Selling Signals

Zhuoran Lu*

June 18, 2022

Abstract

This paper studies a signaling model in which a strategic player can manipulate the signaling cost. A seller chooses a price schedule for a good, and a buyer with a hidden type chooses how much to purchase as a signal to receivers. When receivers observe the price schedule, the seller charges monopoly prices, and the buyer purchases less than the first-best. In contrast, when receivers do not observe the price schedule, the demand for signals is more elastic. In equilibrium, the seller charges lower prices, and the buyer purchases more than when receivers observe the price schedule; the highest types purchase more than the first-best. The model suggests that price transparency benefits the seller but harms the buyer. The model can be applied to schools choosing tuition, retailers selling luxury goods and media companies selling advertisements.

Keywords: Signaling, Screening, Signal Jamming, Price Transparency

JEL Codes: D82, H21, L12

*School of Management, Fudan University, luzhuoran@fudan.edu.cn. I am deeply indebted to my advisors, Simon Board and Moritz Meyer-ter-Vehn, for their guidance and encouragement. I am grateful to Tomasz Sadzik for highly valuable discussions. I have received helpful comments from Yingni Guo, Jianpei Li, Jay Lu, Stephen Morris, Ichiro Obara, Joseph Ostroy, Marco Ottaviani, Alessandro Pavan, Michael Powell, Marek Pycia, Luis Rayo, John Riley, Zhiyong Yao and William Zame. I also thank the seminar audiences at Cornell, ECNU, Fudan, PKU, SUFE, Tsinghua, UCLA, Northwestern IO Theory Conference 2016, NASMES 2017, Nanjing International Conference on Game Theory 2018 and CMES 2021. All errors are mine.

1 Introduction

Signaling is prevalent in various markets. Whereas in classic signaling models the sender's preference depends only on his intrinsic type, in many vertical markets in which signaling prevails, the signaling cost—thus the sender's preference—also depends on the choice made by an upstream strategic player. For example, when a student obtains education to signal his ability, the university sets the tuition; when a consumer purchases a luxury good to signal his wealth, the retailer chooses the price; when a firm incurs advertising expenses to signal its product's quality, the media company determines the costs of advertising messages.

A key observation is that since the signaling cost is endogenous, how receivers interpret and respond to the sender's signal depends on whether they observe the upstream player's choice. Consider a seller choosing the price of a good that generates signaling value for the customers, as in the above instances. How does receivers' information about the price affect the seller's pricing strategy? How does such information affect the degree of signaling and social welfare?

In this paper, we characterize the optimal price schedule for a seller facing a buyer who is endowed with a hidden type and chooses how much to purchase as a signal to receivers. The equilibrium depends critically on whether receivers observe the price schedule. When receivers observe the price schedule, the seller internalizes the buyer's signaling behavior when screening the buyer, leading to a downward distortion in quantity. In contrast, when receivers do not observe the price schedule, the buyer is more sensitive to price changes, because receivers will attribute differences in choice to buyer preference heterogeneity. This indicates that the demand for the good is more elastic, and thus, the seller lowers prices. In equilibrium, the buyer chooses a larger quantity and obtains higher utility, whereas the seller gains lower profits than when receivers observe the price schedule.

This paper has meaningful implications for the price transparency of signaling goods. In the case of job market signaling, our model suggests that education is more costly and students are worse off when employers observe the net prices for school than otherwise. This implies that policies that improve the transparency of the net prices at colleges and universities, e.g., U.S. Code § 1015a,[1] may *unintentionally* raise education expenses and harm students. This is because these policies allow schools to commit to high prices and not dilute the signaling value of a high-cost education by means of fee waivers or financial aid.

[1]Since 2011, American colleges and universities have been required to provide reasonable estimates of the net prices, including tuition, miscellaneous fees and personal expenses, that students will pay for school. See "U.S. Code § 1015a - Transparency in college tuition for consumers" for details.

In addition, our model suggests that a signaling good yields higher profits if the price is more transparent. This is consistent with real-world business practices. For example, luxury brands, such as Louis Vuitton, Tiffany and Hermes, strive for a reputation of never or very rarely being on sale. These strategies help the sellers better commit to high prices, thereby maintaining the signaling values of luxury goods. In the advertising industry, the high costs of each year's Super Bowl commercials are widely reported, thereby enhancing the signaling value of these costly commercials; in China, the TV station CCTV broadcasts the auctions for its popular TV show commercials to accentuate their signaling values.

For expository purposes, we present our model à la Spence (1973) with a school selling productive education to a worker. As a reference point, we revisit Spence's model by fixing tuition at zero, as if schools were competitive and set the price at the marginal cost. In the least-cost separating equilibrium (the Riley outcome), all types except the lowest one choose more education than the first-best, as they attempt to separate themselves from lower types.

In Section 3, we introduce the school and consider the case in which employers observe the tuition scheme. Following the literature, we focus on the seller-optimal equilibrium, in which all types except the highest one choose less education than the first-best. This result contrasts with that of Spence's model. The downward distortion is due to screening. With a cost advantage in education, a higher type can secure higher utility than a lower type by imitating the latter, meaning that the worker can extract information rents from the school. This induces the school to under-supply education.

While this mechanism is similar to screening models such as Mussa and Rosen (1978), our model also incorporates signaling, which can mitigate the downward distortion caused by screening. To illustrate, suppose that employers can observe the worker's ability, thereby eliminating signaling. When a higher type imitates a lower type, he not only incurs a lower total cost than the latter but also obtains a higher wage due to his higher ability. The second effect means that the worker can extract more information rents from the school; thus, the screening distortion is worse compared to when signaling is present.

In Section 4, we turn to the case in which employers do not observe the tuition scheme. We propose a new equilibrium refinement, the *Quasi-Divinity*. The unique equilibrium that satisfies the Quasi-Divinity is the seller-optimal separating equilibrium (the Riley outcome), in which the school charges lower tuition and the worker chooses more education than when employers observe the tuition scheme. This difference is driven by a signal jamming effect. Because employers cannot observe the actual cost of education, they do not know whether a difference in education level is due to a tuition change or worker cost heterogeneity. For example, suppose that the school lowers tuition so that the worker obtains more education than in the initial state. When employers observe the tuition scheme, they cut wages, since any education level now corresponds to a lower-ability worker. This dampens the worker's demand for additional education. In contrast, when employers do not observe the tuition scheme, they do not adjust wages despite that tuition changes. Therefore, the demand for education is more elastic, making the price cut more profitable. In equilibrium, employers correctly anticipate the school's choice and offer lower wages, as education is inflated. This reduces the worker's willingness to pay, and thus, the school achieves lower profits.

Since the school is worse off when employers do not observe the tuition scheme, one may wonder why the school does not disclose tuition to employers. The reason is that the school cannot credibly announce the price absent intervention such as mandatory disclosure, since the school has an incentive to secretly cut prices. Such an observation may explain the fact that whereas the listed tuition at American colleges and universities is rising, these schools offer students various and inclusive forms of financial aid. The rationale is that employers cannot easily observe the details of such financial aid and thus do not know the actual cost of education. By raising the published tuition while simultaneously reducing the undisclosed net prices through stipends, schools persuade employers that their students are smarter than is actually the case, thereby allowing the schools to collect higher revenues from students.

In terms of welfare, we show that when signaling is sufficiently intense, social welfare is higher when the tuition scheme is observed by employers than otherwise. This is because in the observed case signaling mitigates the screening distortion to a large extent, whereas in the unobserved case cheaper tuition leads to a large fraction of higher types who over-invest in education. Moreover, when signaling is intense, both cases yield higher social welfare than Spence's model in which schools are competitive. This implies that promoting competition among the sellers of signals is not necessarily socially beneficial.

In Section 5, we explore some extensions. We first argue that the main results are robust if education is nonessential to the worker's productivity (e.g., in Spence (1973) education is a pure signal). Furthermore, we show that in the observed case, the market may consist of a certification segment and an education segment: A lower interval of types pay a fixed fee for zero education, as if they were certified by the school as having the average productivity, whereas the higher types purchase education to signal their abilities. Finally, we consider when the school's objective is a weighted average of its profit and the worker's utility. We show that our results hold essentially, up to the relative Pareto weight of the two parties. Section 6 concludes our paper. All proofs are in the Appendix.

1.1 Related Literature

This paper is most closely related to the literature on signaling. The paper contributes to the literature on signaling games by allowing a strategic player to affect the signaling cost. In classic signaling models (e.g., Spence 1973; Leland and Pyle 1977; Riley 1979; Milgrom and Roberts 1986; Bagwell and Riordan 1991), with an exogenous cost structure, signaling gives rise to over-investment in costly actions. In this regard, Spence (1974), Ireland (1994) and Andersson (1996) consider the welfare-maximizing tax on signals. In contrast, we solve for the profit-maximizing tax scheme, which "over-taxes" signaling and causes a downward distortion in quantity when receivers observe the tax scheme.

The paper is also closely related to the literature on screening. Screening models, such as Mussa and Rosen (1978) and Maskin and Riley (1984), typically assume that buyers derive intrinsic utility from consuming the seller's product. Our model differs in the sense that the product has further a signaling value, and a buyer's utility depends on the information that the product conveys. Rayo (2013) also considers the optimal monopoly pricing to sell signals, assuming that the seller's mechanism is observed by receivers. Whereas we assume additive separability in receivers' responses (e.g., wages) and the buyer's type (e.g., ability), Rayo's adopts a multiplicative structure, and employs novel screening techniques to characterize which types should be pooled into the same level of signal. The contribution of our paper is to study the case in which receivers cannot observe the seller's mechanism, and comparing this to the observed case and a variety of other benchmarks. This enables us to assess how price transparency affects the degree of signaling and welfare. Calzolari and Pavan (2006) studies information disclosure in a sequential screening model. They show that the upstream principal leaves more rents to the agent if she discloses information about the agent's type to the downstream principal. In our model, the seller leaves more rents to the buyer if receivers can observe the buyer's type, which is perfect information disclosure.

The unobserved tuition case belongs to the class of signal jamming models proposed by Fudenberg and Tirole (1986). For example, in Holmström (1999), the labor market cannot distinguish the impact of the worker's ability from that of his effort on output. Therefore, the worker works harder to improve the market's perception of his ability. In comparison, in our model, the labor market cannot distinguish the impact of the worker's ability from that of tuition on education level. Thus, the school has an incentive to secretly cut tuition, thereby improving the market's perception and stimulating the demand for education.

Our paper is also related to the growing literature on information intermediary initiated by Biglaiser (1993). For example, Lizzeri (1999) studies the design of certification system, and shows that a monopolistic certifier may disclose no information about the agent and capture all the surplus. In Chan et al. (2007), schools design grading systems to place their students in the job market. The authors show that a school has incentives to inflate grades to improve the market's perception of its students. In contrast to these models, our model incorporates screening in addition to signaling, as the designer cannot observe the agent's type. Similarly, Zubrickas (2015) studies a school's optimal grading policy when the students' abilities are private information and the job market has myopic beliefs over the school's grading policies. More recently, Biglaiser and Li (2018) considers an expert market in which a seller privately chooses his effort before going to a middleman who decides whether to purchase the seller's good after receiving a signal about the good's quality. They argue that the informational externality of the middleman can either raise or reduce the seller's effort. Our model differs in that the agent's private information is exogenous and is transmitted through signaling.

Finally, our paper is closely related to the literature on intermediate price transparency. Inderst and Ottaviani (2012) shows how product providers compete through commissions paid to consumer advisers. Commissions bias advice; thus, an increase in a firm's commission reduces consumers' willingness to pay if they observe the commission. Analogously, in our model, cheaper tuition reduces the signaling value of education, and thus, tuition cuts are less effective at stimulating demand than they would be otherwise when employers observe tuition. In Janssen and Shelegia (2015), a manufacturer chooses a wholesale price, retailers choose retail prices, and consumers search for the best deal. They argue that retailers are less sensitive to wholesale price changes when consumers do not observe the price than otherwise, as uninformed consumers are more likely to keep searching when the retail price raises. By contrast, in our model the worker is more sensitive to tuition changes when employers do not observe the tuition scheme than otherwise, as uninformed employers will have better (worse) beliefs over the worker's ability if they observe a higher (lower) education level.

2 The Model

For expositional convenience, in this section we first present our model in conformity with the seminal work of Spence (1973). Then, we describe how to apply our model to different vertical markets, such as luxury goods and advertising, in which signaling prevails.

Players and actions. There is a school (*seller*), a worker (*sender*) and multiple identical and competing firms, also referred to as *the labor market* (*receiver*). At the beginning of the game, the school chooses a tuition scheme $T : \mathbb{R}_+ \to \mathbb{R}_+$, which specifies the tuition fee for each education level z. Then, the worker decides how much education to purchase from the school after observing the tuition scheme. For simplicity, we do not explicitly model firms' actions; rather, we directly assume that they offer the worker a wage equal to the expected productivity of the worker (see below).

The worker's productivity depends on his ability (*type*) θ and his education choice z. Specifically, θ is a random variable, which distributes over the interval $[\underline{\theta}, \bar{\theta}]$, according to a distribution function $F(\theta)$ with a positive density function $f(\theta)$. Denote by $Q(z,\theta)$ the productivity of a type-θ worker with education level z. We assume that $Q(z,\theta)$ is smooth, with Q_z, $Q_\theta > 0$ if $z > 0$, and $Q_{zz} \leq 0$. In addition, a worker with no education has zero productivity irrespective of his ability; that is, $Q(0,\theta) \equiv 0$. We consider this assumption realistic, as many jobs require a minimal education level. For example, a lawyer candidate must graduate from a law school, and medical school education is prerequisite for being a licensed practitioner of medicine. In Section 5, as extensions, we consider the case in which education is nonessential, i.e., $Q(0,\theta) \geq 0$ and $Q_\theta > 0$ for all (z,θ), and the case in which education is a pure signal, i.e., $Q_z \equiv 0$ for all θ.

Information. The worker's education level is publicly observed. However, neither the school nor the labor market observes the worker's ability, but both know its distribution. In this paper, we mainly study two variants of the model: in the *observed* case, the tuition scheme is observed by the labor market; in the *unobserved* case, it is unobserved by the labor market. In each case, based on the available information, the labor market chooses a wage schedule $W : \mathbb{R}_+ \to \mathbb{R}_+$, which specifies the wage for each education level z.

Payoffs. We normalize the school's production cost to zero. Suppose that the school chooses some tuition scheme T, then $z(\theta;T)$ denotes the education level chosen by a type-θ worker under T. Given the tuition scheme T and a wage schedule W, a type-θ worker who chooses education level z has utility

$$U(z,\theta) \coloneqq W(z) - C(z,\theta) - T(z),$$

where $C(z,\theta)$ denotes the worker's effort cost of education. Assume that $C(z,\theta)$ is smooth, with $C_z > 0$ if $z > 0$, and $C_{zz} > k$ for some $k > 0$. Moreover, the standard *single-crossing property* holds: $C_{z\theta} < 0$. This condition reflects the feature that a higher-ability worker has lower marginal effort costs than a lower-ability worker. We further normalize $C(0,\theta)$ to 0 for all $\theta \in [\underline{\theta}, \bar{\theta}]$. This implies that, combined with $C_{z\theta} < 0$, $C_\theta < 0$ if and only if $z > 0$. Finally, we assume that the worker can obtain a zero-utility outside option by acquiring no education and not entering the labor market.

First-best benchmark. Define $S(z,\theta)$ as the social surplus function, i.e.,

$$S(z,\theta) := Q(z,\theta) - C(z,\theta).$$

Note that $S(z,\theta)$ is strictly concave in z, and thus, has a unique maximizer $z^{fb}(\theta)$. Assume that $z^{fb}(\theta) \geq 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$. Then, the first-order condition implies that

$$S_z(z^{fb}(\theta),\theta) = Q_z(z^{fb}(\theta),\theta) - C_z(z^{fb}(\theta),\theta) = 0. \tag{2.1}$$

Assume further that $S_{z\theta} > 0$. Then, by Milgrom and Shannon (1994, Theorem 4), we have that $z^{fb}(\theta)$ is increasing on $[\underline{\theta}, \bar{\theta}]$.

2.1 Direct Mechanisms

We use *perfect Bayesian equilibrium* as the solution concept throughout the paper. Similar to classic signaling games, there may exist multiple equilibria in our model. To this end, in the observed case we focus on the *seller-optimal equilibrium*, i.e., the equilibrium that yields the highest payoff for the school. This is essentially allowing the school to communicate with the worker and the labor market as to which signaling equilibrium to play in each subgame. In the unobserved case, in contrast, because the labor market cannot see the school's choice, we shall rule out such communications. Instead, we propose a new equilibrium refinement, the *Quasi-Divinity*, which uniquely selects the *seller-optimal separating equilibrium*, that is, the most profitable equilibrium for the school, provided that $z(\theta)$ is one-to-one if $z > 0$. In Section 4, we will discuss the equilibrium selection of the unobserved case in greater details.

Appealing to the revelation principle, we consider direct mechanisms between the school and the worker in both the observed and unobserved cases. It is without loss of generality to adjust the timing as follows. First, the school offers a contract $\{z(\theta), T(\theta)\}$ to the worker. Then, the labor market chooses a wage schedule W based on the available information: in the observed case, the labor market sees the contract; in the unobserved case, it does not. Finally, the worker reports a type to the school privately.[2] Reporting a type $\hat{\theta}$, the worker obtains education level $z(\hat{\theta})$, pays tuition $T(\hat{\theta})$ and then receives wage $W(z(\hat{\theta}))$.

Worker's problem. In both cases, given a contract $\{z(\theta), T(\theta)\}$ and the associated wage schedule W, a type-θ worker chooses a report $\hat{\theta}$ to maximize his utility

$$U(\hat{\theta},\theta) := W(z(\hat{\theta})) - C(z(\hat{\theta}),\theta) - T(\hat{\theta}).$$

[2]If reports are public, then the set of outcomes that can be implemented by a truthful direct mechanism is smaller than what can be obtained by an indirect mechanism in which the worker only chooses education.

A contract with the associated wage schedule is *incentive compatible* if the worker is willing to truthfully report his type, and *individually rational* if the worker obtains a nonnegative utility level. A type-θ worker's equilibrium payoff is represented by $U(\theta) \coloneqq U(\theta, \theta)$.

School's problem. In the observed case, the school chooses a contract to maximize its expected profit subject to incentive compatibility (IC), individual rationality (IR), and that the labor market's posterior belief about the worker's ability, or simply *the market belief*, is updated using Bayes' rule whenever possible, i.e., the market belief is correct. In contrast, in the unobserved case, given some wage schedule, the school chooses a contract to maximize its expected profit subject to only IC and IR constraints.

Preliminaries. In both cases, an allocation $\{z(\theta), U(\theta)\}$ is *implementable* if it is incentive compatible and individually rational. Appealing to Mas-Colell, Whinston, and Green (1995, Proposition 23.D.2), we characterize all implementable allocations by the following lemma.

Lemma 1. *In both cases, an allocation $\{z(\theta), U(\theta)\}$ is implementable if and only if*

(i) $z(\theta)$ is nondecreasing.

(ii) Define $\theta_0 \coloneqq \inf\{\theta | z(\theta) > 0\}$; then, for any $\theta \geq \theta_0$,

$$U(\theta) = U(\theta_0) + \int_{\theta_0}^{\theta} -C_\theta(z(s), s) ds \geq 0.$$

From Lemma 1, we can rewrite the school's problem for both cases. Note that IC means that $T(\theta) = W(z(\theta)) - C(z(\theta), \theta) - U(\theta)$ and that $U(\theta_0)$ is optimally set to 0. Substituting and integrating by parts, the school's problem can be stated as

$$\max_{\{z, \theta_0\}} \int_{\theta_0}^{\bar{\theta}} \left[W(z(\theta)) - C(z(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)} C_\theta(z(\theta), \theta) \right] dF(\theta) \tag{2.2}$$

subject to $z(\theta)$ being nondecreasing.

In the observed case, the market belief correctness means that $W(z(\theta)) = \mathbb{E}[Q(z(\theta), \theta)]$ for any implementable allocation $\{z, \theta_0\}$ that the school chooses. Then, by the law of total expectation, Program (2.2) is equivalent to

$$\max_{\{z, \theta_0\}} \int_{\theta_0}^{\bar{\theta}} \left[S(z(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)} C_\theta(z(\theta), \theta) \right] dF(\theta) \tag{2.3}$$

subject to $z(\theta)$ being nondecreasing. Thus, to characterize the equilibrium of the observed case is essentially to solve Program (2.3). Let $\{z^o(\theta), T^o(\theta)\}$ denote the optimal contract.

In the unobserved case, the market's inference is independent of the actual contract but is conditional on a *conjectured* contract; in equilibrium, the conjecture is correct. Without loss of generality, let the school choose an allocation $\{z, \theta_0\}$, while simultaneously, the labor market choose a wage schedule W. As such, an equilibrium consists of an allocation $\{z^u, \theta_0^u\}$ and a wage schedule W^u, satisfying the following conditions: (i) Given W^u, $\{z^u, \theta_0^u\}$ solves the school's problem in (2.2); (ii) $W^u(z) = \mathbb{E}[Q(z, \theta)]$ such that the market belief is updated using Bayes' rule whenever possible. Let $\{z^u(\theta), T^u(\theta)\}$ denote the equilibrium contract.

To ensure that an equilibrium exists in both cases, we introduce a regularity condition. Let $h(\theta) := f(\theta)/[1 - F(\theta)]$ be the hazard rate of $F(\theta)$. We assume throughout the paper:

Assumption 1. *$C_{zz\theta} \leq 0$, $C_{z\theta\theta} \geq 0$, and $h'(\theta) > -[h(\theta)]^2$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$.*

Assumption 1 is less restrictive than the typical regularity condition in classic screening models such as Mussa and Rosen (1978), as it does not require the monotone hazard rate property; instead, it requires that the slope of the hazard rate should not be too negative. Indeed, there are some common distributions that fails the monotone hazard rate property but satisfies the hazard rate property in Assumption 1.[3]

Remark. In this paper, we assume that the school maximizes its tuition revenue, whereas in reality schools are usually nonprofit. We consider the profit-maximizing assumption justified, as schools with abundant revenues have great competitive advantages to fulfill high quality instruction, for instance, by maintaining outstanding faculty. In the U.S., college tuition is excessively higher than the instruction costs, let alone the tuition at top business schools. As Professor Jeffrey Pfeffer of Stanford Graduate School of Business has once put it:

The only difference between Stanford and Google is we have a higher profit margin.[4]

In Section 5, as an extension, we consider the case in which the school's objective assigns a positive Pareto weight to both the school's profit and the worker's utility.

2.2 Applications of the Model

Before the formal analysis of the model, we discuss how our framework might be applied to other cases such as conspicuous consumption and advertising. In the following, we present, in a parallel manner to the case of job market signaling, an adapted version of the model for each case respectively. We start with the case of conspicuous consumption.

[3]For example, a log-normal distribution with parameters $\mu = 0$ and $\sigma = 1$ has a non-monotone hazard rate but satisfies the hazard rate property in Assumption 1.

[4]See "How Non-Profit Business Schools Maximize Profits," StretegyTMT (www.strategytmt.com).

Conspicuous consumption. A luxury good retailer (*seller*) first chooses a price scheme $T : \mathbb{R}_+ \to \mathbb{R}_+$, which specifies the price for each level of quality z. Then, à la Bagwell and Bernheim (1996), a consumer (*sender*) chooses a level of quality to signal his unobserved wealth (*type*) θ to a representative social contact (*receiver*). The social contact observes z; in the observed case, she observes T as well; in the unobserved case, she does not observe T.

In the spirit of the classic work of Veblen (1899), the social contact rewards the consumer according to z. The reward $W(z)$ is given by the social contact's expected benefit $\mathbb{E}[Q(z,\theta)]$ from the consumer. The benefit function $Q(z,\theta)$ is increasing in both arguments: Q_z, $Q_\theta > 0$ if $z > 0$. This is because a social contact benefits more from establishing relationships with wealthier people, and from interacting with people who consume higher quality goods (e.g., the good is non-rivalrous). We also assume that $Q_{zz} \leq 0$ and $Q(0,\theta) \equiv 0$.[5] Moreover, the consumer derives intrinsic utility $V(z,\theta)$ from the luxury good, which is increasing in the quality z. Similarly, the single-crossing property holds: $V_{z\theta} > 0$. This condition states that a wealthier individual has higher marginal utility of a luxury good. For example, a consumer of yacht can voyage more often if he is richer, since he is better able to afford the fuel costs and maintenance fees. Thus, a type-θ consumer who chooses quality z has utility

$$U(z,\theta) \coloneqq W(z) + V(z,\theta) - T(z).$$

The retailer's profit equals the revenue $T(z)$ minus the cost $C(z)$, with C', $C'' > 0$ and $C(0) = 0$. The social surplus function is thus given by $S(z,\theta) \coloneqq Q(z,\theta) + V(z,\theta) - C(z)$.

Advertising. A media company (*seller*) first chooses a price scheme $T : \mathbb{R}_+ \to \mathbb{R}_+$, which specifies the price for each advertising level z. Then, à la Milgrom and Roberts (1986), a producer (*sender*) of a new product chooses the advertising level to signal the unobserved product quality (*type*) θ to consumers. The consumers observe z; in the observed case, they observe T as well; in the unobserved case, they do not observe T.

The producer's total revenue consists of two parts: the purchase in the introductory stage and the repeat purchase in the post-introductory stage. In the introductory stage, $D(z) \geq 0$ consumers become aware of the product and each purchases one unit at a price equal to the expected quality $\mathbb{E}[\theta|z]$. The demand function $D(z)$ is increasing, as more advertising leads to higher consumer awareness. We also assume that $D_{zz} \leq 0$ and $D(0) = 0$. Then, in the post-introductory stage, the product's actual quality θ reveals, and thus, the consumers are willing to purchase the good again at a price equal to θ. We assume that the consumers

[5]The second condition captures the idea that the consumer may need at least an entry-level luxury good to meet the social contact. For example, to join a yacht club, one usually has to own a yacht.

who were unaware of the product do not purchase the good in the post-introductory stage. Thus, the producer's total revenue equals $(\mathbb{E}[\theta|z]+\theta)D(z)$, and his net payoff is given by

$$U(z,\theta) := (\mathbb{E}[\theta|z]+\theta)D(z) - T(z).$$

Note that the single-crossing property holds: $U_{z\theta} > 0$. This is due to the complementarity between advertising and quality, i.e., the marginal revenue of the introductory advertising is higher if the product is of higher quality thereby allowing the producer to charge a higher price in the post-introductory stage. The media company's profit equals the revenue $T(z)$ minus the production cost $C(z)$, with C', $C'' > 0$ and $C(0) = 0$. The social surplus function is thus given by $S(z,\theta) := 2\theta D(z) - C(z)$.

2.3 Spencian Job Market Signaling

Now, let us return to the case of job market signaling. As a reference point, we briefly revisit Spence's signaling game in which tuition is fixed at zero. One can interpret this benchmark as the case in which schools are competitive and thus choose tuition equal to the marginal cost. An equilibrium thus consists of an education function $z^s(\theta)$ and a wage schedule W^s, satisfying that (*i*) Given W^s, $z^s(\theta)$ maximizes $U(z,\theta)$; (*ii*) $W^s(z) = \mathbb{E}[Q(z,\theta)]$, with the market belief updated by Bayes' rule whenever possible. We study the least-cost separating equilibrium (Riley, 1979). By Mailath and von Thadden (2013), such an equilibrium exists, in which $z^s(\theta)$ is given by the initial value problem (IVP):

$$Q_z(z^s(\theta),\theta) + Q_\theta(z^s(\theta),\theta)\theta^{s\prime}(z^s(\theta)) - C_z(z^s(\theta),\theta) = 0, \tag{2.4}$$

with $z^s(\underline{\theta}) = z^{fb}(\underline{\theta})$, where $\theta^s(z)$ is the inverse function of $z^s(\theta)$. Moreover, it can be verified that $z^s(\theta)$ is increasing over $[\underline{\theta},\bar{\theta}]$, and thus, $W^s(z^s(\theta)) = Q(z^s(\theta),\theta)$ for all $\theta \in [\underline{\theta},\bar{\theta}]$.

Note that the first two terms on the left-hand side (LHS) of (2.4) are the total derivative of $W^s(z^s(\theta))$. In particular, the second term is nonnegative given the monotonicity of $z^s(\theta)$. Since $S(z,\theta)$ is strictly concave, comparing (2.4) with (2.1) implies that $z^s(\theta) \geq z^{fb}(\theta)$ for all $\theta \geq \underline{\theta}$, with equality holding at $\underline{\theta}$ only. This result means that in Spence's signaling game, the worker chooses more education than the first-best; i.e., the worker's signaling behavior leads to over-education. The intuition is well-understood. Under complete information, the marginal benefit of education is its marginal contribution to human capital. In contrast, when ability is privately known, in addition to the human capital effect, there is a *signaling effect*; that is, a higher education level makes the labor market regard the worker as having higher ability. Thus, the marginal benefit of education is higher than under complete information.

3 Labor Market Observes Tuition

Starting with this section, we take the school's strategic action into account. Here, we study the case in which the labor market observes the tuition scheme. From Section 2.1, it suffices to solve Program (2.3) for the equilibrium characterization. It is heuristic to interpret the integrand in (2.3) as the school's marginal profit in the observed case. Define

$$MP^o(z,\theta) \coloneqq S(z,\theta) - \left[-\frac{1-F(\theta)}{f(\theta)} C_\theta(z,\theta) \right].$$

Inspired by Martimort and Stole (2009), we say that the school's marginal profit in the observed case is *regular* if $MP^o(z,\theta)$ is strictly quasiconcave in z and $MP^o_z(z,\theta)$ is increasing in θ. Whereas Assumption 1 does not ensure the regularity of $MP^o(z,\theta)$, it is easy to verify that if further $Q_{z\theta} \geq 0$, then $MP^o(z,\theta)$ is regular. This means that $MP^o(z,\theta)$ has a unique maximizer on $\mathbb{R}+$, denoted $z^*(\theta)$, which is continuous and increasing on $[\underline{\theta}, \bar{\theta}]$. As a result, the school's problem reduces to pointwise maximization for $MP^o(z,\theta)$.

In contrast, when the regularity does not hold, $z^*(\theta)$ might be decreasing in some region. In this case, we adopt the generalized ironing technique developed by Toikka (2011). Define

$$J(z,\theta) \coloneqq \int_{\underline{\theta}}^{\theta} MP^o_z(z,s)ds.$$

Because $MP^o_z(z,\theta)$ is continuous, $J(z,\theta)$ is continuously differentiable on $[\underline{\theta}, \bar{\theta}]$ given z. Let $I(z,\cdot) \coloneqq \text{conv } J(z,\cdot)$ be the convex hull of $J(z,\cdot)$; thus, $I(z,\theta)$ is continuously differentiable on $[\underline{\theta}, \bar{\theta}]$, and $I_\theta(z,\theta)$ is nondecreasing in θ. Define the *generalized marginal profit* as

$$\overline{MP}^o(z,\theta) \coloneqq MP^o(0,\theta) + \int_0^z I_\theta(x,\theta)dx.$$

Our first proposition characterizes the equilibrium outcome of the observed case.

Proposition 1. *In the observed case, the seller-optimal equilibrium exists, such that*

$$z^o(\theta) = \begin{cases} \bar{z}(\theta) & \text{if } \theta \geq \theta^o_0 \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

where $\bar{z}(\theta)$ is the unique maximizer of $\overline{MP}^o(\cdot,\theta)$ and θ^o_0 is either the maximal root of $\bar{z}(\theta) = 0$ if it exists, or $\underline{\theta}$ otherwise. In particular, if $MP^o(z,\theta)$ is regular, then $z^o(\theta) = z^*(\theta)$ on $[\underline{\theta}, \bar{\theta}]$. Moreover, $z^o(\theta)$ is continuous on $[\underline{\theta}, \bar{\theta}]$. Thus, for each $z^o(\theta) > 0$, we have

$$T^o(z^o(\theta)) = W^o(z^o(\theta)) - C(z^o(\theta),\theta) + \int_{\theta^o_0}^{\theta} C_\theta(z^o(s),s)ds, \tag{3.2}$$

where $W^o(z^o(\theta)) \coloneqq \mathbb{E}[Q(z^o(\theta),\theta)]$ denotes the equilibrium wage.

Proposition 1 states that the optimal allocation rule $z^o(\theta)$ is continuous, and coincides with the unconstrained optimizer $z^*(\theta)$ under the regularity of $MP^o(z, \theta)$. In the proof, we also show that whenever $z^o(\theta) \neq z^*(\theta)$, θ belongs to a *pooling* interval. The corollary below indicates that the optimal allocation exhibits a downward distortion.

Corollary 1. *In the observed case, the worker acquires less education than the first-best. Specifically, $z^o(\theta) \leq z^{fb}(\theta)$ on $[\underline{\theta}, \bar{\theta}]$, with strict inequality on $(\underline{\theta}, \bar{\theta})$ and $z^o(\bar{\theta}) = z^{fb}(\bar{\theta})$.*

Corollary 1 states that when the labor market observes the tuition scheme, education is under-invested. This result stands in stark contrast to that of Spence's model. The altered equilibrium prediction results from the school's screening. Specifically, with a cost advantage in education, a higher-ability worker can secure higher utility than a lower-ability worker by imitating the latter. To incentivize truth-telling, the school has to leave information rents to the worker. This means that the marginal profit of education is less than the social surplus generated; thus, the school under-supplies education. In particular, an interval of types at the low end of the support will be excluded from education if it is too costly to serve them.

3.1 Screening vs. Signaling

Whereas the equilibrium outcome of the observed case is due to the mechanism of monopoly screening, our model also contains signaling, which has critical impacts on the equilibrium and welfare. First, signaling may lead to discontinuity in the price scheme T^o. Specifically, if $z^o(\theta)$ contains a pooling interval on $[\theta_0^o, \bar{\theta}]$, then due to the continuity of $z^o(\theta)$, W^o has a discontinuity at such $z^o(\theta)$ that is constant on an interval. Thus, by the continuity of $U(\theta)$, T^o is discontinuous at such $z^o(\theta)$ too. In comparison, recall that in classic screening models such as Maskin and Riley (1984) in which signaling is absent, the associated price scheme is continuous even if ironing occurs.

Whenever $z^o(\theta)$ is increasing, T^o is differentiable. Given T^o, the subgame is essentially Spence's signaling game as if the worker had a cost function in the form of $C(z, \theta) + T^o(z)$. According to the same argument as in Section 2.3, the education levels in the observed case are distorted, due to signaling, above the full-information level with respect to the total cost of education. These points reveal that the equilibrium outcome of the observed case results from the interaction between screening and signaling.

Corollary 1 indicates that when both screening and signaling are present and exert the opposite effects—screening induces under-education, but signaling induces over-education—screening outweighs signaling. This is because as a Stackelberg leader, the school internalizes the worker's signaling behavior when screening his type. To show this, assume that $z^o(\theta)$ is increasing, then by the worker's best response, for each $z \in [z^o(\theta_0^o), z^o(\bar{\theta})]$, we have

$$T^{o\prime}(z) = W^{o\prime}(z) - C_z(z, \theta^o(z)) = \frac{d}{dz}\left[Q(z, \theta^o(z))\right] - C_z(z, \theta^o(z)).$$

Substituting this equation into the first-order condition of $MP^o(z, \theta)$, we have

$$T^{o\prime}(z) = Q_\theta(z, \theta^o(z))\theta^{o\prime}(z) + \frac{1 - F(\theta^o(z))}{f(\theta^o(z))}\left[-C_{z\theta}(z, \theta^o(z))\right]. \tag{3.3}$$

On the right-hand side (RHS) of (3.3), the first term is the signaling effect, and the second term is the marginal information rent extracted by the worker. Note that signaling induces over-education, which reduces the school's profit in two ways: on the one hand, it reduces total surplus; on the other hand, it provides the worker with more information rents. Thus, the optimal tuition scheme must undo these two effects, as indicated by (3.3). In contrast, if the school were a welfare-maximizing social planner, it would only undo the signaling effect by levying Pigovian taxes (e.g., Spence 1974). Denote by T^{fb} the welfare-maximizing tax on education. The marginal tax is equal to the signaling effect at the first-best, i.e.,

$$T^{fb\prime}(z) = Q_\theta(z, \theta^{fb}(z))\theta^{fb\prime}(z). \tag{3.4}$$

Because the second term on the RHS of (3.3) is positive, we have that the profit-maximizing scheme "over-taxes" signaling activity and thus leads to under-investment.

To see further how signaling makes a difference, consider the situation where the labor market also observes the worker's ability without changing any other element of the model. Thus, the wage equals the actual productivity, and signaling is eliminated. This means that the worker's willingness to pay for education is the social surplus $S(z, \theta)$. Because $S_\theta > 0$, a higher type can be seen as a higher-value buyer of education. Thus, the school has the same screening problem as in Mussa and Rosen (1978). Analogously, an allocation $\{z(\theta), U(\theta)\}$ is implementable if and only if (i) $z(\theta)$ is nondecreasing; (ii) $U(\theta_0) \geq 0$ and for $\theta > \theta_0$,

$$U(\theta) = U(\theta_0) + \int_{\theta_0}^{\theta} S_\theta(z(s), s)ds.$$

Thus, the school's problem in such a Mussa and Rosen's screening game can be stated as

$$\max_{\{z, \theta_0\}} \int_{\theta_0}^{\bar{\theta}} \left[S(z(\theta), \theta) - \frac{1 - F(\theta)}{f(\theta)} S_\theta(z(\theta), \theta)\right] dF(\theta)$$

subject to $z(\theta)$ being nondecreasing.

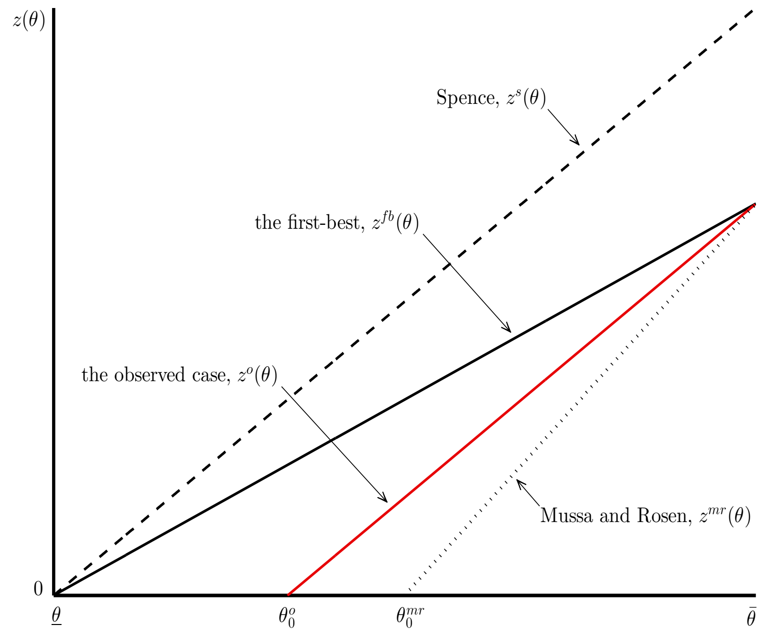

Figure 1: **Screening vs. Signaling.** This figure compares $z^{mr}(\theta)$ with $z^o(\theta)$ over $[\underline{\theta}, \bar{\theta}]$ along with $z^{fb}(\theta)$ and $z^s(\theta)$. This figure assumes that $Q(z,\theta) = \theta z + z$, $C(z,\theta) = z^2 + z - \theta z$, and $\theta \sim U[0,1]$. Therefore, $z^{fb}(\theta) = \theta$, $z^s(\theta) = 3\theta/2$, $z^o(\theta) = (3\theta - 1)/2$, and $z^{mr}(\theta) = 2\theta - 1$.

Define analogously the school's marginal profit as

$$MP^{mr}(z,\theta) \coloneqq S(z,\theta) - \frac{1 - F(\theta)}{f(\theta)} S_\theta(z,\theta).$$

Denote $z^{mr}(\theta)$ and θ_0^{mr} the optimal allocation and cutoff type, respectively. For simplicity, assume that both $MP^o(z,\theta)$ and $MP^{mr}(z,\theta)$ are regular.[6] Thus, $z^{mr}(\theta)$ and θ_0^{mr} can be characterized by pointwise maximization for $MP^{mr}(z,\theta)$.

We shall examine how the allocation in Mussa and Rosen's model differs from that in the observed case. On the extensive margin, because $S_\theta > -C_\theta$, $MP^{mr}(z,\theta) \leq MP^o(z,\theta)$, with strict inequality for $\theta < \bar{\theta}$. Hence, if $\theta_0^o > \underline{\theta}$, then $\theta_0^{mr} > \theta_0^o$; that is, more types are excluded in Mussa and Rosen's model. On the intensive margin, if $Q_{z\theta} > 0$ on $[0, z^{fb}(\bar{\theta})]$,[7] then $z^{mr}(\theta) \leq z^o(\theta)$, with strict inequality on $[\theta_0^o, \bar{\theta})$, meaning that under-education is more significant in Mussa and Rosen's model. These findings are illustrated in Figure 1.

For welfare comparison, note that education is already under-supplied in the observed case, yet the downward distortion is larger in Mussa and Rosen's model; thus, the observed case has higher social welfare. Moreover, since $MP^{mr}(z^{mr}(\theta),\theta) \leq MP^o(z^o(\theta),\theta)$ with strict

[6] Given Assumption 1, $MP^{mr}(z,\theta)$ is regular if $Q_{z\theta\theta} \leq 0$.

[7] This condition is not restrictive; indeed, given that $Q_\theta > 0$ if $z > 0$ and $Q(0,\theta) \equiv 0$, we have $Q_{z\theta} > 0$ for $z \in [0,k]$ for some $k > 0$. Given this condition, $MP_z^{mr}(z,\theta) < MP_z^o(z,\theta)$ on $[0, z^{fb}(\bar{\theta})]$ for $\theta < \bar{\theta}$.

inequality on $[\theta_0^o, \bar{\theta})$ and $\theta_0^{mr} \geq \theta_0^o$, it is readily confirmed that the school's expected profit is also higher in the observed case. In summary, we have the following proposition:

Proposition 2. *If both $MP^o(z,\theta)$ and $MP^{mr}(z,\theta)$ are regular, and $Q_{z\theta} > 0$ on $[0, z^{fb}(\bar{\theta})]$, then under-education is greater when signaling is eliminated. Specifically, $z^{mr}(\theta) \leq z^o(\theta)$, with strict inequality on $[\theta_0^o, \bar{\theta})$; if $\theta_0^o > \underline{\theta}$, then $\theta_0^{mr} > \theta_0^o > \underline{\theta}$. Consequently, social welfare and the school's expected profit are strictly higher when signaling is present than otherwise.*

Proposition 2 indicates that signaling can mitigate the downward distortion caused by screening. Intuitively, when the labor market observes the worker's ability, if a higher type imitates a lower type by choosing the same education, he not only has a lower total cost than the latter but also obtains a higher wage due to his higher productivity. In contrast, when the labor market does not observe the worker's ability, the higher type can no longer directly reap the benefit from higher productivity, and thus, he acquires more education to signal his ability. The signaling incentive reduces the worker's willingness to imitate lower types. Therefore, the school leaves lower information rents to the worker when signaling is present, as we have the following inequality for all $\theta \in [\underline{\theta}, \bar{\theta}]$,

$$\underbrace{\frac{1-F(\theta)}{f(\theta)}[-C_\theta(z,\theta)]}_{\text{information rents with signaling}} \leq \underbrace{\frac{1-F(\theta)}{f(\theta)}S_\theta(z,\theta)}_{\text{information rents without signaling}}$$

which holds with equality at $\bar{\theta}$ only. Consequently, the screening distortion is mitigated.

Recall that in Spence's game, signaling reduces social welfare, as it causes over-education. In the observed case, by contrast, signaling raises social welfare, relative to Mussa and Rosen, because it mitigates the screening distortion. Thus, any instrument that attenuates signaling is socially beneficial in the Spencian world but harmful in the observed case. For example, students' grades substitute for their education levels in signaling.[8] Suppose grades are less informative, e.g., due to grade inflation, an increasingly common phenomenon at American colleges and universities,[9] then signaling through education will be enhanced, as students will attempt to separate themselves from others (Daley and Green, 2014). This reveals that coarse grading can be socially beneficial in the observed case by alleviating under-education,[10] whereas it is harmful in the Spencian world because it aggravates over-education.

[8]See Appendix A.1 for a detailed discussion.

[9]See, for example, Johnson (2006) and Rojstaczer and Healy (2010).

[10]Alternatively, Boleslavsky and Cotton (2015) shows that coarse grading can improve social welfare by enhancing schools' investments in education quality when schools compete in placing graduates.

4 Labor Market Does Not Observe Tuition

In this section, we turn to the case in which the labor market does not observe the tuition scheme. Given some wage schedule W, the school solves the problem in (2.2). Similar to the observed case, we define the school's marginal profit in the unobserved case as

$$MP^u(z,\theta) := W(z) - \left[C(z,\theta) - \frac{1-F(\theta)}{f(\theta)}C_\theta(z,\theta)\right].$$

It is heuristic to call the terms in the bracket the school's *virtual cost*, and we define

$$G(z,\theta) := C(z,\theta) - \frac{1-F(\theta)}{f(\theta)}C_\theta(z,\theta).$$

Given Assumption 1, we have $G_{z\theta} < 0$; thus, the single-crossing property holds. This means that it is less costly for the school to serve a higher-ability worker.

In contrast to the observed case, here we use a more permissive equilibrium concept, i.e., a continuous equilibrium, such that on the equilibrium path, $z(\theta)$ is continuous if $z > 0$.[11] The following theorem indicates that there are at most two types of continuous equilibrium in the unobserved case. Specifically, there always exists a separating equilibrium such that $z^u(\theta)$ is one-to-one on $[\theta_0^u, \bar{\theta}]$. Moreover, under certain conditions, there also exists a pooling equilibrium such that $z^u(\theta)$ is a positive constant on $[\theta_0^u, \bar{\theta}]$. Formally, we have:

Theorem 1. *In the unobserved case, a continuous equilibrium is either separating or pooling. Specifically, there always exists a separating equilibrium in which $z^u(\theta)$ is increasing on $[\theta_0^u, \bar{\theta}]$ for some $\theta_0^u \geq \theta_0^o$, and satisfies the following differential equation:*

$$Q_z(z^u(\theta),\theta) + Q_\theta(z^u(\theta),\theta)\theta^{u\prime}(z^u(\theta)) - G_z(z^u(\theta),\theta) = 0. \tag{4.1}$$

In particular, the seller-optimal separating equilibrium such that $(\theta_0^u, z^u(\theta_0^u)) = (\theta_0^o, z^*(\theta_0^o))$ always exists, and is the unique equilibrium if $z^*(\underline{\theta}) > 0$ and no pooling equilibrium exists.

Theorem 1 states that the seller-optimal separating equilibrium always exists, and is the unique continuous equilibrium (outcome) under certain conditions. To see the uniqueness, note that if $z^*(\underline{\theta}) > 0$, then exclusion is never profitable, and thus, $\theta_0^u = \underline{\theta}$ in any equilibrium. It thus follows that the unique separating equilibrium is the seller-optimal one. Moreover, as we show in the proof, if there does not exist any positive education level such that both the lowest and highest types prefer this education level to their respective best alternatives, then a pooling equilibrium does not exist, and thus, the continuous equilibrium is unique.

[11] We do not impose the restriction of continuity on $z(\theta)$ off the equilibrium path.

However, in view of the possibility of multiple equilibria, we propose a novel equilibrium refinement, the *Quasi-Divinity*.[12] The definition of the Quasi-Divinity is as follows:

Definition 1. *In the unobserved case, an equilibrium satisfies the Quasi-Divinity refinement if there does not exist an out-of-equilibrium signal $\hat{z}$, a receiver's response $\hat{w}$, and a positive-measure subset $\hat{\Theta} \subset [\underline{\theta}, \bar{\theta}]$ such that*

(i) The following allocation of signals

$$z^d(\theta) = \begin{cases} \hat{z} & \text{if } \theta \in \hat{\Theta} \\ z^u(\theta) & \text{otherwise} \end{cases}$$

is nondecreasing in θ.

(ii) $\hat{w} - G(\hat{z}, \theta) > W^u(z^u(\theta)) - G(z^u(\theta), \theta)$ *if and only if* $\theta \in \hat{\Theta}$.

(iii) $\hat{w} < \mathbb{E}^{\hat{z}}[Q(\hat{z}, \theta) | \theta \in \hat{\Theta}]$*, with the expectation formed under any quasi-divine belief, i.e., any receiver's posterior belief that has a distribution function $F^{\hat{z}}$ with* $\text{supp}(F^{\hat{z}}) = \hat{\Theta}$.

Definition 1 states that an equilibrium fails the Quasi-Divinity if there exists an off-path education level and a wage, such that the school can make a profitable deviation by choosing some tuition scheme under which a subset of worker types is willing to choose the off-path education level for that wage whereas all the other types prefer their equilibrium allocations, and that the employers are willing to offer that wage for that education level, so long as they believe that they are facing a type from that subset, no matter how pessimistic this belief is.

The idea of the Quasi-Divinity is that by choosing the off-path signal $\hat{z}$, the worker sends an implicit message to the employers: "Although you cannot observe the tuition scheme, it would be strictly profitable for you to offer me $\hat{w}$. This is because under the tuition scheme that the school *actually* offers, the set of types who prefer $(\hat{z}, \hat{w})$ to any other pair $(z, W^u(z))$ is $\hat{\Theta}$, and for any quasi-divine belief you may have, you will receive a payoff strictly higher than your equilibrium payoff, i.e., $\mathbb{E}^{\hat{z}}[Q(\hat{z}, \theta) | \theta \in \hat{\Theta}] - \hat{w} > 0$." Anticipating the hypothetical speech by the worker and that the employers would think it through, the school indeed has incentives to choose some tuition scheme to implement $z^d(\theta)$ given the new wage, as doing so can yield a strictly higher profit, while making the worker's statement credible. Specifically, the monotonicity of $z^d(\theta)$ ensures that the new allocation is implementable, and that $\hat{\Theta}$ has a positive measure, combined with condition (*ii*) of Definition 1, ensures that the school will be strictly better-off than in the equilibrium, should the employers offer $\hat{w}$ for $\hat{z}$.

[12] I am indebted to Prof. John Riley for his exceptional doctoral class at UCLA that inspired this concept.

The Quasi-Divinity is closely related to the classic refinements for signaling games, such as the Divinity (Banks and Sobel, 1987) and the Intuitive Criterion (Cho and Kreps, 1987). The key difference is that the Quasi-Divinity examines directly the receiver's responses off the equilibrium path, instead of restricting the off-path beliefs. The critical test is whether there is an off-path signal and a subset of sender types, so that the receiver has a dominant response in terms of quasi-divine beliefs, and the seller can strictly profit by inducing only those types to choose the off-path signal. Furthermore, quasi-divine belief is less restrictive than divine belief, as the latter includes beliefs that place probability one on a single type. But in our continuous-type setup, any response that yields a strictly higher marginal profit than the equilibrium level at some type also makes it more profitable at each type in a small neighborhood; thus, the Quasi-Divinity is more applicable to the current model.

The next theorem indicates that the only continuous equilibrium of the unobserved case that satisfies the Quasi-Divinity is the seller-optimal separating equilibrium.

Theorem 2. *In the unobserved case, the unique continuous equilibrium satisfying the Quasi-Divinity is the seller-optimal separating equilibrium in which* $(\theta_0^u, z^u(\theta_0^u)) = (\theta_0^o, z^*(\theta_0^o))$.

Intuitively, due to the single-crossing property, in a pooling equilibrium for any education level $\hat{z}$ that is slightly higher than the equilibrium level, there exists a dominant wage $\hat{w}$ for the employers, such that the school can profitably deviate to $\hat{z}$ if and only if the worker's type is sufficiently high. Thus, those types can generate a profitable and credible deviation. Moreover, in any separating equilibrium other than the seller-optimal one, the participation constraint is binding at some $\theta_0^u > \theta_0^o$ with $z^u(\theta_0^u) > z^*(\theta_0^u)$. Thus, for any $\hat{z}$ that is slightly lower than $z^u(\theta_0^u)$, there exists a dominant wage such that the school can profitably deviate to $\hat{z}$ in some neighborhood of θ_0^u. By contrast, in the seller-optimal separating equilibrium, that is the Riley outcome, the cutoff type has achieved the "full-information" optimum, i.e., $z^u(\theta_0^u) = z^*(\theta_0^u)$. This implies that for sufficiently pessimistic quasi-divine beliefs, there does not exist an off-path signal to which the school can profitably deviate. In what follows, we shall focus on the seller-optimal separating equilibrium.

Our third theorem presents the paper's main result. In contrast with the observed case, the worker acquires more education in the unobserved case. In particular, any worker type above the cutoff type chooses a strictly higher education level than in the observed case.

Theorem 3. *In contrast with the observed case, the worker acquires more education in the unobserved case. Specifically,* $z^u(\theta) \geq z^o(\theta)$ *on* $[\underline{\theta}, \bar{\theta}]$*, with strict inequality for* $\theta > \theta_0^u$.

As an immediate result of Theorem 3, the school's equilibrium payoff in the unobserved case, denoted Π^u, is lower than that in the observed case, denoted Π^o. Thus, we have:

Corollary 2. *In contrast with the observed case, the school gains a strictly lower expected profit in the unobserved case. That is,* $\Pi^u < \Pi^o$.

The difference between the observed and unobserved cases is driven by a signal jamming effect. The worker's signal is jammed in the unobserved case because the labor market does not observe the actual cost of education. Specifically, the labor market cannot distinguish the impact of a change in tuition from that of cost heterogeneity on the change in education. To illustrate, suppose that the school lowers tuition so that the worker chooses more education than in the initial state. When the labor market observes the tuition change, it cuts wages, as any education level now corresponds to a lower-ability worker. In contrast, when the labor market does not observe the tuition change, it does not adjust wages despite that tuition changes; thus, the worker is willing to pay more for additional education. Conversely, if the school raises tuition such that education decreases, then the labor market will *raise* wages in the observed case; thus, the worker's willingness to pay is lower in the unobserved case. This reveals that the worker is more sensitive to tuition changes in the unobserved case.

From the school's perspective, the demand is more elastic in the unobserved case. Note that the LHS of (4.1) represents the marginal profit of education in the unobserved case; the second term represents the signal jamming effect and is positive. In comparison, in the observed case, the marginal profit of education is given by

$$MP^o_z(z, \theta) = Q_z(z, \theta) - G_z(z, \theta).$$

Thus, the school's marginal profit is higher in the unobserved case than in the observed case. This provides the school with an incentive to "fool" the labor market with secret price cuts; that is, the school secretly supplies more education and persuades the labor market that the worker is more productive than is actually the case. In equilibrium, the labor market correctly anticipates the school's incentive and offers lower wages, as education is inflated. This reduces the worker's willingness to pay, and thus, the school achieves lower profits.

4.1 Implications for Price Transparency

We have shown that tuition cuts lead to smaller increases in demand in the observed case than in the unobserved case. This is because when the tuition cuts are publicly observed, the increase in demand is mitigated by the cheaper tuition reducing the signaling value of education. Thus, tuition cuts are less profitable in the observed case. Furthermore, we show that education is more expensive in the observed case. Specifically, the tuition scheme in the

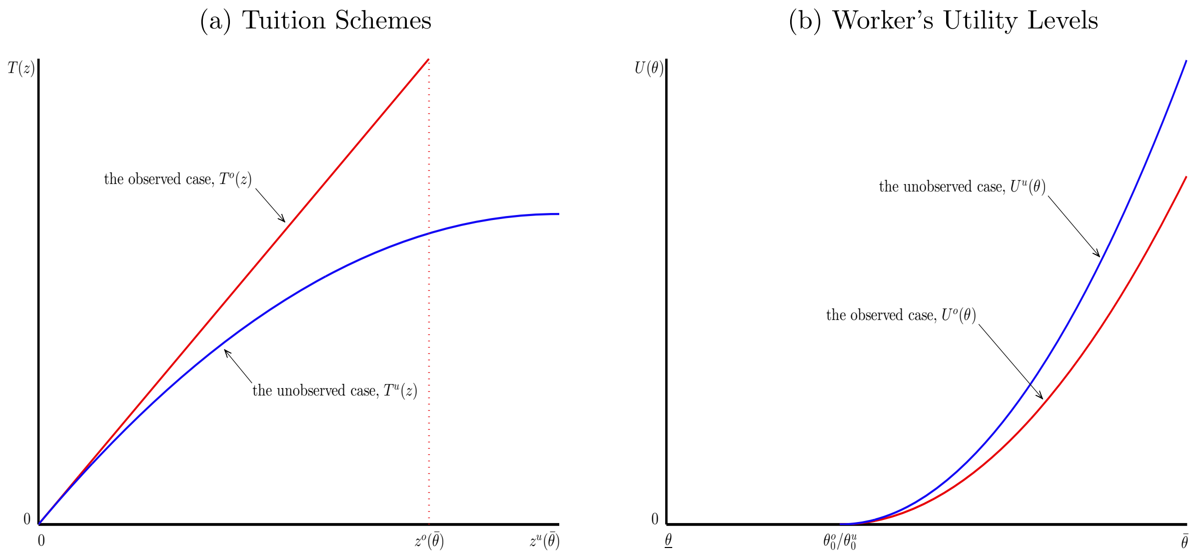

Figure 2: **Implications for Tuition Transparency.** This figure compares the tuition and the worker's utility level between the observed and unobserved cases. This figure considers the same numerical example as Figure 1, such that *(a)* $T^o(z) = \frac{2z}{3}$ and $T^u(z) = -\frac{z^2}{4} + \frac{2z}{3}$; *(b)* $U^o(\theta) = \frac{3}{4}(\theta - \frac{1}{3})^2$ and $U^u(\theta) = (\theta - \frac{1}{3})^2$.

unobserved case is uniformly lower than that in the observed case on the common interval of education, i.e., $[z^*(\theta_0^o), z^*(\bar{\theta})]$. This is illustrated in Panel (a) of Figure 2. To summarize:

Proposition 3. *$T^o(z) \geq T^u(z)$ on $[z^*(\theta_0^o), z^*(\bar{\theta})]$, with strict inequality for $z > z^*(\theta_0^o)$.*

Furthermore, by the worker's best response in the unobserved case, we have

$$T^{u\prime}(z) = W^{u\prime}(z) - C_z(z, \theta^u(z)).$$

Substituting this equation into (4.1), and noticing that $W^u(z) = Q(z, \theta^u(z))$, we obtain

$$T^{u\prime}(z) = \frac{1 - F(\theta^u(z))}{f(\theta^u(z))} \left[-C_{z\theta}(z, \theta^u(z))\right]. \tag{4.2}$$

Equation (4.2) states that in the unobserved case, the marginal tuition equals the marginal information rent extracted by the worker. In contrast to the observed case, as indicated by the comparison between (4.2) and (3.3), the optimal tuition scheme in the unobserved case does not undo the signaling effect. The reason is that the loss in the total surplus caused by over-education will be compensated by the labor market overpaying the worker, as the labor market will overestimate the worker's ability if the school secretly cuts tuition. In addition, (4.2) states that the marginal tuition vanishes at the highest education level. This implies that the school offers quantity discounts (i.e., $T(z)/z$ is declining) for higher education levels in the unobserved case. This echoes the classic screening model of Maskin and Riley (1984), in which quantity discounts are also optimal at the right tail of the distribution.

In terms of the worker's payoff, note that in both cases the market belief about tuition is correct in equilibrium; thus, given the equilibrium tuition scheme, the continuation game is indeed Spence's signaling game as if the worker's cost function were given by the total cost. Because the tuition scheme is uniformly lower in the unobserved case, the signaling costs are lower in this case. Consequently, the worker has a higher utility level in the unobserved case than in the observed case. To be formal, let $U^o(\theta)$ and $U^u(\theta)$ be type θ's equilibrium payoff in the observed and unobserved cases, respectively. By Theorem 3, for any $\theta \in (\theta_0^o, \bar{\theta}]$,

$$U^o(\theta) - U^u(\theta) = \int_{\theta_0^o}^{\theta} [C_\theta(z^u(s), s) - C_\theta(z^o(s), s)]ds < 0.$$

This is illustrated in Panel (b) of Figure 2. To summarize:

Proposition 4. *$U^o(\theta) \leq U^u(\theta)$ on $[\underline{\theta}, \bar{\theta}]$, with strict inequality for $\theta > \theta_0^o$.*

Propositions 3 and 4 indicate that policies that improve the transparency of net prices at colleges and universities through mandatory disclosure may unintentionally induce more expensive education and harm students. These policies, such as U.S. Code § 1015a, require colleges to publicly disclose their net prices, which are usually not previously observed by employers. On the one hand, this reduces the search costs of students, thereby stimulating the competition between schools and lowering prices; on the other hand, this also allows schools to commit to high prices and not dilute the signaling value of a high-cost education by means of fee waivers, financial aid and so forth. It is thus possible that such policies ultimately raise education costs and harm students. Hence, policymakers should not overlook the unintended negative effects of these mandatory disclosure policies.

Consider the application of conspicuous consumption. The previous results suggest that the demand for luxury good is more elastic when the social contact cannot observe the price scheme than otherwise. In equilibrium, the retailer charges lower prices, and the consumer chooses a higher end of the good than in the observed case. Whereas the consumer obtains a higher utility level than in the observed case, the retailer receives lower profits. In reality, many luxury brands strive to build a reputation of never being on sale by establishing own stores, going direct-to-consumer and so forth. Such strategies help the sellers better commit to high prices, thereby maintaining the signaling values of luxury goods.

Analogously, in the case of advertising, the media company charges lower prices, and the producer chooses a higher advertising level, thereby obtaining a larger market share when consumers cannot observe the price scheme than otherwise. Whereas the producer obtains a higher payoff, the media company gains lower profits than in the observed case. In reality, some media companies run open auctions for their advertising services and widely report the winning prices afterward. By doing so, the media companies accentuate the signaling value of their costly advertising messages.

4.2 Welfare Analysis

We have shown that in the unobserved case, the worker chooses more education than in the observed case. For completeness, we compare the education function in the unobserved case with other benchmarks in this paper. The next proposition states that the education levels in the unobserved case are bounded above by that of Spence's signaling game. This result is illustrated in Figure 3.

Proposition 5. *In the unobserved case, the worker acquires less education than in Spence's signaling game. Specifically,* $z^u(\theta) \leq z^s(\theta)$ *on* $[\underline{\theta}, \bar{\theta}]$*, with strict inequality for* $\theta > \underline{\theta}$*.*

The intuition is clear: The unobserved case is essentially Spence's signaling game with higher costs, and thus, it leads to lower education levels than Spence's model.

To compare $z^u(\theta)$ with $z^{fb}(\theta)$, note that $z^u(\theta_0^o) = z^*(\theta_0^o) \leq z^{fb}(\theta_0^o)$ and $z^u(\bar{\theta}) > z^o(\bar{\theta}) = z^{fb}(\bar{\theta})$. Then, continuity implies that $z^u(\theta)$ intersects $z^{fb}(\theta)$ at least once. Moreover, under some mild conditions—the following Assumption 2, for example—$z^u(\theta)$ is single-crossing $z^{fb}(\theta)$; that is, there exists a cutoff type, such that all the lower types obtain less education than the first-best, whereas the others obtain more than the first-best (see Figure 3).

Assumption 2. *The function*

$$Q_\theta(z^{fb}(\theta), \theta)\theta^{fb\prime}(z) + \frac{1 - F(\theta)}{f(\theta)} C_{z\theta}(z^{fb}(\theta), \theta)$$

is single-crossing in θ*, with a unique root* $\theta^* \in (\underline{\theta}, \bar{\theta})$*.*[13]

Proposition 6. *Suppose that Assumption 2 holds, then there exists a type* $\theta^w \in (\theta^*, \bar{\theta})$ *such that* $z^u(\theta) < z^{fb}(\theta)$ *on* $[\underline{\theta}, \theta^w)$ *and* $z^u(\theta) > z^{fb}(\theta)$ *on* $(\theta^w, \bar{\theta}]$*.*

In the unobserved case, there are two competing forces that pull the education function away from the first-best level. On the one hand, the signal jamming effect induces the school to supply more education. On the other hand, more education means more rent extraction by the worker. Because the information rent ultimately vanishes as type approaches the top, the school unambiguously over-supplies education on some upper interval of the spectrum.

[13] A function $g(x)$ is single-crossing in x if given some x^*, $g(x) < 0$ for $x < x^*$ and $g(x) > 0$ for $x > x^*$.

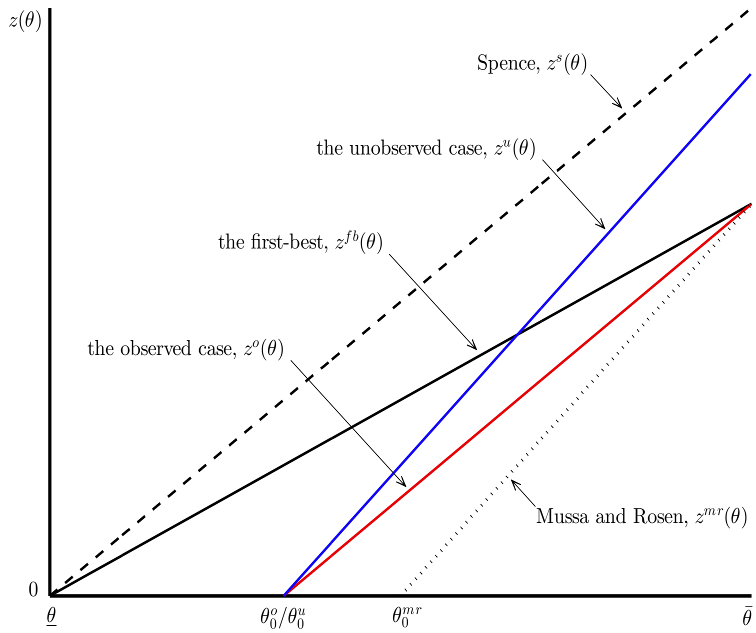

Figure 3: **All Education Functions.** This figure illustrates all the equilibrium education functions that we have discussed in this paper. This figure considers the same numerical example as Figure 1, such that $z^u(\theta) = 2\theta - \frac{2}{3}$; also recall that $z^{fb}(\theta) = \theta$, $z^s(\theta) = \frac{3}{2}\theta$, $z^o(\theta) = \frac{3\theta-1}{2}$ and $z^{mr}(\theta) = 2\theta - 1$.

Assumption 2 ensures that the relative significance of the two forces alters only once, thus it rules out the possibility of multiple intersections between $z^u(\theta)$ and $z^{fb}(\theta)$. Proposition 6 means that under-education is slighter on a lower interval of the spectrum in the unobserved case than in the observed case; it also provides a lower bound for the length of this interval. However, because over-education also occurs in the unobserved case, whether the observed or unobserved case yields higher social welfare is ambiguous in general.

Recall that in the observed case, signaling can mitigate the downward distortion caused by screening. Will Signaling being more intense lead to more distortion cuts in the observed case, whereas more over-investment in the unobserved case? Indeed, it can be shown that if signaling is sufficiently intense (e.g., there is significant over-investment in Spence's model), then the observed case yields higher social welfare than the unobserved case. Furthermore, both cases yield higher social welfare than Spence's model if signaling is intense.

This finding thus has welfare implications for the market structure of signals. When the market is served by perfectly competitive sellers of signals, the equilibrium is predicted by Spence's model. In contrast, when the market is served by a monopolist, the equilibrium is predicted by the current model, which suggests that when the buyer's signaling incentive is relatively strong, monopoly can yield higher social welfare than a competitive market. That is, promoting competition in a signaling good market is not necessarily socially beneficial.

5 Extensions

In this section, we examine how the main results would be affected if we change some of the assumptions. Specifically, we will analyze the cases of nonessential signals and unproductive signals, and the case in which the seller is partially profit-maximizing. To be consistent, we shall focus on the application of job market signaling.

5.1 Nonessential Signals

As a first extension, we relax the model assumption that $Q(0, \theta) \equiv 0$, and assume instead that $Q_\theta > 0$ for all $(z, \theta) \in \mathbb{R}_+ \times [\underline{\theta}, \bar{\theta}]$, with $Q(0, \underline{\theta})$ normalized to 0; all the other elements of the model remain unchanged. We start with the observed case.

Suppose that $z^*(\underline{\theta}) > 0$, then the allocation given by (3.1) constitutes the seller-optimal equilibrium of this case as well. To see this, assume that for all z outside the range of $z^o(\theta)$, the school charges exorbitantly high prices and the labor market has the worst belief; thus, the worker's outside option has a utility $Q(0, \underline{\theta}) = 0$. It follows that $z^o(\theta)$ with $\theta_0^o = \underline{\theta}$ solves the school's problem, and yields the highest possible equilibrium payoff for the school.

Now consider the case in which $\inf\{\theta | z^*(\theta) > 0\} > \underline{\theta}$. Denote the inferior $\underline{\theta}_0$, and define

$$\bar{\Pi}^o := F(\underline{\theta}_0)\mathbb{E}[Q(0, \theta)|\theta \leq \underline{\theta}_0] + \int_{\underline{\theta}_0}^{\bar{\theta}} MP^o(\bar{z}(\theta), \theta)dF(\theta),$$

where $\bar{z}(\theta)$ is given as in Section 3. It can be shown that for any fixed $\delta > 0$, there exists a subgame associated with some tuition scheme such that the school's equilibrium payoff in this subgame is larger than $\bar{\Pi}^o - \delta$ (see Lemma 4 in Section A.4). It follows that if an equilibrium exists, then there is an equilibrium such that $\Pi^o \geq \bar{\Pi}^o - \delta$.

Thus, to achieve the seller-optimal equilibrium (as $\delta \to 0$), we allow the school to charge a fixed fee for zero education, as if the school could sell the worker a certification with no information disclosure as in Lizzeri (1999). Observing that the worker paid the fixed fee, the labor market receives a simple message that the worker is certified by the school, and thus, offers a wage equal to the average productivity of those who paid the fee. Assume that the labor market regards a worker with neither education nor a certification as the lowest type. Thus, the optimal cutoff type is $\theta_0^o = \underline{\theta}_0$, and the optimal fixed fee equals $\mathbb{E}[Q(0, \theta)|\theta \leq \underline{\theta}_0]$ such that any type $\theta \leq \underline{\theta}_0$ will be indifferent. Then, the allocation given by (3.1), combined with the fixed fee, leads to the seller-optimal equilibrium such that $\Pi^o = \bar{\Pi}^o$. Note that in the equilibrium, the market is fully covered and consists of two segments:

(a) Certification segment, $[\underline{\theta}, \theta_0^o]$, in which the worker pays a fixed fee for zero education.

(b) Education segment, $(\theta_0^o, \bar{\theta}]$, in which the worker purchases a positive education level.

As in Lizzeri (1999), the certifier reveals nothing about the agent's type and extracts all the information rents. However, whereas this extreme equilibrium outcome is sustained by a particular belief in our model, it is the unique equilibrium outcome under certain conditions in Lizzeri's model (Lizzeri, 1999, Theorem 3). This is because unlike in our model, in Lizzeri's the certifier can truthfully reveal the agent's type without cost, and thus, can always induce higher types to participate by revealing the highest type with relatively high probability and the other types with sufficiently low probability.

We then turn to the unobserved case. Given the nature of the information structure, we shut down signaling through certification (money burning); thus again, $T(0) = 0$. Suppose that an equilibrium exists, then the school's profit is given by

$$\Pi^u = \int_{\theta_0^u}^{\bar{\theta}} MP^u(z^u(\theta), \theta) dF(\theta) - [1 - F(\theta_0^u)]\mathbb{E}[Q(0, \theta)|\theta \leq \theta_0^u],$$

where the worker's reservation utility, $\mathbb{E}[Q(0, \theta)|\theta \leq \theta_0^u]$, is constant given the market belief. We argue that if $\theta_0^u > \underline{\theta}$, then $z^u(\theta)$ is discontinuous at θ_0^u with $z^u(\theta_0^u) > \lim_{\theta \uparrow \theta_0^u} z^u(\theta) = 0$; in particular, $z^u(\theta_0^u) > z^*(\theta_0^u)$ (see Lemma 5 in Section A.4). This implies that the unique continuous equilibrium that satisfies the Quasi-Divinity is the separating equilibrium with the initial point $(\underline{\theta}, z^*(\underline{\theta}))$, and that such an equilibrium exists. To summarize:

Theorem $2'$**.** *In the unobserved case, there exists a unique continuous equilibrium outcome satisfying the Quasi-Divinity, such that* $z^u(\theta)$ *is increasing on* $[\underline{\theta}, \bar{\theta}]$*, and satisfies (4.1) with the initial condition* $(\theta_0^u, z^u(\theta_0^u)) = (\underline{\theta}, z^*(\underline{\theta}))$*.*

Similar to Theorem 2, Theorem $2'$ indicates that the unique continuous equilibrium that satisfies the Quasi-Divinity is the Riley outcome, that is, the separating equilibrium in which the cutoff type θ_0^u chooses his "full-information" level $z^*(\theta_0^u)$. With nonessential education, if equilibrium exhibits exclusion, i.e., $\theta_0^u > \underline{\theta}$, then type θ_0^u must choose a higher education level than $z^*(\theta_0^u)$, so that the school has no incentive to induce lower types to mimic θ_0^u. It follows that in the Riley outcome, $(\theta_0^u, z^u(\theta_0^u)) = (\underline{\theta}, z^*(\underline{\theta}))$.

According to Theorem $2'$, $\theta_0^u \leq \theta_0^o$ and $z^u(\theta_0^u) = z^*(\theta_0^u)$. It follows that $z^u(\theta_0^u) \geq z^o(\theta_0^u)$. Then, by the same argument of Theorem 3, we have the following similar theorem.

Theorem $3'$**.** *In contrast with the observed case, the worker acquires more education in the unobserved case. Specifically,* $z^u(\theta) \geq z^o(\theta)$ *on* $[\underline{\theta}, \bar{\theta}]$*, with strict inequality for* $\theta > \underline{\theta}$*.*

Thus, one can show analogously that in the unobserved case, the school receives a lower expected profit and charges lower tuition fess, and the worker obtains a higher utility level than in the observed case (as stated by, respectively, Corollary 2, Propositions 3 and 4).

5.2 Unproductive Signals

Then, we consider the case when education is unproductive, i.e., $Q_z(z, \theta) \equiv 0$, which can be regarded as a limit case of nonessential education. Accordingly, we rewrite the productivity function as $Q(\theta)$, and assume that $Q_\theta > 0$ for all (z, θ), with $Q(\underline{\theta})$ normalized to 0.

In the observed case, because $z^*(\theta) \equiv 0$, the seller-optimal equilibrium contains only the certification segment: $z^o(\theta) \equiv 0$, and the school charges a certification fee equal to $\mathbb{E}[Q(\theta)]$. That is, the school extracts fully the surplus with no information disclosure at all.

In the unobserved case, we have that both Theorems $2'$ and $3'$ also hold here; thus, the implications for tuition, the school's and the worker's payoffs remain unchanged. Moreover, because $z^u(\theta) > z^o(\theta) = z^{fb}(\theta)$ on $(\underline{\theta}, \bar{\theta}]$, the unobserved case unambiguously yields a lower social welfare than the observed case. In addition, it is easy to show that $z^u(\theta) \leq z^s(\theta)$ with a strict inequality on $(\underline{\theta}, \bar{\theta}]$. Thus, Spence's model yields the lowest social welfare, compared with the observed and unobserved cases.

5.3 Partially Profit-Maximizing Seller

In the application of job market signaling, we assume that the school maximizes its expected profit. In reality, however, schools are typically not pure profit-maximizers. To this end, we study the school's pricing strategy when its objective is a weighted average of its profit and the worker's utility. Formally, given some wage schedule W, the school solves

$$\max_{z(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} [W(z) - C(z, \theta) - U(\theta)]\, dF(\theta) + \mu \int_{\underline{\theta}}^{\bar{\theta}} U(\theta) dF(\theta),$$

where $\mu \in (0, 1]$ denotes the relative Pareto weight. In particular, $\mu = 0$ leads to the original model in which the school maximizes its profit; $\mu = 1$ means that the school maximizes the joint surplus of the two parties. Assume that the rest parts of the model remain unchanged. Thus, by Lemma 1, the school's problem can be rewritten as

$$\max_{\{z, \theta_0\}} \int_{\theta_0}^{\bar{\theta}} \left[W(z(\theta)) - C(z(\theta), \theta) + (1 - \mu) \frac{1 - F(\theta)}{f(\theta)} C_\theta(z(\theta), \theta) \right] dF(\theta)$$

subject to $z(\theta)$ being nondecreasing.

In the observed case, analogously, the marginal profit from type θ is given by

$$MP^o(z,\theta;\mu) := S(z,\theta) + (1-\mu)\frac{1-F(\theta)}{f(\theta)}C_\theta(z,\theta).$$

Given Assumption 1, $MP^o(z,\theta;\mu)$ is strictly concave in z and $MP^o_z(z,\theta;\mu) \le S_z(z,\theta)$ with strict inequality for $\theta < \bar{\theta}$ and $\mu < 1$. Thus, all the results in Section 3 remain essentially. Denote by $z^o(\theta;\mu)$ and $\theta^o_0(\mu)$ the optimal allocation and cutoff type, respectively. We have that $z^o(\theta;\mu) \le z^{fb}(\theta)$, with strict inequality if $\mu < 1$ and $\theta \in (\underline{\theta},\bar{\theta})$. Moreover, if $MP^o(z,\theta)$ is regular, then so is $MP^o(z,\theta;\mu)$. It follows that $z^o(\theta;\mu) \ge z^o(\theta)$, with equality holding on the boundary only. That is, under-provision is mitigated with a partially profit-maximizing school. In particular, $z^o(\theta;1) \equiv z^{fb}(\theta)$, meaning that the tuition undoes the signaling effect. Thus, substituting $z^o(\theta;1)$ into (3.2), we obtain the welfare-maximizing tuition scheme.

Then, we turn to the unobserved case. Analogously, define the virtual cost as

$$G(z,\theta;\mu) := C(z,\theta) - (1-\mu)\frac{1-F(\theta)}{f(\theta)}C_\theta(z,\theta).$$

Given Assumption 1, $G_z > 0$ if $z > 0$, $G_{zz} > 0$ and $G_{z\theta} < 0$ for all $\mu \in (0,1]$. It follows that all the results in Section 4 remain essentially. In particular, the only continuous equilibrium that satisfies the Quasi-Divinity is the seller-optimal separating equilibrium with respect to $G(z,\theta;\mu)$. Let $z^u(\theta;\mu)$ and $\theta^u_0(\mu)$ be the equilibrium allocation and cutoff type, respectively. It follows from Theorem 3 that $z^u(\theta;\mu) \ge z^o(\theta;\mu)$, with strict inequality for $\theta > \theta^o_0(\mu)$. That is, the worker acquires more education in the unobserved case than in the observed case. In addition, because $G_z(z,\theta;\mu) \le G_z(z,\theta)$ on any common interval of (z,θ), $\theta^u_0(\mu) \le \theta^u_0$ and $z^u(\theta;\mu) > z^u(\theta)$ on $[\theta^u_0(\mu),\bar{\theta}]$; again, the worker acquires more education when the school is not a pure profit-maximizer. In particular, $z^u(\theta;1) \equiv z^s(\theta)$, meaning that the tuition scheme must be flat on $[z^s(\underline{\theta}), z^s(\bar{\theta})]$, and thus, screening is eliminated and signaling prevails.

It is worth noting that the extent to which the school maximizes its profit, measured by μ, has different welfare implications for the observed and unobserved cases. In the observed case, as μ rises, the screening distortion decreases and social welfare increases. In particular, when the school maximizes the joint surplus of itself and the worker ($\mu = 1$), the outcome is socially optimal. In the unobserved case, however, as μ rises from 0 to 1, the equilibrium switches from that of the original unobserved case to that of Spence's model, and thus, the welfare implication is ambiguous. In particular, if the original unobserved case yields higher social welfare than Spence's model, then a profit-maximizing school might be more socially beneficial than a school that maximizes the joint surplus of itself and its students, because the latter school charges such low tuition that there is significant over-education.

6 Conclusion

In this paper, we developed classic signaling models by letting a strategic player affect the signaling cost. A seller chooses a price schedule for a good, and a buyer with a hidden type chooses how much to purchase as a signal to receivers. The equilibrium depends critically on whether receivers observe the price schedule. In the observed case, the seller internalizes signaling in screening, causing a downward distortion. However, such a distortion is slighter than that of the case in which receivers observe the buyer's type. In the unobserved case, the buyer is more sensitive to price changes than in the observed case. This leads to a more elastic demand for signals and provides the seller with an incentive to lower prices. To refine the set of equilibria, we proposed a new refinement, the Quasi-Divinity. In equilibrium, the buyer chooses a higher quantity and obtains higher utility than in the observed case, whereas the seller gains lower profits than in the observed case. We showed that price transparency can be socially beneficial, and social welfare can be higher in a monopoly market than in a competitive market, when the buyer's signaling incentive is relatively strong. Our framework can be applied to schools choosing tuition, retailers selling luxury goods, media companies selling advertising messages, and to other vertical markets in which signaling prevails.

A Appendix

A.1 Proofs for Section 3

Proof of Proposition 1

Proof. Given Assumption 1, $MP^o(z, \theta)$ is strictly concave in z. It follows from Toikka (2011, Theorem 4.4) that the school's problem reduces to pointwise maximization for $\overline{MP}^o(z, \theta)$. Moreover, $\overline{MP}^o(z, \theta)$ is also strictly concave in z and thus has a unique maximizer on $\mathbb{R}+$, denoted $\bar{z}(\theta)$, which is continuous and nondecreasing on $[\underline{\theta}, \bar{\theta}]$. In addition, if $\bar{z}(\theta) \neq z^*(\theta)$ for some θ, then $J(\bar{z}(\theta), \theta)$ and $I(\bar{z}(\theta), \theta)$ differ in a neighborhood of θ, in which $I_\theta(\bar{z}(\theta), \theta)$ is flat; thus, $\bar{z}(\theta)$ is constant in this neighborhood. That is, if $\bar{z}(\theta) \neq z^*(\theta)$, then θ belongs to a *pooling* interval; outside these intervals, $\bar{z}(\theta) = z^*(\theta)$. Note that $MP^o_\theta(z, \theta) > 0$ if $z > 0$, and that $MP^o(0, \theta) \equiv 0$. Then by the envelope theorem, $MP^o(\bar{z}(\theta), \theta)$ is nondecreasing and nonnegative. Therefore, the cutoff type θ^o_0 is either the maximal root of $MP^o(\bar{z}(\theta), \theta) = 0$ if it exists, or $\underline{\theta}$ otherwise. In particular, if $\theta^o_0 > \underline{\theta}$, then $\bar{z}(\theta^o_0) = z^*(\theta^o_0) = 0$. Suppose not, then $\bar{z}(\theta^o_0) > z^*(\theta^o_0) > 0$, as $MP^o(z, \theta)$ is strictly concave in z; thus, $MP^o(\cdot, \theta^o_0) > 0$ on $(0, \bar{z}(\theta^o_0))$. Then, the school can profit by assigning any $z \in (0, \bar{z}(\theta^o_0))$ to some neighborhood of θ^o_0 with the monotonicity still holding, a contradiction. It follows that θ^o_0 is either the maximal root of $\bar{z}(\theta) = 0$ if it exists, or $\underline{\theta}$ otherwise, and that if $z^*(\underline{\theta}) > 0$, then $\theta^o_0 = \underline{\theta}$ and $\bar{z}(\underline{\theta}) \leq z^*(\underline{\theta})$. In summary, irrespective of whether $MP^o(z, \theta)$ is regular, $z^o(\theta)$ is given by (3.1). Then, the characterizations of T^o and W^o follow immediately. Thus, the proposition is proven. □

Proof of Corollary 1

Proof. Note that $MP^o_z(z, \theta)$ is less than $S_z(z, \theta)$, holding weakly on the boundary. Thus, we have $z^*(\theta) \leq z^{fb}(\theta)$ on $(\underline{\theta}, \bar{\theta}]$, with equality holding at $\bar{\theta}$ only. Moreover, $MP^o_z(z, \theta)$ must be increasing for θ close to $\bar{\theta}$, as $MP^o_{z\theta}(z, \bar{\theta}) > 0$. Thus, $J(z, \theta)$ coincides with $I(z, \theta)$ in some neighborhood of $\bar{\theta}$, meaning that $z^o(\theta)$ is increasing near $\bar{\theta}$, and thus, $z^o(\bar{\theta}) = z^{fb}(\bar{\theta})$. Note further that any θ satisfying $z^*(\theta) < z^o(\theta)$ belongs to some pooling interval $[\alpha, \beta]$ such that $z^o(\theta) = \min\{z^*(\alpha), z^*(\beta)\}$. Because $z^{fb}(\theta)$ is increasing over $[\underline{\theta}, \bar{\theta}]$, $z^o(\theta) \leq z^{fb}(\theta)$ on $(\underline{\theta}, \bar{\theta}]$, with equality holding at $\bar{\theta}$ only. Thus, the corollary is proven. □

Proof of a Generalized Version of Proposition 2

Here, we consider a more general information structure based on Mussa and Rosen's model in Section 3.1. Suppose that after the worker chooses his education level, the labor market can observe the worker's ability with a probability $p \in (0,1]$. For example, the worker takes some test in school; with a probability p, the test is perfectly informative, and reveals the worker's ability; otherwise, the test is completely uninformative about the worker's ability. The probability p thus measures the informativeness of the test; in particular, if $p = 1$, then we return to Mussa and Rosen's model. For simplicity, assume that p is independent of the worker's ability and education level. Thus, from the worker's (and the school's) perspective, his expected wage from choosing education level z is given by

$$\mathbb{E}[W(z)] = pQ(z,\theta) + (1-p)\mathbb{E}_\theta[Q(z,\theta)],$$

where the second expectation stands for the labor market's expectation about the worker's productivity. Then, given some tuition scheme T, the worker's expected utility is equal to

$$U(z,\theta) = \mathbb{E}[W(z)] - C(z,\theta) - T(z).$$

Analogous to Section 3.1, an allocation $\{z(\theta), U(\theta)\}$ is implementable if and only if (i) $z(\theta)$ is nondecreasing; (ii) $U(\theta_0) \geq 0$ and for $\theta \geq \theta_0$,

$$U(\theta) = U(\theta_0) + \int_{\theta_0}^{\theta} [pQ_\theta(z(s),s) - C_\theta(z(s),s)]\, ds.$$

Substituting and integrating by parts, the school's problem can be stated as

$$\max_{\{z,\theta_0\}} \int_{\theta_0}^{\bar{\theta}} \left\{ S(z(\theta),\theta) - \frac{1-F(\theta)}{f(\theta)} [pQ_\theta(z(s),s) - C_\theta(z(s),s)] \right\} dF(\theta)$$

subject to $z(\theta)$ being nondecreasing. Thus, the school's marginal profit is given by

$$MP^{mr}(z,\theta) := S(z,\theta) - \frac{1-F(\theta)}{f(\theta)} [pQ_\theta(z(\theta),\theta) - C_\theta(z(\theta),\theta)].$$

Assume that both $MP^o(z,\theta)$ and $MP^{mr}(z,\theta)$ are regular, so $z^{mr}(\theta)$ and θ_0^{mr} can be solved by pointwise maximization for $MP^{mr}(z,\theta)$. As a result, on the extensive margin, if $\theta_0^o > \underline{\theta}$, then $\theta_0^{mr} > \theta_0^o$; on the intensive margin, if $Q_{z\theta} > 0$ on $[0, z^{fb}(\bar{\theta})]$, then $z^{mr}(\theta) \leq z^o(\theta)$, with strict inequality on $[\theta_0^o, \bar{\theta})$. Clearly, these differences are reinforced when p increases.

Indeed, the optimal allocation rule $z^{mr}(\theta)$ varies continuously from that of the observed case to that of Mussa and Rosen's model, as p increases from 0 to 1. That is, the downward distortion caused by screening is intensified when signaling is attenuated. In summary, we have the following generalized version of Proposition 2.

Proposition 2′. *If both $MP^o(z,\theta)$ and $MP^{mr}(z,\theta)$ are regular, and $Q_{z\theta} > 0$ on $[0, z^{fb}(\bar{\theta})]$, then under-education is greater when signaling is attenuated. Specifically, $z^{mr}(\theta) \leq z^o(\theta)$, with strict inequality on $[\theta_0^o, \bar{\theta})$; if $\theta_0^o > \underline{\theta}$, then $\theta_0^{mr} > \theta_0^o > \underline{\theta}$. Consequently, social welfare and the school's expected profit are strictly lower when the incentive of signaling is weaker.*

A.2 Proofs for Section 4

Proof of Theorem 1

Proof. We first prove that a separating equilibrium exists. Let $MP(\theta, \hat{\theta}, z)$ be the marginal profit of type θ in the unobserved case when he obtains education level z and is believed as type $\hat{\theta}$. In particular, $MP(\theta, \theta, z) \equiv MP^o(z, \theta)$. It follows that $MP_2(\theta, \hat{\theta}, z) = Q_\theta(z, \hat{\theta}) > 0$ and is bounded, $MP_{13}(\theta, \hat{\theta}, z) = -G_{z\theta}(z, \theta) > 0$, and $MP_3(\theta, \hat{\theta}, z)/MP_2(\theta, \hat{\theta}, z)$ is increasing in θ on the relevant subset $[\theta_0^u, \bar{\theta}]^2 \times (0, \zeta]$ for some sufficiently large ζ. Then, appealing to Mailath and von Thadden (2013), we have that for any admissible initial point $(\theta_0^u, z^u(\theta_0^u))$, there exists a unique separating equilibrium outcome in which $z^u(\theta)$ maximizes $MP^u(z, \theta)$, satisfies (4.1), and is increasing on $[\theta_0^u, \bar{\theta}]$. Consider $(\theta_0^u, z^u(\theta_0^u)) = (\theta_0^o, z^*(\theta_0^o))$ as a candidate for equilibrium initial point. Assume that the labor market regards any off-path education level as chosen by the lowest type. There are two cases. First, if $\theta_0^o = \underline{\theta}$, then we have

$$MP^o(z^*(\underline{\theta}), \underline{\theta}) = \max_{z \geq 0} MP^o(z, \underline{\theta}) \geq MP^o(z^o(\underline{\theta}), \underline{\theta}) \geq 0.$$

Because types reveal in the equilibrium, we have $MP^u(z, \underline{\theta}) \leq MP^o(z^*(\underline{\theta}), \underline{\theta})$, with equality holding at $z^*(\underline{\theta})$ only. Thus, the school will not deviate from $(\theta_0^o, z^*(\theta_0^o))$. Second, if $\theta_0^o > \underline{\theta}$, then $z^o(\theta_0^o) = z^*(\theta_0^o) = 0$ and $MP^o(z^*(\theta_0^o), \theta_0^o) = 0$. Clearly, the school will not deviate from this initial point. This implies that the seller-optimal separating equilibrium always exists. Furthermore, in Appendix A.3, we characterize all the separating equilibrium outcomes of the unobserved case, and prove that in every separating equilibrium, $\theta_0^u \geq \theta_0^o$.

Then, consider pooling equilibria. For each θ, define a function of z on $\mathbb{R}_+$ as follows:

$$\Delta(z, \theta) := \mathbb{E}\{Q(z, \theta) | \theta \in [\underline{\theta}, \bar{\theta}]\} - G(z, \theta) - \max_{y \geq 0} \{Q(y, \underline{\theta}) - G(y, \theta)\}.$$

Intuitively, $\Delta(z, \theta)$ is the net gain in marginal profit of type θ if in equilibrium all types pool at education level z, compared to type θ's optimal deviation under the worst off-path belief. Because $Q_{zz} \leq 0$ and $G_{zz} > 0$, $\Delta(z, \theta)$ is strictly concave in z for all $\theta \in [\underline{\theta}, \bar{\theta}]$. Moreover, it is clear that $\Delta(0, \theta) \leq 0$, and that if $\Delta(0, \theta) < 0$, then $\Delta(z, \theta) > 0$ for some $z > 0$. Thus, for each θ, $\Delta(z, \theta)$ has a root and at most two roots on $\mathbb{R}_+$. Let $\underline{x}(\theta)$ and $\bar{x}(\theta)$ be the minimal and maximal roots of $\Delta(z, \theta) = 0$, respectively, with the possibility that $\underline{x}(\theta) = \bar{x}(\theta) = 0$.

Now, we prove that a pooling equilibrium exists if $\bar{x}(\underline{\theta}) > \underline{x}(\bar{\theta})$. For each θ, define

$$y(\theta) := \operatorname*{argmax}_{z \geq 0} Q(z, \underline{\theta}) - G(z, \theta).$$

It is clear that $y(\theta)$ is nondecreasing on $[\underline{\theta}, \bar{\theta}]$. By the envelope theorem, for any $z > 0$,

$$\Delta_\theta(z, \theta) = G_\theta(y(\theta), \theta) - G_\theta(z, \theta).$$

Because $G_{z\theta} < 0$, $\Delta_\theta(z,\theta) > 0$ if and only if $y(\theta) < z$. It follows that for any fixed $z > 0$, $\Delta(z,\theta)$ is strictly quasiconcave on $[\underline{\theta}, \bar{\theta}]$. Assume that the labor market regards any off-path education level as chosen by the lowest type. Therefore, a pooling equilibrium exists if there exists some $\tilde{z} > 0$ such that $\Delta(\tilde{z},\theta) \geq 0$ on $[\underline{\theta}, \bar{\theta}]$. This reduces to that $\Delta(\tilde{z},\underline{\theta}), \Delta(\tilde{z},\bar{\theta}) \geq 0$, as $\Delta(\tilde{z},\cdot)$ is strictly quasiconcave. We consider two cases. First, if $y(\underline{\theta}) > 0$, then $y(\bar{\theta}) > y(\underline{\theta})$, and thus, $\Delta(y(\underline{\theta}),\underline{\theta}), \Delta(y(\bar{\theta}),\bar{\theta}) > 0$. It follows that $\underline{x}(\theta) < \bar{x}(\bar{\theta})$. Because $\bar{x}(\underline{\theta}) > \underline{x}(\bar{\theta})$, there exists a $\tilde{z}$ such that $\max\left\{\underline{x}(\underline{\theta}), \underline{x}(\bar{\theta})\right\} < \tilde{z} < \min\left\{\bar{x}(\underline{\theta}), \bar{x}(\bar{\theta})\right\}$, and thus, $\Delta(\tilde{z},\underline{\theta}), \Delta(\tilde{z},\bar{\theta}) > 0$. Second, if $y(\underline{\theta}) = 0$, then $y(\underline{\theta}) = \underline{x}(\underline{\theta})$. Because $\bar{x}(\underline{\theta}) > 0$, $\Delta(z,\underline{\theta}) > 0$ on $(0, \bar{x}(\underline{\theta}))$. Therefore, if $y(\bar{\theta}) > 0$, then $0 < \underline{x}(\bar{\theta}) < \bar{x}(\bar{\theta})$, and thus, for any $\tilde{z}$ such that $\underline{x}(\bar{\theta}) < \tilde{z} < \min\left\{\bar{x}(\underline{\theta}), \bar{x}(\bar{\theta})\right\}$, $\Delta(\tilde{z},\underline{\theta}), \Delta(\tilde{z},\bar{\theta}) > 0$; if $y(\bar{\theta}) = 0$, then $\Delta(z,\bar{\theta}) \geq \Delta(z,\underline{\theta})$ because $G_\theta < 0$, and thus, for any $\tilde{z} \in (0, \bar{x}(\underline{\theta}))$, $\Delta(\tilde{z},\underline{\theta}), \Delta(\tilde{z},\bar{\theta}) > 0$. In summary, a pooling equilibrium exists if $\bar{x}(\underline{\theta}) > \underline{x}(\bar{\theta})$.

Moreover, if there exists a third type of continuous equilibrium, then $z^u(\theta)$ must include a pooling interval and a separating interval that are connected at some point $\tilde{\theta} \in (\theta_0^u, \bar{\theta})$. It follows from the continuity of $z^u(\theta)$ that $W^u(z)$ is discontinuous at $z^u(\tilde{\theta})$. Then, the school can make a profitable deviation by choosing $z(\theta) = \lim_{\theta \downarrow \tilde{\theta}} z^u(\theta)$ in a left neighborhood of $\tilde{\theta}$, a contradiction. Thus, a continuous equilibrium is either separating or pooling.

Finally, we show by Corollary 3 in Appendix A.3 that when $z^*(\underline{\theta}) > 0$ and $\bar{x}(\underline{\theta}) < \underline{x}(\bar{\theta})$, the unique continuous equilibrium is the seller-optimal separating equilibrium that satisfies the initial condition $(\theta_0^u, z^u(\theta_0^u)) = (\underline{\theta}, z^*(\underline{\theta}))$. Thus, the theorem is proven. □

Proof of Theorem 2

Proof. We first prove that no pooling equilibrium satisfies the Quasi-Divinity. To show this, fix a type α that is close to $\bar{\theta}$, and choose some $\hat{z} > \tilde{z}$ such that

$$Q(\hat{z},\alpha) - G(\hat{z},\alpha) = W^u(\tilde{z}) - G(\tilde{z},\alpha),$$

where $\tilde{z} > 0$ is the equilibrium education level for $\theta \geq \theta_0^u$, and $W^u(\tilde{z}) = \mathbb{E}[Q(\tilde{z},\theta)|\theta \geq \theta_0^u]$. It is clear that such $\hat{z}$ exists. Let $\hat{w} = Q(\hat{z},\alpha)$. For any $\theta > \alpha$, we have

$$\hat{w} - G(\hat{z},\theta) - [W^u(\tilde{z}) - G(\tilde{z},\theta)] > \hat{w} - W^u(\tilde{z}) - [G(\hat{z},\alpha) - G(\tilde{z},\alpha)] = 0.$$

The inequality is due to that $G_{z\theta} < 0$ and $\hat{z} > \tilde{z}$. Analogously, for any $\theta \leq \alpha$, we have

$$\hat{w} - G(\hat{z},\theta) - [W^u(z^u(\theta)) - G(z^u(\theta),\theta)] \leq \hat{w} - G(\hat{z},\theta) - [W^u(\tilde{z}) - G(\tilde{z},\theta)] \leq 0.$$

The first inequality is due to the optimality of $z^u(\theta)$.

Then, define $\hat{\Theta} \coloneqq (\alpha, \bar{\theta}]$; thus, we have

$$\hat{w} - G(\hat{z}, \theta) > W^u(z^u(\theta)) - G(z^u(\theta), \theta) = MP^u(z^u(\theta), \theta)$$

if and only if $\theta \in \hat{\Theta}$. Moreover, note that for any quasi-divine belief,

$$\mathbb{E}^{\hat{z}}[Q(\hat{z}, \theta) | \theta > \alpha] > Q(\hat{z}, \alpha) = \hat{w}.$$

Consider the allocation $z^d(\theta)$ such that $z^d(\theta) = \hat{z}$ if $\theta \in \hat{\Theta}$; $z^d(\theta) = z^u(\theta)$ otherwise. Because $\hat{z} > \tilde{z}$, $z^d(\theta)$ is nondecreasing. Therefore, it follows immediately from Definition 1 that any pooling equilibrium fails the Quasi-Divinity.

Now, we show that a separating equilibrium with an initial point other than $(\theta_0^o, z^*(\theta_0^o))$ fails the Quasi-Divinity. It follows from Lemma 2 in Appendix A.3 that such an equilibrium satisfies that $\theta_0^u > \theta_0^o$ and $z^u(\theta_0^u)$ is the maximal root of $MP^o(z, \theta_0^u) = 0$. Hence, $z^u(\theta_0^u) > 0$. Then, fix a type $\alpha < \theta_0^u$, which is close to θ_0^u, and choose some $\hat{z} \in (0, z^u(\theta_0^u))$ such that

$$Q(\hat{z}, \alpha) - G(\hat{z}, \alpha) = 0.$$

Clearly, such $\hat{z}$ exists. Let $\hat{w} = Q(\hat{z}, \alpha)$. Because $G_{z\theta} < 0$, we have $\hat{w} - G(\hat{z}, \theta_0^u) > 0$, and

$$\frac{d}{d\theta} \left[\hat{w} - G(\hat{z}, \theta) - MP^u(z^u(\theta), \theta)\right] = G_\theta(z^u(\theta), \theta) - G_\theta(\hat{z}, \theta) \tag{A.1}$$

is positive if and only if $\theta < \theta_0^u$. This implies that there exists a type $\beta \in (\theta_0^u, \bar{\theta})$ such that

$$\hat{w} - G(\hat{z}, \theta) - MP^u(z^u(\theta), \theta) > 0$$

if and only if $\theta \in (\alpha, \beta)$. Then, define $\hat{\Theta} \coloneqq (\alpha, \beta)$. Note that for any quasi-divine belief,

$$\mathbb{E}^{\hat{z}}[Q(\hat{z}, \theta) | \theta \in \hat{\Theta}] > Q(\hat{z}, \alpha) = \hat{w}.$$

Consider the allocation $z^d(\theta)$ such that $z^d(\theta) = \hat{z}$ if $\theta \in \hat{\Theta}$; $z^d(\theta) = z^u(\theta)$ otherwise. Clearly, $z^d(\theta)$ is nondecreasing. Then by Definition 1, such an equilibrium fails the Quasi-Divinity.

Finally, we show that a separating equilibrium with the initial point $(\theta_0^o, z^*(\theta_0^o))$ satisfies the Quasi-Divinity. Suppose not, then there exists an off-path education level $\hat{z}$, a wage $\hat{w}$ and a positive-measure subset $\hat{\Theta}$ satisfy conditions (i) to (iii) of Definition 1. There are two possibilities. First, if $\hat{z} > z^u(\bar{\theta})$, then by the proof of Theorem 3, $\hat{z} > z^*(\theta)$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$. Note that $\hat{w} < Q(\hat{z}, \bar{\theta})$, as required by condition (iii). Thus, for any θ, we have

$$\hat{w} - G(\hat{z}, \theta) < Q(\hat{z}, \bar{\theta}) - G(\hat{z}, \theta) < Q(z^u(\bar{\theta}), \bar{\theta}) - G(z^u(\bar{\theta}), \theta) \leq MP^u(z^u(\theta), \theta).$$

The second inequality is due to that $Q(z, \bar{\theta}) - G(z, \theta)$ is strictly concave in z and thus has a unique maximizer that is smaller than $z^*(\bar{\theta})$, and that $\hat{z} > z^u(\bar{\theta}) > z^*(\bar{\theta})$. The last inequality is due to the optimality of $z^u(\theta)$. Hence, $\hat{\Theta}$ is empty, a contradiction. Second, if $\hat{z} < z^u(\theta_0^u)$, then by Lemma 2, $\theta_0^u = \underline{\theta}$ and $z^u(\underline{\theta}) = z^*(\underline{\theta})$. Because $\hat{\Theta}$ has a positive measure, there exists a type $\alpha > \underline{\theta}$ such that $\hat{w} - G(\hat{z}, \alpha) > MP^u(z^u(\alpha), \alpha)$. Moreover, because $\hat{z} < z^u(\underline{\theta})$, (A.1) is negative for all $\theta \in [\underline{\theta}, \bar{\theta}]$. This means that $\underline{\theta} \in \hat{\Theta}$, i.e.,

$$\hat{w} - G(\hat{z}, \underline{\theta}) > MP^u(z^u(\underline{\theta}), \underline{\theta}) = MP^o(z^*(\underline{\theta}), \underline{\theta}).$$

But for sufficiently pessimistic quasi-divine beliefs, we have $|\hat{w} - Q(\hat{z}, \underline{\theta})| < \varepsilon$, for any $\varepsilon > 0$; that is, the LHS of the inequality is bounded above by $MP^o(\hat{z}, \underline{\theta}) + \varepsilon$, a contradiction. Thus, the equilibrium satisfies the Quasi-Divinity. In summary, the theorem is proven. □

Proof of Theorem 3

Proof. We first show that in each separating equilibrium, $z^u(\theta) \geq z^*(\theta)$ on $[\theta_0^u, \bar{\theta}]$, with strict inequality for $\theta > \theta_0^u$. Because types reveal in equilibrium, $MP^u(z^u(\theta), \theta) = MP^o(z^u(\theta), \theta)$. Moreover, $MP^o(z, \theta)$ is strictly concave in z. It thus suffices to show that $MP_z^o(z^u(\theta), \theta) \leq 0$, with strict inequality for $\theta > \theta_0^u$. This is given by the following:

$$\begin{aligned} MP_z^o(z^u(\theta), \theta) &= Q_z(z^u(\theta), \theta) - G_z(z^u(\theta), \theta) \\ &\leq Q_z(z^u(\theta), \theta) + Q_\theta(z^u(\theta), \theta)\theta^{u\prime}(z^u(\theta)) - G_z(z^u(\theta), \theta) = 0. \end{aligned}$$

The inequality is because $z^u(\theta)$ is nondecreasing; the last equality is due to (4.1). Moreover, for $\theta > \theta_0^u$, the second term in (4.1) is positive; thus, the above inequality becomes strict.

Then, we prove that in each separating equilibrium, $z^u(\theta) \geq z^o(\theta)$ on $[\theta_0^u, \bar{\theta}]$, with strict inequality on $(\theta_0^u, \bar{\theta}]$. By Lemma 2 in Appendix A.3, $z^u(\theta_0^u) \geq \max\{z^*(\theta_0^u), z^o(\theta_0^u)\}$. Suppose there exists a type $\tilde{\theta} \in (\theta_0^u, \bar{\theta}]$ such that $z^u(\tilde{\theta}) \leq z^o(\tilde{\theta})$, then $\tilde{\theta}$ belongs to a pooling interval for $z^o(\tilde{\theta})$, as $z^u(\theta) > z^*(\theta)$ on $(\theta_0^u, \bar{\theta}]$. Let α be the left end of this interval. Because $z^u(\theta)$ is increasing on $[\theta_0^u, \bar{\theta}]$, $z^u(\theta) < z^o(\theta)$ on $[\alpha, \tilde{\theta})$. Moreover, because $z^u(\theta_0^u) \geq z^o(\theta_0^u)$, $\alpha \in (\theta_0^u, \tilde{\theta})$. It follows from the continuity of $z^o(\theta)$ that $z^*(\alpha) = z^o(\alpha) > z^u(\alpha)$, a contradiction. Hence, we have $z^u(\theta) \geq z^o(\theta)$ on $[\theta_0^u, \bar{\theta}]$, with strict inequality on $(\theta_0^u, \bar{\theta}]$.

Finally, we consider the seller-optimal separating equilibrium. Because $\theta_0^u = \theta_0^o$, we have $z^u(\theta) \geq z^o(\theta)$ on $[\underline{\theta}, \bar{\theta}]$, with strict inequality for $\theta > \theta_0^u$. Thus, the theorem is proven. □

Proof of Proposition 3

Proof. With a bit abuse of notation, for each $z \in [z^*(\theta_0^o), z^*(\bar{\theta})]$, let $\theta^o(z)$ be some type who chooses z in the observed case; $\theta^o(z)$ is single-valued if it is outside a pooling interval.

According to Lemma 1, in the unobserved case, the tuition fee at z is given by

$$T^u(z) = W^u(z) - C(z, \theta^u(z)) + \int_{\theta_0^o}^{\theta^u(z)} C_\theta(z^u(s), s)ds.$$

Subtracting this equation from (3.2) at z, we have that $T^o(z) - T^u(z)$ is equal to

$$W^o(z) - W^u(z) - [C(z, \theta^o(z)) - C(z, \theta^u(z))] + \int_{\theta_0^o}^{\theta^o(z)} C_\theta(z^o(s), s)ds - \int_{\theta_0^o}^{\theta^u(z)} C_\theta(z^u(s), s)ds$$

$$\geq W^o(z) - W^u(z) - [C(z, \theta^o(z)) - C(z, \theta^u(z))] + \int_{\theta^u(z)}^{\theta^o(z)} C_\theta(z^o(s), s)ds$$

$$= W^o(z) - W^u(z) + \int_{\theta^u(z)}^{\theta^o(z)} [C_\theta(z^o(s), s) - C_\theta(z, s)]ds$$

$$\geq W^o(z) - W^u(z).$$

The first inequality is due to Theorem 3 and that $C_{z\theta} < 0$; the second inequality is due to that $\theta^o(z) \geq \theta^u(z)$ and that $z^o(\theta) \leq z$ for all $\theta \in [\theta^u(z), \theta^o(z)]$. Moreover, because wage is equal to the worker's expected productivity, we have $W^o(z) \geq W^u(z)$ on $[z^*(\theta_0^o), z^*(\bar{\theta})]$, with strict inequality for $z > z^*(\theta_0^o)$. Thus, the proposition is proven. □

Proof of Proposition 5

Proof. We only need to prove that $z^u(\theta) < z^s(\theta)$ on $(\theta_0^u, \bar{\theta}]$. Rearranging (2.4) and (4.1), we can derive $z^s(\theta)$ and $z^u(\theta)$ through the IVPs:

$$z^{s\prime} = \frac{Q_\theta(z, \theta)}{C_z(z, \theta) - Q_z(z, \theta)} \quad \text{and} \quad z^{u\prime} = \frac{Q_\theta(z, \theta)}{G_z(z, \theta) - Q_z(z, \theta)}$$

with the initial points $(\theta_0^u, z^s(\theta_0^u))$ and $(\theta_0^u, z^u(\theta_0^u))$, respectively. Because $C_{z\theta} < 0$, we have $C_z(z, \theta) \leq G_z(z, \theta)$ on any common plain of (θ, z), with strict inequality for $\theta < \bar{\theta}$. Moreover, note that $z^s(\theta_0^u) \geq z^u(\theta_0^u)$. Then, by Hartman (1964, Corollary 4.2, page 27), $z^s(\theta) \geq z^u(\theta)$ on $[\theta_0^u, \bar{\theta}]$, with strict inequality for $\theta > \theta_0^u$. Thus, the proposition is proven. □

Proof of Proposition 6

Proof. We only need to consider the interval $(\theta_0^u, \bar{\theta})$. Note that $z^u(\theta_0^u) = z^*(\theta_0^u) \leq z^{fb}(\theta_0^u)$ and $z^u(\bar{\theta}) > z^o(\bar{\theta}) = z^{fb}(\bar{\theta})$. Therefore, if $z^u(\theta)$ intersects $z^{fb}(\theta)$ multiple times, then there are at least three intersections. Denote by $W^{fb}(z)$ the wage schedule in the first-best benchmark. By the monotonicity of $z^u(\theta)$ and $z^{fb}(\theta)$, both $W^u(z)$ and $W^{fb}(z)$ are increasing on $(\theta_0^u, \bar{\theta})$. Thus, it suffices to show that $W^u(z)$ and $W^{fb}(z)$ have a unique intersection. Suppose that $z^u(\theta)$ intersects $z^{fb}(\theta)$ at some θ^w, then $W^u(z)$ intersects $W^{fb}(z)$ at $z^{fb}(\theta^w)$.

Then, differentiating $W^u(z)$ and $W^{fb}(z)$ at $z^{fb}(\theta^w)$, and substituting $G(z,\theta)$, we have

$$W^{u\prime}(z^{fb}(\theta^w)) = C_z(z^{fb}(\theta^w),\theta^w) - \frac{1-F(\theta^w)}{f(\theta^w)}C_{z\theta}(z^{fb}(\theta^w),\theta^w),$$
$$W^{fb\prime}(z^{fb}(\theta^w)) = Q_z(z^{fb}(\theta^w),\theta^w) + Q_\theta(z^{fb}(\theta^w),\theta^w)\theta^{fb\prime}(z^{fb}(\theta^w)).$$

Rearranging and substituting (2.1), we have that $W^{fb\prime}(z^{fb}(\theta^w)) - W^{u\prime}(z^{fb}(\theta^w))$ equals

$$Q_\theta(z^{fb}(\theta^w),\theta^w)\theta^{fb\prime}(z^{fb}(\theta^w)) + \frac{1-F(\theta^w)}{f(\theta^w)}C_{z\theta}(z^{fb}(\theta^w),\theta^w).$$

Given Assumption 2, the above function changes sign only once with different values of θ^w. Suppose that $W^u(z)$ intersects $W^{fb}(z)$ multiple times, then the above function changes sign more than once, a contradiction. Thus, $z^u(\theta)$ intersects $z^{fb}(\theta)$ only once and from below. It follows that $W^{fb\prime}(z^{fb}(\theta^w)) - W^{u\prime}(z^{fb}(\theta^w)) > 0$, meaning that the above function is positive at θ^w. Then by Assumption 2, we have $\theta^w > \theta^*$. Thus, the proposition is proven. □

A.3 Continuous Equilibrium of the Unobserved Case

We first characterize all the separating equilibrium outcomes of the unobserved case. From the proof of Theorem 1, it suffices to characterize each equilibrium initial point $(\theta_0^u, z^u(\theta_0^u))$. The lemma below provides necessary conditions for an equilibrium initial point.

Lemma 2. *Every separating equilibrium of the unobserved case satisfies that $\theta_0^u \geq \theta_0^o$, and that $z^u(\theta_0^u) \geq \max\{z^*(\theta_0^u), z^o(\theta_0^u)\}$: If $\theta_0^u = \theta_0^o$, then $z^u(\theta_0^u) = z^*(\theta_0^u) \geq z^o(\theta_0^u)$; if $\theta_0^u > \theta_0^o$, then $z^u(\theta_0^u)$ is the maximal root of $MP^o(z,\theta_0^u) = 0$, and thus, $z^u(\theta_0^u) > \max\{z^*(\theta_0^u), z^o(\theta_0^u)\}$.*

Proof. We have shown that $(\theta_0^u, z^u(\theta_0^u)) = (\theta_0^o, z^*(\theta_0^o))$ is an equilibrium initial point. Then, we prove that when $\theta_0^u = \theta_0^o$, $z^u(\theta_0^u) = z^*(\theta_0^u)$. We consider two cases. First, if $\theta_0^u = \theta_0^o = \underline{\theta}$, then by the proof of Theorem 3, $z^u(\underline{\theta}) \geq z^*(\underline{\theta})$. Suppose $z^u(\underline{\theta}) > z^*(\underline{\theta})$, then the school can make a profitable deviation by assigning $z^*(\underline{\theta})$ to a neighborhood of $\underline{\theta}$. This is because the employers cannot punish such a deviation by having a worse belief than $\underline{\theta}$, and $MP^o(z,\theta)$ is continuous and strictly concave in z. Thus, $z^u(\underline{\theta}) = z^*(\underline{\theta})$. Second, if $\theta_0^u = \theta_0^o > \underline{\theta}$, then we have $z^u(\theta_0^u) = z^o(\theta_0^u) = z^*(\theta_0^u) = 0$. Moreover, by the proof of Proposition 1, $z^*(\theta_0^o) \geq z^o(\theta_0^o)$. It follows that if $\theta_0^u = \theta_0^o$, then $z^u(\theta_0^u) = z^*(\theta_0^u) \geq z^o(\theta_0^u)$.

Now suppose $\theta_0^u \neq \theta_0^o$. If $\theta_0^u < \theta_0^o$, then $MP^u(z^u(\theta_0^o),\theta_0^o) \leq MP^o(z^*(\theta_0^o),\theta_0^o) = 0$, leading to a contradiction. In contrast, if $\theta_0^u > \theta_0^o$, then $MP^o(z^u(\theta_0^u),\theta_0^u) = 0$. Because $MP^o(z,\theta)$ is strictly concave in z and $z^u(\theta_0^u) \geq z^*(\theta_0^u) > 0$, $z^u(\theta_0^u)$ is the maximal root of $MP^o(z,\theta_0^u) = 0$, and thus, $z^u(\theta_0^u) > \max\{z^*(\theta_0^u), z^o(\theta_0^u)\}$. In summary, in each separating equilibrium of the unobserved case, $\theta_0^u \geq \theta_0^o$ and $z^u(\theta_0^u) \geq \max\{z^*(\theta_0^u), z^o(\theta_0^u)\}$. Thus, the lemma is proven. □

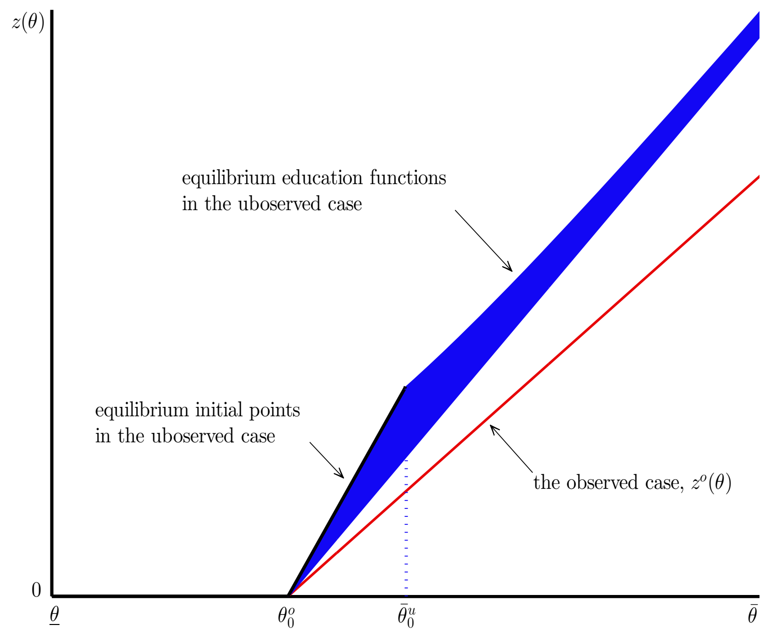

Figure 4: **The Set of Separating Equilibria.** This figure illustrates the set of separating equilibria of the unobserved case given that $\theta_0^o > \underline{\theta}$. The shaded area depicts the set of equilibrium education functions. This region is uniformly above the equilibrium education function of the observed case $z^o(\theta)$. The bold line is the set of equilibrium initial points with the cutoff type ranging from θ_0^o to $\bar{\theta}_0^u$. Each point uniquely determines an equilibrium education function $z^u(\theta)$, and thus, an equilibrium outcome. This figure considers the same numerical example as Figure 1, such that the set of the initial points is $\{(\theta, z)|z(\theta) = 3\theta - 1; \frac{1}{3} \leq \theta \leq \frac{1}{2}\}$.

Lemma 2 indicates that given an equilibrium cutoff type, his education level is uniquely determined: If $\theta_0^u = \theta_0^o$, then $z^u(\theta_0^u) = z^*(\theta_0^u)$; otherwise, $\theta_0^u > \theta_0^o$ and $z^u(\theta_0^u)$ is the maximal root of $MP^o(z, \theta_0^u) = 0$. Therefore, we can characterize the set of equilibrium initial points by characterizing the set of cutoff types. Note that the lower bound of θ_0^u is θ_0^o. Moreover, we can determine the upper bound of θ_0^u by assuming the worst off-path belief. Specifically, denote by $\bar{\theta}_0^u$ the upper bound of θ_0^u, which is the maximal root of

$$\max_{z \geq 0}\{Q(z, \underline{\theta}) - G(z, \theta)\} = 0$$

if it exists, or $\underline{\theta}$ otherwise. Therefore, if $\bar{\theta}_0^u = \underline{\theta}$, then the cutoff type and thus the separating equilibrium outcome is unique. In contrast, if $\bar{\theta}_0^u > \underline{\theta}$, then any type between θ_0^o and $\bar{\theta}_0^u$ can be an equilibrium cutoff type for some proper off-path belief. Hence, the set of cutoff type is given by $[\theta_0^o, \bar{\theta}_0^u]$. It follows from Lemma 2 that $z^u(\theta_0^u)$ is continuous and increasing in θ_0^u. Figure 4 illustrates the set of education functions of separating equilibria. As depicted, each education function satisfies that $z^u(\theta) \geq z^o(\theta)$ on $[\theta_0^u, \bar{\theta}]$, with strict inequality for $\theta > \theta_0^u$.

The next lemma shows that the separating equilibrium associated with the initial point $(\theta_0^o, z^*(\theta_0^o))$ yields the highest profit for the school among all separating equilibria.

Lemma 3. *The seller-optimal separating equilibrium satisfies that* $(\theta_0^u, z^u(\theta_0^u)) = (\theta_0^o, z^*(\theta_0^o))$.

Proof. Suppose in addition to θ_0^o there exists an equilibrium cutoff type $\hat{\theta}_0^u > \theta_0^o$. Let $z_1^u(\theta)$ and $z_2^u(\theta)$ be the equilibrium education functions associated with θ_0^o and $\hat{\theta}_0^u$, respectively. It follows from Lemma 2 that $z^u(\hat{\theta}_0^u) > z^*(\theta_0^o)$. Because both $z_1^u(\theta)$ and $z_2^u(\theta)$ satisfy (4.1), by Hartman (1964, Corollary 4.2, page 27), $z_2^u(\theta) > z_1^u(\theta)$ on the common interval $[\hat{\theta}_0^u, \bar{\theta}]$. Note that $MP^u(z^u(\theta), \theta) = MP^o(z^u(\theta), \theta)$ and $z_2^u(\theta) > z_1^u(\theta) \geq z^*(\theta)$ on $[\hat{\theta}_0^u, \bar{\theta}]$. It follows from the strict concavity of $MP^o(z, \theta)$ that $MP^u(z_1^u(\theta), \theta) > MP^u(z_2^u(\theta), \theta)$ on $(\theta_0^o, \bar{\theta}]$. Denote by Π_1^u and Π_2^u the school's equilibrium payoff associated with θ_0^o and $\hat{\theta}_0^u$, respectively. Thus,

$$\Pi_1^u - \Pi_2^u = \int_{\theta_0^o}^{\bar{\theta}} MP^u(z_1^u(\theta), \theta) dF(\theta) - \int_{\hat{\theta}_0^u}^{\bar{\theta}} MP^u(z_2^u(\theta), \theta) dF(\theta) > 0.$$

That is, the separating equilibrium given by the initial point $(\theta_0^o, z^*(\theta_0^o))$ leads to the highest payoff for the school, among all the separating equilibria. Thus, the lemma is proven. □

We now study pooling equilibrium of the unobserved case. The next proposition provides a necessary and sufficient condition for the existence of pooling equilibrium when $z^*(\underline{\theta}) > 0$.

Proposition 7. *Suppose that* $z^*(\underline{\theta}) > 0$*, then a pooling equilibrium exists in the unobserved case if and only if* $\bar{x}(\underline{\theta}) \geq \underline{x}(\bar{\theta})$.

Proof. We proved the sufficiency in the proof of Theorem 1. Hence, it remains to prove the necessity. Suppose a pooling equilibrium exists. Because $z^*(\underline{\theta}) > 0$ and $MP^o(z, \underline{\theta})$ is strictly concave in z, for any $z \in (0, z^*(\underline{\theta}))$, $MP^u(z, \underline{\theta}) \geq MP^o(z, \underline{\theta}) > 0$, meaning that $\theta_0^u = \underline{\theta}$. In addition, $y(\bar{\theta}) > y(\underline{\theta}) = z^*(\underline{\theta}) > 0$; thus, both $\Delta(z, \underline{\theta})$ and $\Delta(z, \bar{\theta})$ have two zeroes. Let $\tilde{z} > 0$ be the equilibrium education level. We claim that $\tilde{z} \leq \bar{x}(\underline{\theta})$. Suppose not, then $\Delta(\tilde{z}, \underline{\theta}) < 0$, meaning that the school can profit by decreasing $z(\theta)$ in some neighborhood of $\underline{\theta}$, with the monotonicity still holding. But note that $\Delta(z, \theta)$ posits the worst off-path belief, and thus, any different off-path belief makes such a deviation weakly more tempting, a contradiction. Similarly, if $\tilde{z} < \underline{x}(\bar{\theta})$, then $\Delta(\tilde{z}, \bar{\theta}) < 0$; thus, the school can profit by increasing $z(\theta)$ in some neighborhood of $\bar{\theta}$, with the monotonicity still holding, irrespective of the off-path belief, a contradiction. In summary, we must have $\underline{x}(\bar{\theta}) \leq \tilde{z} \leq \bar{x}(\underline{\theta})$. However, because $\bar{x}(\underline{\theta}) < \underline{x}(\bar{\theta})$, such $\tilde{z}$ does not exist. Thus, the proposition is proven. □

Note that when $z^*(\underline{\theta}) > 0$, $\bar{\theta}_0^u = \underline{\theta} = \theta_0^o$; thus, there exists a unique separating equilibrium outcome such that $(\theta_0^u, z^u(\theta_0^u)) = (\underline{\theta}, z^*(\underline{\theta}))$. If further $\bar{x}(\underline{\theta}) < \underline{x}(\bar{\theta})$, then there does not exist any pooling equilibrium. Therefore, the unique continuous equilibrium is the seller-optimal separating equilibrium in which $(\theta_0^u, z^u(\theta_0^u)) = (\underline{\theta}, z^*(\underline{\theta}))$. To summarize:

Corollary 3. *Suppose that $z^*(\underline{\theta}) > 0$ and $\bar{x}(\underline{\theta}) < \underline{x}(\bar{\theta})$, then the unobserved case has a unique continuous equilibrium, that is, the seller-optimal separating equilibrium.*

Example. To illustrate, we provide an example in which there exists a unique continuous equilibrium. Assume that $Q(z,\theta) = \theta z + 2z$, $C(z,\theta) = 1.4z^2 + 2.8z - 1.4\theta z$, and $\theta \sim U[1,2]$. It follows that $G(z,\theta) = 1.4z^2 + 5.6z - 2.8\theta z$ and $\mathbb{E}\{Q(z,\theta)|\theta \in [1,2]\} = 3.5z$. By simple calculation, $z^*(1) = y(1) = 1/14$ and $y(2) = 15/14$. Thus,

$$Q(y(1),1) - G(y(1),1) = \frac{1}{140},$$
$$Q(y(2),1) - G(y(2),2) = \frac{45}{28}.$$

Substituting, we have

$$\Delta(z,1) = -1.4z^2 + 0.7z - \frac{1}{140},$$
$$\Delta(z,2) = -1.4z^2 + 3.5z - \frac{45}{28}.$$

It follows that $\bar{x}(1) \approx 0.49$ and $\underline{x}(2) \approx 0.61$. Thus, from Corollary 3, the unique continuous equilibrium is the seller-optimal separating equilibrium such that $(\theta_0^u, z^u(\theta_0^u)) = (1, 1/14)$.

A.4 Proofs for Section 5

Lemma 4. *In the observed case, for any $\delta > 0$, there is a subgame associated with some T, in which a signaling equilibrium exists, with the school's payoff greater than $\bar{\Pi}^o - \delta$.*

Proof. Because $\underline{\theta}_0 > \underline{\theta}$, $\bar{z}(\theta)$ is increasing in a right neighborhood of $\underline{\theta}_0$. Moreover, note that $MP^o(z,\theta) \to Q(0,\theta)$ as $z \to 0$. Thus, given δ, there exists a small enough $\varepsilon > 0$ such that

$$\int_{\underline{\theta}}^{\theta^o(\varepsilon)} MP^o(\varepsilon,\theta)dF(\theta) + \int_{\theta^o(\varepsilon)}^{\bar{\theta}} MP^o(\bar{z}(\theta),\theta)dF(\theta) \geq \bar{\Pi}^o - \delta, \tag{A.2}$$

where $\theta^o(\varepsilon)$ is the preimage of $\bar{z}(\theta)$ at ε. Clearly, $\theta^o(\varepsilon) > \underline{\theta}_0$. Then, let the allocation be

$$z^o(\theta) = \begin{cases} \bar{z}(\theta) & \text{if } \theta \geq \theta^o(\varepsilon) \\ \varepsilon & \text{otherwise,} \end{cases}$$

and assume again that for all z outside the range of $z^o(\theta)$, the school charges exorbitantly high prices and the labor market holds the worst belief. It is easy to verify that $z^o(\theta)$ and the wage schedule resulted from the above market belief constitute a signaling equilibrium in the corresponding subgame, such that the school's payoff is equal to the LHS of (A.2). Thus, the lemma is proven. □

Lemma 5. *In the unobserved case, if $\theta_0^u > \underline{\theta}$, then $z^u(\theta_0^u) > z^*(\theta_0^u) > \lim_{\theta\uparrow\theta_0^u} z^u(\theta) = 0$.*

Proof. Given Π^u, the marginal profit of θ_0^u should satisfy that

$$-f(\theta_0^u)\{MP^u(z^u(\theta_0^u),\theta_0^u) - \mathbb{E}[Q(0,\theta)|\theta \le \theta_0^u]\} \le 0.$$

In particular, if $\theta_0^u > \underline{\theta}$, then we must have

$$MP^u(z^u(\theta_0^u),\theta_0^u) - \mathbb{E}[Q(0,\theta)|\theta \le \theta_0^u] = 0. \tag{A.3}$$

This implies that $z^u(\theta)$ is discontinuous at θ_0^u. Suppose not, then $W^u(z)$ must have a jump at 0 because $z^u(\theta)$ is increasing in any small right neighborhood of θ_0^u, whereas $z^u(\theta) \equiv 0$ for $\theta \le \theta_0^u$. Hence, the school can make a profitable deviation by allocating a positive but sufficiently small education level to the types in some neighborhood of θ_0^u, thereby receiving a discrete wage gain. Thus, we have a contradiction. It follows that in any equilibrium such that $\theta_0^u > \underline{\theta}$, $z^u(\theta)$ is discontinuous at θ_0^u with $z^u(\theta_0^u) > \lim_{\theta\uparrow\theta_0^u} z^u(\theta) = 0$. Because $z^u(\theta)$ is increasing at θ_0^u, we have $MP^u(z^u(\theta_0^u),\theta_0^u) = MP^o(z^u(\theta_0^u),\theta_0^u)$. It follows from the concavity of $MP^o(\cdot,\theta)$ and (A.3) that $z^u(\theta_0^u) > z^*(\theta_0^u) > 0$. Thus, the lemma is proven. □

Proof of Theorem $2'$

Proof. Because $MP^o(z^*(\underline{\theta}),\underline{\theta}) \ge Q(0,\underline{\theta}) = 0$ and $Q_\theta > 0$ for all (z,θ), $MP^o(z^*(\theta),\theta) \ge 0$, with strict inequality for $\theta > \underline{\theta}$. Thus, analogous to the proof of Theorem 1, there exists a separating equilibrium such that $(\theta_0^u, z^u(\theta_0^u)) = (\underline{\theta}, z^*(\underline{\theta}))$. Then, by the proof of Theorem 2, we have that this equilibrium satisfies the Quasi-Divinity. Moreover, a pooling equilibrium, if it exists, fails the Quasi-Divinity. It remains to show that any other separating equilibrium outcome, if it exists, fails the Quasi-Divinity. Suppose there exists a separating equilibrium outcome with $(\theta_0^u, z^u(\theta_0^u)) \neq (\underline{\theta}, z^*(\underline{\theta}))$, then according to Lemmas 2 and 5, we have $\theta_0^u > \underline{\theta}$ and $z^u(\theta_0^u) > 0$. Analogous to the proof of Theorem 2, fix a type α that is lower than and close to θ_0^u, and choose some $\hat{z} \in (0, z^u(\theta_0^u))$ such that

$$Q(\hat{z},\alpha) - G(\hat{z},\alpha) = \mathbb{E}[Q(0,\theta)|\theta \le \theta_0^u].$$

Clearly, such $\hat{z}$ exists. Let $\hat{w} = Q(\hat{z},\alpha)$, and define $\hat{\Theta}$ analogously. By a similar argument, we have that this equilibrium fails the Quasi-Divinity. Thus, the theorem is proven. □

References

Andersson, F. "Income Taxation and Job-Market Signaling." *Journal of Public Economics*, Vol. 59(2) (1996), pp. 277–298.

Bagwell, K. and Riordan, M.H. "High and Declining Prices Signal Product Quality." *American Economic Review*, Vol. 81(1) (1991), pp. 224–239.

Bagwell, L.S. and Bernheim, B.D. "Veblen Effects in a Theory of Conspicuous Consumption." *American Economic Review*, Vol. 86(3) (1996), pp. 349–373.

Banks, J.S. and Sobel, J. "Equilibrium Selection in Signaling Games." *Econometrica*, Vol. 55(3) (1987), pp. 647–661.

Biglaiser, G. "Middlemen as experts." *The RAND journal of Economics* (1993), pp. 212–223.

Biglaiser, G. and Li, F. "Middlemen: the Good, the Bad, and the Ugly." *The RAND Journal of Economics*, Vol. 49(1) (2018), pp. 3–22.

Boleslavsky, R. and Cotton, C. "Grading Standards and Education Quality." *American Economic Journal: Microeconomics*, Vol. 7(2) (2015), pp. 248–79.

Calzolari, G. and Pavan, A. "On the Optimality of Privacy in Sequential Contracting." *Journal of Economic Theory*, Vol. 130(1) (2006), pp. 168–204.

Chan, W., Li, H., and Suen, W. "A Signaling Theory of Grade Inflation." *International Economic Review*, Vol. 48(3) (2007), pp. 1065–1090.

Cho, I.K. and Kreps, D.M. "Signaling Games and Stable Equilibria." *The Quarterly Journal of Economics* (1987), pp. 179–221.

Daley, B. and Green, B. "Market Signaling with Grades." *Journal of Economic Theory*, Vol. 151 (2014), pp. 114–145.

Fudenberg, D. and Tirole, J. "A 'Signal-Jamming' Theory of Predation." *RAND Journal of Economics*, Vol. 17(3) (1986), pp. 366–376.

Hartman, P. *Ordinary Differential Equations*. New York: Wiley, 1964.

Holmström, B. "Managerial Incentive Problems: A Dynamic Perspective." *The Review of Economic Studies*, Vol. 66(1) (1999), pp. 169–182.

Inderst, R. and Ottaviani, M. “Competition through Commissions and Kickbacks.” *American Economic Review*, Vol. 102(2) (2012), pp. 780–809.

Ireland, N.J. “On Limiting the Market for Status Signals.” *Journal of Public Economics*, Vol. 53(1) (1994), pp. 91–110.

Janssen, M. and Shelegia, S. “Consumer Search and Double Marginalization.” *American Economic Review*, Vol. 105(6) (2015), pp. 1683–1710.

Johnson, V.E. *Grade Inflation: A Crisis in College Education*. Springer Science & Business Media, 2006.

Leland, H.E. and Pyle, D.H. “Informational Asymmetries, Financial Structure, and Financial Intermediation.” *Journal of Finance*, Vol. 32(2) (1977), pp. 371–387.

Lizzeri, A. “Information Revelation and Certification Intermediaries.” *The RAND Journal of Economics* (1999), pp. 214–231.

Mailath, G.J. and von Thadden, E.L. “Incentive Compatibility and Differentiability: New Results and Classic Applications.” *Journal of Economic Theory*, Vol. 148(5) (2013), pp. 1841–1861.

Martimort, D. and Stole, L. “Market Participation in Delegated and Intrinsic Common-Agency Games.” *RAND Journal of Economics*, Vol. 40(1) (2009), pp. 78–102.

Mas-Colell, A., Whinston, M.D., and Green, J.R. *Microeconomic Theory*. New York: Oxford University Press, 1995.

Maskin, E. and Riley, J. “Monopoly with Incomplete Information.” *RAND Journal of Economics*, Vol. 15(2) (1984), pp. 171–196.

Milgrom, P. and Roberts, J. “Price and Advertising Signals of Product Quality.” *Journal of Political Economy*, Vol. 94(4) (1986), pp. 796–821.

Milgrom, P. and Shannon, C. “Monotone Comparative Statics.” *Econometrica*, Vol. 62(1) (1994), pp. 157–180.

Mussa, M. and Rosen, S. “Monopoly and Product Quality.” *Journal of Economic Theory*, Vol. 18(2) (1978), pp. 301–317.

Rayo, L. "Monopolistic Signal Provision." *The BE Journal of Theoretical Economics*, Vol. 13(1) (2013), pp. 27–58.

Riley, J.G. "Informational Equilibrium." *Econometrica*, Vol. 47(2) (1979), pp. 331–359.

Rojstaczer, S. and Healy, C. "Grading in American Colleges and Universities." *Teachers College Record*, Vol. 4 (2010).

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

Spence, M. "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution." *Journal of Economic Theory*, Vol. 7(3) (1974), pp. 296–332.

Toikka, J. "Ironing without Control." *Journal of Economic Theory*, Vol. 146(6) (2011), pp. 2510–2526.

Veblen, T. *The Theory of the Leisure Class: An Economic Study of Institutions.* London: Unwin Books, 1899.

Zubrickas, R. "Optimal Grading." *International Economic Review*, Vol. 56(3) (2015), pp. 751–776.