(q_{H}-w)+\lambda(w-q_{L})}{q_{H}-w}$, this implies that total non-selective hires over an interval of wages $[w,w+\epsilon]$ are

$$\epsilon\frac{\lambda\left(q_{H}-q_{L}\right)}{\left(q_{H}-w\right)^{2}}\left[i^{r}\left(w\right)-\left(i_{S}-\lambda\right)\right]\bar{f}'\left(i^{r}\left(w\right)\right).$$

Hence, the total measure of demand of each firm $\theta\geq\theta^{*}$ using the non-selective hiring rule $\chi(i)=1\forall i$ placed on any set of markets M where $e(m)=0$ for all $m\in M$ and $w(M)=$

66

$(w_0, w_1) \subset [\underline{w}(0, i^*), \underline{w}(0, i_H)]$ must be

$$\delta_\theta(M, \chi) = \frac{\int_{w_0}^{w_1} \frac{\lambda(q_H - q_L)}{(q_H - w)^2} [i^r(w) - (i_S - \lambda)] \bar{f}'(i^r(w)) \, dw}{F(1) - F(\theta^*)}.$$

Finally, $\delta_\theta(m, \chi) = 0$ otherwise.

(d) Beliefs: for selection rule $\chi(i) = \mathbb{I}(i \geq \theta)$,

$$g(i; m, \chi) = \begin{cases} \frac{\mathbb{I}(i \geq \max\{\theta, i_S\})}{1 - \max\{\theta, i_S\}} & \text{if } e(m) = 0, \ w(m) \geq q_H \\ \frac{\mathbb{I}(i \in [\max\{\theta, i_S\}, i_H))}{i_H - \max\{\theta, i_S\}} & \text{if } e(m) = 0, \ w(m) \in [\underline{w}(0, i_H), q_H), \theta < i_H \\ \frac{\mathbb{I}(i \in [\max\{i_S, \theta\}, i^r(w(m))])}{i^r(w(m)) - \max\{i_S, \theta\}} & \text{if } e(m) = 0, \ w(m) \in [\underline{w}_S, \underline{w}(0, i_H)), \ \theta < i_H \\ \frac{\mathbb{I}(i \in [\theta, i^r(w(m) - c_H e(m))))}{i^r(w(m) - c_H e(m)) - \theta} & \text{if } \begin{cases} e(m) > 0, \\ w(m) - c_H e(m) \in [\tilde{w}(e(m), \lambda), \tilde{w}(e(m), i_H)), \theta < i_H \end{cases} \\ \frac{\mathbb{I}(i \geq \theta)}{1 - \theta} & \text{if } e(m) > 0, \ w(m) - c_H e(m) \geq \tilde{w}(e(m), i_H) \\ \emptyset & \text{otherwise} \end{cases}$$

and for selection rule $\chi(i) = 1 \forall i$,

$$g(i; m, \chi) = \begin{cases} \frac{\mathbb{I}(i < \lambda) + \mathbb{I}(i \geq i_S)}{1 - i_S + \lambda} & \text{if } e(m) = 0, \ w(m) \geq q_H \\ \frac{\mathbb{I}(i < \lambda) + \mathbb{I}(i \in [i_S, i_H))}{\lambda + i_H - i_S} & \text{if } e(m) = 0, \ w(m) \in [\underline{w}(0, i_H), q_H) \\ \frac{\mathbb{I}(i < \lambda) + \mathbb{I}(i \in [i_S, i^r(w(m))])}{\lambda + i^r(w(m)) - i_S} & \text{if } e(m) = 0, \ w(m) \in [\underline{w}_S, \underline{w}(0, i_H)) \\ \frac{1}{\lambda} \mathbb{I}(i < \lambda) & \text{if } e(m) = 0, \ w(m) \in [q_L, \underline{w}_S) \\ \frac{1}{\lambda} \mathbb{I}(i < \lambda) & \text{if } e(m) \in (0, e_S), \ w(m) \geq \tilde{w}(e(m), 0) \\ \frac{\mathbb{I}(i \in [\lambda, i_S])}{i_S - \lambda} & \text{if } e(m) \geq e_S, \ w(m) \geq \tilde{w}(e(m), \lambda) \\ \emptyset & \text{otherwise} \end{cases}$$

To see that the proposed $\{\pi, \delta, \mu, G\}$ is an equilibrium, note first that the rationing function defined in (b) implies that low types are indifferent between any $e \in [0, e_S]$ and high types are indifferent between any e, so the education decisions defined in (a) solve the workers' problem. The beliefs defined in (d) imply that it is profit maximizing for firms $\theta \leq i^*$ to hire selectively in market $\{e = 0, w = \underline{w}_S\}$ and for firms $\theta \in (i^*, i_H)$ to hire in market $\{e = 0, w = \underline{w}(0, \theta)\}$. Firms $\theta \geq i_H$ make zero profits by hiring selectively in market $\{e = 0, w = q_H\}$. Moreover, firms $\theta \geq i_H$ make zero profits by hiring non-selectively in markets $\{e = 0, w = q_L\}$, $\{e = e_S, w = q_H\}$ or $\{e = 0, w = \underline{w}(0, i), i \in [i^*, i_H)\}$. Any other market has either $G(\cdot; m, \chi_\theta) = \emptyset$ or results in losses. Therefore the demands defined in (c) are an optimal choice. They also satisfy $\delta_\theta(M, X) \leq 1$ for all θ (with equality for $\theta \leq i_H$), so no firm hires more than one worker. Finally, supply is:

$$S(i, m) = \begin{cases} \mathbb{I}(i < \lambda) & \text{if } e(m) = 0, \ w(m) \in [q_L, \underline{w}_S) \\ \mathbb{I}(i < \lambda) + \mathbb{I}(i \in (i_S, i^r(w(m))]) & \text{if } e(m) = 0, \ w(m) \in [\underline{w}_S, \underline{w}(0, i_H)) \\ \mathbb{I}(i < \lambda) + \mathbb{I}(i \in (i_S, i_H)) & \text{if } e(m) = 0, \ w(m) \in [\underline{w}(0, i_H), q_H) \\ \mathbb{I}(i < \lambda) + \mathbb{I}(i \geq i_S) & \text{if } e(m) = 0, \ w(m) \geq q_H \\ \mathbb{I}(i \in [\lambda, i_S)) & \text{if } e(m) = e_S, w(m) \geq q_H \\ 0 & \text{otherwise} \end{cases}$$

and aggregate demand $D(m, \chi)$ follows directly from the individual firms' demands defined in (c).

Using (8), hires are:

$$H\left(w;e,i\right)=\begin{cases}\mathbb{I}\left(i\geq i_H\right) & \text{if }e=0,\ w\in(\underline{w}(0,i_H),q_H]\\ \mathbb{I}\left(i\geq i^r\left(w\right)\right)+(1-\bar{f}(i^r\left(w\right)))\mathbb{I}\left(i\in[0,\lambda)\cup[i_S,i^r\left(w\right))\right) & \text{if }e=0,\ w\in(\underline{w}_S,\underline{w}(0,i_H)]\\ \mathbb{I}\left(i\geq i_S\right)+(1-\bar{f}(i^*))\mathbb{I}(i<\lambda) & \text{if }e=0,\ w\in(q_L,\underline{w}_S]\\ \mathbb{I}\left(i<\lambda\right)+\mathbb{I}\left(i\geq i_S\right) & \text{if }e=0,\ w\leq q_L\\ \mathbb{I}(i\in[\lambda,i_S)) & \text{if }e=e_S,\ w\leq q_H\\ 0 & \text{otherwise}\end{cases}$$

Hence, (9) holds. Finally, $S(i,m)$ implies that the beliefs defined in (d) satisfy the consistency condition (ivb) from Definition 1 in nonempty markets. In zero supply markets, beliefs are also constructed to satisfy condition (ivb) when they are well defined, and they are not well defined only at wages where $\mu\left(w;e,i\right)=0$ for all i such that $\chi(i)=1$, so condition (ivc) is satisfied as well.

Corner equilibrium.

(a) Education decisions:

$$\pi_i\left(e\right)=\begin{cases}1 & \text{if }\begin{cases}i\geq i_H-F(i_H)\text{ and }e=0\\ i\in[\lambda,i_H-F\left(i_H\right))\text{ and }e=e^*\\ i<\lambda\text{ and }e=0\end{cases}\\ 0 & \text{otherwise}\end{cases}$$

(b) Demand decisions:

$$\delta_\theta\left(m,\chi\right)=\begin{cases}1 & \text{if }\begin{cases}\theta<i_H,\ w(m)=w^*,\ e(m)=0\text{ and }\chi(i)=\mathbb{I}(i\geq\theta)\\ \theta\in[i_H,\theta^*),\ w(m)=q_H,\ e(m)=0\text{ and }\chi(i)=\mathbb{I}(i\geq\theta)\end{cases}\\ \frac{\lambda}{F(1)-F(\theta^*)} & \text{if }\theta\geq\theta^*,\ w(m)=q_L,\ e(m)=0\text{ and }\chi(i)=1\forall i\\ \frac{i_H-F(i_H)-\lambda}{F(1)-F(\theta^*)} & \text{if }\theta\geq\theta^*,\ w(m)=q_H,\ e(m)=e^*\text{ and }\chi(i)=1\forall i\\ 0 & \text{otherwise}\end{cases}$$

where θ^* is such that $F(\theta^*)-F(i_H)=1-i_H$

(c) Rationing function:

$$\mu\left(w,e,i\right)=\begin{cases}\mathbb{I}\left(w\geq q_H\right) & \text{if} & e=0,\ i>i_H\\ \mathbb{I}\left(w\geq w^*\right) & \text{if} & e=0,\ i\in[\lambda,i_H]\\ \mathbb{I}\left(w\geq q_L\right) & \text{if} & e=0,\ i<\lambda\\ \mathbb{I}(w\geq\tilde{w}(e,i)) & \text{if} & e>0,\ i\geq\lambda\\ \mathbb{I}\left(w\geq\min\{\tilde{w}(e,i),\tilde{w}(e,\lambda)\}\right) & \text{if} & e>0,\ i<\lambda\end{cases}$$

where $\tilde{w}(e,i)$ is given by (24) and

$$u\left(i\right)=\begin{cases}q_H & \text{if} & i>i_H\\ w^* & \text{if} & i\in[\lambda,i_H]\\ q_L & \text{if} & i<\lambda\end{cases}$$

(d) Beliefs: for selection rule $\chi(i) = \mathbb{I}(i \geq \theta)$,

$$
g(i; m, \chi) = \begin{cases}
\frac{\mathbb{I}(i \geq \max\{\theta, i_H - F(i_H)\})}{1 - \max\{\theta, i_H - F(i_H)\}} & \text{if } e(m) = 0, \ w(m) \geq q_H \\
\frac{\mathbb{I}(i \in [\max\{i_H - F(i_H), \theta\}, i_H])}{\min\{F(i_H), i_H - \theta\}} & \text{if } e(m) = 0, \ w(m) \in [w^*, q_H), \ \theta < i_H \\
\frac{\mathbb{I}(i \in [\theta, i_H))}{i_H - \theta} & \text{if } e(m) > 0, w(m) \in [\tilde{w}(e(m), \lambda), \tilde{w}(e(m), i_H)), \ \theta < i_H \\
\frac{\mathbb{I}(i \geq \theta)}{1 - \theta} & \text{if } e(m) > 0, \ w(m) \geq \tilde{w}(e(m), i_H) \\
\emptyset & \text{otherwise}
\end{cases}
$$

and for selection rule $\chi(i) = 1 \forall i$,

$$
g(i; m, \chi) = \begin{cases}
\frac{\mathbb{I}(i < \lambda) + \mathbb{I}(i \geq i_H - F(i_H))}{\lambda + 1 - i_H + F(i_H)} & \text{if } e(m) = 0, \ w(m) \geq q_H \\
\frac{\mathbb{I}(i < \lambda) + \mathbb{I}(i \in [i_H - F(i_H), i_H))}{\lambda + F(i_H)} & \text{if } e(m) = 0, \ w(m) \in [w^*, q_H) \\
\frac{1}{\lambda} \mathbb{I}(i < \lambda) & \text{if } e(m) = 0, w(m) \in [q_L, w^*) \\
\frac{1}{\lambda} \mathbb{I}(i < \lambda) & \text{if } e(m) \in (0, e^*), \ w(m) \geq \tilde{w}(e(m), 0) \\
\frac{\mathbb{I}(i \in [\lambda, i_H - F(i_H)])}{i_H - F(i_H) - \lambda} & \text{if } e(m) \geq e^*, \ w(m) \geq \tilde{w}(e(m), \lambda) \\
\emptyset & \text{otherwise}
\end{cases}
$$

Verifying that the proposed $\{\pi, \delta, \mu, G\}$ is an equilibrium is analogous to the interior equilibrium case.