Signaling to Experts*

Pablo Kurlat
University of Southern California

Florian Scheuer
University of Zurich

August 2020

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 that can be Pareto ranked. We characterize wage dispersion and the extent of signaling as a function of the distribution of expertise among firms. 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.

*Email: kurlat@usc.edu, florian.scheuer@uzh.ch. We thank Adrien Auclert, Alex Bloedel, Gabriel Carroll, Veronica Guerrieri, Patrick Kehoe, Guido Menzio, Nick Netzer, Venky Venkateswaran as well as numerous seminar and conference participants for helpful comments and suggestions.

[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 present 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 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 asset managers' heterogeneous pricing techniques 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. 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. Markets do not necessarily clear: in any given market, workers can apply for jobs and not get them and firms may not find acceptable workers. 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 "partial signaling" 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^P 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^P, and a market clearing condition requiring high-type workers who forgo education to indeed find jobs at wage w^P. This tractable structure allows us, for instance, to study comparative statics. We find that signaling decreases if the cost is higher, if the demand for workers increases, or if firms' expertise improves, intuitive properties that, somewhat unappealingly, cannot be obtained in the standard signaling model with uninformed firms.

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 improves. 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—an extreme form of coordination failure. More generally, 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.

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. It also predicts that higher demand for workers leads to polarization in signaling: fewer high types signal, but those who signal do so more intensely.

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 (2016), all of which are based on the idea that different prices define different markets and the probability of trade 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 (2016). However, Kurlat (2016) 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 (2016). 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. This 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. Fishman and Parker (2015), Bolton et al. (2016) and Kurlat (2019) study environments where buyers can differ in the quality of their information but where sellers do not have a way to signal. Their focus is on the efficiency of buyers' information acquisition decision.

Board et al. (2017) 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. (2017) 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 and in Section 6 the case of false negatives. Finally, Section 7 concludes. Various extensions and 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 $q_L < q_H$. Workers with $i < \lambda$ and $i \geq \lambda$ are 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 Spence (1973) signaling model. Our innovation is to introduce firms' 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. 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 $[\lambda,1]$. For instance, firms can be interpreted as being "bold" in the first and "cautious" in the second case.[2] Clearly, (2) is a restrictive model of how well informed firms are: in general, firms could make both kinds of mistakes in arbitrarily correlated ways. This formulation has the advantage of providing a natural measure of a firm's expertise since 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

[2]See Farboodi and Kondor (2018) for a model that links these two cases to the business cycle.

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.[3] For example, some managers have better judgement in assessing the talent of job applicants, as in Board et al. (2017), 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 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 $u(e,w) = w - c(i)\,e$ and computing the marginal rate of substitution: $-\frac{\partial u(e,w)/\partial e}{\partial u(e,w)/\partial w} = c(i)$, 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.

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.

[3] 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.

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 $u(e, w) = (Y - e)\, w + \alpha \int \max\{y + e - Y, 0\}\, dG_k(y)$, with $k = L$ if $i < \lambda$ and $k = H$ if $i \geq \lambda$. The marginal rate of substitution is $-\frac{\partial u(e,w)/\partial e}{\partial u(e,w)/\partial w} = \frac{w - \alpha[1 - G_k(Y-e)]}{Y-e}$. 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 $q_k(e) \equiv \int \min\left\{\frac{y}{Y-e}, 1\right\} dG_k(y)$, 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 in their data were sophisticated investors or intermediaries. They conclude that this variability is due to differences in pricing technology (see also Eisfeldt et al., 2019). 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 $\alpha q(i)$, 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 $w(1-e) + \alpha q(i)\, e$. The marginal rate of substitution is $-\frac{\partial u(e,w)/\partial e}{\partial u(e,w)/\partial w} = \frac{w - \alpha q(i)}{1-e}$ which, again, is higher for low types. If an investor buys one unit of firm i at a unit price w his profits are $q(i) - w$. Heterogeneous information among investors could be the result of differential experience in this particular industry, differential contacts with company insiders, or differential access to analyst reports, which make some investors better than others at distinguishing good from bad projects.[4]

[4] 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. 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 (e,w), its utility is $u(e,w) = qv(X-(1-w)(d-e)) + (1-q)v(X-(1-w)(d-e)-e)$, where $v(\cdot)$ is the household's von Neumann-Morgenstern utility function. The marginal rate of substitution is $-\frac{\partial u(e,w)/\partial e}{\partial u(e,w)/\partial w} = \frac{1}{d-e}\left(w - q\left(q + (1-q)\frac{v'(X-(1-w)(d-e)-e)}{v'(X-(1-w)(d-e))}\right)^{-1}\right)$. 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 $1 - w - (1-q(i)) = q(i) - w$.[5] 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 $G_H(y)$ 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 $u(e,w) = w - \tau e - \beta\int_0^e (e-y)\,dG_k(y)$ with $k = L$ if $i < \lambda$ and $k = H$ if $i \geq \lambda$. The marginal rate of substitution is $-\frac{\partial u(e,w)/\partial e}{\partial u(e,w)/\partial w} = \tau + \beta G_k(e)$. 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

insurance application we sketch below.

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

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

We adopt a Walrasian approach similar to the notion of competitive search equilibrium. There are many (nonexclusive) markets open simultaneously, each defined by a required signal $e \in \mathbb{R}^+$ and a wage $w \in [0, q_H]$, and there is no guarantee for either workers or firms of finding a counterparty in a market they visit.

3.1 Worker's Problem

Worker i first chooses a signal e and then applies for jobs. 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, potentially in multiple markets with different unit prices.

A worker is allowed to apply to all the markets that require his chosen signal e. We assume that, for any given signal e, markets at different wages clear sequentially, starting from the highest wage, as in the "buyer's equilibrium" studied by Wilson (1980). Therefore a worker starts by applying to market (e, q_H) and, as long as he hasn't been hired, continues to apply to lower-wage markets. Eventually, he gets hired in market (e, w), and does not apply to markets with lower wages. The worker understands that each choice of e is associated with some probability distribution over wage offers, with c.d.f. denoted by $\mu(\cdot; e, i)$. The worker's problem is therefore:

$$\max_e \bar{w}(e, i) - c(i)\, e$$

where $\bar{w}(e, i) = \int w d\mu(w; e, i)$ is the expected wage. We denote the choice of worker i by e_i.

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, $\chi(i) = \chi(i')$ whenever $x(i, \theta) = x(i', \theta)$. A firm will reject applicants with $\chi(i) = 0$ and hire workers (which we describe as χ-acceptable) from the set $I_\chi = \{i \in [0, 1] : \chi(i) = 1\}$. Let X denote the set of possible hiring rules.

A firm must decide what market to hire from and what hiring rule to apply (it is without loss of generality to assume the firm hires only in one market and uses only one rule). To make this decision, the firm needs to form beliefs $G(\cdot; e, w, \chi)$ about what workers it will be drawing from should it choose to hire in market (e, w) with hiring rule χ. If the firm thinks it will find χ-acceptable workers in market (e, w), then $G(\cdot; e, w, \chi)$ is a probability measure on I_χ; otherwise beliefs integrate to zero: $G(I_\chi; e, w, \chi) = 0$. Let g denote the density or p.m.f. of G, which we assume is well-defined. Firm θ's problem is:

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

Note that a firm has the choice not to hire workers by simply directing its search to a market/hiring rule where $G(I_\chi; e, w, \chi) = 0$. We denote the choices of firm θ by $(e_\theta, w_\theta, \chi_\theta)$.

3.3 Consistency of Wage Distributions

We define demand as a measure on the set of wages, signals and hiring rules $[0, q_H] \times \mathbb{R}^+ \times X$. For any set of wages $W_0 \subseteq [0, q_H]$, signals $E_0 \subseteq \mathbb{R}^+$ and hiring rules $X_0 \subseteq X$, demand is the total number of firms who make those choices:

$$D(E_0, W_0, X_0) \equiv \int \mathbb{I}(e_\theta \in E_0)\, \mathbb{I}(w_\theta \in W_0)\, \mathbb{I}(\chi_\theta \in X_0)\, dF(\theta)\,. \tag{3}$$

We then impose the following consistency condition on firms' hiring and the distribution of wage offers received by workers:

Condition 1.

$$\mu(w; e, i)\, \mathbb{I}(e_i = e) = \int\limits_{\tilde{w} \le w, X} g(i; e, \tilde{w}, \chi)\, dD(e, \tilde{w}, \chi) \qquad \textit{for all } e, w, i.$$

The indicator $\mathbb{I}(e_i = e)$ takes the value 1 if worker i chooses signal e and zero otherwise; $\mu(w; e, i)$ is his probability of getting a wage at most w. Hence, the left-hand side of Condition 1 is the total number of i-type workers with signal e who will obtain wages at most w. Moreover, since beliefs are rational, a firm imposing hiring rule χ in market (e, w) will hire $g(i; e, w, \chi)$ workers of type i. Adding these hires across all hiring rules and wages below w using the demand measure results in the right-hand side of Condition 1, which is the total number of i-type workers hired in markets with signal e and wages up to w.

Condition 1 simplifies when i-type workers choose signal e (so $\mathbb{I}(e_i = e) = 1$), and they have strictly positive probability of finding a job at wage w, so the c.d.f. μ makes a discrete step of size $d\mu(w; e, i)$. Then Condition 1 can be written as:

$$d\mu(w; e, i) = \frac{\int_X g(i; e, w, \chi) dD(e, w, \chi)}{\mathbb{I}\left(e_i = e\right)}. \tag{4}$$

This is the standard rationing rule under frictionless matching, by which the probability $d\mu$ for an i-type worker of finding a job at wage w is equal to the ratio of i-type workers demanded by firms in that market over their supply, which is equal to 1.[6] The more general formulation of Condition 1 also deals with cases where μ may increase continuously over some interval of wages, so the probability of being hired in any single market is zero but there is an associated probability density. Both situations will occur in the equilibria we find below.

Example 1. To illustrate the meaning of equations (3) and (4), consider the following example. There are three types of workers $i \in \{A, B, C\}$, each of mass one, who choose signal e, and three firm types $\theta \in \{\alpha, \beta, \gamma\}$ who hire in markets that require e. The measures of each type of firm, the wage at which they recruit, their hiring rules, and beliefs are:

Firm	θ	α	β	γ
Measure	$f\left(\theta\right)$	0.5	1	1.5
Wage	w_θ	w_H	w_H	w_L
Hiring Rule	$\chi_\theta\left(A\right)$	1	1	1
	$\chi_\theta\left(B\right)$	0	1	1
	$\chi_\theta\left(C\right)$	0	0	1
Beliefs	$g(A; e, w_\theta, \chi_\theta)$	1	0.5	0
	$g(B; e, w_\theta, \chi_\theta)$	0	0.5	1/3
	$g(C; e, w_\theta, \chi_\theta)$	0	0	2/3

Firms of type α and β hire at wage w_H. Type-α firms only hire $i = A$ workers while type-β firms hire type $i = A$ and $i = B$ workers. Type-γ firms hire at wage w_L, and they accept everyone. For now, we take firms' beliefs simply as given and will show below how they are derived. Given beliefs and firms' demand decisions, we can compute workers' probabilities of

[6]Note that equation (4) and Condition 1 rule out a situation with excess demand for a type of worker because it would imply a job-finding probability higher than one. In other words, in any given active market, the firm-to-worker ratio (or market tightness), which generally differs across worker types, never exceeds one.

being hired at wages w_H and w_L using (3) and (4):

$$
\begin{aligned}
d\mu\left(w_H; e, A\right) &= \underbrace{g\left(A; e, w_H, \chi_\alpha\right)}_{1}\underbrace{D\left(e, w_H, \chi_\alpha\right)}_{0.5} + \underbrace{g\left(A; e, w_H, \chi_\beta\right)}_{0.5}\underbrace{D\left(e, w_H, \chi_\beta\right)}_{1} = 1 \\
d\mu\left(w_H; e, B\right) &= \underbrace{g\left(B; e, w_H, \chi_\alpha\right)}_{0}\underbrace{D\left(e, w_H, \chi_\alpha\right)}_{0.5} + \underbrace{g\left(B; e, w_H, \chi_\beta\right)}_{0.5}\underbrace{D\left(e, w_H, \chi_\beta\right)}_{1} = 0.5 \\
d\mu\left(w_L; e, B\right) &= \underbrace{g\left(B; e, w_L, \chi_\gamma\right)}_{\frac{1}{3}}\underbrace{D\left(e, w_L, \chi_\gamma\right)}_{1.5} = 0.5 \\
d\mu\left(w_L; e, C\right) &= \underbrace{g\left(C; e, w_L, \chi_\gamma\right)}_{\frac{2}{3}}\underbrace{D\left(e, w_L, \chi_\gamma\right)}_{1.5} = 1.
\end{aligned}
$$

In words, type-A workers get hired for sure at wage w_H, type-B workers get hired with probability 0.5 at wage w_H and probability 0.5 at wage w_L, and type-C workers get hired for sure at wage w_L.

Condition 1 imposes no constraints on μ when $\mathbb{I}\left(e_i = e\right) = 0$, i.e. no constraints on i-workers' chances of being hired in markets where there are no i-applicants. For theses markets, we impose the condition:

Condition 2. *μ is weakly decreasing in i*

Condition 2 says that higher-i workers expect higher wages in a FOSD sense. This rules out low types being more optimistic than high types about the wages they would obtain for some off-equilibrium signals. This condition can be derived from more primitive assumptions. Suppose workers believe that firms which hire in markets with off-equilibrium signals use optimal hiring rules. These are weakly monotonic: no firm finds it optimal to accept worker i while rejecting worker $i' > i$. If firms draw workers randomly from those that satisfy their hiring rule (as discussed in the next section), applying Condition 1 to off-equilibrium signals implies that higher types will be hired at weakly higher rates than lower types, resulting in FOSD higher wages.

3.4 Consistency of Beliefs

Consider a firm that hires in market (e, w) with hiring rule χ. The pool of workers available for hire in this market includes i-type workers only if they choose education e and have not already been hired at higher wages. Therefore it includes $\mathbb{I}\left(e_i = e\right)\mu\left(w; e, i\right)$ i-type workers. If firms simply chose at random from the χ-acceptable subset of this pool, then Bayes' Rule would imply that rational beliefs should be:

$$g(i;e,w,\chi)=\frac{\mathbb{I}(e_i=e)\,\chi(i)\,\mu(w;e,i)}{\int_i \mathbb{I}(e_i=e)\,\chi(i)\,\mu(w;e,i)\,di},\tag{5}$$

However, if firms with different hiring rules hire sequentially in the same market, firms that hire earlier skew the pool that later firms face, so rational beliefs depend on the order in which firms hire within a market. Kurlat (2016) assumes that there exist separate markets for each wage combined with each possible way of ordering rules, and firms and workers choose which markets to trade in, making the order endogenous. He shows that under "false positives" information, less selective firms hire first. This implies that no one's sample is skewed by earlier firms, so it is as if all firms were drawing from the entire pool of χ-acceptable applicants, and (5) applies.

Example 2. Continuing on Example 1, we illustrate how to derive firms' beliefs using (5) and workers' hiring probabilities computed above. In market (e,w_H), both firms of type α and β hire, but β-firms get to pick first. Since all three worker types apply in this market but firms of type β reject type-C workers, their beliefs are

$$g(i;e,w_H,\chi_\beta)=\frac{1}{1+1+0}=0.5\text{ for }i=A,B,\quad g(C;e,w_H,\chi_\beta)=\frac{0}{1+1+0}=0.$$

This means that the unit measure of type-β firms hire a total of half a type-A and half a type-B worker. Firms of type α hire next in this market, and they reject all but type-A workers, so their beliefs are trivially:

$$g(A;e,w_H,\chi_\alpha)=\frac{1}{1+0+0}=1,\quad g(i;e,w_H,\chi_\alpha)=\frac{0}{1+0+0}=0\text{ for }i=B,C.$$

Hence, the measure 0.5 of type-α firms hires the remaining half type-A worker. Finally, firms of type γ hire in market (e,w_L) accepting all applicants. Since type-A workers have been hired for sure at the higher wage w_H, they are no longer applying at w_L, so $g(A;e,w_L,\chi_\gamma)=0$. γ-firms thus face an applicant pool of half a type-B worker and one type-C worker, so Bayes' Rule implies beliefs

$$g(B;e,w_L,\chi_\gamma)=\frac{\frac{1}{2}}{0+\frac{1}{2}+1}=\frac{1}{3},\ g(C;e,w_L,\chi_\gamma)=\frac{1}{0+\frac{1}{2}+1}=\frac{2}{3}.$$

As a result, the 1.5 measure of type-γ firms hire all the remaining workers.

Instead, under "false negatives" information, Kurlat (2016) shows that there may be markets where more selective firms hire first, and (5) does not apply. However, this possibility only arises among firms who only accept high types. Therefore, even if early firms skew the pool, later firms still hire only high types, just as if they had been the first in line. Therefore, the following weaker condition still holds:

Condition 3. *If* $\int_i \mathbb{I}(e_i = e)\,\chi(i)\,\mu(w;e,i)\,di > 0$*, then*

$$\int_i q(i)\,dG(i;e,w,\chi) = \frac{\int_i q(i)\,\mathbb{I}(e_i = e)\,\chi(i)\,\mu(w;e,i)\,di}{\int_i \mathbb{I}(e_i = e)\,\chi(i)\,\mu(w;e,i)\,di}.$$

Condition 3 says that beliefs must be such that the average productivity that firms expect to get if they hire in market (e, w) with hiring rule χ must be the same as if they were drawing from the entire pool. For the false positives case, it holds because it is implied by (5). For the false negatives case, it holds because the only cases where (5) might not hold are when firms only accept high type workers. Rather than explicitly allowing for endogenous ordering of firms' trades and re-deriving these results, we incorporate them directly into our definition of equilibrium by imposing Condition 3.

Example 3. Continuing further on Examples 1 and 2, suppose that workers of type A and B have productivity q_H while worker type C has productivity q_L. If less selective firms hire first, as assumed in Example 2, then by Bayes' rule, firms α and β expect average quality q_H, while firm γ expects average quality $(2q_L + q_H)/3$, so Condition 3 applies. Suppose now, by contrast, that more selective firms hire first, so firm α gets to pick workers before firm β in market (e, w_H). Since firm α hires half a type-A worker first, this skews firm β's remaining applicant sample, and hence its beliefs are as follows:

$$g(A;e,w_H,\chi_\beta) = \frac{\frac{1}{2}}{0+\frac{1}{2}+1} = \frac{1}{3},\; g(B;e,w_H,\chi_\beta) = \frac{1}{0+\frac{1}{2}+1} = \frac{2}{3},$$

which violates equation (5). Nonetheless, since both workers A and B have quality q_H, firm β expects average quality q_H, so Condition 3 remains satisfied.

Condition 3 only applies to markets where the denominator is positive, i.e. where there are χ-acceptable workers. The key challenge in constructing a tractable equilibrium notion is how to discipline firms' beliefs in markets that are empty of χ-acceptable workers. We propose a refinement which guarantees equilibrium uniqueness in the no-information benchmark, and at the same time preserves equilibrium existence throughout.[7] For markets in which no χ-acceptable workers apply, there are two possibilities: either the firm nevertheless believes it could find χ-acceptable workers and $G(\cdot;e,w,\chi)$ is a well-defined probability measure, or the firm believes the market is empty and $G(I_\chi;e,w,\chi) = 0$.

[7] This issue does not arise in Kurlat (2016). He only studies unidimensional contracts, where the price is the only contract dimension. This rules out signaling, and thus there are no markets corresponding to off-equilibrium signals, and no need to specify how beliefs react to these off-equilibrium signals. In his setup, beliefs must satisfy Condition 3, but there is no equivalent to the condition we impose in the following.

For the first case, we require that beliefs only place weight on χ-acceptable workers that would in fact be willing to look for a job in market (e, w). In other words, a firm can never expect to find in market (e, w) a worker who could obtain higher utility by choosing a different signal, or who can find a job for sure with the same signal but a higher wage. Formally, we require:

Condition 4. *For any i in the support of $G(\cdot; e, w, \chi)$:*

1. $\chi(i) = 1$
2. *e solves worker i's problem*
3. $\mu(w; e, i) > 0$

The alternative is that a firm is certain that it cannot find χ-acceptable workers in market (e, w). We impose that a firm can only reach that conclusion if guaranteeing χ-acceptable workers a job with a wage at least w is not enough to persuade them to choose signal e. Formally, we impose:

Condition 5. *If $G(I_\chi; e, w, \chi) = 0$, then $\mu(w; e, i) = 0$ for all i such that $\chi(i) = 1$.*

Conditions 4 and 5 are 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. Conditions 4 and 5 generalize this idea by imposing it separately for each χ-acceptance group. For each χ, it has to be the case that either some χ-acceptable worker finds visiting this market optimal (in which case this worker can be in the support of well-defined beliefs) or all χ-acceptable workers are guaranteed jobs. Within a given market, which of these possibilities applies can be different across different hiring rules χ.

3.5 Equilibrium Definition

We summarize the above discussion in the following equilibrium definition:

Definition 1. An equilibrium consists of (i) a signal e_i for each worker i; (ii) a hiring decision $(e_\theta, w_\theta, \chi_\theta)$ for each firm θ; (iii) wage distributions $\mu(\cdot; e, i)$; and (iv) beliefs $G(\cdot; e, w, \chi)$ that satisfy:

1. *Worker optimization.* e_i solves worker i's problem, taking μ as given.
2. *Firm optimization.* $(e_\theta, w_\theta, \chi_\theta)$ solves firm θ's problem, taking G as given.
3. *Consistency.* μ, G, $(e_\theta, w_\theta, \chi_\theta)$ and e_i satisfy Conditions 1 to 5.

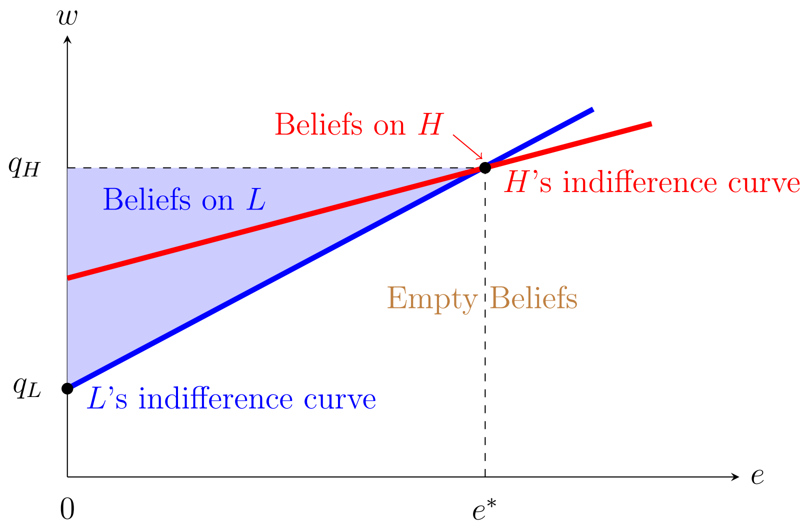

Figure 1: The least-cost separating allocation.

4 Pure Signaling

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

$$e^* = \frac{q_H - q_L}{c_L}, \tag{6}$$

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:*

1. *Worker decisions:*
$$e_i = \begin{cases} 0 & \text{if } i < \lambda \\ e^* & \text{if } i \geq \lambda \end{cases} \tag{7}$$

2. *Firm decisions:*
$$\begin{aligned} &e = e^*, w = q_H, \chi(i) = 1 && \text{for a measure } \lambda \text{ of firms} \\ &e = 0, w = q_L, \chi(i) = 1 && \text{for a measure } 1-\lambda \text{ of firms} \end{aligned} \tag{8}$$

3. *Wage distributions:*
$$\mu(w; e, i) = \mathbb{I}\left(w \geq \min\left\{q_L + c_L e, q_H + c_H\left(e - e^*\right)\right\}\right) \tag{9}$$

4. Beliefs:

$$g\left(i;e,w,\chi\right)=\begin{cases}\frac{1}{\lambda}\mathbb{I}\left(i<\lambda\right) & \textit{if } e<e^{*},w\geq q_{L}+c_{L}e\\ \frac{1}{1-\lambda}\mathbb{I}\left(i\geq\lambda\right) & \textit{if } e\geq e^{*},w\geq q_{H}+c_{H}\left(e-e^{*}\right)\\ 0 & \textit{otherwise}\end{cases} \tag{10}$$

The equilibrium is constructed by setting the distribution μ as a point mass 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, and firms are trivially optimizing. A measure $F(1) - 1$ remain inactive, for instance by choosing a market with $e = 0$ and $w < q_L$ (and selection rule $\chi(i) = 1$ for all i).

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 of a pure-strategy equilibrium in Rothschild and Stiglitz (1976) are not profitable because each firm perceives itself to be small.[8] 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

We now consider the case where F has full support on $[0, \lambda]$ and is continuous. We start by characterizing the equilibrium when workers don't have a way of signaling ("no signaling"). Next, we characterize a class of possible equilibria ("partial signaling") that involve signaling by a fraction of the high types. We then show that any equilibrium must be either the pure signaling equilibrium described in Section 4, a no signaling equilibrium or a partial signaling

[8] Rosenthal and Weiss (1984) and Dasgupta and Maskin (1986) show that mixed-strategy equilibria do exist.

equilibrium, and find conditions for each of them to arise.

5.1 No Signaling

We now characterize the equilibrium for the case where workers are constrained to choose $e = 0$, which is the case studied by Kurlat (2016). Let θ^N and w^N be defined as the solutions to:

$$\int_{\theta^N}^{\lambda} \frac{1}{1-\theta} dF(\theta) = 1. \tag{11}$$

$$\theta^N = \frac{w^N - (\lambda q_L + (1-\lambda) q_H)}{w^N - q_L} \tag{12}$$

In equilibrium, all high-type workers (and some low-type workers) are hired at some wage w^N, and the low-type workers who fail to find a job at wage w^N are hired at wage q_L.

To understand the meaning of conditions (11) and (12), observe first that if firm θ hires at wage w^N and imposes hiring rule $\chi_\theta(i) = \mathbb{I}(i \geq \theta)$, it hires randomly from the interval $[\theta, 1]$. Therefore it ends up hiring a low type with probability $\frac{\lambda-\theta}{1-\theta}$ and a high type with probability $\frac{1-\lambda}{1-\theta}$. Its expected profits will be: $\Pi(\theta) = \frac{(\lambda-\theta)q_L + (1-\lambda)q_H}{1-\theta} - w^N$. Profits are increasing in θ: firms whose information enables them to screen out a higher proportion of low types will be hiring from a better pool of workers. Only firms that are sufficiently confident in their ability to tell workers apart will be willing to hire in this market; they will make profits if and only if they are above the cutoff θ^N defined by equation (12), which satisfies $\Pi(\theta^N) = 0$.

For all $1 - \lambda$ high-type workers to be hired at wage w^N, it must be that there are enough firms in the range $[\theta^N, \lambda]$ to hire all of them. Given that in expectation firm θ hires $\frac{1-\lambda}{1-\theta}$ high-type workers, this means that θ^N must satisfy (11).[9]

The following result, proved by Kurlat (2016), establishes that this is the unique equilibrium.

Proposition 2. *If workers are constrained to choose $e = 0$, there is a unique equilibrium where:*

1. *Firms with $\theta \geq \theta^N$ hire at wage w^N and other firms are indifferent between hiring at wage q_L or not hiring.*

[9] Note that low-type workers do *not* guarantee themselves a job at wage w^N since some of the firms hiring in this market will reject them; only firms $\theta \in [\theta^N, i]$ hire in market at wage w^N and accept worker i. Therefore, his probability of finding a job at wage w^N is

$$d\mu(w^N; i) = \int_{\theta^N}^{i} \frac{1}{1-\theta} dF(\theta) \quad \text{for } i \in [\theta^N, \lambda]. \tag{13}$$

This probability is increasing in i since higher-i low types mislead more firms into hiring them at wage w^N. It is equal to zero for workers $i < \theta^N$ since no firm that would accept them hires at wage w^N.

2. *High types are hired at* w^N *with probability* 1.

3. *Beliefs follow* (5) *for all* $w \geq q_L$ *and are zero for lower wages.*

Conditions (11) and (12) are the analogues of conditions (19) and (20) in Kurlat (2016). There are four minor differences. First, Kurlat (2016) assumes that assets are divisible and the law of large numbers applies, so he has exact pro-rata rationing instead of probabilistic rationing. Under risk neutrality, this distinction does not matter. Second, he allows some sellers to have a positive value for retaining the good, while we assume it to be zero so workers sell all their labor endowment in equilibrium. Third, Kurlat (2016) assumes $q_L = 0$, so the one-price equilibrium he finds is equivalent to the two-price equilibrium we have, where some low types trade at price q_L. Finally, he models buyers' capacity constraints in terms of dollars rather than in terms of quantities, so the price appears in the market-clearing condition.

Note that as F approaches a point mass at $\theta = 0$, then equations (11)-(13) imply that $\theta^N \to 0$, $w^N \to \lambda q_L + (1-\lambda) q_H$ and $d\mu\left(w^N; i\right) \to 1$. If firms are uninformed and workers cannot signal, then all workers get hired for sure at a wage equal to the average productivity. This is the pure Akerlof (1970) outcome: all workers have the same reservation wage (zero) so there is no adverse selection at the pooled price.

5.2 Partial Signaling

In a partial signaling equilibrium, low-type workers choose $e = 0$. They are hired with some probability in at wage w^P, defined by:

$$w^P = q_H - c_H e^*. \tag{14}$$

and otherwise at wage q_L. High-type workers choose either $e = 0$ (and are hired for sure at wage w^P) or $e = e^*$ (and are hired for sure at wage q_H, which gives them the same utility).

Let π^P be the fraction of high types that choose $e = 0$. If firm θ hires in market $\left(0, w^P\right)$ with hiring rule $\chi_\theta(i) = \mathbb{I}[i \geq \theta]$, it will hire a high type with probability $\frac{\pi^P(1-\lambda)}{\lambda - \theta + \pi^P(1-\lambda)}$, so its profits will be $\Pi(\theta) = \frac{(\lambda-\theta)q_L + \pi^P(1-\lambda)q_H}{\lambda-\theta+\pi^P(1-\lambda)} - w^P$. This defines a cutoff θ^P such that firms can make profits in market $\left(0, w^P\right)$ if and only if $\theta > \theta^P$:

$$\theta^P = \lambda - \pi^P (1-\lambda) \frac{q_H - w^P}{w^P - q_L}. \tag{15}$$

Firms with $\theta < \theta^P$ are indifferent between hiring in market $(0, q_L)$ (with low-type applicants only), in market (e^*, q_H) (with high-type applicants only), or not hiring at all, since they make zero profits in any case.

The calculations above assume that some high types are indeed willing to apply to market $\left(0, w^P\right)$. For this to be true, it must be the case that they are sure they will find a job, since

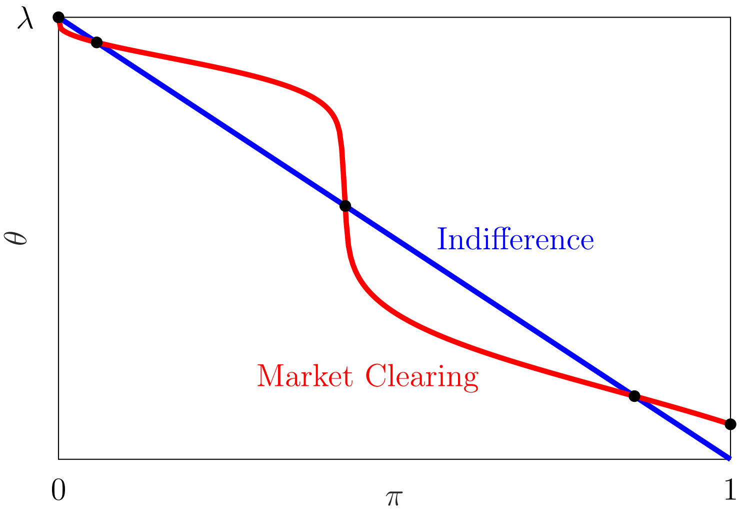

Figure 2: Indifference and market clearing conditions for the false positives case.

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^P to hire all $\pi^P (1 - \lambda)$ high types who forgo education and apply to market $\left(0, w^P\right)$. By the arguments above, each firm $\theta \geq \theta^P$ hires $\frac{\pi^P(1-\lambda)}{\lambda-\theta+\pi^P(1-\lambda)}$ high types, so in equilibrium we need

$$\int_{\theta^P}^{\lambda} \frac{1}{\lambda - \theta + \pi^P (1 - \lambda)} dF(\theta) = 1. \tag{16}$$

By the same reasoning as in the no-signaling case, low types are hired in market $\left(0, w^P\right)$ with probability $d\mu \left(w^P; 0, i\right) = \int_{\theta^P}^{i} \frac{1}{\lambda-\theta+\pi^P(1-\lambda)} dF(\theta)$.

The indifference condition (15) and the market clearing condition (16) define two relationships between the cutoff firm θ^P and the fraction of high types π^P that forgo signaling. Both of these relationships are downward sloping, as shown in Figure 2.

The indifference condition (15) is downward-sloping because if more high types decide to forgo education, they improve the pool of workers available for hire in market $\left(0, w^P\right)$, allowing less-informed firms to earn profits. The same is true for the market clearing condition (16) because if more high types decide to forgo education, they can only find jobs in market $\left(0, w^P\right)$ if additional firms decide to hire there. In other words, there is a *complementarity* between entry into market $\left(0, w^P\right)$ by firms and by high-type workers. The more high types forgo education, the more profitable it is for any given firm to hire in $\left(0, w^P\right)$; the more firms hire in $\left(0, w^P\right)$, the more high type workers can refrain from signaling.

The strategic complementarity implies that there can be multiple intersections of (15) and (16), and possibly multiple partial signaling equilibria. This source of multiplicity is different

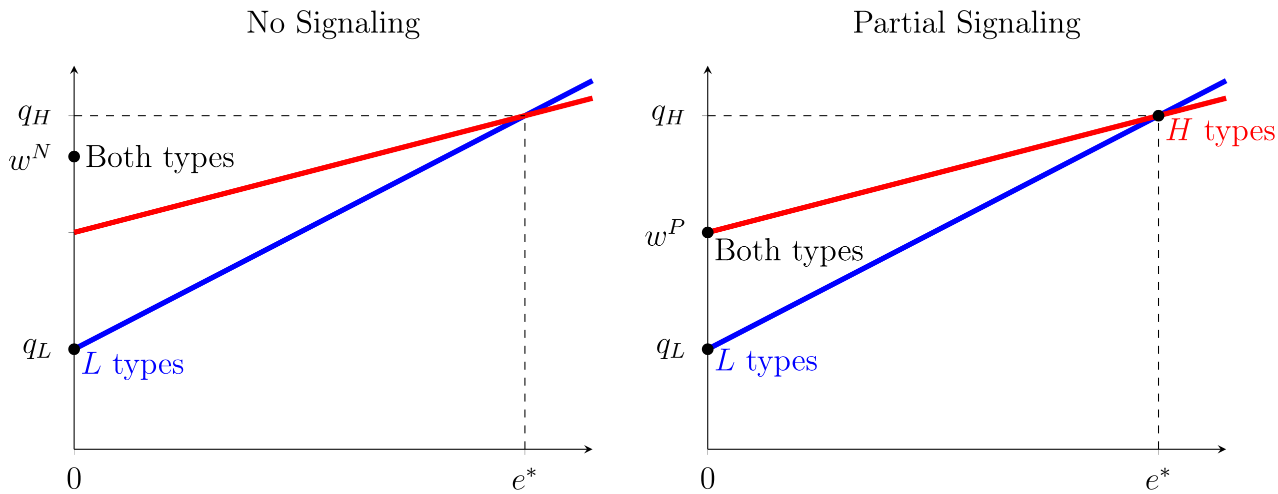

Figure 3: Active markets in each class of equilibrium.

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-firms benchmark has a unique equilibrium (Proposition 1), as does the no-signaling case (Proposition 2). The multiplicity we identify here relies on the presence of both signaling and heterogeneous information among firms.

Note that a no-signaling equilibrium corresponds to a situation where the market clearing condition is above the indifference condition at $\pi = 1$, as in Figure 2. This means that if $\pi = 1$ (no high types signal) there are more firms willing to hire at w^P than the total mass of workers they would accept. As a result, high-θ firms "bid up" the wage to $w^N > w^P$, leading some firms to drop out until the number of firms willing to pay this wage equals the number of high-type workers. Moreover, a pure signaling equilibrium is a special case of a partial signaling equilibrium, with $\pi^P = 0$ and $\theta^P = \lambda$. Figure 3 shows which markets are active in each class of equilibrium.

5.3 Candidate Equilibria

The following result establishes that any equilibrium must belong to one of the three cases described above.

Proposition 3. *Any equilibrium is of one of the three following types:*

1. *Pure signaling. Low types choose $e = 0$ and high types choose $e = e^*$.*

2. *No signaling. All workers choose $e = 0$. Firms hire in market $\left(0, w^N\right)$ if and only if $\theta \geq \theta^N$. θ^N and w^N satisfy (11) and (12); and $w^N \geq w^P$.*

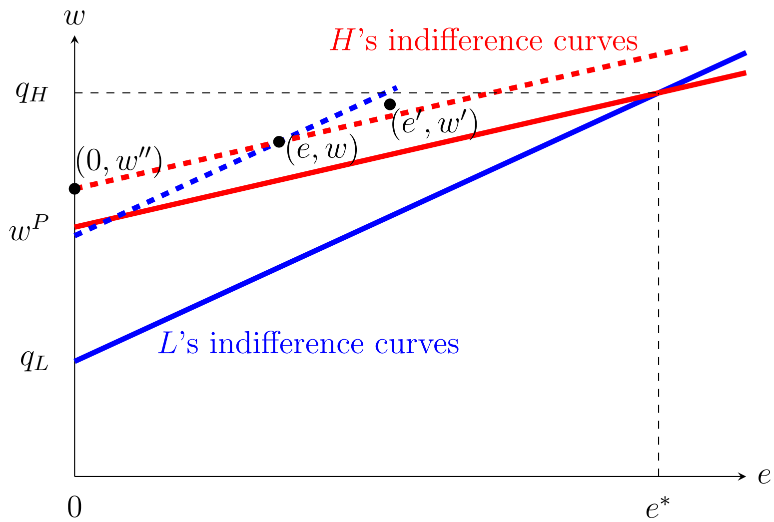

Figure 4: Ruling out pooling at $e > 0$.

3. *Partial signaling. Low types choose $e = 0$; a fraction π^P of high types choose $e = 0$ and the rest choose $e = e^*$. Firms hire in market $\left(0, w^P\right)$ if and only if $\theta \geq \theta^P$. π^P and θ^P satisfy (15) and (16).*

The key to proving Proposition 3 is to establish that high and low types cannot coexist at any level of education other than $e = 0$, so there is no pooling at positive signaling levels. The logic is similar, though somewhat subtler, to that in the uninformed-firms benchmark.

Suppose that low and high types coexisted in some market (e, w) with $e > 0$, as illustrated in Figure 4. With differentially informed firms, the standard argument that rules this out by considering market (e', w') does not go through. Some firms hiring in market (e, w) may be screening out low types, so low types' expected wage with signal e could be lower than that of high types. Thus, it is possible that both high and low types find (e', w') more attractive than (e, w). Instead, we can rule out pooling in market (e, w) by contradiction, as follows. There are two possibilities: either the highest firm that hires in market (e, w) has $\theta = \lambda$ (i.e., it can tell workers apart perfectly), or the highest firm to hire in this market has $\theta < \lambda$. In the first case, we arrive at a contradiction by considering the beliefs of firm $\theta = \lambda$ about market $(0, w'')$. This firm's beliefs can only include high types so it only cares about the wage it pays. Therefore this firm finds market $(0, w'')$ preferable over market (e, w), leading to a contradiction. If instead the highest firm that hires in market (e, w) has $\theta < \lambda$, then any low-type worker in the range $i \in (\theta, \lambda)$ has the same chance of getting a job in market (e, w) as a high type. If this is so, then firm θ can apply the standard cream-skimming deviation by hiring in market (e', w'). Firm θ can reject all low types who prefer (e', w') over (e, w), so it can guarantee itself high types by hiring in this market, which contradicts the premise that it hires in market (e, w).

5.4 Existence

So far we have described the possible candidates for equilibrium but we have not proved that any of them is actually an equilibrium. We now show that the candidate equilibria described above may or may not actually be equilibria. We construct a class of possible deviations and derive an easy-to-verify condition to determine whether these deviations are profitable. We then show that checking this condition is sufficient to establish an equilibrium, and prove that at least one equilibrium always exists.

Consider first a candidate partial signaling equilibrium. Define $\left(e_\theta^D, w_\theta^D\right)$ as the lowest-wage market where equilibrium requires that the beliefs of firm $\theta \in \left(\theta^P, \lambda\right)$ only include high types. A necessary condition for equilibrium is that firm θ cannot increase its profits by recruiting in market $\left(e_\theta^D, w_\theta^D\right)$ instead of market $\left(0, w^P\right)$.

The location of market $\left(e_\theta^D, w_\theta^D\right)$ is illustrated in Figure 5. Worker $i = \theta$ is the lowest-i low-type worker that firm θ cannot filter out. In equilibrium, this worker obtains expected utility:

$$u(\theta) = \ q_L + \int_{\theta^P}^{\theta} \frac{1}{\pi^P(1-\lambda)+\lambda-t} dF(t) \left(w^P - q_L\right) \tag{17}$$

by getting a wage of either w^P or q_L with the equilibrium probabilities. For small but positive levels of education e, it is consistent with equilibrium for firm θ to believe that it will only encounter low types in e-markets. The reason is that since $u(\theta) < w^P$, worker $i = \theta$ will be willing to choose e for a lower wage than high types would. Hence one can specify beliefs such that firm θ does not want to recruit at education level e. However, for large e that is no longer the case because education is more costly for low types. $\left(e_\theta^D, w_\theta^D\right)$ is defined by the intersection of the equilibrium indifference curves of worker $i = \theta$ and high types. At education levels higher than e_θ^D, firm θ can only believe that it will encounter exclusively high types, because high types would be willing to choose these education levels for a lower wage than worker $i = \theta$.

Hiring in a market like $\left(e_\theta^D, w_\theta^D\right)$ 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 $\left(0, w^P\right)$, where they are pooled with low types. Just like in the benchmark, the possible deviation involves peeling off high types 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 signaling. Since low types are hired from the $\left(0, w^P\right)$ 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,

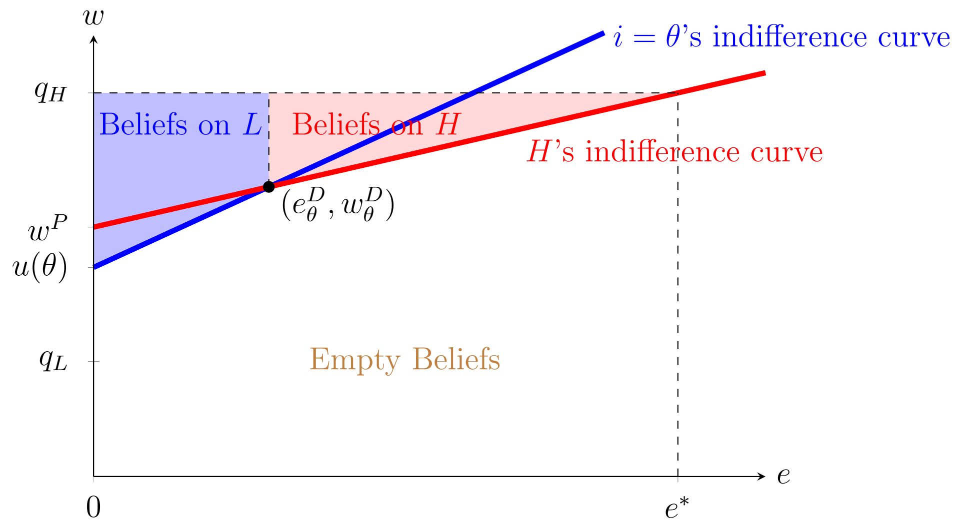

Figure 5: Beliefs for firm θ

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 completely uninformed firm cannot profitably deviate because in order to repel the lowest-i low types (who cannot get jobs at $\left(0, w^P\right)$ 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 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.

A candidate partial pooling equilibrium can only be an equilibrium if, for every $\theta \in (\theta^P, \lambda)$, the profits firm θ can obtain in market $\left(e^D_\theta, w^D_\theta\right)$ by hiring only high types are weakly lower than those it obtains in market $\left(0, w^P\right)$ by hiring a mixture of workers at a lower wage. A similar logic applies to the case of a no-signaling equilibrium. The following result determines when this condition is satisfied in either case and establishes that checking against this possible deviation is a sufficient condition for equilibrium existence.

Proposition 4. *1. The pure signaling candidate equilibrium described in Proposition 3 part 1 is always an equilibrium.*

2. Suppose θ^N and $w^N \geq w^P$ satisfy equations (11) and (12) for a no-signaling candidate equilibrium. Then the worker and firm decisions described in Proposition 3 part 2 are part of an equilibrium if and only if:

$$\frac{\lambda - \theta}{1 - \theta} \leq \frac{c_H}{c_L - c_H} \frac{1 - \lambda}{1 - \theta^N} \int_{\theta}^{\lambda} \frac{1}{1 - t} dF(t) \quad \text{for all } \theta \in \left(\theta^N, \lambda\right). \tag{18}$$

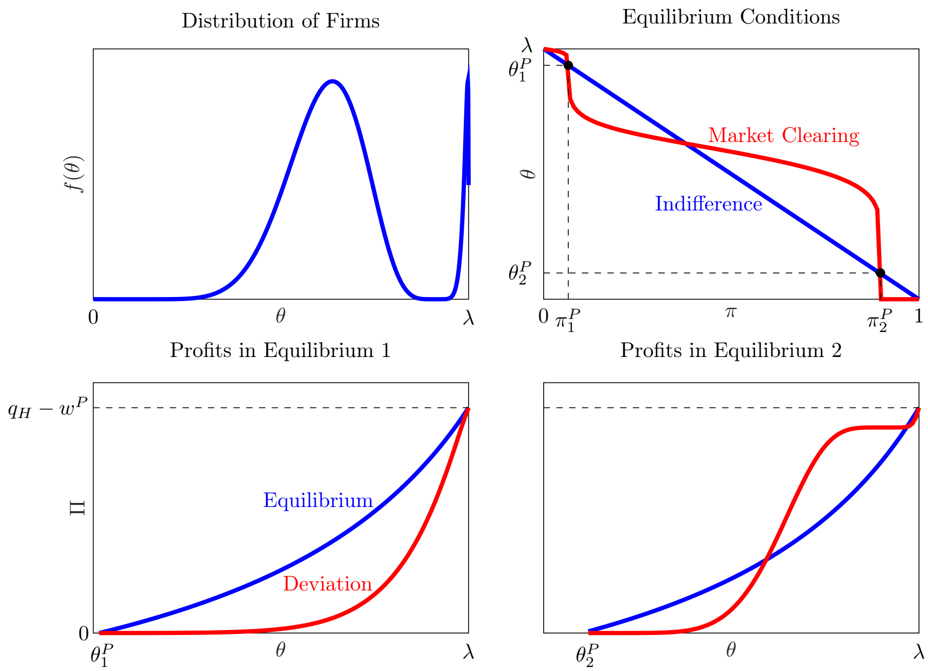

Figure 6: Candidate equilibria that do or do not satisfy condition (19).

3. *Suppose π^P and θ^P satisfy equations (15) and (16) for a partial signaling candidate equilibrium. Then the worker and firm decisions described in Proposition 3 part 3 are part of an equilibrium if and only if*

$$\frac{\lambda - \theta}{\pi^P(1-\lambda) + \lambda - \theta} \leq \frac{c_H}{c_L} \int_{\theta}^{\lambda} \frac{1}{\pi^P(1-\lambda)+\lambda - t} dF(t) \quad \textit{for all } \theta \in \left(\theta^P, \lambda\right). \tag{19}$$

In sum, the pure signaling equilibrium, which coincides with the no-information benchmark, always exists in our model. The reason is that $\pi^P = 0$, $\theta^P = \lambda$ always satisfies equations (15) and (16) and condition (19) holds for $\theta = \lambda$. Depending on parameters, additional equilibria may exist where firms use their expertise.

It is easy to construct examples where a partial or no-signaling equilibrium does exist. Figure 6 shows an economy with multiple candidate equilibria. For the candidate equilibrium $\left(\pi_1^P, \theta_1^P\right)$, condition (19) holds, so it is indeed a partial signaling equilibrium. Instead, for candidate equilibrium $\left(\pi_2^P, \theta_2^P\right)$, condition (19) fails for some $\theta > \theta_2^P$, so it is not an equilibrium.[10]

[10] The example uses $q_H = 1, q_L = 0.4, c_H = 0.9, c_L = 0.15, \lambda = 0.6, f(\theta) = 0.5[\sin(13.3(\theta-\lambda)^{0.4} + 2.2\pi) + 1]^{2.8}$.

5.5 Properties of the Equilibrium

Equilibrium Regions. Figure 7 illustrates what type of equilibrium arises in different regions of the parameter space. As we change parameters, the possible outcomes of the model span the range from pure signaling, via partial signaling, up to the no-signaling allocations in Kurlat (2016).

Both panels plot the equilibrium regions as a function of a parameter A on the vertical axis that shifts the distribution of firms F towards more expertise.[11] We know from Proposition 4 that the pure signaling equilibrium always exists, and it is indeed the only equilibrium for low enough levels of expertise as captured by the parameter A. As the distribution of expertise improves (in a FOSD sense), holding the other parameters fixed, first a partial signaling and finally a no-signaling equilibrium emerges in addition. Hence, as firms become better informed, less costly signaling is required. Moreover, we show formally in Appendix A that, in the region with a partial signaling equilibrium, the share of high types $1-\pi^P$ who signal also decreases with a FOSD shift in expertise. Better tools for directly evaluating job applicants, firm shares, asset-backed securities or insurance applicants reduce the need to signal through education, dividends, retained equity tranches, or high deductibles, respectively. In this way, direct information substitutes for traditional signaling.

A FOSD increase in F is isomorphic to an increase in demand where each firm hires Δ workers instead of just one. This is because making firms more expert is equivalent to letting the more expert firms hire more workers.[12] Our model thus generates the plausible prediction that more high types forgo signaling through education (or that the amount of retained equity falls) in boom times (see Gee (2018) for descriptive evidence of this effect). This intuitive property is absent when buyers are uninformed: in that case, pure signaling is always the only equilibrium independent of demand. It is also absent in the no-signaling equilibrium where higher demand just translates into higher wages.

The left panel shows that increasing the relative cost of signaling c_H/c_L has the same effects as improving expertise on the type of equilibrium we find, holding the other parameters fixed (including A).[13] Moreover, we show in Appendix A that, within the class of partial signaling equilibria, the amount of signaling $1-\pi^P$ decreases with signaling costs. Hence, as signaling gets more expensive, fewer high types signal. 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

[11] Specifically, we use the linear density $f(\theta) = A(\theta - \lambda/2) + B/\lambda$, so the total measure of firms always equals B, with $B = 1.2$.

[12] Mechanically, this is because condition (16) becomes $\int_{\theta^P}^{\lambda} \frac{\Delta f(\theta)}{\lambda-\theta+\pi^P(1-\lambda)} d\theta = 1$, so changing Δ is equivalent to a change in $f(\theta)$.

[13] The example in the graph uses $q_H = 1, q_L = 0.4, \lambda = 0.55$ in addition to the linear specification of $f(\theta)$ from above.

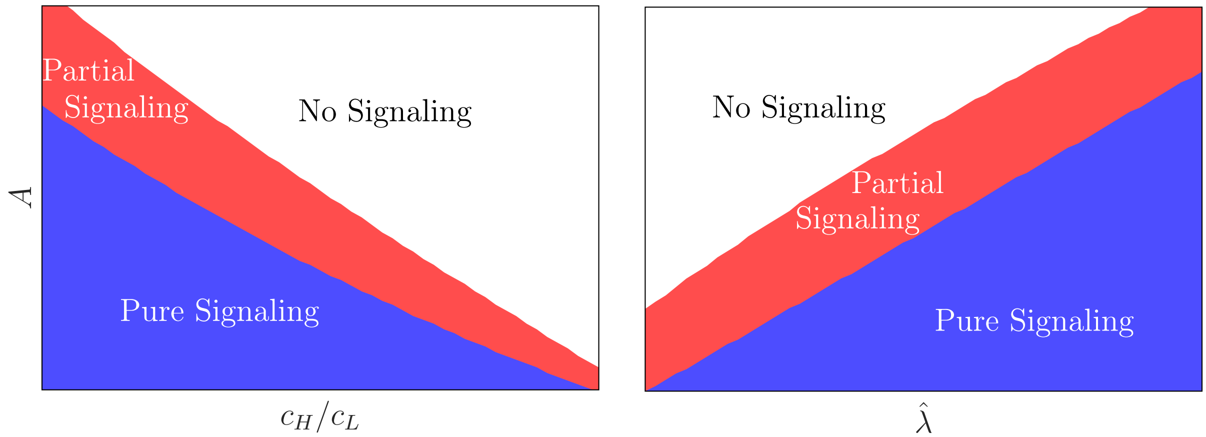

Figure 7: Types of equilibria depending on parameters

for heterogeneously informed firms overturns this counterintuitive feature: equilibrium forces do lead workers to respond on the extensive margin.

Finally, in the right panel, we vary the share of low types on the horizontal axis. To do so in a clean way, we reparametrize the model by assuming 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 types, leaving their relative detectability in the eyes of firms constant. We see that reducing the share of low types this way moves the equilibrium from pure signaling to partial equilibrium and finally to no signaling.[14] Indeed, we show formally below that, as the share of low types becomes sufficiently small, a no-signaling equilibrium must always emerge.

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, as in Daley and Green (2014) where the presence of exogenous information also avoids the discontinuity.

Proposition 5. *1. Let F be any continuous measure with full support on $[0, \lambda]$. For low $\hat{\lambda}$ there is a no-signaling equilibrium with $\lim_{\hat{\lambda}\to 0} w^N = q_H$.*

2. Let F^* be a mass point at $\theta = \lambda$. For any continuous F sufficiently close to F^* (under the total variation distance), there exists a no-signaling equilibrium, and $\lim_{F\to F^*} w^N = q_H$.

[14]The example uses $q_H = 1, q_L = 0.4, \lambda = 0.55, c_H = 0.1, c_L = 0.3$.

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 no-signaling equilibrium, and $w^N \to q_H$. Hence, this equilibrium smoothly approaches the symmetric information outcome. Pure signaling is also an equilibrium for any positive $\hat{\lambda}$, so the discontinuity does not go away entirely, but the set of equilibria is lower 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 no-signaling equilibrium always exists sufficiently close to the limit, so the set of equilibria is lower hemi-continuous in this dimension as well.[15]

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 lower 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. *1. For c_H/c_L sufficiently close to 1, there is a no-signaling equilibrium.*

2. For c_H/c_L sufficiently close to 0, only the pure signaling equilibrium exists.

Part 1 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 2 establishes that if signaling is sufficiently cheap, then the only equilibrium allocation is the benchmark least-cost separating allocation and firms' expertise is not used.

5.6 Welfare

The only reason why allocations in the model are not first-best efficient is that signaling is socially wasteful. This does not immediately imply that equilibria with less signaling 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 partial signaling equilibria can indeed be Pareto-ranked against each other (and against the pure signaling equilibrium), but cannot be Pareto-ranked against a no-signaling equilibrium if it exists:

[15] By contrast, in the degenerate case where F has full mass at some $\theta < \lambda$, a no-signaling equilibrium never exists. The right-hand side of (18) 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 a standard signaling model and only the pure signaling equilibrium exists. Hence, heterogeneity of information is crucial to obtain the continuity results in this section.

Proposition 7. *1. Suppose there is a partial signaling equilibrium with $\pi_1^P > 0$.*

(a) It Pareto dominates the pure signaling equilibrium in the same economy.

(b) If there is another partial signaling equilibrium with $\pi_1^P > \pi_2^P$ in the same economy, the first equilibrium Pareto dominates the second.

2. Suppose there is a no-signaling equilibrium.

(a) It Pareto dominates the pure signaling equilibrium in the same economy.

(b) If there is also a partial signaling equilibrium with $\pi^P > 0$ in the same economy, neither equilibrium Pareto dominates the other.

In comparing partial signaling equilibria, it is straightforward to show that firms are better off in the higher-π^P equilibrium, since wages are the same and there is a better pool of workers at $\left(0, w^P\right)$. High-type workers are indifferent because their payoff is w^P. The 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 $\left(0, w^P\right)$, which (other things being equal) increases their chances of earning w^P but lose from the fact that there are more high-type workers looking for work at $\left(0, w^P\right)$, 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^P equilibrium.

A no-signaling equilibrium (if it exists) cannot be Pareto ranked against partial signaling equilibria. Since the wage is higher and the cutoff firm is lower, workers are better off in the no-signaling 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 $\theta \in \left(\theta^N, \theta^P\right)$ are better off in the no-signaling equilibrium while they would make zero profits in the partial signaling equilibrium.

The model also makes it possible to ask, assuming there is a technology for firms to choose θ at some cost, whether they have the right incentives to invest in acquiring expertise, such as improving assessment models for job applicants, risk scoring models in insurance markets or pricing models for stocks and financial derivatives. In Appendix B, following the approach in Kurlat (2019), we show that in general the answer is ambiguous: firms may have incentives to either over-invest or under-invest in expertise. We also provide a simple formula to quantify the ratio of the social and private returns to expertise based on observable properties of the equilibrium.

6 False Negatives

We now turn to the case with "false negative" mistakes, where F is continuous with support $[\lambda, 1]$. 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. For expositional purposes, assume that the density of firms $f(\theta)$ is strictly increasing, meaning that there is a higher density of less informed firms. The general case, which requires working with an "ironed" density, is treated in Appendix D.

Unlike the false positives case, firms 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 it knows that in equilibrium some high-type workers with $i \in [\lambda, \theta)$ do apply, so it may want to hire from the pool of *all* applicants. We refer to this as non-selective hiring.

Description. In equilibrium, only the least transparent high-type workers signal. Letting u_L denote the low types' payoff, there is a cutoff i_S such that workers in the interval $i \in [\lambda, i_S]$ signal by choosing:

$$e_S = \frac{q_H - u_L}{c_L}, \tag{20}$$

while everyone else chooses $e = 0$. Signaling markets with $e = e_S$ are straightforward: all the applicants are high types, so less informed firms compete for them and hire them (non-selectively) at a wage $w = q_H$.

No-signaling markets, with $e = 0$, are more interesting. Define i_H by:

$$f(i_H) = 1. \tag{21}$$

Since $f(\theta)$ is assumed to be increasing, this means that for all $i > i_H$ there are more firms who can detect high-type workers than there are workers. Hence, firms compete for them and hire them (selectively) at wage $w = q_H$.

Conversely, for $i \in (i_S, i_H)$, there are more workers than firms who can identify them as high types. Therefore, some of them have to be hired non-selectively, at wages sufficiently low to attract non-selective firms. At each wage $w \in (q_L, q_H)$ where there is active hiring, two types of hiring take place: some workers are hired non-selectively, and in addition all the highest remaining i-types are hired selectively and thus drop out of subsequent, lower-wage markets. Let $\underline{w}(0, i)$ be the wage such that all worker types above i have already been hired. The pool of applicants at $\underline{w}(0, i)$ consists of all the low types plus high types in the interval $(i_S, i]$ who have not been hired non-selectively at higher wages. As a result, non-selective firms break even at a wage of:[16]

$$\underline{w}(0,i) = \frac{(i - i_S)\, q_H + \lambda q_L}{i - i_S + \lambda}. \tag{22}$$

Firms with $\theta = i$ hire $f(i)$ workers selectively in this market since it involves the cheapest wage at which they can identify high-type workers. Therefore, it must be that the remaining $1 - f(i)$ workers of type i were already hired non-selectively at wages above $\underline{w}(0,i)$. Since this is true for any i, the probability density for any worker of being hired non-selectively in market $(0, \underline{w}(0,i))$ must be $f'(i)$. Hence, the expected utility obtained by worker i is:

$$u(i) = f(i)\, \underline{w}(0,i) + \int_i^{i_H} \underline{w}(0,i')\, df(i'). \tag{23}$$

This defines a cutoff worker i^* who is indifferent between signaling (which gives a payoff $q_H - c_H e_S$) and not signaling:

$$f(i^*)\, \underline{w}(0,i^*) + \int_{i^*}^{i_H} \underline{w}(0,i')\, df(i') = q_H - c_H e_S. \tag{24}$$

Market $(0, \underline{w}(0,i^*))$ is the lowest-wage market at which there is a chance of being hired non-selectively. Low-type workers who have not found a job at or above this wage end up getting hired at $w = q_L$. Therefore, the expected utility of low types is

$$u_L = f(i^*)\, q_L + \int_{i^*}^{i_H} \underline{w}(0,i')\, df(i'). \tag{25}$$

Replacing (20), (22) and (25) into (24) and simplifying gives the following indifference condition for the marginal worker i^*:

$$\Gamma(i^*, i_S) = f(i^*) \left(\frac{c_H}{c_L} - \frac{\lambda}{i^* - i_S + \lambda} \right) - \lambda \left(1 - \frac{c_H}{c_L} \right) \int_{i^*}^{i_H} \frac{1}{i - i_S + \lambda} df(i) = 0. \tag{26}$$

Equation (26) defines a positive relationship between i^* (the worker who is indifferent between signaling and not signaling) and i_S (the cutoff for actually signaling). In general, i^* and i_S are not equal; there is a range of workers $i \in (i_S, i^*)$ who are indifferent between signaling and not signaling but choose not to. It is straightforward to show that i^* is increasing in i_S.

[16] Accordingly, the nonselective firms' beliefs are given by

$$g(i; 0, w, \chi) = \frac{\mathbb{I}(i < \lambda) + \mathbb{I}(i \in (i_S, i^r(w)])}{\lambda + i^r(w) - i_S}$$

for $\chi(i) = 1 \forall i$, where $i^r(w)$ is the inverse of $\underline{w}(0,i)$.

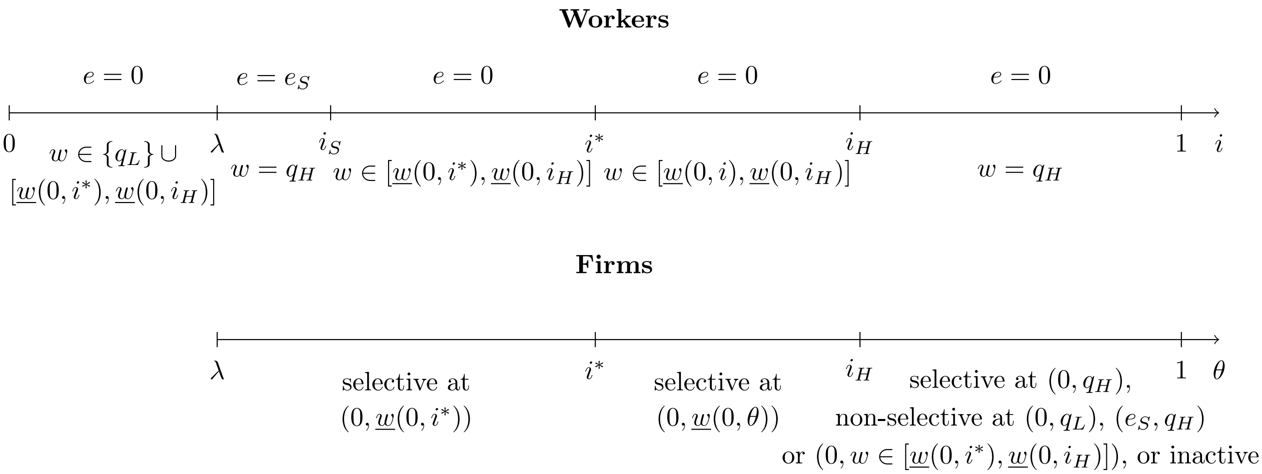

Figure 8: Signals, wages and hiring decisions in an interior equilibrium.

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 (22)). In turn, this means that the utility of both high- and low-type workers falls (equations (23) and (25)). It falls more for high types because low types are hired with positive probability in market $(0, q_L)$, where the wage is unaffected by higher i_S. Hence, other things equal, higher i_S makes signaling more attractive, so the indifferent type i^* rises.

A fraction $1 - f(i^*)$ of workers in the range $i \in (i_S, i^*)$ are hired at wages above $\underline{w}(0, i^*)$, so the remaining $f(i^*)(i^* - i_S)$ workers must be hired at wage $\underline{w}(0, i^*)$. For any $i \in (i_S, i^*)$, the measure of firms who are capable of identifying i as being a high type is $F(i)$, so we need $f(i^*)(i - i_S) \leq F(i)$. By monotonicity of f, this is implied by the market clearing condition:

$$F(i^*) = f(i^*)(i^* - i_S). \tag{27}$$

Equation (27) 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. Figure 8 summarizes the equilibrium signals, wages and hiring decisions.

Corner Equilibrium. Equations (26) and (27) 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, which would result in a corner equilibrium with $i^* = i_H$. For this corner equilibrium, the market clearing condition (27) and definition (21) 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:

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

so, using (6),

$$1 - \frac{c_H}{c_L} > \frac{F(i_H)}{F(i_H) + \lambda}, \tag{28}$$

which is equivalent to $\Gamma(i_H, i_H - F(i_H)) < 0$. The following proposition summarizes these results:

Proposition 8. *Under false negatives, there exists a generically unique equilibrium:*

1. *All high types $i \in [i_H, 1]$ choose $e = 0$ and are hired at $w = q_H$.*
2. *For $i \in [0, i_H)$, the equilibrium takes one of the following two possible forms:*
   (a) *An interior equilibrium where i_S and i^* solve (26) and (27) and:*
      i. *A measure $i_S - \lambda$ of high types with $i \in [\lambda, i^*)$ choose $e = e_S$, given by (20) and are hired at $w = q_H$.*
      ii. *All other high types with $i \in [\lambda, i^*)$ choose $e = 0$ and are hired at $w \in [\underline{w}(0, i^*), \underline{w}(0, i_H)]$.*
      iii. *All high types with $i \in [i^*, i_H)$ choose $e = 0$ and are hired at $w \in [\underline{w}(0, i), \underline{w}(0, i_H)]$.*
      iv. *All low types $i \in [0, \lambda)$ choose $e = 0$ and are hired at $w = q_L$ or $w \in [\underline{w}(0, i^*), \underline{w}(0, i_H)]$.*

   (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 are hired at $w = w^P$.*
      ii. *All other high types with $i \in [\lambda, i_H)$ choose $e = e^*$ and are hired at $w = q_H$.*
      iii. *All low types $i \in [0, \lambda)$ choose $e = 0$ and are hired at $w = q_L$.*

The proof is in Appendix D, which also describes all firms' decisions. Moreover, we show that the equilibrium behaves continuously in the symmetric information and expensive signaling limits, and we deal with the general case in which the density of firm types $f(\theta)$ is not necessarily monotone.

Properties. This model generates dispersion in expected wages among workers who are equally productive *and educated*, depending on how transparent they are. In particular, the expected wages of high types $i \in [i^*, i_H)$, who all select $e = 0$, are increasing in i.[17] Similarly, the model can explain, for instance, different prices for asset-backed securities for which both the

[17]In contrast, in the partial or no-signaling equilibria in the "false positives" case there is dispersion in expected wages among low types depending on their chances of being hired at w^P or w^N versus q_L.

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.

The structure of equilibrium is similar to the 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.

We can also ask how the intensive and extensive margins of signaling, measured by e_S and $i_S - \lambda$ respectively, depend (locally) on parameters around an interior equilibrium. In Appendix A, we show that an increase in the ratio c_H/c_L reduces signaling along both margins. For example, an increase in dividend taxes leads to both a smaller fraction of firms paying dividends and a lower dividends per dividend-paying firm. We also show that an increase in demand leads to *polarization* in signaling: fewer workers choose positive education but those who do choose a higher quantity.

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 implies 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," 2017. UCLA Working paper.

Bolton, Patrick, Tano Santos, and Jose A. Scheinkman, "Cream-Skimming in Financial Markets," *The Journal of Finance*, 2016, *71* (2), 709–736.

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.

Dasgupta, Partha and Eric Maskin, "The Existence of Equilibrium in Discontinuous Economic Games, II: Applications," *Review of Economic Studies*, 1986, *53* (1), 27–41.

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.

Eisfeldt, Andrea L., Hanno Lustig, and Lei Zhang, "Complex Asset Markets," *UCLA Working Paper*, 2019.

Farboodi, Maryam and Peter Kondor, "Heterogeneous Global Cycles," *MIT Working Paper*, 2018.

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

Fishman, Michael J. and Jonathan A. Parker, "Valuation, Adverse Selection, and Market Collapses," *The Review of Financial Studies*, 04 2015, *28* (9), 2575–2607.

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

Gee, Kelsey, "Employers Eager to Hire Try a New Policy: 'No Experience Necessary'," 2018. Wall Street Journal, July 29, 2018.

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," *American Economic Review*, 2019, *109* (2), 556–590.

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.

Rosenthal, Robert W. and Andrew Weiss, "Mixed-strategy equilibrium in a market with asymmetric information," *The Review of Economic Studies*, 1984, *51* (2), 333–342.

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.

_ , "The Nature of Equilibrium in Markets with Adverse Selection," *The Bell Journal of Economics*, 1980, *11* (1), 108–130.

A Comparative Statics

A.1 False Positives

We compute how the amount of signaling $1 - \pi^P$ in a partial signaling equilibrium depends on parameters. We focus on cases where the locus of the market clearing condition (16) is steeper than of the indifference condition (15), which corresponds to a heuristic notion of stability of the equilibrium.

Proposition 9. 1. *Signaling decreases with the cost ratio c_H/c_L.*

2. *Signaling decreases with a FOSD increase in the expertise distribution F or an increase in the demand for workers Δ.*

3. *Signaling does not change with productivities q_H and q_L.*

The logic of part 1 is as follows. The ratio c_H/c_L governs how much utility high types obtain if they separate by choosing e^*. Since w^P is the wage that makes them indifferent, higher c_H/c_L means a lower wage. This attracts lower-θ firms, so more high types can forgo signaling and still find a job. As for part 2, 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 (and an increase in Δ has the same effect). Therefore, more high types are able to forgo education and still find a job. Finally, productivities have no effect on equilibrium signaling. The wage w^P is a weighted average of q_H and q_L. Therefore, no matter what these productivities are, the indifferent firm θ^P will be the one whose pool of acceptable workers includes a proportion of exactly c_H/c_L low types. If, say, the productivity of low types was lower, the wage w adjusts exactly so as to leave firm θ^P indifferent and the fraction of high types who signal unchanged.

Proof. Using the reparametrization of the model where each firm demands Δ workers rather than just one, it is straightforward to show that equations (15) and (16) become

$$\theta^P = \lambda - \pi^P (1-\lambda) \frac{q_H - w^P}{w^P - q_L} \tag{29}$$

$$\int_{\theta^P}^{\lambda} \frac{\Delta}{(\lambda - \theta) + \pi^P (1-\lambda)} dF(\theta) = 1. \tag{30}$$

Replacing (6) and (14) into (29), the indifference condition reduces to:

$$\theta^P = \lambda - \pi^P (1-\lambda) \frac{1}{\frac{c_L}{c_H} - 1}. \tag{31}$$

Let $\theta^I\left(\pi^P, p\right)$ and $\theta^M\left(\pi^P, p\right)$ represent the solutions to (31) and (30) respectively, where p is a parameter. The equilibrium value of π^P is given by a solution to the equation $\theta^I\left(\pi^P, p\right) - \theta^M\left(\pi^P, p\right) = 0$. Using the implicit function theorem, the derivative of π^P with respect to parameter p is given by:

$$\frac{d\pi^P}{dp} = \frac{\frac{\partial \theta^M}{\partial p} - \frac{\partial \theta^I}{\partial p}}{\frac{\partial \theta^I}{\partial \pi^P} - \frac{\partial \theta^M}{\partial \pi^P}}. \tag{32}$$

By assumption, the denominator of (32) is positive, and equation (31) implies that $\frac{\partial \theta^I}{\partial \pi^P}$ is negative.

1. (31) implies that θ^I is decreasing in c_H/c_L, whereas c_H/c_L does not appear in equation (30). Using this in equations (32) gives the result.

2. The distribution F does not appear in equation (31). The term inside the integral in equation (30) is an increasing function of θ. Therefore a FOSD increase in F implies that the left-hand side of (30) increases, so θ^P must rise to maintain equality. Using this in equation (32) gives the first part of the result. (30) implies that θ^M is increasing in Δ, and Δ does not appear in equation (31). Using this in equation (32) gives the second part of the result.

3. This follows because neither q_H nor q_L appear in equation (32).

□

A.2 False Negatives

We compute how the intensive and extensive margins of signaling depend on parameters around an interior equilibrium.

Proposition 10. *1. Both the intensive and extensive margins of signaling decrease with the cost ratio c_H/c_L.*

2. The intensive margin of signaling increases but the extensive margin decreases with the demand for workers Δ.

3. The extensive margin of signaling is invariant with respect to productivities q_H and q_L; the intensive margin e_S increases with $q_H - q_L$.

Higher c_H/c_L makes separation more costly, so fewer high types signal. This improves the pool of workers in no-signaling markets, so non-selective firms pay higher wages. This raises the utility of low types, so less intense signaling is required to separate from them. 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 nonselectively. This makes low types worse off; therefore a more intense signal is needed to successfully separate, so fewer high types do so. As in the false-positives case, q_H and q_L drop out of equations (26) and (27), so the extensive margin is unchanged. However, 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.

Proof. Let $i^{*I}(i_S, p)$ and $i^{*M}(i_S, p)$ represent the solutions to (26) and (27) 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:

$$\frac{di_S}{dp} = \frac{\frac{\partial i^{*M}}{\partial p} - \frac{\partial i^{*I}}{\partial p}}{\frac{\partial i^{*I}}{\partial i_S} - \frac{\partial i^{*M}}{\partial i_S}} \tag{33}$$

$$\frac{di^*}{dp} = \frac{\partial i^{*I}}{\partial p} + \frac{\partial i^{*I}}{\partial i_S}\frac{\partial i_S}{\partial p}. \tag{34}$$

The denominator of (33) is negative, and equation (26) implies that $\frac{\partial i^{*I}}{\partial i_S}$ is positive. Furthermore, the implicit function theorem implies that

$$\frac{\partial i^{*I}}{\partial p} = -\frac{\frac{\partial \Gamma(i^*, i_S; p)}{\partial p}}{\frac{\partial \Gamma(i^*, i_S; p)}{\partial i^*}} \tag{35}$$

and equation (26) implies that $\frac{\partial \Gamma(i^*, i_S; p)}{\partial i^*}$ is positive.

1. Using (26),

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

   so using (35) $\frac{\partial i^{*I}}{\partial \frac{c_H}{c_L}} < 0$. Since $\frac{\partial i^{*M}}{\partial \frac{c_H}{c_L}} = 0$, using this in (33) and (34) 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, (25) implies that:

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

   and (20) then implies that $\frac{\partial e_S}{\partial \frac{c_H}{c_L}} < 0$.

2. Introducing variable demand Δ leaves equations (26) and (27) unchanged except that equation (21) generalizes to $\Delta f(i_H) = 1$, so

$$\left.\frac{\partial i_H}{\partial \Delta}\right|_{\Delta=1} = -\frac{1}{f'(i_H)} \tag{36}$$

   Therefore

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

so using (35) $\frac{\partial i^{*I}}{\partial \Delta} < 0$. Since $\frac{\partial i^{*M}}{\partial \Delta} = 0$, using this in (33) and (34) 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 high type i^*, given by $u(i^*) = q_H - c_H e_S$, must rise. Equation (23) generalizes to:

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

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

$$\frac{du(i^*)}{d\Delta} = \frac{\partial u(i^*)}{\partial \Delta} + \frac{\partial u(i^*)}{\partial i^*}\frac{\partial i^*}{\partial \Delta} \tag{37}$$

with

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

(using (36) to replace $\frac{\partial i_H}{\partial \Delta}$), and $\partial u(i^*)/\partial i^* = f(i^*)\, \partial \underline{w}(0, i^*)/\partial i^* > 0$. Replacing in (37) and using the assumption that $\frac{du(i^*)}{d\Delta} \geq 0$, this implies

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

which contradicts the first part of the result.

3. 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 (26) and (27). Using (20), (24) and (25):

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

which is increasing in $q_H - q_L$.

□

B Expertise Acquisition

Following the approach in Kurlat (2019), we ask whether firms have the correct incentives to acquire expertise. Consider an individual firm j 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. Denote by θ_j the level of expertise that firm j chooses to acquire.

Let

$$\Pi(\theta_j, F) = \frac{(\lambda - \theta_j) q_L + \pi^P(F)(1 - \lambda) q_H}{\lambda - \theta_j + \pi^P(F)(1 - \lambda)} - w^P \tag{38}$$

denote the individual firm's profits, where we have made explicit that these depend on the firm's choice θ_j and the distribution of expertise of all other firms F, which this firm takes as given. Furthermore, let $W(\theta_j, F)$ denote the equilibrium total payoffs (ignoring their distribution across workers and firms):

$$W(\theta_j, F) = \lambda q_L + (1 - \lambda) q_H - (1 - \lambda)\left(1 - \pi^P(\theta_j, F)\right) c_H e^*. \tag{39}$$

W depends on θ_j because firm j's choice of expertise affects equilibrium allocations.

Assume the firm's cost of acquiring its screening technology is $c_j(\theta_j)$, where $c_j(\cdot)$ is increasing and sufficiently convex such that $\Pi(\theta_j, F) - c_j(\theta_j)$ is concave in θ_j. The function $c_j(\cdot)$ can be different for different firms, leading to different equilibrium expertise choices. Taking the equilibrium as given, firm j will invest until the marginal cost of better screening equals the marginal benefit: $c_j'(\theta_j) = \partial\Pi(\theta_j, F)/\partial\theta_j$. A social planner interested in minimizing deadweight costs would instead want the firm to invest up to the point where $c_j'(\theta_j) = \partial W(\theta_j, F)/\partial\theta_j$. Using the model, we can compute the ratio

$$r(\theta_j) = \frac{\partial W(\theta_j, F)/\partial\theta_j}{\partial\Pi(\theta_j, F)/\partial\theta_j}. \tag{40}$$

If $r(\theta_j) > 1$, the marginal social value of better screening is greater than the marginal cost, which would provide a rationale for subsidizing investments in expertise. Conversely, if $r(\theta_j) < 1$, there would be a case for taxing those investments.

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

$$\eta \equiv \frac{1}{\pi^P} \left.\frac{\partial\pi^P}{\partial\Delta}\right|_{\Delta=1} \tag{41}$$

the elasticity of the share of high types who do not signal (in a partial signaling equilibrium) with respect to an increase in demand.

Proposition 11. *1. The ratio of social to private marginal value of expertise is* $r(\theta_j) = \frac{c_H}{c_L}\eta$.

2. The elasticity η *is greater than* 1.

First, Proposition 11 establishes, perhaps surprisingly, that $r(\theta_j)$ does not depend on θ_j. One might have conjectured that the misalignment of incentives would be different for firms that, e.g. due to different cost functions $c_j(\cdot)$, 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 11 shows that r can be written as the product of the signaling cost ratio and the demand elasticity of π^P. The ratio c_H/c_L enters the formula because, by equation (39), it governs the deadweight cost of signaling for a high type that chooses e^*.

To understand the role of the elasticity of π^P with respect to demand, observe that, again by (39), $\partial W(\theta_j, F)/\partial\theta_j$ crucially depends on how the equilibrium π^P changes in response to an individual firm's screening technology θ_j. 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 5.5 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^P. In particular, we show in the proof of Proposition 11 that the overall effect on π^P 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^P 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^P 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 11 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 $(0, w^P)$ remained constant, each firm in $\left[\theta^P, \lambda\right]$ would hire $\Delta\%$ more high types, implying an elasticity of 1. However, precisely because π^P increases, the mix of workers available in market $(0, w^P)$ improves, so each firm increases its hiring of high types by more than $\Delta\%$. Furthermore, higher π^P means that marginal firms enter market $(0, w^P)$, further increasing demand. The strength of this last effect depends on the density $f\left(\theta^P\right)$ of firms near the cutoff θ^P. 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^P 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 partial signaling equilibrium, improved screening always helps other firms, since it leads more high types to forgo education and improves the mix of workers available at $(0, w^P)$. Second, it is neutral for high-type workers since they get a payoff of w^P 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 $(0, w^P)$ and lose from having more high type workers looking for work in $(0, w^P)$. 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 a partial signaling equilibrium the economy is at a no-signaling equilibrium, 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 $(0, w^N)$ but drives up the wage w^N.

Proof. 1. Using (38) yields

$$\frac{\partial \Pi\left(\theta_j, F\right)}{\partial \theta_j} = \frac{\pi^P (1-\lambda)}{\left[\lambda - \theta_j + \pi^P (1-\lambda)\right]^2} \left(q_H - q_L\right) \tag{42}$$

and using (39)

$$\frac{\partial W\left(\theta_j, F\right)}{\partial \theta_j} = (1-\lambda)\, c_H e^* \frac{\partial \pi^P}{\partial \theta_j}. \tag{43}$$

Around a partial signaling equilibrium, equations (15) and (16) imply

$$\frac{\partial \pi^P}{\partial \theta_j} = \frac{\overbrace{\frac{\left(\lambda - \theta^P\right) + \pi^P (1-\lambda)}{\left[\lambda - \theta_j + \pi^P (1-\lambda)\right]^2 f\left(\theta^P\right)}}^{\text{Sensitivity of market clearing to expertise of one firm}}}{\underbrace{\frac{\int_{\theta^P}^{\lambda} (1-\lambda) \left[(\lambda - \theta) + \pi^P (1-\lambda)\right]^{-2} dF(\theta) \left[\left(\lambda - \theta^P\right) + \pi^P (1-\lambda)\right]}{f\left(\theta^P\right)}}_{\text{Slope of market clearing}} - \underbrace{(1-\lambda) \frac{1}{\frac{c_L}{c_H} - 1}}_{\text{Slope of indifference}}}. \tag{44}$$

Replacing (42), (43) and (44) into (40), we obtain

$$r\left(\theta_j\right) = \frac{c_H/c_L}{\pi^P\left(1-\lambda\right)\left[\int_{\theta^P}^{\lambda}\left[\left(\lambda-\theta\right)+\pi^P\left(1-\lambda\right)\right]^{-2}dF\left(\theta\right)-\frac{f\left(\theta^P\right)}{\pi^P\left(1-\lambda\right)+\lambda-\theta^P}\frac{1}{\frac{c_L}{c_H}-1}\right]}. \tag{45}$$

Applying formula (32) and definition (41) yields

$$\eta = \frac{\left[\lambda-\theta^P+\pi^P\left(1-\lambda\right)\right]/\left[\pi^P f\left(\theta^P\right)\right]}{\int_{\theta^P}^{\lambda}\left(1-\lambda\right)\left[\left(\lambda-\theta\right)+\pi^P\left(1-\lambda\right)\right]^{-2}dF\left(\theta\right)\left[\left(\lambda-\theta^P\right)+\pi^P\left(1-\lambda\right)\right]/f\left(\theta^P\right)-\left(1-\lambda\right)/\left(\frac{c_L}{c_H}-1\right)}. \tag{46}$$

Replacing (46) into equation (45) and simplifying gives the result.

2. Rearranging (46) and using that $f\left(\theta^P\right)\geq 0$:

$$\eta \geq \frac{1}{\pi^P\left(1-\lambda\right)\int_{\theta^P}^{\lambda}\left[\left(\lambda-\theta\right)+\pi^P\left(1-\lambda\right)\right]^{-2}dF\left(\theta\right)} \tag{47}$$

Now rearrange the market clearing condition (16) as

$$\int_{\theta^P}^{\lambda}\left[\pi^P\left(1-\lambda\right)\right]^{-1}\left[\left(\lambda-\theta\right)+\pi^P\left(1-\lambda\right)\right]^{-1}dF\left(\theta\right)=\left[\pi^P\left(1-\lambda\right)\right]^{-1}$$

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

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

Replacing in equation (47) gives the result.

□

C Omitted Proofs

Proof of Proposition 1

1. The proposed $\{e_i, (e_\theta, w_\theta, \chi_\theta), \mu, G\}$ is an equilibrium.

   (9) 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 (7) solve the workers' problem. (10) implies that firms can make zero profits by hiring in market $(0, q_L)$ (where there are only low types) or (e^*, q_H) (where there are only high types), and any other market has either $G\left(I_\chi; e, w, \chi\right) = 0$ or results in losses. Therefore (8), which places demand only in markets $(0, q_L)$ and (e^*, q_H) and yields zero profits, is an optimal choice. Furthermore, (8) implies that no firm hires more than one worker. Replacing (8) in (3) implies that demand is:

$$D(e,w) = \begin{cases} \lambda & \text{if } e=0, w=q_L \\ 1-\lambda & \text{if } e=e^*, w=q_H \\ 0 & \text{otherwise.} \end{cases} \tag{48}$$

Equations (7), (9), (10) and (48) imply that Condition 1 holds. Condition 2 is trivially satisfied because (9) is independent of i. Finally, (7) and (9) imply that beliefs (10) satisfy Condition 3 in nonempty markets. Since low types find $e \in [0, e^*]$ optimal and high types find $e \geq e^*$ optimal, (9) implies that beliefs satisfy Condition 4 when they are well defined, and $G(I_\chi; e, w, \chi) = 0$ only at wages where $\mu(w; e, i) = 0$ for all i, so Condition 5 is satisfied as well.

2. The above equilibrium is unique.

   (a) In any equilibrium, each firm makes zero profits. If there was a firm that made strictly negative profits, it could increase profits by setting $\chi(i) = 0$ for all i. On the other hand, suppose there is a firm that makes strictly positive profits in some market (e, w). Recall that $F(1) > 1$, so there must exist a strictly positive measure of firms that do not hire. Any such firm could increase its profits by directing its search to market (e, w), so it cannot be optimizing.

   (b) In any equilibrium, there does not exist a market (e, w) such that $\mathbb{I}(e_i = e)\,\mu(w; e, i) > 0$ both for some $i < \lambda$ and some $i' \geq \lambda$. Otherwise, consider a market (e', w') with $e' = e + \epsilon$ and $w \in (\bar{w}(e, i') + c_H\epsilon, q_H)$. Suppose type $i < \lambda$ is in the support of $G(\cdot; e', w', \chi)$. This requires

   $$\bar{w}\left(e', i\right) - c_L e' \geq \bar{w}(e, i) - c_L e.$$

   Rearranging gives $\bar{w}(e', i) - \bar{w}(e, i) \geq c_L\epsilon$. Since firms cannot discriminate, it follows that $\bar{w}(e, i)$ is the same for all i. Also, since $\mu(w'; e, i)$ must be weakly decreasing in i by Condition 2, $\bar{w}(e', i)$ is weakly increasing in i. Therefore:

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

   which implies $\bar{w}(e', i') - \bar{w}(e, i') > c_H\epsilon$. This contradicts the premise that type i' finds e optimal. Hence, no $i < \lambda$ can be in the support of $G(\cdot; e', w', \chi)$. If the support of $G(\cdot; e', w', \chi)$ only includes $i \geq \lambda$, then firms could make profits by hiring in market (e', w'), which contradicts part (2a). Therefore it must be that $G(I_\chi; e', w', \chi) = 0$. This implies that $\mu(w'; e', i') = 0$ by Condition 5, which in turn implies $\bar{w}(e', i') > \bar{w}(e, i') + c_H\epsilon$, which contradicts the premise that type i' finds e optimal.

   (c) In any equilibrium, all low types obtain a payoff of q_L. Suppose that they obtain a payoff $q_L' > q_L$. This implies that they are hired with positive probability in a market with $w > q_L$. By part (2b), the supply in this market only includes low types, which implies negative profits for firms, contradicting part (2a). Suppose that they obtain a payoff $q_L' < q_L$ and consider a market with $e = 0$ and $w \in (q_L', q_L)$. If $G(I_\chi; e, w, \chi) > 0$, then firms can make profits by hiring in this market; otherwise, $\mu(w; e, i) = 0$, which means low-type workers can obtain a payoff $w > q_L'$ by choosing $e = 0$.

   (d) In any equilibrium, all high types obtain payoff $q_H - c_H e^*$. Suppose first that they obtain a payoff $u_H > q_H - c_H e^*$. If they do so by selecting $e' < e^*$, then this implies $\bar{w}(e', i) - c_H e' > q_H - c_H e^*$, which in turn 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, contradicting part (2c). 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 part (2a). Second, suppose they obtain 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 with $e = e^*$ and $w \in (u_H + c_H e^*, q_H)$. $G(I_\chi; e^*, w, \chi) > 0$ because otherwise $\mu(w; e^*, i) = 0$

by Condition 5, so high types can obtain a payoff of at least $w - c_H e^* > u_H$ by choosing education e^*. But the support of $G(\cdot; e^*, w, \chi)$ cannot include low types because choosing e^* implies a payoff of $\bar{w}(e^*, i) - c_L e^* < q_H - c_L e^* < q_L$, contradicting part (2c); and the support of $G(\cdot; e^*, w, \chi)$ cannot include only high types because then firms could make profits by hiring in market (e^*, w), contradicting part (2a).

(e) Step (2c) implies that all low types select $e = 0$ and get hired for sure in market $(0, q_L)$. Step (2d) implies that all high types must select $e = e^*$ and get hired for sure in market (e^*, q_H). This determines (7) as well as (9) and (10) in these markets. It also requires that there is total demand λ in market $(0, q_L)$ and demand $1 - \lambda$ in (e^*, q_H), thus (8) must hold. For all other markets, (9) and (10) then follow from Conditions 1 to 5.

Proof of Proposition 3

We first show that, in any equilibrium, all low types choose $e = 0$ and get hired at least at wage $w = q_L$. Some fraction $\pi \in [0, 1]$ of the high types choose $e = 0$ and find a job for sure at wage $w = w^P$ (if $\pi < 1$) or $w \geq w^P$ (if $\pi = 1$). The rest of the high types choose $e = e^*$ and get hired with certainty at wage $w = q_H$. We prove this claim based on the following sequence of steps:

1. By the same argument as in the proof of Proposition 1, all firms make non-negative profits.

2. 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 consider the market and hiring rule $(e_\theta, w_\theta, \chi_\theta)$ chosen by firm θ. By hiring in market (e_θ, w_θ) and setting $\chi_{\theta'}(i) = \mathbb{I}(i \geq \theta')$, firm θ' could make profits at least as high as firm θ since it accepts all the high types but rejects more low types than firm θ possibly can.

3. 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 do not hire, which implies zero profits. The claim then follows from the monotonicity of profits in θ.

4. In any equilibrium, low types obtain a payoff of at least q_L. Suppose that they obtain a payoff $q'_L < q_L$ and consider a market with $e = 0$ and $w \in (q'_L, q_L)$. If $G(I_\chi; 0, w, \chi) > 0$, then firms $\theta < \bar{\theta}$ can make profits by hiring in this market; otherwise, $\mu(w; 0, i) = 0$ by Condition 1, which means low-type workers can obtain a payoff $w > q'_L$ by choosing $e = 0$.

5. In any equilibrium, high types obtain a payoff of at least $w^P = q_H - c_H e^*$. Suppose they obtain 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 with $e = e^*$ and $w \in (u_H + c_H e^*, q_H)$. $G(I_{\chi_\theta}; e^*, w, \chi_\theta) > 0$ for all θ because otherwise $\mu(w; e^*, \lambda) = 0$ by Condition 5, so high types can obtain a payoff of at least $w - c_H e^* > u_H$ by choosing education e^*. But the support of $G(\cdot; e^*, w, \chi_\theta)$ 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 $G(\cdot; e^*, w, \chi_\theta)$ cannot include only high types for $\theta < \bar{\theta}$ because then firms $\theta < \bar{\theta}$ could make profits by hiring in market (e^*, w).

6. For any $i \geq \lambda$ and any e, $\mu(\cdot; e, i)$ has a point mass at a single wage. To see this, consider two wage levels $w' > w$ and suppose that high types are hired with positive probability in both of them if they choose e. Let θ' be the highest-type firm that hires at wage w'. Conditions 1 and 3 imply that the expected productivity of workers that firm θ' will find in markets (e, w') and (e, w) is the same, and therefore it cannot be optimal for firm θ' to hire at wage w'. Therefore it must be that all high types are hired at the same wage, which implies that $\mu(\cdot; e, i)$ is a step function for every i.

7. In any equilibrium, all low types get education $e = 0$. To see this, assume to the contrary that some $i < \lambda$ chooses $e = \tilde{e} > 0$. By step (4), we have that $\bar{w}(\tilde{e}, i) \geq q_L + c_L\tilde{e} > q_L$. Together with step (1), this implies that in every market $(w, \tilde{e})$ with $w > q_L$ where type i has some chance of being hired, there are also high-type applicants, because otherwise firms would make losses by paying more than q_L. Step (6) implies that there can be only one such market; label it $(\tilde{e}, \tilde{w})$. Letting u_H be the utility obtained by high types in equilibrium, this implies $\tilde{w} = u_H + c_H\tilde{e}$. Let $\underline{\theta}$ be the lowest firm type that hires in market $(\tilde{e}, \tilde{w})$ and π_H be the measure of high types that choose $e = \tilde{e}$. Using the fact that all high types that choose $e = \tilde{e}$ are hired in market $(\tilde{e}, \tilde{w})$, the probability that type $i < \lambda$ is hired in market $(\tilde{e}, \tilde{w})$ is bounded above by $\int_{\underline{\theta}}^{i} \frac{1}{\pi_H} dF(\theta)$. Since not being hired in market $(\tilde{e}, \tilde{w})$ implies getting a wage q_L, this implies that the payoff from choosing $e = \tilde{e}$ is bounded above by:

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

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 $e = \tilde{e}$. We know that $\underline{i} > \underline{\theta}$.

Assume that some firm $\tilde{\theta} \in [\underline{\theta}, \underline{i}]$ prefers to hire in some market $(e'', w'') \neq (\tilde{e}, \tilde{w})$. This implies firms $\theta \in \left(\tilde{\theta}, \underline{i}\right)$ also prefer (e'', w'') over $(\tilde{e}, \tilde{w})$, since they hire from the same pool of workers as firm $\tilde{\theta}$ in market $(\tilde{e}, \tilde{w})$ but from a more selected pool in other markets. But then the fact that worker $i = \tilde{\theta}$ does not choose $\tilde{e}$ implies that worker $\underline{i}$ does not want to choose $\tilde{e}$ either, since he obtains the same payoff as worker $i = \tilde{\theta}$ upon choosing $\tilde{e}$ but weakly higher in every other market. This contradicts the assumption that worker $\underline{i}$ chooses $\tilde{e}$. Therefore it must be that all firms in the interval $[\underline{\theta}, \underline{i}]$ hire in market $(\tilde{e}, \tilde{w})$.

Since there are no workers with $i < \underline{i}$ in market $(\tilde{e}, \tilde{w})$, then upon hiring in market $(\tilde{e}, \tilde{w})$, 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 (3) implies that all $\theta \leq \underline{i}$ firms must make zero profits by hiring in market $(\tilde{e}, \tilde{w})$. 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 $(\tilde{e}, \tilde{w})$. Therefore, all firms in the interval $(\underline{i}, \underline{i} + \delta_1]$ hire in market $(\tilde{e}, \tilde{w})$. This in turn implies that if worker $\underline{i}$ is willing to choose $\tilde{e}$, then worker $\underline{i} + \delta_1$ *strictly* prefers $\tilde{e}$, since, compared to worker $\underline{i}$, he has a higher chance of being hired in market $(\tilde{e}, \tilde{w})$ 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 $e = \tilde{e}$. 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 $\tilde{e}$. Therefore all workers $i \in [\underline{i}, \lambda]$ choose $e = \tilde{e}$.

Let $\hat{\theta}$ be the highest firm type that hires in market $(\tilde{e}, \tilde{w})$.

(a) Assume $\hat{\theta} = \lambda$. Consider market $(0, u_H + \epsilon)$. If $G\left(I_{\chi_{\hat{\theta}}}; 0, u_H + \epsilon, \chi_{\hat{\theta}}\right) > 0$, 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 $(0, u_H + \epsilon)$ than in market $(\tilde{e}, \tilde{w})$, a contradiction. Instead, if $G\left(I_{\chi_{\hat{\theta}}}; 0, u_H + \epsilon, \chi_{\hat{\theta}}\right) = 0$, this requires $\mu(u_H + \epsilon; 0, i) = 0$ for all $i \geq \lambda$, which implies that $e = 0$ is a better choice than $\tilde{e}$ for high types, again a contradiction.

(b) Assume instead that $\hat{\theta} \leq \lambda$. This implies that all firms that hire in market $(\tilde{e}, \tilde{w})$ accept workers $i \in \left[\hat{\theta}, \lambda\right]$. Since high-type workers are hired for sure in market $(\tilde{e}, \tilde{w})$, this implies that workers $i \in \left[\hat{\theta}, \lambda\right]$ are hired for sure as well, and therefore obtain utility $u_H - (c_L - c_H)\,\tilde{e}$. Now consider a market with $e' = \tilde{e} + \epsilon$ and $w' \in (\tilde{w} + c_H\epsilon, \tilde{w} + c_L\epsilon)$. Suppose type $i' \in \left[\hat{\theta}, \lambda\right]$ is in the support of $G\left(\cdot; e', w', \chi_\theta\right)$. This requires:

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

Since $\bar{w}\left(e', i\right)$ must be increasing in i by Condition 2, for any high-type worker i'':

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

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

8. In any equilibrium, the high types select either $e = 0$ or $e = e^*$. If some types selected $e > e^*$, then by step (5) this would require paying them $w > q_H$ and therefore involve negative profits for firms. On the other hand, suppose some high type i' sets $e' \in (0, e^*)$ and let $\underline{w}$ be the highest wage such that $\mu\left(\underline{w}; e', i'\right) = 0$. For any $w \geq \underline{w}$, beliefs can only place weight on high types since by part (7), 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 (1). 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 $e_i = e'$ 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 $(0, w')$ and therefore obtain a payoff of q_L. Let $\Theta \subseteq [0, \theta']$ be the set of firms who hire workers in market (e', 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 $e' < e^*$), $(1 - \mu\left(q_H; e', i'\right))q_H - c_L e' > q_L$. Thus, there is a low type who would prefer $e = e'$ to $e = 0$, which contradicts step (7).

To complete the proof, let u_H be the equilibrium payoff of high types.

1. If $u_H > w^P$, 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 (12) holds. Furthermore, since all high types must be hired at this wage, (11) must hold.

2. If $u_H = w^P$, then high types are indifferent between choosing $e = 0$ and getting hired at wage w^P 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 $\left(0, w^P\right)$ iff

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

This defines the cutoff $\bar{\theta}$, so (15) holds. Furthermore, since all high types who choose $e = 0$ must be hired at w, (16) must hold too.

Proof of Proposition 4

Partial Signaling Equilibrium.

1. Necessity of condition (19).

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

$$\tilde{w}\,(e,i)=\begin{cases} u\,(i)+c_L e & \text{if } i<\lambda \\ u\,(i)+c_H e & \text{if } i\geq\lambda, \end{cases} \tag{49}$$

   where $u(i)$ is given by (17). Suppose firm θ considers hiring in market (e, w). For it to believe that it will find χ_θ-acceptable low types, i.e. workers with $i \in [\theta, \lambda)$, it must be that:

$$\tilde{w}\,(e,\theta)\leq\tilde{w}\,(e,i)=\bar{w}\,(e,i)\leq\bar{w}\,(e,\lambda)\leq\tilde{w}\,(e,\lambda)\,. \tag{50}$$

   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 Condition 4: if beliefs place weight on type i, then i must be indifferent between e and his equilibrium choice. The third follows from Condition 2, which implies that $\bar{w}$ is monotonic in i. The last inequality follows from the fact that otherwise worker λ could exceed his equilibrium payoff by choosing e. By Condition 4, 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.

   Moreover, for firm θ not to have well-defined beliefs about market (e, w) it must be that:

$$w\leq\tilde{w}\,(e,\theta)\,, \tag{51}$$

   since otherwise Condition 5 requires $\mu\,(w; e, \theta) = 0$, so some χ_θ-acceptable worker could exceed his equilibrium payoff by choosing e.

   Together, conditions (50) and (51) imply that for any market (e, w) such that $\tilde{w}\,(e, \lambda) < \tilde{w}\,(e, \theta)$ and $w > \tilde{w}\,(e, \lambda)$, firm θ's beliefs $G\,(\cdot; e, w, \chi_\theta)$ can only place weight on high types.

   Denote by (e^D_θ, w^D_θ) the lowest-wage market where firm θ's beliefs are guaranteed to only include high types, which satisfies

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

Using (6), (14), (17) and (49) and rearranging, the profits that firm θ can obtain by hiring in market (e_θ^D, w_θ^D) are

$$\Pi^D(\theta) = q_H - w_\theta^D = \frac{c_H}{c_L} \int_{\theta^P}^{\theta} \frac{1}{\pi^P(1-\lambda) + \lambda - \theta} dF(\theta)\,(q_H - q_L).$$

By (38), profits in market (e_θ^D, w_θ^D) exceed those that firm θ obtains in equilibrium if condition (19) is violated, which implies it cannot be an equilibrium.

2. Sufficiency of condition (19). We construct the equilibrium objects $\{e_i, (e_\theta, w_\theta, \chi_\theta), \mu, G\}$.

   (a) Worker decisions:

$$e_i = \begin{cases} 0 & \text{if } i < \lambda + \pi^P(1-\lambda) \\ e^* & \text{if } i \geq \lambda + \pi^P(1-\lambda) \end{cases} \tag{52}$$

   (b) Firm decisions:

$$(e_\theta, w_\theta, \chi_\theta) = \begin{cases} (0, w^P, \chi(i) = \mathbb{I}(i \geq \theta)) & \text{for } \theta \geq \theta^P \\ (0, q_L, \chi(i) = 1 \forall i) & \text{for a measure } \lambda - \varphi^P \text{ of firms } \theta < \theta^P \\ (e^*, q_H, \chi(i) = 1 \forall i) & \text{for a measure } (1-\lambda)\left(1-\pi^P\right) \text{ of firms } \theta < \theta^P \\ (0, 0, \chi(i) = 1 \forall i) & \text{otherwise} \end{cases} \tag{53}$$

   where

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

   (c) Probabilities:

$$\mu(w; e, i) = \begin{cases} 1 & \text{if } e = 0,\ w \geq w^P \\ 1 - \int_{\theta^P}^{\min\{i,\lambda\}} \frac{1}{\lambda - \theta + \pi^P(1-\lambda)} dF(\theta) & \text{if } e = 0,\ w^P > w \geq q_L \\ \mathbb{I}\left(w \geq \min\left\{\tilde{w}(e,i), \tilde{w}(e,\lambda)\right\}\right) & \text{otherwise} \end{cases} \tag{54}$$

   where we used (49).

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

$$g(i; e, w, \chi) = \begin{cases} \frac{\mathbb{I}(i \in [\theta, \lambda)) + \pi^P \mathbb{I}(i \geq \lambda)}{\pi^P(1-\lambda) + \lambda - \theta} & \text{if } e = 0, w \geq w^P \\ \frac{\mathbb{I}(i \in [\theta, \lambda])\mu(w; e, i)}{\int_\theta^\lambda \mu(w; e, i) di} & \text{if } e = 0, w^P > w \geq q_L \\ \mathbb{I}(i = \theta) & \text{if } e \in \left(0, e_\theta^D\right), w \geq \tilde{w}(e, \theta) \\ \frac{\mathbb{I}(i \geq \lambda)}{1-\lambda} & \text{if } e \geq e_\theta^D, w \geq \tilde{w}(e, \lambda) \\ 0 & \text{for any other } (e, w) \end{cases} \tag{55}$$

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

$$g\left(i; e, w, \chi\right) = \begin{cases} \frac{\mathbb{I}(i<\lambda)+\pi^P \mathbb{I}(i\geq\lambda)}{\pi^P(1-\lambda)+\lambda} & \text{if } e=0, w \geq w^P \\ \frac{\mu(w;e,i)}{\int_0^1 \mu(w;e,i)di} & \text{if } e=0, w^P > w \geq q_L \\ \frac{\mathbb{I}\left(i<\theta^P\right)}{\theta^P} & \text{if } e \in \left(0, e^*\right), w \geq \tilde{w}\left(e, 0\right) \\ \frac{\mathbb{I}(i\geq\lambda)}{1-\lambda} & \text{if } e \geq e^*, w \geq \tilde{w}\left(e, \lambda\right) \\ 0 & \text{for any other } (e, w) \end{cases} \tag{56}$$

We now verify that $\{e_i, (e_\theta, w_\theta, \chi_\theta), \mu, G\}$ satisfies all the equilibrium conditions from Definition 1. (54) implies that low types $i \in [0, \lambda)$ are indifferent between any $e \in \left[0, e_i^D)\right]$ and high types are indifferent between any $e \geq 0$, so the education decisions (52) solve the workers' problem. The beliefs (55) and (56) together with the fact that condition (19) holds implies that firms $\theta \geq \theta^P$ maximize profits by hiring selectively in market $(0, w^P)$. All other firms make zero profits by hiring non-selectively either in market $(0, q_L)$ or (e^*, q_H), and any other market has either $G\left(I_{\chi_\theta}; e, w, \chi_\theta\right) = 0$ or results in losses. Therefore the demands (53) are an optimal choice. Furthermore, replacing (53) in (3) implies that demand in market (e, w) for a set of selection rules $X_0\left(\theta'\right) = \{\chi\left(i\right) = \mathbb{I}\left(i \geq \theta\right) : \ \theta \in [0, \theta']\}$ is:

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

Together with (55) and (56), this implies that Condition 1 holds. Condition 2 is satisfied because, by (54), $\mu\left(\cdot; e, i\right)$ is weakly decreasing in i. Finally, (52) and (54) imply that beliefs (55) and (56) satisfy Condition 3 in nonempty markets. Since low types i find $e \in \left[0, e_i^D)\right]$ optimal and high types find any $e \geq 0$ optimal, beliefs satisfy Condition 4 when they are well defined, and $G(I_\chi; e, w, \chi) = 0$ only at wages where $\mu\left(w; e, i\right) = 0$ for all i such that $\chi(i) = 1$, so Condition 5 is satisfied as well.

Pure Signaling Equilibrium. The above analysis applies for the special case with $\pi^P = 0$.

No-Signaling Equilibrium. Necessity and sufficiency of condition (18) are proved by the same steps as for the Partial Signaling Equilibrium. For completeness, we state the equilibrium objects $\{e_i, (e_\theta, w_\theta, \chi_\theta), \mu, G\}$.

(a) Worker decisions:

$$e_i = 0 \ \forall i \tag{57}$$

(b) Firm decisions:

$$\left(e_\theta, w_\theta, \chi_\theta\right) = \begin{cases} (0, w^N, \chi(i) = \mathbb{I}(i \geq \theta)) & \text{for } \theta \geq \theta^N \\ (0, q_L, \chi(i) = 1 \forall i) & \text{for a measure } 1 - F(\lambda) + F(\theta^N) \text{ of firms } \theta < \theta^N \\ (0, 0, \chi(i) = 1 \forall i) & \text{otherwise} \end{cases} \tag{58}$$

(c) Probabilities:

$$\mu(w;e,i)=\begin{cases} 1 & \text{if } e=0,\ w\geq w^N \\ 1-\int_{\theta^N}^{\min\{i,\lambda\}}\frac{1}{1-\theta}dF(\theta) & \text{if } e=0,\ w^N>w\geq q_L \\ \mathbb{I}\left(w\geq \min\left\{\tilde{w}(e,i),\tilde{w}(e,\lambda)\right\}\right) & \text{otherwise}\end{cases} \tag{59}$$

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

$$g(i;e,w,\chi)=\begin{cases} \frac{\mathbb{I}(i\geq\theta)}{1-\theta} & \text{if } e=0, w\geq w^N \\ \frac{\mathbb{I}(i\in[\theta,\lambda])\mu(w;e,i)}{\int_\theta^\lambda \mu(w;e,i)di} & \text{if } e=0, w^N>w\geq q_L \\ \mathbb{I}(i=\theta) & \text{if } e\in\left(0,e_\theta^D\right), w\geq\tilde{w}(e,\theta) \\ \frac{\mathbb{I}(i\geq\lambda)}{1-\lambda} & \text{if } e\geq e_\theta^D, w\geq\tilde{w}(e,\lambda) \\ 0 & \text{for any other } (e,w)\end{cases} \tag{60}$$

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

$$g(i;e,w,\chi)=\begin{cases} 1\forall i & \text{if } e=0, w\geq w^N \\ \frac{\mu(w;e,i)}{\int_0^1 \mu(w;e,i)di} & \text{if } e=0, w^N>w\geq q_L \\ \frac{\mathbb{I}\left(i<\theta^N\right)}{\theta^N} & \text{if } e\in\left(0,e^N\right), w\geq\tilde{w}(e,0) \\ \frac{\mathbb{I}(i\geq\lambda)}{1-\lambda} & \text{if } e\geq e^N, w\geq\tilde{w}(e,\lambda) \\ 0 & \text{for any other } (e,w)\end{cases} \tag{61}$$

with $e^N=(w^N-q_L)/(c_L-c_H)$

Proof of Proposition 5

1. Using the reparametrization of the model in terms of $\hat{\lambda}$, equation (12) generalizes to

$$w^N=\frac{\left(\lambda-\theta^N\right)\frac{\hat{\lambda}}{\lambda}q_L+\left(1-\hat{\lambda}\right)q_H}{\left(\lambda-\theta^N\right)\frac{\hat{\lambda}}{\lambda}+\left(1-\hat{\lambda}\right)}. \tag{62}$$

For $\hat{\lambda}$ low enough, $w^N>w^P$ so there is a candidate corner equilibrium. Condition (18) generalizes to

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

which cannot hold for sufficiently low $\hat{\lambda}$, so the candidate equilibrium is indeed an equilibrium. Furthermore, taking the limit in (62) we obtain $\lim_{\hat{\lambda}\to 0}w^N=q_H$.

2. Equation (11) implies that $\lim_{F\to F^*}\theta^N=\lambda$, which implies, using (12), that $\lim_{F\to F^*}w^N=q_H$ for F sufficiently close to F^*, so a candidate equilibrium with the desired properties exists. Furthermore, as $\theta^N\to\lambda$, condition (18) cannot hold so the candidate equilibrium is indeed an equilibrium.

Proof of Proposition 6

1. Using (14) and (6):
$$w^P = \left(1 - \frac{c_H}{c_L}\right) q_H + \frac{c_H}{c_L} q_L,$$
and therefore $\lim_{\frac{c_H}{c_L}\to 1} w^P = q_L$. Using (12) we have $w^N > w^P$, so there is a candidate corner equilibrium. Furthermore, as $c_H/c_L \to 1$, condition (18) holds so the candidate equilibrium is indeed an equilibrium.

2. Taking the limit, $\lim_{\frac{c_H}{c_L}\to 0} w^P = q_H$. Using (15), this implies $\theta^P \to \lambda$, so condition (16) cannot hold for any $\pi^P > 0$.

Proof of Proposition 7

1. It is sufficient to prove claim (b) because claim (a) is a special case with $\pi_2^P = \theta_2^P = 0$. By equation (38), firm profits are increasing in π^P, and since w^P is the same across equilibria, firms are better off in the higher-π^P equilibrium. High-type workers obtain a payoff of w^P in both equilibria, so they are indifferent. Using (17), workers with $i \leq \theta_2^P$ get a payoff of q_L in both equilibria, so they are also indifferent. Workers with $i \in (\theta_2^P, \theta_1^P]$ 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 $i \in \left(\theta_1^P, \lambda\right)$, their payoff is:
$$\begin{aligned} u(i) &= q_L + \int_{\theta^P}^{i} \frac{1}{\pi^P(1-\lambda)+\lambda-\theta} dF(\theta)\left(w^P - q_L\right) \\ &= w^P - \int_{i}^{\lambda} \frac{1}{\pi^P(1-\lambda)+\lambda-\theta} dF(\theta)\left(w^P - q_L\right) \end{aligned}$$
where we used (16). This is increasing in π^P, so they are also better off in the second equilibrium.

2. (a) 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 $i > \theta^N$ get more in the second equilibrium. High-type workers get a payoff of w^P in the first equilibrium but w^N in the second, so they are also better off.

   (b) By equation (38), for θ sufficiently close to λ, firm θ's profits approach $q_H - w$, so $w^P < w^N$ implies they are higher in the first equilibrium. High-type workers get a payoff of w^P in the first equilibrium but w^N in the second, so they are better off in the second.

D False Negatives

Uniqueness in case $f(\theta)$ is strictly increasing

Proposition 12. *If condition (28) holds, the system of equations (26), (27) has no solution. Otherwise, it has a unique solution.*

Proof. Solving (27) for i_S and replacing in (26), a solution requires:

$$\Delta\left(i^*\right) \equiv f\left(i^*\right)\left[\frac{c_H}{c_L} - \lambda\left(\frac{F\left(i^*\right)}{f\left(i^*\right)} + \lambda\right)^{-1}\right] - \left(1 - \frac{c_H}{c_L}\right)\int_{i^*}^{i^H} \frac{\lambda}{i - i^* + \frac{F(i^*)}{f(i^*)} + \lambda} df(i) = 0 \tag{63}$$

Taking the derivative and rearranging:

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

where the inequality follows because $i > i^*$. If i^* satisfies (63), then:

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

so the function $\Delta\left(i^*\right)$ is increasing at any i^* such that $\Delta\left(i^*\right) = 0$. Condition (28) is equivalent to $\Delta\left(i_H\right) < 0$. 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 (28) holds, there can be no $i^* \in [\lambda, i_H]$ that satisfies $\Delta\left(i^*\right) = 0$ because $\Delta(\lambda) < 0$ and $\Delta\left(i_H\right) < 0$ and Δ must be increasing at any solution. Instead, if condition (28) does not hold, $\Delta\left(i_H\right) \geq 0$, 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 $\Delta\left(i^*\right) = 0$. □

Continuity in the Limit

Proposition 13. *1.* $\lim_{\hat{\lambda}\to 0}\underline{w}(0,i) = q_H$, $\lim_{\hat{\lambda}\to 0} i^* = \lambda$, *and* $\lim_{\hat{\lambda}\to 0} i_S = \lambda$, *for all* $i \in [i^*, i_H)$.

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

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

Proof.

1. Letting the fraction of low types be $\hat{\lambda}$, equations (21), (22) and (26) generalize, respectively, to:

$$f\left(i^{H}\right)=\frac{1-\hat{\lambda}}{1-\lambda}, \qquad \underline{w}\left(0,i\right)=\frac{\frac{1-\hat{\lambda}}{1-\lambda}\left(i-i_S\right)q_H+\hat{\lambda}q_L}{\frac{1-\hat{\lambda}}{1-\lambda}\left(i-i_S\right)+\hat{\lambda}}, \text{ and} \tag{64}$$

$$\Gamma\left(i^*,i_S\right)=\frac{c_H}{c_L}f\left(i^*\right)-\hat{\lambda}\left(1-\frac{c_H}{c_L}\right)\int_{i^*}^{i_H}\left[\frac{1}{\frac{1-\hat{\lambda}}{1-\lambda}\left(i-i_S\right)+\hat{\lambda}}\right]df\left(i'\right)-f\left(i^*\right)\hat{\lambda}\frac{1}{\frac{1-\hat{\lambda}}{1-\lambda}\left(i^*-i_S\right)+\hat{\lambda}}=0 \tag{65}$$

while the market clearing condition is unchanged. The first statement follows directly from equation (64), the second from equation (65) and the last from equation (69).

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

3. Equation (26) implies $\lim_{\frac{c_H}{c_L}\to 1} i^*-i_S=0$. Using this and equation (27), we have $\lim_{\frac{c_H}{c_L}\to 1} F\left(i^*\right)=0$, which implies $\lim_{\frac{c_H}{c_L}\to 1} i^*=\lambda$ and therefore $\lim_{\frac{c_H}{c_L}\to 1} i_S=\lambda$.

□

By part 1, as the fraction of low types goes to zero, nobody signals and everyone is paid q_H. Part 2 says that as firms become fully informed, again nobody signals and all high types are paid q_H. Finally, part 3 shows that if signaling is sufficiently expensive, no workers signal in equilibrium. In all cases the equilibrium allocations are continuous in the limit.

General Case

For the case where $f\left(i\right)$ is not monotone, the argument in Section 6 needs to be modified. Consider two workers, i and i' with $i^*<i<i'<i_H$ and assume $f\left(i\right)>f\left(i'\right)$. The argument above, unmodified, implies that any worker will be able to sell a fraction $1-f\left(i'\right)$ of his labor to non-selective firms at wages above $\underline{w}\left(0,i'\right)$. This means that only $f\left(i'\right)$ of i-workers will be available for hire in market $\left(0,\underline{w}\left(0,i\right)\right)$, which is less than the $f\left(i\right)$ 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 $\bar{F}(\theta)\leq F(\theta)$:

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

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 extends to this general case, replacing F with $\bar{F}$. Let i_H be defined as

$$i_H\equiv\min_{i\in[\lambda,1]}\left\{i:\bar{f}(i)\geq 1\right\}.$$

This generalizes the definition of i_H in (21), 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

$$\underline{w}(0,i)=\max_{i'}\frac{(i'-i_S)q_H+\lambda q_L}{i'-i_S+\lambda} \quad s.t. \quad \bar{f}\left(i'\right)=\bar{f}\left(i\right). \tag{66}$$

Hence, when $\bar{f}$ is strictly increasing, this coincides with (22), but in a flat region (due to ironing), $\underline{w}(0, i)$ equals the value for 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 remaining workers who are not hired non-selectively at higher wages are hired at the same wage $\bar{w}(0, i_1)$ by firms $\theta \in [i_0, i_1]$. Based on the same steps as underlying (26) but using (66) instead of (22), we obtain

$$\Gamma(i, i_S) = \bar{f}(i) \left(\frac{c_H}{c_L} - \frac{\lambda}{i^b(i) - i_S + \lambda} \right) - \lambda \left(1 - \frac{c_H}{c_L} \right) \int_i^{i_H} \frac{1}{i' - i_S + \lambda} d\bar{f}(i') \tag{67}$$

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

$$i^* = \min_{i \in [\lambda, 1]} \{ i : \Gamma(i, i_S) \geq 0 \} \tag{68}$$

and

$$\bar{F}(i^*) = \bar{f}(i^*)(i^* - i_S). \tag{69}$$

Equation (68) generalizes the indifference condition (26) to account for the fact that, with bunching, the reservation wage function (66) and hence $\Gamma(i, i_S)$ can be discontinuous in i. Note that, by (68), 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, of which Proposition 8 in Section 6 is a special case.

Proposition 14. *There exists a generically unique equilibrium:*

1. *All low types $i \in [0, \lambda)$ choose $e = 0$.*
2. *All high types $i \in [i_H, 1]$ choose $e = 0$.*
3. *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 (68) and (69) and:*
      i. *A measure $i_S - \lambda$ of high types with $i \in [\lambda, i^*)$ choose $e = e_S$.*
      ii. *All other high types with $i \in [\lambda, i^*)$ choose $e = 0$.*
   (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 $e = 0$.*
      ii. *All other high types with $i \in [\lambda, i_H)$ choose $e = e^*$.*

The corner equilibrium is of the same form as described in Section 6. 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 are hired at least at 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. The wages at which workers are hired and all firms' decisions are specified in the proof below.

When there is bunching at the bottom (i.e. on the interval $[\lambda, i^*]$), the market clearing condition (69) implies $i_S = \lambda$, so there is no signaling whatsoever in equilibrium. This occurs when there is a high density of precisely informed buyers relative to less informed ones.

We now provide a proof of Proposition 14, establishing first the uniqueness and then the existence of the stated equilibrium.

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

1. By the same arguments as in the proof of Proposition 3:
   (a) all firms make non-negative profits
   (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 $w^P = q_H - c_H e^*$.
2. Because all low types are indistinguishable for all firms, all low types must obtain the same utility. Denote this by u_L.
3. Utility for workers is weakly increasing in i. This follows immediately from Condition 2.
4. In any equilibrium, all low types choose $e = 0$. Suppose some low types choose $e' > 0$. By step (1d), we have $\bar{w}(e', i) \geq q_L + c_L e'$ for all $i < \lambda$. Consider all markets with $e = e'$ and $w > q_L$ where low types are hired with positive probability. For low types to be hired, in any such market there must be firms that hire non-selectively, setting $\chi(i) = 1$ for all i. By step (1a), there must be high-type applicants in all these markets.
   Let w' be the highest wage where anyone choosing e' is hired with positive probability. Suppose first that some high types $i' \geq \lambda$ are hired in market (e', w') 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 $(0, \tilde{w})$ with $\tilde{w} \in (w' - c_H e', w')$. Then $G(I_{\chi_\theta}; 0, \tilde{w}, \chi_\theta) > 0$ since otherwise $\mu(\tilde{w}; 0, i') = 0$ by Condition 5, so type i' could obtain a payoff of at least $\tilde{w} > u'_H$ by choosing $e = 0$. The support of $G(\cdot; 0, \tilde{w}, \chi_\theta)$ 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 $(0, \tilde{w})$ instead of market (e', w') at wage $\tilde{w} < w'$. Hence, everyone in market (e', w') must be hired by non-selective firms. Since this is feasible for any firm and by step (1a), all firms must make zero profits in market (e', w'). This implies $w' < q_H$.
   Because all firms hire non-selectively in market (e', w'), $\mu(w'; e', i) = \mu'$ is the same for all i. Suppose first that $\mu' > 0$. Consider a market $(e', w' - \epsilon)$. Then for sufficiently small $\epsilon > 0$, the applicant pool is the same in markets $(e', w' - \epsilon)$ and (e', w'). Since profits are zero in market (e', w'), all firms could make positive profits by hiring in market $(e', w' - \epsilon)$, contradicting step (1c). 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' is $u'_H = w' - c_H e'$. Consider a market (e'', w'') such that $e'' = e' + \epsilon$ and $w'' \in (w' + c_H \epsilon, w' + c_L \epsilon)$. Suppose $\chi(i) = 1$ for all i. Then $G(I_\chi; e'', w'', \chi) > 0$ since otherwise $\mu(w''; e'', i) = 0$ for all i, so all high types who choose e' could obtain payoff $w'' - c_H e'' > w' - c_H e' = u'_H$, a contradiction. The support of $G(\cdot; e'', w'', \chi)$ cannot include low types since $w'' - c_L e'' < w' - c_L e' = u'_L$. The support of $G(\cdot; e'', w'', \chi)$ cannot include only high types since then any firm $\theta > \bar{\theta}$ could make strictly positive profits in market (e'', w'') for $\epsilon \in (0, (q_H - w')/c_L)$. This delivers the final contradiction.
5. Any high type who chooses $e > 0$ is hired with probability 1 at $w = q_H$. Suppose otherwise, then there exists a market (e, w) with $w < q_H$ such that there are high-type applicants. Since there are no low types in market (e, w) by step (4), any firm $\theta > \bar{\theta}$ could then make positive profits by hiring non-selectively in market (e, w), contradicting step (1c).
6. No firm hires high types selectively at any $e > 0$. Suppose there was a high type $i > \lambda$ who is hired in market (e, q_H) with $e > 0$ by a firm $\theta < i$ that sets selection rule $\chi_\theta(i) = \mathbb{I}(i \geq \theta)$.

Consider market $(0, w')$ with $w' \in (q_H - c_H e, q_H)$. Then $G(I_{\chi_\theta}; e, q_H, \chi_\theta) > 0$ since otherwise $\mu(q_H, 0, i) = 0$ by Condition 5, so type i could obtain a payoff $w' > q_H - c_H e$ by choosing $e = 0$. Since by construction the support of $G(\cdot; e, q_H, \chi_\theta)$ only includes high types and since $w' < q_H$, firm θ can increase its profits by hiring high types in market $(0, w')$ rather than (e, q_H), a contradiction. Hence, all high types selecting $e > 0$ are hired by firms using selection rule $\chi(i) = 1$ for all i.

7. If any high types choose some education $e_S > 0$, it must satisfy $q_H - c_L e_S = u_L$. Suppose first that some high types choose $e \in (0, e_S)$. By step (5), they are hired at wage q_H and by step (6) 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 $(e_S, q_H - \epsilon)$ and selection rule $\chi(i) = 1$ for all i, which is feasible for all firms. For sufficiently small ϵ, $G(I_\chi; e_S, q_H - \epsilon, \chi) > 0$ since otherwise $\mu(q_H - \epsilon; e_S, i) = 0$ and those high types choosing e could do better by choosing education e_S. By Condition 4, the support cannot include low types because $q_H - c_L e < u_L$. Hence, firms $\theta \geq \bar{\theta}$ could make strictly positive profits in market $(e_S, q_H - \epsilon)$, contradicting step (1c).

8. Define

$$\bar{w}_S \equiv q_H - c_H e_S = \left(1 - \frac{c_H}{c_L}\right) q_H + \frac{c_H}{c_L} u_L. \tag{70}$$

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 (5) and (7) 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 (3).

9. For workers $i \geq i^*$ (who choose $e = 0$) the minimum wage in their support $\underline{w}(0, i)$ 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 by Condition 2.

10. If some type $i \geq 0$ who chooses $e = 0$ is hired by a selective firm, this can only occur at the minimum wage in worker $i's$ support $\underline{w}(0, i)$. To see this, consider a market $(0, w)$ 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 that there are i-type applicants in some market $(0, w - \epsilon)$. As a result, firm θ could increase its profits by shifting demand to market $(0, w - \epsilon)$ using the same selection rule.

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

12. All types $i > i_H$, who select $e = 0$ by step (8), must be hired with probability 1 at $w = q_H$. They cannot be hired with positive probability above q_H because no firm would hire at such a wage. Suppose some $\tilde{i} > i_H$ is not hired with probability 1 at $w = q_H$. This implies that all

firms $\theta \in \left(i_H, \tilde{i}\right)$ maximize profits by hiring selectively at the lower bound of the support of the wages of worker $i = \theta$, which is below q_H by step (9). The total number of workers these firms would hire is $F\left(\tilde{i}\right) - F\left(i_H\right) \geq \bar{F}\left(\tilde{i}\right) - \bar{F}\left(i_H\right) \geq \tilde{i} - i_H$. The first inequality follows from the fact that $F\left(\theta\right) \geq \bar{F}\left(\theta\right)$ for all θ by construction of $\bar{F}$, and $F\left(i_H\right) = \bar{F}\left(i_H\right)$ by definition of i_H. The second follows from the fact that $\bar{f}\left(\theta\right) \geq 1$ for all $\theta \geq i_H$. Moreover, generically the second inequality is strict. By Condition 1, this implies $\mu\left(w; 0, i\right) < 0$ for some worker i in this interval, a contradiction.

13. 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 $i \in [\lambda, i_H]$, the lower bound on their wage distribution $\underline{w}_S$ must satisfy:

$$\bar{f}(i^*)\underline{w}_S + \int_{i^*}^{i_H} \underline{w}(0, i')d\bar{f}(i') = \bar{w}_S \tag{71}$$

with the cutoff i^* defined in step (8) given by (68).

Suppose first that there exist workers in $[\lambda, i_H]$ with lower bounds on wages lower than those defined by (71), and let $\tilde{i}$ be the highest worker such that for some $\epsilon > 0$, the lower bound is higher for all $i \in \left(\tilde{i} - \epsilon, \tilde{i}\right)$.

(a) If $\tilde{i} \in (i^*, i_H]$ is in a region where $\bar{f}\left(i\right)$ is strictly increasing, let $\underline{\tilde{w}}(0, i)$ be the lower bounds on the wages of $i \in (\tilde{i} - \epsilon, \tilde{i})$. Define markets $M\left(\tilde{i}\right) = \{(e, w) : e = 0, w = \underline{\tilde{w}}(0, i), i \in (\tilde{i} - \epsilon, \tilde{i})\}$. 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 $(e, w) \in M\left(\tilde{i}\right)$. Therefore total selective hiring in markets $(e, w) \in M\left(\tilde{i}\right)$ will be $\int_{\tilde{i}-\epsilon}^{\tilde{i}} d\bar{F}\left(i\right)$. By construction of $\tilde{i}$, all workers $\tilde{i} + \varepsilon$ for $\varepsilon > 0$ have lower bounds on wages $\underline{w}\left(0, \tilde{i} + \varepsilon\right)$ given by (66). By step (11), a fraction $\bar{f}\left(\tilde{i} + \varepsilon\right)$ of them are hired by selective firms, and step (10) implies that the selective hiring occurs at the lower bound of their 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 $1 - \bar{f}\left(\tilde{i} + \varepsilon\right)$ 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 $1 - \bar{f}\left(\tilde{i}\right)$ 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 $i' > i$ have lower bounds on wages $\underline{\tilde{w}}(0, i)$, so the supply of workers in market $(0, \underline{\tilde{w}}(0, i))$ includes $i - i_S$ high types and λ low types. Therefore (66) and the fact that $\underline{\tilde{w}}(0, i) > \underline{w}\left(0, i\right)$ imply that no firms want to hire non-selectively in any market $(0, \underline{\tilde{w}}(0, i))$. 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 $1 - \bar{f}\left(\tilde{i}\right)$ 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 $\bar{f}\left(\tilde{i}\right)\epsilon$. Since $\bar{f}\left(i\right)$ is strictly increasing, $\bar{f}\left(\tilde{i}\right)\epsilon > \int_{\tilde{i}-\epsilon}^{\tilde{i}} d\bar{F}\left(i\right)$, which implies that $\mu(\underline{\tilde{w}}(0, i); 0, i) > 0$ for some workers $i \in (\tilde{i} - \epsilon, \tilde{i})$, and the lower bound on wages must be lower than $\underline{\tilde{w}}(0, i)$.

(b) If $\tilde{i} \in (i^*, i_H]$ is in a region where $\bar{f}\left(i\right)$ is constant or if $\tilde{i} < i^*$ then this implies that the lower bound on the wages of worker $\tilde{i}$ is higher than that of some worker $i' > \tilde{i}$, which would violate step (9).

(c) If $\tilde{i} = i^*$, this would imply that some workers $i \in [\lambda, i^*]$ have a lower bound on their wage $\underline{\tilde{w}}(0, i) > \underline{w}_S$. This can only occur without violating step (9) 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 (71). 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 (13a) would apply), this implies that $\mu(\underline{\tilde{w}}(0,i), 0, i) > 0$ for some workers $i \in (i', i^*)$, and therefore the lower bound on wages must be lower than $\underline{\tilde{w}}(0,i)$.

Suppose next that there exist workers with lower bounds on wages lower than those defined by (71), and let $\tilde{i}$ be the highest worker such that for some $\epsilon > 0$, lower bounds on wages are lower than those defined by (71) for all $i \in \left(\tilde{i} - \epsilon, \tilde{i}\right)$.

(a) If $\tilde{i} \in [i^*, i_H]$ is in a region where $\bar{f}(i)$ is strictly increasing, take some $i' \in \left(\tilde{i} - \epsilon, \tilde{i}\right)$ with a lower bound on wages $w' < [(i' - i_S)q_H + \lambda q_L] / [i' - i_S + \lambda]$ and consider the market $(0, w')$. 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 $(0, w')$ 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 (1c).

(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 non-selective firms at wages at least $\underline{w}(0, i_1)$. The measure of firms in $\left(i_0, \tilde{i}\right)$ is $F\left(\tilde{i}\right) - F(i_0) \geq \bar{F}(\tilde{i}) - \bar{F}(i_0) = \bar{f}(i_1)\left(\tilde{i} - i_0\right)$, 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 lower bound on wages 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 (68) implies that the utility of all workers $i \leq \tilde{i}$ is below $\bar{w}_S$. By (71) and (70), they would be better off choosing $e = e_S$.

14. In any interior equilibrium, the cutoff i^* defined in step (8) must also satisfy (69). To see this, observe first that not all workers $i \in [\lambda, i^*]$ can possibly signal. If this were the case, consider the beliefs of a firm $\theta \in [\lambda, i^*)$ in a market $(0, \underline{w}_S + \varepsilon)$. Then $G(I_{\chi\theta}; 0, \underline{w}_S + \varepsilon, \chi_\theta) > 0$ since otherwise $\mu(\underline{w}_S + \varepsilon; 0, i) = 0$ for $i \in [\lambda, i^*)$ by Condition 5, so these workers could get payoff in excess of $\bar{w}_S$ by choosing $e = 0$ instead of signaling. Moreover, the support of $G(\cdot; 0, \underline{w}_S + \varepsilon, \chi_\theta)$ cannot include high types since otherwise some firms could increase their profits by shifting demand to market $(0, \underline{w}_S + \varepsilon)$ for ε sufficiently small. Hence, for all firms $\theta \in [\lambda, i^*)$, it is profit maximizing to hire selectively in market $(0, \underline{w}_S)$, so total selective hires will be $F(i^*) = \bar{F}(i^*)$, where the equality follows from the fact that, by (68), 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 non-selective firms at higher wages is $\bar{f}(i^*)(i^* - i_S)$. Hence, if $\bar{F}(i^*) > \bar{f}(i^*)(i^* - i_S)$, then $\mu(\underline{w}_S; 0, i) < 0$ for some i in this interval, which is a contradiction. If $\bar{F}(i^*) < \bar{f}(i^*)(i^* - i_S)$, then $\mu(\underline{w}_S; 0, i) > 0$ for some i in this interval, so $\underline{w}_S$ cannot be the lower bound on wages.

15. For the case of a corner equilibrium, note first that $\Gamma(i_H, i_H - F(i_H)) < 0$ implies $\underline{w}(0, i_H) < w^P$, where we used $\bar{f}(i_H) = 1$ (abstracting from the trivial case $i_H = \lambda$, which is fully characterized by step (12)). Together with steps (1e) and (9), this means that there cannot be any non-selective 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 have a lower bound on wages w^P. The bound cannot be lower than w^P by step (1e). Suppose for some workers the bound is higher and let i' be the lowest $i \in [\lambda, i_H)$ such that $\underline{w}(0, i) > w^P$ 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 $(0, w^P)$. Hence, total selective hires in (i', i_H) are given by $F(i_H) - F(i') \le \bar{F}(i_H) - \bar{F}(i') < i_H - i'$. 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 wage $\underline{w}(0, i)$ cannot be the lower bound on their wages.

16. By the same argument as in step (14), in the corner equilibrium not every worker with $i < i_H$ can signal. Moreover, again by the same argument as in step (14), 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 $i_S = i_H - F(i_H)$.

17. Finally, we show that there is a unique solution to equations (68) and (69). The argument in the proof of Proposition (12) applies, except that, with bunching, the function $\Delta(i^*)$ is no longer continuous. From (67) 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 (68) and (69) is $i^* = \min_{i \in [\lambda, i_H]} \{i | \Delta(i) \ge 0\}$. Together with the result from Proposition 12 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 (68) and (69) and the corner equilibrium is the unique equilibrium.

    (b) If $\Delta(i_H) \ge 0$, 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 established the existence of a solution to equations (68) and (69). We now provide the equilibrium decisions, probabilities and beliefs, and verify the conditions in Definition 1.

Interior Equilibrium.

(a) Education decisions:

$$e_i = \begin{cases} 0 & \text{if } i < \lambda \text{ or } i \ge i_S \\ e_S & \text{otherwise} \end{cases}$$

(b) Probabilities:

$$\mu(w;e,i)=\begin{cases}\mathbb{I}(w\geq q_H) & \text{if} \quad e=0,\ i\geq i_H\\ \bar{f}(i^r(w)) & \text{if} \quad \begin{cases} e=0,\ i\in[i^*,i_H),\ w\in[\underline{w}(0,i),\underline{w}(0,i_H))\\ e=0,\ i\in[\lambda,i^*),\ w\in[\underline{w}_S,\underline{w}(0,i_H))\\ e=0,\ i\in[0,\lambda),\ w\in[q_L,\underline{w}(0,i_H))\end{cases}\\ 0 & \text{if} \quad \begin{cases} e=0,\ i\in[i^*,i_H],\ w<\underline{w}(0,i)\\ e=0,\ i\in[\lambda,i^*),\ w<\underline{w}_S\\ e=0,\ i\in[0,\lambda),\ w<q_L\end{cases}\\ 1 & \text{if} \quad e=0,\ i\leq i_H,\ w\geq\underline{w}(0,i_H)\\ \mathbb{I}(w\geq\tilde{w}(e,i)) & \text{if} \quad e>0,\ i\geq\lambda\\ \mathbb{I}(w\geq\min\{\tilde{w}(e,i),\tilde{w}(e,\lambda)\}) & \text{if} \quad e>0,\ i<\lambda\end{cases}$$

where

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

$\tilde{w}(e,i)$ is given by (49) and

$$u(i)=\begin{cases} q_H & \text{if} \quad i>i_H\\ \bar{f}(i)\underline{w}(0,i)+\int_i^{i_H}\underline{w}(0,i')d\bar{f}(i') & \text{if} \quad i\in[i^*,i_H]\\ \bar{f}(i^*)\underline{w}_S+\int_{i^*}^{i_H}\underline{w}(0,i)d\bar{f}(i) & \text{if} \quad i\in[\lambda,i^*)\\ \bar{f}(i^*)q_L+\int_{i^*}^{i_H}\underline{w}(0,i)d\bar{f}(i) & \text{if} \quad i<\lambda\end{cases}$$

(c) Demand decisions:

$$(e_\theta,w_\theta,\chi_\theta)=\begin{cases}(0,\underline{w}_S,\chi(i)=\mathbb{I}(i\geq\theta)) & \text{if } \theta\in[\lambda,i^*)\\ (0,\underline{w}(0,\theta),\chi(i)=\mathbb{I}(i\geq\theta)) & \text{if } \theta\in[i^*,i_H)\\ (0,q_H,\chi(i)=\mathbb{I}(i\geq\theta)) & \text{if } \theta\in[i_H,\theta^*)\\ (0,q_L,\chi(i)=1\forall i) & \text{for a measure } \lambda(1-\bar{f}(i^*)) \text{ of firms } \theta\geq\theta^*\\ (e_S,q_H,\chi(i)=1\forall i) & \text{for a measure } i_S-\lambda \text{ of firms } \theta\geq\theta^*\end{cases}$$

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

The non-selective demand in markets $(0,\underline{w}(0,i))$ with $i\in[i^*,i_H)$ and $\bar{f}'(i)>0$ remains to be specified. For a small interval of types $[i_0,i_0+\Delta]$ the change in the probability of being hired non-selectively is:

$$\bar{f}(i_0+\Delta)-\bar{f}(i_0)\approx\bar{f}'(i_0)\Delta$$

Using that in a no-bunching region $i^r(w)=\frac{i_S(q_H-w)+\lambda(w-q_L)}{q_H-w}$, this implies that total non-selective hires over an interval of wages $[w,w+\epsilon]$ are

$$\epsilon\frac{\lambda(q_H-q_L)}{(q_H-w)^2}[i^r(w)-(i_S-\lambda)]\bar{f}'(i^r(w)).$$

Hence, the total measure of demand from firms $\theta\geq\theta^*$ using the non-selective hiring rule $\chi(i)=1\forall i$ placed on any set of markets $(E_0,W_0)=\{(e,w):e=0,w\in(w_0,w_1)\subset[\underline{w}(0,i^*),\underline{w}(0,i_H)]\}$ must be

$$D(E_0,W_0,\chi)=\int_{w_0}^{w_1}\frac{\lambda(q_H-q_L)}{(q_H-w)^2}[i^r(w)-(i_S-\lambda)]\bar{f}'(i^r(w))dw.$$

All firms $\theta \geq \theta^*$ are indifferent between hiring in any of these markets and remaining inactive, for instance by setting $(e_\theta, w_\theta, \chi_\theta) = (0, 0, \chi(i) = 1\forall 1)$.

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

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

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

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

To see that the proposed $\{e_i, (e_\theta, w_\theta, \chi_\theta), \mu, G\}$ is an equilibrium, note first that the probabilities defined in (b) imply 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 $(0, w = \underline{w}_S)$ and for firms $\theta \in (i^*, i_H)$ to hire in market $(0, \underline{w}(0, \theta))$. Firms $\theta \geq i_H$ make zero profits by hiring selectively in market $(0, q_H)$. Moreover, firms $\theta \geq i_H$ make zero profits by hiring non-selectively in markets $(0, q_L)$, (e_S, q_H) or $(0, \underline{w}(0, i))$, $i \in [i^*, i_H)$. Any other market has either $G\left(I_{\chi_\theta}; e, w, \chi_\theta\right) = 0$ or results in losses. Therefore the demands defined in (c) are an optimal choice. Finally, using the above-specified demand and beliefs, Condition 1 holds. It is straightforward to verify that $\mu(w; e, i)$ given in (b) is weakly decreasing in i, so Condition 2 is also satisfied. Beliefs satisfy Condition 3 in nonempty markets. In zero supply markets, beliefs are also constructed to satisfy Condition 4 when they are well defined, and $G\left(\cdot; e, w, \chi_\theta\right) = 0$ only at wages where $\mu\left(w; e, i\right) = 0$ for all i such that $\chi(i) = 1$, so Condition 5 is satisfied as well.

Corner equilibrium. We state the equilibrium objects $\{e_i, (e_\theta, w_\theta, \chi_\theta), \mu, G\}$. Verifying that this is an equilibrium is analogous to the interior equilibrium case.

(a) Education decisions:

$$e_i = \begin{cases} 0 & \text{if } i < \lambda \text{ or } i \geq i_H - F(i_H) \\ e^* & \text{otherwise} \end{cases}$$

(b) Demand decisions:

$$(e_\theta, w_\theta, \chi_\theta) = \begin{cases} (0, w^P, \chi(i) = \mathbb{I}(i \geq \theta)) & \text{for } \theta < i_H \\ (0, q_H, \chi(i) = \mathbb{I}(i \geq \theta)) & \text{for } \theta \in [i_H, \theta^*) \\ (0, q_L, \chi(i) = 1 \forall i) & \text{for a measure } \lambda \text{ of firms } \theta \geq \theta^* \\ (e^*, q_H, \chi(i) = 1 \forall i) & \text{for a measure } i_H - F(i_H) - \lambda \text{ of firms } \theta \geq \theta^* \\ (0, 0, \chi(i) = 1 \forall i) & \text{otherwise} \end{cases}$$

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

(c) Probabilities:

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

where $\tilde{w}(e, i)$ is given by (49) and

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

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

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

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

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