

# Competitive Screening under Heterogeneous Information\*

**Daniel Garrett**

Toulouse School of Economics  
daniel.garrett@tse-fr.eu

**Renato Gomes**

Toulouse School of Economics  
renato.gomes@tse-fr.eu

**Lucas Maestri**

Toulouse School of Economics  
lucas.maestri@tse-fr.eu

June 9, 2014

## **Abstract**

We study competition in price-quality menus when consumers privately know their valuation for quality (type), and are heterogeneously informed about the offers available in the market. While firms are ex-ante identical, the menus offered in equilibrium are ordered so that more generous menus leave more surplus uniformly over types. More generous menus provide quality more efficiently, serve a larger range of consumers, and generate a greater fraction of profits from sales of low-quality goods. By varying the mass of competing firms, or the level of informational frictions, we span the entire spectrum of competitive intensity, from perfect competition to monopoly.

**JEL Classification:** D82

**Keywords:** competition, screening, heterogeneous information, price discrimination, adverse selection

---

\*We are grateful to Jean Tirole for extensive comments at the early stages of this project. We also thank Yeon-Koo Che, Andrew Clausen, Jacques Crémer, Wouter Dessen, Renaud Foucart, Ed Hopkins, Bruno Jullien, Tatiana Kornienko, Wojciech Olszewski, Michael Peters, Patrick Rey, Andrew Rhodes, Mike Riordan, Maher Said, Ron Siegel, Bruno Strulovici, André Veiga, and Yaron Yehezkel for very helpful conversations. For useful feedback, we thank seminar participants at Bilkent University, Columbia University, FGV-EPGE, the LMU of Munich, Northwestern University, Oxford University, the University of Bonn, the University of Edinburgh, the International Conference on Game Theory at Stony Brook, the Toulouse-Northwestern IO conference, the UBC-UHK Theory conference, and the CEPR Applied IO conference. The usual disclaimer applies.

# 1 Introduction

Price discrimination through menus of products at different prices is a widespread practice across many markets. Examples include flight seats with different classes of service, electronic printers with various processing speeds, and automobiles in standard or deluxe versions.

The benchmark setting where the seller is a monopolist was first studied by Mussa and Rosen (1978). This paper considers consumers who differ in their appreciation of quality and shows that the monopolist’s profit-maximizing policy involves downward distortion of quality for all consumers, except for those who value it the most. Thus purchasers of printers with the highest value for speed enjoy the efficient speed; all others’ printers are inefficiently slow. The optimal distortions resolve a trade-off between extracting rents from consumers with high willingness to pay, and providing more efficient qualities to the others.

Much of the price discrimination we observe in practice occurs, however, not in the textbook monopoly setting, but in the presence of competition. A range of models have therefore been developed to study price discrimination in settings where market power is limited by competition; see, among others, Champsaur and Rochet (1989), Rochet and Stole (1997, 2002) and Armstrong and Vickers (2001) (the next subsection gives a detailed account of this literature).<sup>1</sup> An assumption maintained throughout this body of work, as well as in much of the large empirical literature on competition with differentiated products following Berry, Levinsohn and Pakes (1995), is that consumers enjoy *perfect* (and therefore homogeneous) information about the offers available in the market.

The aim of our paper, in contrast, is to study competitive price discrimination in settings where consumers are *heterogeneously* informed about the price/quality menus offered by firms. That is, due to information frictions, consumers are able to consider only a sample of firms’ offers. That each consumer’s choice is often restricted to a “information set” narrower than the entire spectrum of firms’ offers has long been recognized in the marketing literature, and has been widely documented empirically (see, for example, De los Santos, Hortaçsu and Wildenbeest (2012) and the references therein). Its importance for empirical work studying consumer demand and industry conduct is increasingly recognized (see, for instance, Sovinsky Goeree (2008) and Draganska and Klapper (2011)). Theoretical work on competitive price discrimination with heterogeneous information sets, however, has to date been missing.

Following the canonical work of Mussa and Rosen, the only heterogeneity we permit in consumer tastes relates to their valuation for quality. That is, consumers have no “brand” preferences, and so evaluate offers from different firms symmetrically (any consumer is indifferent between two contracts with different firms that have the same price and quality). This assumption contrasts with works such as Rochet and Stole (1997, 2002) and Armstrong and Vickers (2001) who capture imperfect

---

<sup>1</sup>See also Stole (2007) for a comprehensive survey.

competition by allowing consumer heterogeneity not only over “vertical” preferences (for quality) but also over “horizontal” preferences (for brands). Our approach is not only distinct from these earlier works, but it leads to a tractable theory of competitive price discrimination, with new and distinctive empirical implications.

While we confine attention to canonical Mussa-Rosen type preferences for quality, we permit consumer heterogeneity over information sets to take a general form. In particular, we do not restrict ourselves to a particular “matching” process determining which firms belong to the information set of each consumer. To achieve this level of generality, we introduce *sales functions*, which capture in reduced form the information heterogeneity among consumers about firms’ offers. For each type of consumer, the sales function determines the mass of sales of a firm as a function of the ranking (or *quantile*) occupied by the indirect utility induced by its contract relative to the cross-section distribution of indirect utilities in the market (as induced by the contracts of all other firms). The sales functions introduced in this paper play a role similar to that of *matching functions* in the macroeconomics literature.<sup>2</sup>

Importantly, we consider a broad class of sales functions, requiring only that sales are bounded away from zero at *any* quantile (capturing the idea that each firm is the only firm that *some* consumer is aware of), and that sales strictly increase in the ranking of the indirect utility induced by the firm’s contract. These two mild assumptions, together with the ranking property alluded to above, are satisfied by natural random matching models, such as the sample-size search model of Burdett and Judd (1983), the urn-ball matching model of Butters (1977), and the on-the-job search model of Burdett and Mortensen (1998). It is worth reiterating that, because sales in our model depend on *ordinal* properties of indirect utilities, our approach is distinct (and in a sense orthogonal) to the horizontal differentiation approach followed by most of the literature (as mentioned above) where sales depend on *cardinal* properties of indirect utilities.

An equilibrium in our economy consists of a distribution of menus such that every menu in its support is a profit-maximizing response to that distribution. As consumer preferences are unobserved, the menus offered by firms have to satisfy the self-selection constraints inherent to price discrimination. Such constraints create a link between the contracts designed for each consumer type.

The equilibrium distribution over menus that firms offer is non-degenerate (and in fact atomless). In equilibrium, firms are indifferent among a continuum of menus. The cross-section distribution over menus (or, if we follow a mixed strategy interpretation, the firms’ randomization procedure) is determined so as to guarantee that all equilibrium menus generate the same expected profits.

Any profit-maximizing menu balances sales volume, rent extraction, and efficiency considerations. As in the case of monopoly, a menu trades off efficiency and rent extraction across consumer

---

<sup>2</sup>See Petrongolo and Pissarides (2001) for a survey of matching functions, and their associated micro-foundations.

types (as implied by the self-selection constraints). Competition introduces another trade-off: For each consumer type, rent extraction must be traded off against sales volume.

A firm's trade-offs can best be understood by considering how profits depend on the indirect utilities left to consumers. Importantly, we find that a firm's profit function satisfies increasing differences in the indirect utility left to low and high-type consumers. Intuitively, leaving more indirect utility to high types relaxes incentive constraints, and enables firms to decrease the distortions in quality provision present in the low-type contracts. This, in turn, increases the firms' marginal profit associated with increasing the indirect utility left to low types, as marginal sales generate greater surplus.

Our main result characterizes an equilibrium of this economy, which, under mild qualifications, is the unique one. This equilibrium, which we call the *ordered equilibrium*, displays three important properties. First, all menus offered by firms are ordered in the sense that, for any two menus, one of them leaves more indirect utility *uniformly* across types. Second, more generous menus (i.e., menus that leave more indirect utility to all types) provide more efficient (or less distorted) quality levels. Third, more generous menus cover a larger fraction of the consumer population, and generate a greater fraction of profits from sales of low-quality goods.

Our model is also amenable to natural comparative statics exercises. We can use two related measures to capture the degree of competition in a market. The first, and more conventional one, is the total mass of competing firms. The second measure is the degree of informational frictions faced by consumers, which captures how large their information sets are likely to be. As one should expect, we show that as the degree of competition increases, the equilibrium distribution of menus assigns higher mass to menus that generate more indirect utility to consumers and offer more efficient quality provision.

In the limit as competition becomes perfect, the equilibrium distribution converges to the Bertrand outcome, in which quality provision is efficient for all types of consumers, and marginal-cost pricing prevails. In the opposite limiting case, as competitive pressures vanish, the equilibrium distribution approaches the monopolistic outcome of Mussa and Rosen (1978). Our model, therefore, spans the entire spectrum of competitive intensity, from perfect competition to monopoly.

The results above deliver interesting empirical implications. Consider a market where multi-product retailers offer two-part tariffs (where an upgrade fee is summed to a baseline price in case the consumer decides to buy the superior version of the product). In equilibrium, the price of the superior version of the product goes down as menus become generous. In turn, when competition is not too intense, the baseline price, which targets low-type consumers, increases as menus become more generous (this follows because quality provision increases faster than indirect utilities in equilibrium). As a consequence, the empirical correlation between baseline prices and the prices of the superior version of the product should be *negative*. Moreover, as the market becomes more competitive

(due to a reduction on informational frictions, or an increase in the mass of competing firms), the distributions of qualities *and* prices of the baseline good offered by firms *increases* in the sense of first-order stochastic dominance. In fine, low-valuation consumers benefit from competition, as prices increase less than qualities as competition becomes fiercer.

Another implication pertains to data about consumer satisfaction and comparison across retailers (which online markets made increasingly available). The ordered property of equilibrium implies that more popular retailers derive a greater profit share from low-priced products.

The baseline model described above takes as exogenous the process that determines the information sets of consumers. We develop two extensions that endogenize the amount of information possessed by consumers. The first allows consumers to engage in information acquisition. In this setting, consumers can undertake costly investment to increase the size of their sample of offers in the sense of first-order stochastic dominance. The proportions of sales to high and low types is then endogenous: high types have more to gain by investing, collect larger sample sizes, and are thus over-represented in terms of sales relative to their proportion in the population. The second extension allows for an endogenous choice of advertising by firms. In this setting, we are able to revisit a question raised by Butters (1977) regarding the efficiency of advertising. We show that the equilibrium level of advertising is inefficiently low relative to that which would be chosen by a planner able to control the intensity of advertising, but not the offers chosen by firms (i.e., that a planner would choose to subsidize advertising).

The rest of the paper is organized as follows. Below, we close the introduction by briefly reviewing the most pertinent literature. Section 2 describes the model. Section 3 describes our results in the context of a binary type model, while Section 4 analyzes the case of a continuum of types. Section 5 develops the extensions, and Section 6 concludes. All proofs are in the Appendix at the end of the document.

## 1.1 Related Literature

This paper brings the theory of nonlinear pricing under asymmetric information (Mussa and Rosen (1978), Maskin and Riley (1984) and Goldman, Leland and Sibley (1984)) to a competitive setting where consumers are heterogeneously informed about the offers made by firms. Other related literature includes:

**Competition in Nonlinear Pricing.** This article primarily contributes to the literature that studies imperfect competition in nonlinear pricing schedules when consumers make exclusive purchasing decisions (exclusive agency). In one strand of this literature, firms' market power stems from comparative advantages for serving consumer segments. In Stole (1995) such comparative advantages are exogenous, whereas in Champsaur and Rochet (1989) they are endogenous, as firms can commit to a range of qualities before choosing prices.

Another strand of this literature generates market power by assuming that consumers have preferences over brands - see Spulber (1989) for a one-dimensional model where consumers are distributed in a Salop circle, and Rochet and Stole (1997, 2002), Armstrong and Vickers (2001), and Yang and Ye (2008) for multi-dimensional models where brand preferences enter utility additively. These papers study symmetric equilibria, and show that (i) the equilibrium outcome under duopoly lies between the monopoly and the perfectly competitive outcome, and that (ii) when brand preferences are narrowly dispersed, quality provision is efficient and cost-plus-fixed-fee pricing prevails.<sup>3</sup>

Our model offers an alternative to the aforementioned papers, as market power in our model originates from the heterogeneity of consumer information regarding the firms' offers. Our results differ in many respects. First, there is menu dispersion in equilibrium. Second, although ex-ante identical, firms are endogenously segmented with respect to the generosity of their menus, quality provision, and market coverage. Our model is tractable and amenable to comparative statics, leading to empirical implications incompatible with models where consumers avail themselves of all offers in the market.

There is, of course, other work recognizing that consumers may not be perfectly informed about offers in competitive settings.<sup>4</sup> The works of Verboven (1999) and Ellison (2005) depart from the benchmark of perfect consumer information by assuming that consumers observe the baseline prices offered by *all* firms, but have to pay a search cost to observe the price of upgrades (or add-on prices). The focus of these papers is on the strategic consequences of the holdup problem faced by consumers *once* their store choices are made. By taking quality provision as exogenous, these papers ignore the mechanism design issues that are at the core of the present article. Katz (1984) studies a model of price discrimination where a measure of low-value consumers are uninformed about prices while other consumers are perfectly informed. Heterogeneity of information thus takes a very particular form in this model, and price dispersion does not arise (when quantity discounting is permitted, a unique price schedule is offered in equilibrium).

Assuming perfect consumer information, Stole (1991) and Ivaldi and Martimort (1994) study duopolistic competition in nonlinear price schedules when consumers can purchase from more than one firm (common agency).<sup>5</sup> In a related setting, Calzolari and Denicolo (2013) study the welfare effects of contracts for exclusivity and market-share discounts (i.e., discounts that depend on the seller's share of a consumer's total purchases). The analysis of these papers is relevant for markets where goods are divisible and/or exhibit some degree of complementarity, whereas our analysis is relevant for markets where exclusive purchases are prevalent (e.g., most markets for durable goods).

---

<sup>3</sup>See Borenstein (1985), Wilson (1993) and Borenstein *et al* (1994) for numerical results in closely related settings.

<sup>4</sup>There is also work where consumers have imperfect information about offers in the absence of competition. Most closely related to our paper, Villas-Boas (2004) studies *monopoly* price discrimination where consumers randomly observe either some or all elements of the menu.

<sup>5</sup>See Stole (2007) for a comprehensive survey of the common agency literature.

**Price Dispersion.** We borrow important insights from the seminal papers of Butters (1977), Salop and Stigitz (1977), Varian (1980) and Burdett and Judd (1983), that study oligopolistic competition in settings where consumers are differently informed about the prices offered by firms. In these papers, there is complete information about consumer preferences, and firms compete only on prices.<sup>6</sup> Relative to this literature, we introduce asymmetric information about consumers' tastes, and allow firms to compete on price *and* quality.

**Competing Auctioneers.** McAfee (1993), Peters (1997), Peters and Severinov (1997) and Pai (2012) study competition among principals who propose auction-like mechanisms. These papers assume that buyers perfectly observe the sellers' mechanisms, and that the meeting technology between buyers and sellers is perfectly non-rival. This last assumption is relaxed by Eeckhout and Kircher (2010), who show that posted prices prevail in equilibrium if the meeting technology is sufficiently rival. A key ingredient of these papers is that sellers face capacity constraints (each seller has *one* indivisible good to sell), and offer homogenous goods whose quality is exogenous. Our paper differs from this literature in three important respects. First, sellers in our model control both the price and the quality of the good to be sold. Second, we assume away capacity constraints. Third, buyers are heterogeneously informed about the offers made by sellers.

**Search and Matching.** Inderst (2001) embeds the setup of Mussa and Rosen (1978) in a dynamic matching environment, where sellers and buyers meet pairwise and, in each match, each side may be chosen to make a take-it-or-leave offer. His main result shows that inefficiencies vanish when frictions (captured by discounting) are sufficiently small, thus providing a foundation for perfectly competitive outcomes.<sup>7</sup> Frictions in our model have a different nature (they are informational). Yet, we obtain a convergence result similar to that of Inderst, as efficiency prevails in the limit as consumers become perfectly informed.

Faig and Jerez (2005) study the effect of buyers' private information in a general equilibrium model with directed search. They show that if sellers can use two-tier pricing, private information has no bite, and the equilibrium allocation is efficient. In turn, Guerrieri, Shimer and Wright (2010) show that private information leads to inefficiencies in a directed-search environment with common values. Our model is closer to Faig and Jerez (2005), as we study private values. In contrast to Faig and Jerez (2005), our model leads to menu dispersion and distortions.

Our paper is also related to Moen and Rosén (2011), who introduce private information on match quality and effort choice in a labor market with search frictions. We focus on private information about willingness to pay (which is the same for all firms), while workers have private information about the match-specific shock in their model.

---

<sup>6</sup>See, however, Grossman and Shapiro (1984) where customers not only have heterogeneous information about offers, but also about brand preferences. Firms compete in prices and advertising intensities, but do not price discriminate.

<sup>7</sup>In contrast, Inderst (2004) shows that if frictions affect agents' utilities through type-independent costs of search (or waiting), equilibrium contracts are always first-best.

## 2 Model and Preliminaries

The economy is populated by a unit-mass continuum of consumers with single-unit demands for a vertically differentiated good. If a consumer with valuation per quality  $\theta$  purchases a unit of the good with quality  $q$  at a price  $x$ , his utility is

$$u(q, x, \theta) \equiv \theta \cdot q - x.$$

Consumers are heterogeneous in their valuations per quality: the valuation of each consumer is an iid draw from a discrete distribution with support  $\{\theta_l, \theta_h\}$ , where  $\Delta\theta \equiv \theta_h - \theta_l > 0$ , and associated probabilities  $p_l$  and  $p_h$ . Consumers privately observe their valuations per quality. The utility from not buying the good is normalized to zero.

A continuum of firms with mass  $v > 0$  compete by posting *menus* of contracts with different combinations of quality and price. Firms have no capacity constraints and share a technology that exhibits constant returns to scale. The per-unit profit of a firm who sells a good with quality  $q$  at a price  $x$  is

$$x - \varphi(q),$$

where  $\varphi(q)$  is the per-unit cost to the firm of providing quality  $q$ . We assume that  $\varphi(\cdot)$  is twice continuously differentiable, strictly increasing and strictly convex, with  $\varphi(0) = \varphi'(0) = 0$ . Furthermore, we require that  $\lim_{q \rightarrow \infty} \varphi'(q) = \infty$ , which guarantees that surplus maximizer qualities are interior.

We assume that firms' offers stipulate simply that consumers choose a combination of quality and price from a menu of options. Given the absence of capacity constraints, a consumer is assured to receive his choice. We thus rule out stochastic mechanisms as well as mechanisms which condition on the choices of other buyers or on the offers of other firms.<sup>8</sup> Given our restriction to menus of price-quality pairs, it is without loss of generality to suppose firms' menus include only two pairs:  $\mathcal{M} \equiv ((q_l, x_l), (q_h, x_h)) \subset (\mathbb{R}_+ \times \mathbb{R})^2$ , where  $(q_k, x_k)$  is the contract designed for the type  $k \in \{l, h\}$ .<sup>9</sup> Furthermore, every menu has to satisfy the following incentive compatibility constraints: For each type  $k \in \{l, h\}$ ,

$$IC_k : \quad u(q_k, x_k, \theta_k) \geq \max_{\hat{k} \in \{l, h\}} \theta_k \cdot q_{\hat{k}} - x_{\hat{k}}.$$

---

<sup>8</sup>There is no loss of generality in considering deterministic mechanisms, provided that one assumes that each consumer can contract with at most one firm. The difficulties associated with stochastic mechanisms in environments where consumers can try firms sequentially (e.g., a consumer might look for a second firm if the lottery offered by the first firm resulted in bad outcome) are discussed in Rochet and Stole (2002).

<sup>9</sup>Suppose a seller offers a menu with more than two price-quality pairs and that at least one type chooses two or more options with positive probability. It is easily verified that there exists a menu, with a single option intended for each customer type, which yields the same payoff to each type but strictly increases the seller's profit. See Lemma 1 below.

Menus must also be individually rational (IR), i.e.  $u(q_k, x_k, \theta_k) \geq 0$  for each  $k$ . A menu  $\mathcal{M}$  that satisfies the IC and IR constraints is said to be *implementable*. The set of implementable menus is denoted by  $\mathbb{I}$ .

As will be clear shortly, it is convenient to denote by  $\tilde{F}$  be the (possibly degenerate) cross-section distribution over menus prevailing in the economy. This distribution has support  $\mathbb{S}$  contained in the set of implementable menus  $\mathbb{I}$ . The distribution over menus  $\tilde{F}$  induces, for each type  $k$ , a marginal distribution over indirect utilities

$$F_k(\tilde{u}_k) \equiv \text{Prob}[\mathcal{M} : u(q_k, x_k, \theta_k) \leq \tilde{u}_k].$$

We denote by  $\Upsilon_k$  the support of indirect utilities offered to type- $k$  consumers, and by  $f_k$  the density of  $F_k$ , whenever it exists.

The key feature of our model is that there is heterogeneity in the information possessed by consumers about the menus offered by firms. We take a reduced-form approach to modeling this heterogeneity. In particular, we introduce the *sales function*

$$\Phi(u_k | F_k, v, p_k),$$

which determines the mass of sales to type- $k$  consumers obtained by a firm that (i) offers a contract with indirect utility  $u_k$  when (ii) the cross-section cdf of indirect utilities to  $k$ -types is  $F_k$ , (iii) there is a  $v$ -mass of firms in the market, and (iv) there is a  $p_k$ -mass of type- $k$  consumers. For expositional reasons, we defer to the next subsection a detailed discussion about sales functions. We will then clarify how different matching technologies between firms and consumers lead to different sales functions, and detail the economic and technical assumptions that define the class of sales functions considered in this paper.

A firm that faces a cross-section distribution of menus  $\tilde{F}$  (with marginal cdf over type- $k$  indirect utilities  $F_k$ ) chooses a menu  $((q_l, x_l); (q_h, x_h)) \in \mathbb{I}$  to maximize profits

$$\sum_{k=l,h} \Phi(u(q_k, x_k, \theta_k) | F_k, v, p_k) \cdot (x_k - \varphi(q_k)). \quad (1)$$

The next definition formalizes our notion of equilibrium in terms of the cross-section cdf over menus prevailing in the economy.

**Definition 1** [*Equilibrium*] *An equilibrium is a distribution over menus  $\tilde{F}$  (with marginal cdf over type- $k$  indirect utilities  $F_k$ ) such that  $\mathcal{M} \in \text{supp } \tilde{F} \subset \mathbb{I}$  implies that  $\mathcal{M}$  maximizes (1).*

Accordingly, an equilibrium is described by a distribution over menus such that every menu in the support of this distribution is profit-maximizing to firms.

**Remark 1** *The equilibrium definition above renders itself to multiple interpretations. In one interpretation, firms follow symmetric mixed strategies by randomizing over menus according to the distribution  $\tilde{F}$ . Another interpretation is that each firm follows a pure strategy that consists in posting the menu associated with a given quantile of the distribution  $\tilde{F}$ . Alternatively, firms might randomize over different subsets of the support  $\mathbb{S}$  according to the conditional distributions induced by  $\tilde{F}$ .*

The next subsection is devoted to the sales functions described above.

## 2.1 Sales Functions

A number of consumer search/matching models have been proposed to resolve both the Diamond and the Bertrand paradoxes (according to which the equilibrium outcome in oligopolistic markets coincides with the monopolist and the perfectly competitive solutions, respectively). One key common feature of these approaches is that consumers are differently informed about the offers made by firms. In order to derive robust predictions, we proceed by identifying properties of sales functions that hold across a number of natural matching technologies. To clarify ideas, consider the following examples, where, to simplify the exposition, we assume that the cross-section distribution of indirect utilities is continuous.

**Example 1** [*Generalized Burdett and Judd (1983)*] *Let each consumer observe the menus of a sample of firms independently and uniformly drawn from the set of all firms. For each consumer, the size of the observed sample is  $j \in \{0, 1, 2, \dots\}$  with probability  $\omega_j(v)$ , where  $\omega_1(v), \omega_2(v) > 0$  for all  $v > 0$ . The distribution over sample sizes  $\Omega(v) \equiv \{\omega_j(v) : j = 0, 1, 2, \dots\}$  is indexed  $v$ , so as to allow the mass of firms in the market to affect the amount of information observed by consumers. Consumers select the best contract among all menus in their samples.*

*In this case, the sales function faced by firms has the functional form*

$$\Phi(u_k | F_k, v, p_k) = \frac{p_k}{v} \cdot \sum_{j=1}^{\infty} j \cdot \omega_j(v) \cdot F_k(u_k)^{j-1}.$$

The next example presents an important special case of the Burdett-Judd matching technology.

**Example 2** [*Poisson-Burdett-Judd*] *The Poisson-Burdett-Judd search model adds to the search model of Example 1 the feature that the size of the sample observed by each consumer is distributed according to a Poisson law with mean  $\beta \cdot v$ , where  $\beta > 0$ :*

$$\omega_j(v) = \frac{(\beta \cdot v)^j}{j!} \cdot \exp\{-\beta \cdot v\} \quad \text{for } j = 0, 1, 2, \dots$$

*Accordingly, as the mass of firms  $v$  increases, consumers observe larger samples of menus with higher probability (in the sense of likelihood ratio dominance). The parameter  $\beta$  measures how an increase*

in the mass of firms affects the distribution of sample sizes. The sales function of the Poisson-Burdett-Judd model is:

$$\Phi(u_k|F_k, v, p_k) = p_k \cdot \beta \cdot \exp\{-\beta \cdot v \cdot (1 - F_k(u_k))\}.$$

The next example describes the urn-ball matching model of Butters (1977), widely employed in the macro/labor literature.

**Example 3 [Generalized Butters (1977)]** Let the menu offered by each firm be observed by exactly  $n \geq 1$  consumers. The size- $n$  subset of consumers reached by each firm is uniformly (and independently) drawn from the set of all  $n$ -size subsets of consumers. When the number of firms and consumers in the market is large (with ratio  $v$ ), Butters (1977) shows that the sales function faced by firms has the functional form

$$\Phi(u_k|F_k, v, p_k) = p_k \cdot n \cdot \exp\{-n \cdot v \cdot (1 - F_k(u_k))\}.$$

In the original Butters (1977) model,  $n$  is set to one.

It is interesting to note that the Generalized Butters and the Poisson-Burdett-Judd matching technologies can imply identical sales functions. Another natural model of heterogenous information comes from the labor search literature.

**Example 4 [Burdett and Mortensen (1998)]** The “on-the-job search” model of Burdett and Mortensen (1998) studies a dynamic economy in continuous time in which consumers receive ads (each ad describes the menu of a particular firm) according to independent Poisson processes with arrival rate  $\lambda$ . Consumers must make purchasing decisions as soon as an ad arrives, and there is no recall. Each matched consumer purchases continuously from the seller until the match is dissolved. This can occur exogenously due to an event which arrives at Poisson rate  $\gamma$ . Alternatively, consumers may switch firms if they receive an ad describing a more attractive menu. There is a common discount rate equal to  $r$ .

It follows from the analysis of Burdett and Mortensen (1998) that the steady-state outcome of this economy can be modeled as a static competition game which sales function has the functional form

$$\Phi(u_k|F_k, v, p_k) = p_k \cdot \gamma \cdot \left[ \frac{1}{\gamma + \lambda \cdot v \cdot (1 - F_k(u_k))} \right] \left[ \frac{1}{\gamma + r + \lambda \cdot v \cdot (1 - F_k(u_k))} \right].$$

The examples above share a number of features. First, the mass of sales is linear in the mass consumers in the market. Intuitively, these matching models rule out “externalities” among consumers.<sup>10</sup> Second, sales functions depend on  $u_k$  *only through* the rank in the distribution of indirect

---

<sup>10</sup>Such “externalities” might arise due to “word of mouth” or other peer effects.

payoffs to type  $k$ ,  $F_k(u_k)$ . This “ranking property” corresponds to an assumption that consumers are concerned only for the utility of consumption net of transfers (and thus pick the best offer available based on these features), and not with other characteristics of a firm’s offer such as transportation costs or the firm’s identity. Third, sales *strictly increase* in the ranking occupied a given indirect utility.

These three features, together with some other technical requirements satisfied by the examples above, define the class of sales functions considered in this paper.

**Assumption 1** *Let  $\tilde{F}$  be a distribution over menus with  $\text{supp } \tilde{F} \subset \mathbb{I}$ , and marginal distribution over type- $k$  indirect utilities  $F_k$ , with support  $\Upsilon_k$ .*

*At any continuity point  $u_k \in \Upsilon_k$  of  $F_k$ , the sales function  $\Phi(u_k|F_k, v, p_k)$  can be written as*

$$\Phi(u_k|F_k, v, p_k) \equiv p_k \cdot \Lambda(F_k(u_k)|v), \quad (2)$$

where the kernel  $\Lambda(y|v) : [0, 1] \times \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$ :

1. *is continuously differentiable and bounded,*
2. *for each  $v > 0$ , is strictly increasing in  $y$  with derivative  $\Lambda_1(y|v)$  bounded away from zero at any  $y \in [0, 1]$ .*

*At any point where  $F_k$  is discontinuous (i.e., has an atom), sales are determined according to uniform rationing rule.<sup>11</sup>*

A crucial ingredient of Assumption 1, shared by all examples discussed above, is that, for each consumer type, firms with the lowest indirect utility ranking make a positive number of sales. That is,  $\Lambda(0|v) > 0$ . This assumption reflects the fact that each firm’s offer is observed with positive probability by a consumer who has no other offers. Also important is the assumption that the mass of sales is strictly increasing in the indirect utility ranking, as required by Part 2. This means that better deals lead to more sales. This rules out the Diamond Paradox, according to which all firms offering the monopolistic (Mussa-Rosen) menu constitutes an equilibrium. More generally, this property also implies that no equilibria exist in which a positive mass of firms post the same menu. As a result, equilibria *necessarily* involve dispersion on menus.

---

<sup>11</sup>In the example above, this corresponds to the assumption that consumers evenly randomize across identical offers. Formally, if  $u_k \in \Upsilon_k$  is a mass point of  $F_k$ , then

$$\Phi(u_k|F_k, v, p_k) \equiv p_k \cdot \left( F_k(u_k) - \lim_{\tilde{u}_k \uparrow u_k} F_k(\tilde{u}_k) \right)^{-1} \cdot \int_{\lim_{\tilde{u}_k \uparrow u_k} F_k(\tilde{u}_k)}^{F_k(u_k)} \Lambda(y|v) dy.$$

Finally, set  $\Phi(u_k|F_k, v, p_k) = p_k \cdot \Lambda(1|v)$  if  $u_k > \tilde{u}_k$  for all  $\tilde{u}_k \in \Upsilon_k$ , and  $\Phi(u_k|F_k, v, p_k) = p_k \cdot \Lambda(0|v)$  if  $0 \leq u_k < \tilde{u}_k$  for all  $\tilde{u}_k \in \Upsilon_k$ .

It is worthwhile reiterating that the “ranking property” of sales functions imposed by Assumption 1 distinguishes our model from spatial models of competition (such as Hotelling or differentiated Bertrand). In such models, the mass of sales obtained by each firm is a function of the profile of *cardinal* indirect utilities offered to each consumer type. In contrast, in our model the mass of sales is a function of the quantiles (relative to the cross-section) associated with the indirect utilities offered by a firm (i.e., it depends on *ordinal* properties of indirect utilities).

**Remark 2** *While the literature has found it convenient to model competition with a mass of “infinitesimal” firms, our analysis applies just as well to models with finitely many firms. In such cases, the sales function gives a firm’s expected number of sales to type  $k$ . To give a further example, suppose (abusing slightly notation) that  $v \in \mathbb{N} \setminus \{0, 1\}$  identical firms compete for a unit-mass of consumers. Consumers are aware of each firm independently with probability  $a \in (0, 1)$ . The sales function in this case is given by*

$$\Phi(u_k | F_k, v, p_k) = p_k \cdot a \cdot (a \cdot F_k(u_k) + 1 - a)^{v-1},$$

*which satisfies Assumption 1. With finitely many firms, the solution concept of Definition 1 then corresponds to symmetric Nash equilibria. For consistency, the analysis below considers the case of a continuum of firms.*

In the baseline model described above, the information possessed by consumers is determined by an *exogenous* matching technology. Subsection 5.1 extends this model to allow consumers to engage in information acquisition.

## 2.2 Incentive Compatibility and Indirect Utilities

A key step in our analysis is to formulate the firms’ maximization problem in terms of the of indirect utilities offered to consumers. To this end, denote by

$$q_k^* \equiv \arg \max_q \theta_k \cdot q - \varphi(q),$$

the efficient quality for type- $k$  consumers, and let  $S_k^* \equiv \theta_k \cdot q_k^* - \varphi(q_k^*)$  be the social surplus associated with the efficient quality provision. The next lemma uses the incentive constraints and the optimality of equilibrium contracts to map indirect utilities into quality levels.

**Lemma 1 [Incentive Compatibility]** *Consider a menu  $\mathcal{M} = \{(q_l, x_l), (q_h, x_h)\}$  in the support of the equilibrium distribution over menus,  $\tilde{F}$ , and let  $u_k \equiv u(q_k, x_k, \theta_k)$ . Then, for all  $k \in \{l, h\}$ ,*

$$q_k = \mathbf{1}_k(u_h - u_l) \cdot \frac{u_h - u_l}{\Delta\theta} + (1 - \mathbf{1}_k(u_h - u_l)) \cdot q_k^*, \quad (3)$$

*where  $\mathbf{1}_h(z)$  is an indicator function that equals one if and only if  $z > q_h^* \cdot \Delta\theta$ , and  $\mathbf{1}_l(z)$  is an indicator function that equals one if and only if  $z < q_l^* \cdot \Delta\theta$ .*

The result above is standard in adverse selection models. Consider some menu  $\mathcal{M} \in \text{supp}(\tilde{F})$  offered in equilibrium. If the  $\text{IC}_k$  constraint does not bind under  $\mathcal{M}$ , then profit-maximization by firms implies that the quality provision to the other type of consumer (i.e., type  $-k$ ) is efficient under  $\mathcal{M}$ . However, if the  $\text{IC}_k$  constraint does bind under  $\mathcal{M}$ , then the quality to consumers of type  $-k$  is chosen to make type- $k$  consumers indifferent between either contract. These facts are summarized in equation (3). Using this equation, we may henceforth let  $q_k(u_l, u_h)$  denote the quality supplied to type  $k$  when the indirect utilities offered are  $(u_l, u_h)$ .

In light of Lemma 1, we can describe each menu in the support of  $\tilde{F}$  in terms of the indirect utilities induced by  $\mathcal{M}$ . Accordingly, we shall write  $\mathcal{M} = (u_l, u_h)$  to describe the menu  $\mathcal{M} = ((q_l, x_l), (q_h, x_h))$ , where the map between  $q$ 's and  $u$ 's follows from equation (3). In a similar fashion, for convenience, we will more often refer to the marginal distribution over indirect utilities,  $F_k$ , rather than to the distribution over menus  $\tilde{F}$ .

Two natural benchmarks play an important role in the analysis that follows. The first one is the static monopolistic (or Mussa-Rosen) solution. Under this benchmark, the quality provided to low types, denote it  $q_l^m$ , is implicitly defined by:

$$\varphi'(q_l^m) = \max \left\{ \theta_l - \frac{p_h}{p_l} \cdot \Delta\theta, 0 \right\}. \quad (4)$$

We interpret  $q_l^m = 0$  as meaning that low-type consumers are not served under the monopolistic solution. In turn, quality provision for high types is efficient:  $q_h^m = q_h^*$ . Finally, recall that, in the monopolistic solution, the indirect utility left to low types is zero,  $u_l^m = 0$  (as the IR is binding), and the indirect utility left to high types is  $u_h^m = q_l^m \cdot \Delta\theta$ , as the  $\text{IC}_h$  is binding. Written in terms of indirect utilities, the menu  $\mathcal{M}^m \equiv (0, q_l^m \cdot \Delta\theta)$  is the *monopolist* (or *Mussa-Rosen*) *menu*.

The second benchmark is the competitive (or Bertrand) solution. Under this benchmark, quality provision is efficient to both types, and firms derive zero profits from each contract in the menu. Written in terms of indirect utilities, the menu  $\mathcal{M}^* \equiv (S_l^*, S_h^*)$  is the *competitive* (or *Bertrand*) *menu*. We can now proceed to characterizing the equilibrium of our model.

### 3 Screening and Competition

This section contains our main results. We start by studying the firms' profit-maximization problem. We then characterize equilibrium, study its main properties, and conduct a number of comparative statics exercises. The last subsection discusses equilibrium uniqueness.

#### 3.1 Firm Problem

For each menu  $\mathcal{M} = (u_l, u_h)$  offered in equilibrium, let

$$S_k(u_l, u_h) \equiv \theta_k \cdot q_k(u_l, u_h) - \varphi(q_k(u_l, u_h)) \quad (5)$$

be the social surplus induced by  $\mathcal{M}$  for each consumer type, where the quality levels  $q_k(u_l, u_h)$  are computed according to (3). We can then write the profit from type- $k$  consumers produced by the menu  $\mathcal{M} = (u_l, u_h)$  as  $S_k(u_l, u_h) - u_k$ .

Employing Lemma 1 and Assumption 1, we can rewrite the firm's profit-maximization problem (in response to the cross-section cdf's over indirect utilities  $\{F_l, F_h\}$ ) as that of choosing menus  $(u_l, u_h)$  to maximize

$$\pi(u_l, u_h) \equiv \sum_{k=l,h} p_k \cdot \Lambda(F_k(u_k)|v) \cdot (S_k(u_l, u_h) - u_k), \quad (6)$$

subject to the constraint  $u_h \geq u_l \geq 0$ . This constraint guarantees that menus are individually rational. Together with the definition of the surplus function  $S_k(u_l, u_h)$ , this constraint also guarantees that menus are incentive compatible, as required by implementability.

To better understand the firms' trade-offs, we will now analyze the first-order conditions associated with (6). We will follow the common practice in mechanism design of assuming that  $IC_l$  is slack in equilibrium, in which case  $IC_h$  is the only potentially binding constraint. As will become clear, this is indeed true in any equilibrium of this economy. Assuming differentiability of  $F_k$  for each  $k \in \{l, h\}$  (which will hold in the equilibrium we construct below), the first-order conditions for the firm's problem are

$$\underbrace{p_h \cdot \Lambda_1(F_h(u_h)|v) \cdot f_h(u_h) \cdot (S_h^* - u_h)}_{\text{sales gains}} - \underbrace{p_h \cdot \Lambda(F_h(u_h)|v)}_{\text{profit losses}} + \underbrace{p_l \cdot \Lambda(F_l(u_l)|v) \cdot \frac{\partial S_l}{\partial u_h}(u_l, u_h)}_{\text{efficiency gains}} = 0 \quad (7)$$

for  $u_h$  and

$$\underbrace{p_l \cdot \Lambda_1(F_l(u_l)|v) \cdot f_l(u_l) \cdot (S_l(u_l, u_h) - u_l)}_{\text{sales gains}} - \underbrace{p_l \cdot \Lambda(F_l(u_l)|v)}_{\text{profit losses}} + \underbrace{p_l \cdot \Lambda(F_l(u_l)|v) \cdot \frac{\partial S_l}{\partial u_l}(u_l, u_h)}_{\text{efficiency losses}} = 0 \quad (8)$$

for  $u_l$ . Intuitively, the firms' choice of menus balances sales, profit, and efficiency considerations.

First consider the first-order condition for high types, given by equation (7). The first two terms in (7) are familiar from models without asymmetric information on type. By increasing the indirect utility  $u_h$ , the firm increases sales (the first term), but decreases profits (the second term). The third term captures the effect of an increase in  $u_h$  on the quality offered to low-type consumers. When  $IC_h$  is slack (i.e.,  $u_h > u_l + \Delta\theta \cdot q_l^*$ ), high types have no incentive to imitate low types, and this term is zero. Let us then focus on the complementary case where  $IC_h$  is binding. As implied by profit maximization, the low-type quality is set to satisfy the constraint  $u_h \geq u_l + \Delta\theta \cdot q_l$  with equality. As a consequence, an increase in  $u_h$  relaxes this constraint and allows the firm to marginally increase the quality to low-type consumers by

$$\frac{\partial q_l(u_l, u_h)}{\partial u_h} = \left( \frac{1}{\Delta\theta} \right).$$

Therefore, the efficiency gains from increasing the quality of high types are generated by the decrease in distortions of the contract to low types, and equal

$$p_l \cdot \Lambda(F_l(u_l)|v) \left( \frac{\theta_l - \varphi'(q_l)}{\Delta\theta} \right) > 0, \quad (9)$$

which is the third term in equation (7).

Let us now consider the first-order condition for low-type utilities, given by equation (8). The first two terms are familiar from (7). In contrast to (7), however, increasing  $u_l$  has the effect of tightening the incentive constraint  $IC_h$ , which implies that the quality distortion present in the low types' contract has to increase. This efficiency loss is the third term in equation (8). By the same reasoning as above, this term has the same magnitude as (9), but the opposite sign.

Equations (7) and (8) thus capture the role of private information about consumer preferences in the firms' choice of menus. One way to see this is to contrast the first-order conditions above with the case where information asymmetries are absent. In this case, each firm's problem of determining the indirect utility to leave to each consumer type would be completely separable; we would have  $S_l(u_l, u_h) = S_l^*$ , and the third terms in (7) and (8) would be zero. Instead, when consumer types are private information, the problems of choosing  $u_l$  and  $u_h$  are interdependent (as implied by incentive constraints). Our equilibrium analysis of the next subsections will clarify how firms simultaneously resolve the efficiency-rent-extraction and the rent-extraction-sales-volume trade-offs in equilibrium.

As a step towards characterizing equilibria, we establish the increasing differences property of expected profits  $\pi$  which was discussed in the Introduction.

**Lemma 2** [*Increasing differences*] *Consider any two implementable menus  $(u_l^1, u_h^1)$  and  $(u_l^2, u_h^2)$ , with  $u_l^2 > u_l^1$  and  $u_h^2 > u_h^1$ . Then we have*

$$\pi(u_l^2, u_h^2) - \pi(u_l^2, u_h^1) \geq \pi(u_l^1, u_h^2) - \pi(u_l^1, u_h^1). \quad (10)$$

*If some incentive constraint binds for at least one of these menus (i.e.,  $u_h^m - u_l^m \notin [q_l^* \cdot \Delta\theta, q_h^* \cdot \Delta\theta]$  for some  $m \in \{1, 2\}$ ), then the inequality in (10) is strict. Otherwise, (10) holds with equality.*

The intuition for this result can easily be understood from the first-order conditions derived above. For simplicity, suppose that  $IC_h$  is binding for both menus (i.e.,  $u_h^m - u_l^m < \Delta\theta q_l^*$  for  $m \in \{1, 2\}$ ).<sup>12</sup> In this case,  $q_l = \frac{u_h - u_l}{\Delta\theta}$ , and increasing  $u_h$  from  $u_h^1$  to  $u_h^2$  raises the quality supplied to the low type. This increases the marginal profit of raising  $u_l$  for two reasons. First, the sales gains from raising  $u_l$  (which is the first term in (8)) go up as  $u_h$  increases. Second, the efficiency losses from raising  $u_l$  (which is the third term in (8)) go down (in absolute value) as  $u_h$  increases. This is so because the cost of quality  $\varphi$  is convex, in which case a marginal reduction in low-type quality has less effect on

<sup>12</sup>The intuition for the case where the low types' incentive constraint binds is similar. However, we will show that this constraint does not bind in equilibrium.

surplus when this quality is closer to its first-best level. These effects are summarized by the cross derivative of the profit function  $\pi$  at any menu for which  $IC_h$  is binding. Assuming differentiability of  $F_l$ , this partial derivative is

$$\frac{\partial^2 \pi(u_l, u_h)}{\partial u_h \partial u_l} = p_l \cdot f_l(u_l) \cdot \Lambda_1(F_l(u_l)|v) \left( \frac{\theta_l - \varphi'(q_l)}{\Delta\theta} \right) + \frac{p_l \cdot \Lambda(F_l(u_l)|v) \cdot \varphi''(q_l)}{(\Delta\theta)^2}, \quad (11)$$

as can be directly computed from either (7) or (8).<sup>13</sup> The first term captures the effect of  $u_h$  on the sales gain from raising  $u_l$ , while the second term captures the effect of  $u_h$  on the efficiency loss from raising  $u_l$ . Both terms are positive (and the second is necessarily strictly positive).

In contrast, if no incentive constraints bind at some menu  $(u_l, u_h)$ , the profit function  $\pi$  exhibits constant differences; i.e.,  $\frac{\partial^2 \pi(u_l, u_h)}{\partial u_h \partial u_l} = 0$ . In this case, as established by Lemma 1, optimality requires that qualities are fixed at their efficient levels to both consumer types, and the effects of  $u_l$  and  $u_h$  on profits are completely separable.

Before moving to equilibrium characterization, we will make use of Lemma 2 to establish that, in any equilibrium, the distributions over indirect utilities,  $F_l$  and  $F_h$ , are absolutely continuous, and have support on an interval that starts at the indirect utility associated with the monopolistic (Mussa-Rosen) menu.

**Lemma 3 [Support]** *In any equilibrium of this economy, the marginal cdf over indirect utilities for type  $k \in \{l, h\}$ ,  $F_k$ , is absolutely continuous. Its support is*

$$\Upsilon_k = [u_k^m, \bar{u}_k],$$

where  $\bar{u}_k < S_k^*$ .

The lemma above has a number of important implications. First, because the distributions  $F_k$  are absolutely continuous, no equilibria exist in which a positive mass of firms post the same menu. To see why, consider a distribution over menus  $\tilde{F}$  with a mass point at some menu  $\mathcal{M} = (u_l, u_h)$  with  $u_k < S_k^*$ . For some  $\varepsilon > 0$  sufficiently small, firms strictly increase profits by offering the menu  $\mathcal{M}' = (u_l + \varepsilon, u_h + \varepsilon)$ . The reason is that, relative to  $\mathcal{M}$ , the menu  $\mathcal{M}'$  involves a marginal loss in profits per sales, but a discrete gain in the mass of sales (as implied by the fact that the kernel  $\Lambda(y|v)$  is strictly increasing in the quantile  $y$ ).<sup>14</sup>

Second, the minimum indirect utilities offered in equilibrium are those induced by the monopoly menu. The proof in the Appendix first establishes, employing Lemma 2, that the “infimum” menu  $(\min \Upsilon_l, \min \Upsilon_h)$  must be an optimal menu. It then shows that optimality requires that  $(\min \Upsilon_l, \min \Upsilon_h) = (0, u_h^m)$ . To see why, suppose to the contrary. If  $\min \Upsilon_h > u_h^m$ , then offering

<sup>13</sup>Differentiability of  $F_l$  holds in equilibrium, but is not assumed in the proof of Lemma 2.

<sup>14</sup>This intuition is familiar from models of price dispersion under complete information, e.g., Varian (1980).

$(0, u_h^m)$  instead of  $(\min \Upsilon_l, \min \Upsilon_h)$  does not affect sales (since the quantile occupied by the high-type indirect utility is the same), but it strictly increases the profits per sale (since the Mussa-Rosen menu  $(0, u_h^m)$  uniquely maximizes profits *conditional* on a sale). If  $\min \Upsilon_h < u_h^m$ , then offering  $(0, u_h^m)$  instead of  $(\min \Upsilon_l, \min \Upsilon_h)$  strictly increases profits for two reasons. First, it increases profits on the sales that would be made under  $(\min \Upsilon_l, \min \Upsilon_h)$  (i.e., to consumers with no other offers). Second, it increases sales for high types (which are profitable), since the quantile occupied by the high-type indirect utility is higher under the menu  $(0, u_h^m)$ . A similar argument establishes that  $\min \Upsilon_l = 0$ . The proof contained in the Appendix formalizes the heuristics above, and applies similar ideas to establish that  $\Upsilon_k$  is connected.

### 3.2 Ordered Equilibrium

We construct an equilibrium in which firms that cede high indirect utilities to high types also cede high indirect utilities low types. We say that equilibria that satisfy this property are ordered.

**Definition 2 [Ordered Equilibrium]** *An equilibrium is said to be ordered if, for any two menus  $\mathcal{M} = (u_l, u_h)$  and  $\mathcal{M}' = (u'_l, u'_h)$  offered in equilibrium,  $u_l < u'_l$  if and only if  $u_h < u'_h$ . In this case, the menu  $(u'_l, u'_h)$  is said to be more generous than the menu  $(u_l, u_h)$ .*

As the next proposition establishes, there always exists a unique ordered equilibrium. We then identify below natural conditions under which the ordered equilibrium is the only equilibrium of the economy. Ordered equilibria have the following important property.

**Remark 3 [Support Function]** *In every ordered equilibrium, the support of indirect utilities offered by firms can be described by a strictly increasing and bijective support function  $\hat{u}_l : \Upsilon_h \rightarrow \Upsilon_l$  such that, for every menu  $\mathcal{M} = (u_l, u_h)$  in  $\Upsilon_l \times \Upsilon_h$ ,  $u_l = \hat{u}_l(u_h)$ .*

Remark 3 tells us that there is a strictly increasing function  $\hat{u}_l$  that determines the utility offered to the low type as a function of the utility of the high type. We find it notationally convenient to denote the identity function by  $\hat{u}_h(u_h) = u_h$  for all  $u_h \in \Upsilon_h$ . Proposition 1 characterizes the unique ordered equilibrium of the economy.

**Proposition 1 [Equilibrium Characterization]** *There exists a unique ordered equilibrium. In this equilibrium, the support of indirect utilities offered by firms is described by the support function  $\hat{u}_l : [u_h^m, \bar{u}_h] \rightarrow [0, \bar{u}_l]$  that is the unique solution to the differential equation*

$$\hat{u}'_l(u_h) = \frac{S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)}{S_h^* - u_h} \cdot \frac{1 - \frac{p_l}{p_h} \cdot \frac{\partial S_l}{\partial u_h}(\hat{u}_l(u_h), u_h)}{1 - \frac{\partial S_l}{\partial u_l}(\hat{u}_l(u_h), u_h)} \quad (12)$$

with boundary condition  $\hat{u}_l(u_h^m) = 0$ .

The equilibrium distribution over menus solves

$$\frac{\Lambda(F_h(u_h)|v)}{\Lambda(0|v)} = \frac{\sum_{k=l,h} p_k \cdot (S_k(0, u_h^m) - u_k^m)}{\sum_{k=l,h} p_k \cdot (S_k(\hat{u}_l(u_h), u_h) - \hat{u}_k(u_h))}, \quad (13)$$

and the supremum point  $\bar{u}_h$  is determined by  $F_h(\bar{u}_h) = 1$ .

The existence of an ordered equilibrium is intimately related to the increasing differences property of firms' profit functions established in Lemma 2. Intuitively, if a firm offers a higher payoff to the high type, it should also do so for the low type; i.e., equilibrium offers should be ordered. The differential equation (12) (together with the boundary condition  $\hat{u}_l(u_h^m) = 0$ ) describes precisely the relationship between these payoffs. Equation (13) then describes the marginal distribution over high-type payoffs  $F_h$ , and thus the distribution over the menus offered by firms. We now sketch the main arguments in arriving at Proposition 1.

**Proof Sketch of Proposition 1.** We proceed in three steps. First, we construct the support function  $\hat{u}_l(\cdot)$ . In the second step, we derive the equilibrium distribution over menus. In the last step, we show that firms cannot benefit from deviating to an out-of-equilibrium menu.

### Step 1 Constructing the support function

Because of the ranking property of kernels, it follows that in any ordered equilibrium with support function  $\hat{u}_l(\cdot)$ ,

$$\Lambda(F_h(u_h)|v) = \Lambda(F_l(\hat{u}_l(u_h))|v). \quad (14)$$

The equation above implies that sales to each type  $k$  are proportional to the probability of that type  $p_k$ . Accordingly, the support function  $\hat{u}_l(\cdot)$  describes the locus of indirect utility pairs  $(\hat{u}_l(u_h), u_h)$  such that the proportion of sales to each type is constant.

Differentiating the expression above, we obtain

$$\hat{u}_l'(u_h) = \frac{\Lambda_1(F_h(u_h)|v) \cdot f_h(u_h)}{\Lambda(F_h(u_h)|v)} \cdot \left[ \frac{\Lambda_1(F_l(\hat{u}_l(u_h))|v) \cdot f_l(\hat{u}_l(u_h))}{\Lambda(F_l(\hat{u}_l(u_h))|v)} \right]^{-1}. \quad (15)$$

Intuitively, the slope of the support function,  $\hat{u}_l'(u_h)$ , equals the ratio between the semi-elasticities of sales with respect to indirect utilities for each type of consumer.

The first-order conditions (7) and (8) provide an alternative expression for these semi-elasticities. Evaluated at the locus  $(\hat{u}_l(u_h), u_h)$ , with the help of (14), equations (7) and (8) can be rewritten as

$$p_k \cdot \frac{\Lambda_1(F_k(\hat{u}_k(u_h))|v) \cdot f_k(\hat{u}_k(u_h))}{\Lambda(F_k(\hat{u}_k(u_h))|v)} \cdot (S_k(\hat{u}_l(u_h), u_h) - u_k) = p_k - p_l \cdot \frac{\partial S_l}{\partial u_k}(\hat{u}_l(u_h), u_h), \quad (16)$$

for  $k = h$  and  $k = l$ , respectively. In equilibrium, the optimality of firms' menus requires that the support function  $\hat{u}_l(\cdot)$  simultaneously satisfies the first-order conditions (16) and equation (15).

Combining these two equations leads to the differential equation (12) which describes how the utility of the low type relates to the utility of the high type in the equilibrium menus.

From Lemma 3, we know that the least generous menu in equilibrium is the Mussa and Rosen menu  $(0, u_h^m)$ . Hence, we require that the solution to (12) satisfy the initial condition  $\hat{u}_l(u_h^m) = 0$ . Finally, the increasing differences property of the profit function, established in Lemma 2, implies that the solution to the differential equation (12) satisfies  $\hat{u}_l'(u_h) > 0$ , in which case the menus  $(\hat{u}_l(u_h), u_h)$  are indeed ordered.

We also need to verify that  $IC_l$  is never binding in any menu  $(\hat{u}_l(u_h), u_h)$ . Indeed, we are able to show that, for all  $u_h \in [u_h^m, \bar{u}_h]$ ,

$$u_h - \hat{u}_l(u_h) \leq \bar{u}_h - \hat{u}_l(\bar{u}_h) < S_h^* - S_l^* < \Delta\theta \cdot q_h^*,$$

which, by Lemma 1, implies that  $IC_l$  is slack at any equilibrium menu (see the proof in the Appendix for details).

### Step 2 *Constructing the distribution over menus*

In view of the support function  $\hat{u}_l(\cdot)$ , we can describe the equilibrium distribution over menus in terms of the distribution of indirect utilities to high type consumers,  $F_h(\cdot)$ . The key idea in the construction is to choose, for each  $u_h$ , the quantile  $F_h(u_h)$  in a way that all menus offered in equilibrium lead to the same expected profits as the Mussa-Rosen menu  $\mathcal{M}^m$ . This is reflected in the indifference condition (13). Importantly, we find that  $\Lambda(F_h(u_h)|v) = \Lambda(F_l(\hat{u}_l(u_h))|v)$  is strictly increasing in  $u_h$ ; or equivalently, by the indifference condition, profits conditional on sale

$$p_l(S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)) + p_h(S_h(\hat{u}_l(u_h), u_h) - u_h)$$

are strictly decreasing in  $u_h$ . Together with Assumption 1.2, this guarantees that  $F_h(\cdot)$  is indeed an increasing function.

In order to complete the construction of  $F_h(\cdot)$ , we need to determine the support of high type indirect utilities,  $\Upsilon_h$ . By Lemma 3,  $\Upsilon_h$  is a closed interval of the form  $[u_h^m, \bar{u}_h]$ , so we are only left to compute the upper limit of  $\Upsilon_h$ ,  $\bar{u}_h$ . In the Appendix, we show that the solution to the differential equation (12) satisfies  $\hat{u}_l(S_h^*) = S_l^*$ , that is: When high types receive their Bertrand utility  $S_h^*$ , so do low types. This property implies that the right-hand side of the indifference condition (13) approaches infinity as  $u_h \rightarrow S_h^*$ . This, together with the boundedness of  $\Lambda(\cdot|v)$  for each  $v$  (as required by Assumption 1.3), guarantees that there exists a unique  $\bar{u}_h < S_h^*$  for which  $F_h(\bar{u}_h) = 1$ .

### Step 3 *Verifying the optimality of equilibrium menus*

Finally, we verify that no seller has a profitable deviation. Observe first that no deviation to a menu that leads to indirect utilities outside of the range  $\Upsilon_l \times \Upsilon_h = [u_l^m, \hat{u}_l(\bar{u}_h)] \times [u_h^m, \bar{u}_h]$  can be

optimal. Consider therefore a menu  $(u'_l, u'_h) \in \Upsilon_l \times \Upsilon_h$  with  $u'_l \neq \hat{u}_l(u'_h)$ . We show in the Appendix that the gains from this deviation relative the the equilibrium menu  $(\hat{u}_l(u'_h), u'_h)$  equal

$$\pi(u'_l, u'_h) - \pi(\hat{u}_l(u'_h), u'_h) = - \int_{u'_l}^{\hat{u}_l(u'_h)} \int_{\hat{u}_l^{-1}(\tilde{u}_l)}^{u'_h} \frac{\partial^2 \pi(\tilde{u}_l, \tilde{u}_h)}{\partial u_h \partial u_l} d\tilde{u}_h d\tilde{u}_l,$$

which is non-positive by virtue of the increasing-differences property established in Lemma 2. This completes the proof of Proposition 1. Q.E.D.

It is worth noting some interesting features of the equilibrium characterized in the above proposition. First, the the support function  $\hat{u}_l(\cdot)$  does not depend on the function  $\Lambda(\cdot|v)$ . Accordingly, the function  $\hat{u}_l(\cdot)$  is *invariant* to the matching process that determines the consumers' information sets. However, the support  $\mathbb{S}$  of equilibrium menus *does* depend on the function  $\Lambda(\cdot|v)$ , but *only* through the supremum indirect utility  $\bar{u}_h$  (determined by the indifference condition (13)). As revealed by condition (13), the function  $\Lambda(\cdot|v)$  also plays an important role in determining the cross-section distribution over menus prevailing in the economy.

In what follows, we focus attention on the ordered equilibrium described above. In subsection 3.6, we present a complete characterization of the equilibrium set, and show that little (if anything) is lost by restricting attention to ordered equilibrium.

### 3.3 Equilibrium Properties

Recall from the best-response analysis of subsection 4.1 that, *ceteris paribus*, increasing the indirect utility of high and low types have opposing effects on how much firms optimally distort the quality in low-type contracts. Which of these countervailing effects prevails in equilibrium?

Our characterization in Proposition 1 answers this question. A key property of the support function is that  $\delta(u_h) \equiv u_h - \hat{u}_l(u_h)$  is strictly increasing in  $u_h$ , reaching its maximum at the upper limit of  $\Upsilon_h$ ,  $\bar{u}_h$ . Intuitively, this reflects the fact that competition for high types is fiercer than competition for low types in equilibrium, as high-type consumers “have more surplus to share” with firms. An immediate consequence is that, whenever  $IC_h$  binds, the quality provided to low types,

$$q_l(\hat{u}_l(u_h), u_h) = \frac{\delta(u_h)}{\Delta\theta}, \tag{17}$$

is strictly increasing in  $u_h$ . Proposition 2 below formalizes these claims, as well as providing an implication for firms' profits.

**Proposition 2** [*Equilibrium Properties*] *The following properties hold in the ordered equilibrium.*

1. **Efficiency:** *Menus for which consumers earn higher payoffs are more efficient. In particular, there exists  $u_h^c \in (u_h^m, S_h^*)$  such that the social surplus produced by low-type contracts,*

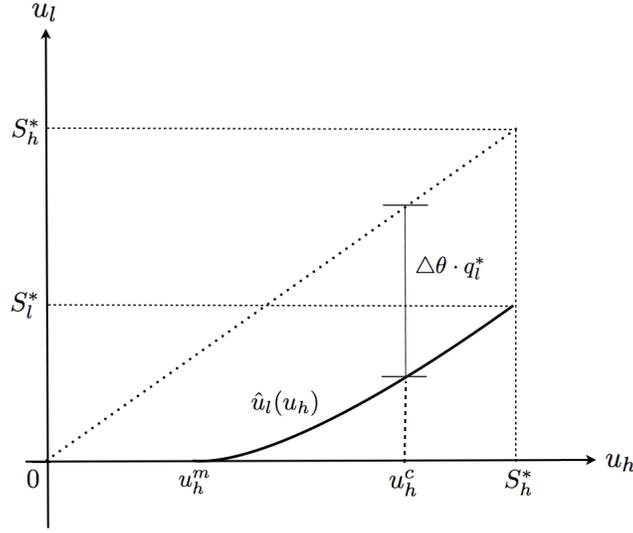


Figure 1: The equilibrium support function  $\hat{u}_l(\cdot)$ . The dotted line is the 45-degree line.

$S_l(\hat{u}_l(u_h), u_h)$ , is strictly increasing in  $u_h$  whenever  $u_h < u_h^c$ , and such that  $S_l(\hat{u}_l(u_h), u_h) = S_l^*$  whenever  $u_h \geq u_h^c$ .<sup>15</sup>

2. **Profits:** Firms which offer more generous menus obtain a larger fraction of their profits from low-type consumers.

The first statement in Proposition 2 establishes the existence of a threshold  $u_h^c$  on the high-type indirect utility above which equilibrium menus are efficient. Recall that, by Lemma 1, efficient quality is supplied to low types if and only if

$$\delta(u_h) \geq \Delta\theta \cdot q_l^*.$$

Therefore, the threshold  $u_h^c$  solves  $\delta(u_h^c) = \Delta\theta \cdot q_l^*$ .

The second statement in Proposition 2 shows that firms sort themselves in equilibrium according to the composition of their profits. It establishes that firms that offer more generous (or equivalently, more efficient) menus derive a higher share of profits from low-type consumers. As menus become more generous, the ratio of profits derived from low and high types approaches the upper bound and is constant at this level for all menus that provide quality efficiently for both types (i.e., those menus for which  $u_h \geq u_h^c$ ).

Figure 1 illustrates the proposition above. This figure represents the entire graph of the support function,  $\{(\hat{u}_l(u_h), u_h) : u_h \in [u_h^m, S_h^*]\}$ ; which of these offers are made in equilibrium depends on the supremum point  $\bar{u}_h$ .

<sup>15</sup>Whether the threshold  $u_h^c$  belongs to the support  $\Upsilon_h = [u_h^m, \bar{u}_h]$  depends on the kernel  $\Lambda(\cdot|v)$ . See Proposition 5 below, which discusses equilibrium uniqueness and provides necessary and sufficient conditions (in terms of primitives) for  $\bar{u}_h \geq u_h^c$ .

While, in an ordered equilibrium, more efficient menus are always better for consumers, pricing patterns are perhaps more subtle. The reason is that price and quality provision are *substitute* instruments for competing for consumers. To see how prices vary across equilibrium menus, define the prices of the low and high quality goods respectively by

$$\begin{aligned} x_l(u_h) &\equiv \theta_l \cdot q_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h), \quad \text{and} \\ x_h(u_h) &\equiv \theta_h \cdot q_h^* - u_h \end{aligned}$$

for  $u_h \in [u_h^m, S_h^*]$ . We find the following.

**Corollary 1 [Equilibrium Prices]** *There exists  $u_h^d \in (u_h^m, u_h^c]$  such that the price of the low quality good  $x_l(u_h)$  is strictly increasing in  $u_h$  if  $u_h \leq u_h^d$ , and strictly decreasing otherwise. In contrast, the price of the high quality good  $x_h(u_h)$  is strictly decreasing in  $u_h$  over  $[u_h^m, S_h^*]$ .*

The corollary above reveals what contract feature (price or quality) different firms use to compete for *marginal* consumers. Firms of low generosity (i.e., those for which  $u_h \leq u_h^d$ ) compete more fiercely on quality provision, while firms of high generosity compete more fiercely on prices. This is reflected by the fact that, for  $u_h \leq u_h^d$ , the price *and* the quality of the low-quality good strictly increase in the generosity of the menu. In turn, for  $u_h > u_h^d$ , the quality of the low-quality good weakly increases, but its price strictly decreases, in the generosity of the menu. These facts are illustrated in Figure 2.

The reason why the price of the low-quality good is initially increasing in the menu's generosity relates to the shape of the support function  $\hat{u}_l(\cdot)$ . Crucially, the support function  $\hat{u}_l(\cdot)$  is *convex* and has zero derivative at  $u_h^m$ , the Mussa-Rosen indirect utility.<sup>16,17</sup> This implies that, for low values of  $u_h$ , the low-type quality increases fast, while low-type payoffs  $\hat{u}_l(u_h)$  increase slowly in  $u_h$ . Necessarily, therefore, the price of the low-quality good has to increase in  $u_h$ , as dictated by incentive compatibility. In turn, for high values of  $u_h$ , quality provision to low-types increases at a lower rate than indirect utilities, implying that prices have to decrease in  $u_h$ . Finally, that the price of the high-quality good  $x_h(u_h)$  is decreasing follows straightforwardly because the high-type quality remains fixed at its efficient level.

<sup>16</sup>A simple intuition for why  $\hat{u}_l(\cdot)$  is convex is as follows. As the generosity of the menu increases, so does the social surplus generated by the low-type contract (as established in Proposition 2). This implies that, relative to high-types, sales to low-types become increasingly attractive for firms (as the surplus to be shared with consumers from each sale increases). Therefore, relative to high types, competition for low types get fiercer as  $u_h$  increases. This is reflected in the fact that as menus become more generous, the indirect utilities left to low-types increases faster in  $u_h$ .

<sup>17</sup>To get intuition on why  $\hat{u}_l'(u_h) \rightarrow 0$  as  $u_h \rightarrow u_h^m$ , consider the case where  $u_h^m > 0$ . By increasing the generosity of the menu, the firm trades off profits per sale against the increased probability of a sale. For menus in a neighborhood of the monopoly menu  $(0, u_h^m)$ , increasing  $u_h$  has only a second-order effect on the profitability of a sale, since  $u_h^m$  is an interior maximizer of these profits. Increasing  $u_l$ , however, leads to a first-order loss in profits per sale. Indifference over menus therefore requires that the increase in high-type indirect utility be an order of magnitude larger than the increase in low-type indirect utility for the same gain in the probability of a sale.

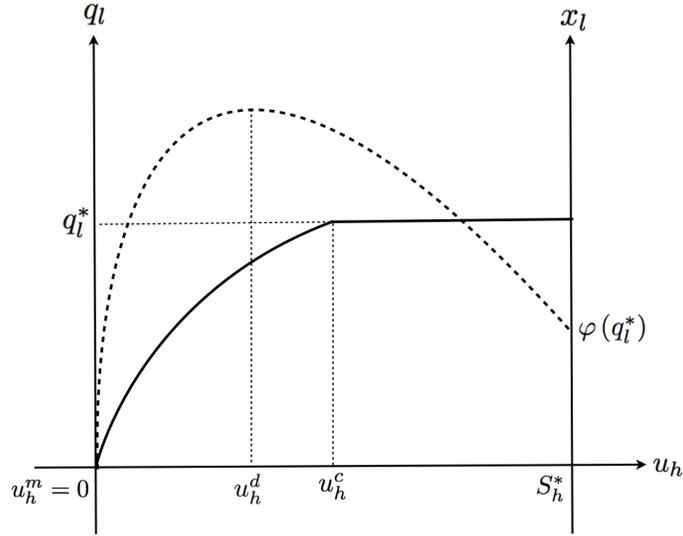


Figure 2: The low-type quality schedule (full curve and left-side Y-axis)) and the low-type price (dotted curve and right-side Y-axis) as a function of the generosity of the menu,  $u_h$ .

The next subsection studies how the distributions over menus (and its support) vary with the degree of competition in the market. In particular, we will show that the highest indirect utility offered to high types in equilibrium,  $\bar{u}_h$ , is increasing in the level of competition in the market. As a consequence, the conclusions of Corollary 1 hold through the *entire* support of equilibrium menus provided competition is not too intense. A testable prediction of our model is therefore that, if competition is not too intense, the low-quality (or baseline) price  $x_l(u_h)$  is negatively correlated with the high-quality (or premium) price  $x_h(u_h)$ . Other testable predictions follow from the comparative statics exercises from the next subsection.

### 3.4 Comparative Statics

We will now investigate how a higher degree of competition affects the equilibrium distribution over menus. Before stating results, we introduce a mild regularity condition on the kernel  $\Lambda(y|v)$ . This condition controls for how sales functions change with the mass of firms  $v$ .

**Condition 1 [VM] V-Monotonicity:** *The kernel ratio*

$$R(y|v) \equiv \frac{\Lambda(y|v)}{\Lambda(0|v)}$$

*is strictly increasing in  $v$  for all  $y \in (0, 1]$ .*

Intuitively, this condition means that, relative to the least generous menu in the cross-section, the proportional gains in sales from offering a contract whose indirect utility lies in some quantile  $y > 0$  increases with the mass of competing firms  $v$ . This captures the idea that, as the number

of competing firms increases, consumers are likely to have larger information sets, in which case increasing the generosity of the offer has a larger impact on sales (relative to the monopolistic offer).

The monotonicity requirement of Condition VM is satisfied by the Generalized Burdett-Judd matching model provided that, for any  $\hat{v} > v$ , the sample size distribution  $\Omega(\hat{v})$  dominates the distribution  $\Omega(v)$  in the likelihood-ratio order. In particular, this assumption is satisfied by the Poisson-Burdett-Judd matching model (and, therefore, by the Butters model, which shares a similar sales function). It is also satisfied by the Burdett-Mortensen matching model.

The next result establishes that, when competition increases, firms more often offer menus that lead to high indirect utilities for both consumer types. As implied by Proposition 2, the mass of firms that offer inefficient qualities in equilibrium decreases as competition gets fiercer.

**Proposition 3** [*Competition and Distortions: Comparative Statics*] *Assume that condition VM holds, and denote by  $F_k$  and  $\hat{F}_k$  (with supports  $\Upsilon_k$  and  $\hat{\Upsilon}_k$ ) the equilibrium distributions over indirect utilities when the mass of firms is  $v$  and  $\hat{v}$ , respectively. If  $v > \hat{v}$ , then*

1.  $F_k$  first-order stochastically dominates  $\hat{F}_k$ , with  $\hat{\Upsilon}_k \subseteq \Upsilon_k$ , for  $k \in \{l, h\}$
2. the fraction of firms offering inefficient qualities is weakly lower for mass  $v$ : i.e.,  $F_h(u_h^c) \leq \hat{F}_h(u_h^c)$ .<sup>18</sup>

The proposition above captures changes in the degree of market competition by varying the mass of firms,  $v$ . An alternative and intimately related notion of competition keeps  $v$  fixed, but varies the level of *frictions* of the random matching technology. This is explored in the next remark.

**Remark 4** [*Frictions and Distortions*] *We say that the matching technology associated with the kernel  $\Lambda(y|v)$  is less frictional than the matching technology associated with the kernel  $\hat{\Lambda}(y|v)$  if for all  $y \in [0, 1]$ ,*

$$\frac{\Lambda(y|v)}{\Lambda(0|v)} \geq \frac{\hat{\Lambda}(y|v)}{\hat{\Lambda}(0|v)}.$$

*This condition describes how sales functions change as the consumers' information sets get larger (in a probabilistic sense). In the Generalized Burdett-Judd model, the matching technology becomes less frictional as the distribution of sample sizes increases in the sense of likelihood ratio dominance. In the Poisson-Burdett-Judd, the level of frictions is captured by the parameter  $\beta$ , which measures how the mass of firms  $v$  impacts the average sample size observed by consumers. In the Butters model, the level of frictions is captured by the parameter  $n$ , which is the number of consumers aware of the menu of each firm.*

---

<sup>18</sup>If the IC-threshold  $u_h^c$  belongs to support  $\Upsilon_h = [u_h^m, \bar{u}_h]$ , an increase in  $v$  can be shown to *strictly* decrease the mass of firms offering inefficient qualities. See Proposition 5 below for necessary and sufficient conditions (in terms of primitives) under which  $\bar{u}_h \geq u_h^c$ .

*Proposition 3 can be recast in terms of the degree of frictions of the matching technology: As the matching technology becomes less frictional, e.g. when  $\beta$  or  $n$  increase, the distributions of indirect utilities increase in the sense of first-order stochastic dominance, and the fraction of firms offering efficient qualities increases.*

Taken together, the results of Propositions 2 and 3 and Corollary 1 deliver a number of empirical implications. First, for markets which are not too competitive, reductions in the informational frictions, or an increase in the mass of competing firms, should lead to higher qualities *and* prices for the baseline good offered by firms (in the sense of first-order stochastic dominance). Second, consider, for each retailer, the distribution of consumer ratings, assuming that such ratings reflect the indirect utilities obtained by different types of consumers. It then follows that the profit share from low-priced products and the distribution of consumer ratings should co-move (the order relation for comparing distributions being, as before, first-order stochastic dominance).

### 3.5 Limiting Cases: Perfect Competition and Monopoly

The next proposition studies limiting properties of equilibria as the mass of firms in the market converges to zero or infinity. These properties hold independently of Condition VM.

#### **Proposition 4** [*Competition and Distortions: Limiting Cases*]

1. *If  $\lim_{v \rightarrow 0} R(1|v) = 1$ , then, as the mass of firms converges to zero,  $v \rightarrow 0$ , the equilibrium distribution over menus converges to a degenerate distribution centered at the monopolistic (Mussa-Rosen) menu  $\mathcal{M}^m$ . In particular, the fraction of firms offering inefficient menus is one for small enough  $v$ .*
2. *If  $\lim_{v \rightarrow \infty} R(y|v) = \infty$  for all  $y \in (0, 1]$ , then, as the mass of firms grows large,  $v \rightarrow \infty$ , the distribution over menus converges to a degenerate distribution centered at the competitive (Bertrand) menu  $\mathcal{M}^*$ . In particular, the fraction of firms offering efficient menus converges to one.*

The first part of Proposition 4 investigates the limit properties of equilibrium when  $v \rightarrow 0$ . It requires that the proportional gains in sales from offering the most generous contract in the cross-section, relative to offering the least generous contract, converges to zero when the mass of competing firms approaches zero. This is a weak condition satisfied by the matching technologies of Examples 2, 3 and 4. It also holds for the Generalized Burdett-Judd matching technology of Example 1 provided that the collection of sample size distributions  $\{\Omega(v) : v > 0\}$  satisfies weak regularity conditions;<sup>19</sup> the Poisson-Burdett-Judd matching technology is a particular case.

---

<sup>19</sup>A sufficient condition is that the  $l_1$ -limit of  $\Omega(v)$  as  $v \rightarrow 0$  has support  $\{0, 1\}$ .

To understand the result, note that, as the mass of firms  $v$  approaches zero, the support of  $h$ -type indirect utilities converges to  $u_h^m$ , the Mussa-Rosen indirect utility. As a consequence, the distribution over menus approaches a degenerate distribution centered at the monopolistic menu. When the parameters of the price-discrimination problem dictate that  $q_l^m = 0$  (see equation (4)), low types are excluded in the limit as  $v \rightarrow 0$ .

The second part of Proposition 4 investigates the limit properties of equilibria when  $v \rightarrow \infty$ . It requires that the proportional gains on sales, relative to the least generous contract, from offering a contract at *any* quantile  $y > 0$ , grows large as  $v \rightarrow \infty$ . This condition is satisfied by the Generalized Burdett-Judd matching technology provided that weak regularity conditions are satisfied;<sup>20</sup> the Poisson-Burdett-Judd matching technology is again a particular case. The condition is also satisfied by the Butters matching technology. However, the condition is not satisfied by the Burdett-Mortensen matching technology. Under this technology, the distribution over the menus that firms offer converges to a non-degenerate distribution. It can be shown however that the distribution over indirect utilities in the buyer-seller relationships that persist in the steady-state equilibrium indeed converges to a degenerate distribution centered at the Bertrand menu.<sup>21</sup>

Importantly, Corollaries 3 and 4 show how our model captures the entire spectrum of industry competitiveness. When  $v$  is small, competition is weak, and we obtain the sensible prediction that firms' behavior is close to that of a firm with complete market power. When  $v$  is large, equilibria approach the outcome of a perfectly competitive market.

**Remark 5 [Vanishing Frictions]** *Similarly to Proposition 3, Proposition 4 can be recast in terms of the degree of frictions of the matching technology. In the case of the Poisson-Burdett-Judd and the Generalized Butters matching technologies (where frictions can be modeled parametrically), we say that frictions vanish as  $\beta \rightarrow \infty$  and  $n \rightarrow \infty$ , respectively. Accordingly, in the limit as frictions vanish, the distribution over menus converges to a degenerate distribution centered at the competitive (Bertrand) menu  $\mathcal{M}^*$ .*

### 3.6 Equilibrium Uniqueness

We will now discuss the important issue of equilibrium uniqueness, and identify the only possible source of equilibrium multiplicity in our model. In a nutshell, the next proposition shows that, when the mass of firms  $v$  is small, the ordered equilibrium is the unique equilibrium. In turn, when the mass

<sup>20</sup>A sufficient condition is that the  $l_1$ -limit of  $\Omega(v)$  as  $v \rightarrow \infty$  has support  $\{2, 3, \dots\}$ .

<sup>21</sup>The proof is available upon request. As described in Example 4, the Burdett-Mortensen model is a dynamic model in which dynamic relationships persist only until the match exogenously terminates or the consumer receives a better offer. Intuitively, as  $v$  becomes large, consumers receive offers very frequently in expectation and hence a relationship in which the consumer earns a payoff bounded below the efficient surplus can be expected to last only a short while. This explains why the distribution over payoffs earned by consumers in the relationships that have formed and not yet broken in the steady state equilibrium of a Burdett-Mortensen economy converges to the Bertrand payoffs.

of firms is large there are equilibria which are not ordered. As will be clear below, the uniqueness of equilibria crucially depends on whether the incentive constraint  $IC_h$  binds for all menus offered in the ordered equilibrium. In the case of multiplicity, all equilibria lead to the same distribution over contracts for each type of consumer as the ordered equilibrium. Therefore, all equilibria induce the same distribution over indirect utilities to each type of consumer, and the same ex-ante profits for firms.

**Proposition 5** [*Incentive Constraints and Equilibrium Uniqueness*] *Assume that condition VM holds, and that  $\lim_{v \rightarrow 0} R(1|v) = 1$  and  $\lim_{v \rightarrow \infty} R(1|v) = \infty$ .<sup>22</sup> Then there exists a threshold  $v^c > 0$  on the mass of competing firms such that:*

1. *if  $v \leq v^c$ , the IC-threshold  $u_h^c$  satisfies  $u_h^c \geq \bar{u}_h$ , and the downward incentive constraint ( $IC_h$ ) is binding for all menus offered in the ordered equilibrium. In this case, the only equilibrium is the ordered equilibrium.*
2. *if  $v > v^c$ , the IC-threshold  $u_h^c$  satisfies  $u_h^c < \bar{u}_h$ , and the downward incentive constraint ( $IC_h$ ) is slack for all menus in the ordered equilibrium with  $u_h > u_h^c$ , and binding for  $u_h \leq u_h^c$ . In this case, there exist multiple equilibria that differ only in the menus for which  $u_h > u_h^c$  (i.e., the efficient menus). However, all equilibria (including the non-ordered ones) lead to the same marginal distributions over indirect utilities  $F_k(\cdot)$ , and the same ex-ante profits for firms.*

The proof, contained in the Appendix, shows that in *any* equilibrium, when the mass of firms is small (i.e.,  $v \leq v^c$ ), the support of utilities of type- $k$  consumers,  $\Upsilon_k$ , is contained in  $[u_k^m, u_k^c]$ . Using the increasing differences property (see Lemma 2) we show that this implies that all equilibria are equal to the ordered equilibrium.

In contrast, when the mass of firms is large (i.e.,  $v > v^c$ ), some menus offered in the ordered equilibrium exhibit non-binding incentive constraints. Consider such a menu  $(\hat{u}_l(u_h), u_h)$ , in which case  $u_h \in (u_h^c, \bar{u}_h]$ . For this menu, the profit function  $\pi(u_l, u_h)$  is locally modular, i.e. its cross-partial derivative is zero. As a result, for some (small)  $\varepsilon > 0$ , both the menus  $(\hat{u}_l(u_h - \varepsilon), u_h)$  and  $(\hat{u}_l(u_h), u_h - \varepsilon)$  are profit-maximizing for the firm. Based on the ordered equilibrium, we can thus construct a non-ordered equilibrium by replacing the menus  $(\hat{u}_l(u_h), u_h)$  and  $(\hat{u}_l(u_h - \varepsilon), u_h - \varepsilon)$  by their non-ordered counterparts  $(\hat{u}_l(u_h - \varepsilon), u_h)$  and  $(\hat{u}_l(u_h), u_h - \varepsilon)$ . Proposition 5 confirms that this is the unique source of multiplicity of equilibria in our economy.

**Remark 6** [*Frictions and Uniqueness of Equilibrium*] *The statements above can be recast in terms of the degree of frictions of the matching technology. Namely, in the case of the Poisson-Burdett-Judd and the Generalized Butters models, the uniqueness result of Proposition 5 holds if and only if the friction parameters  $\beta$  and  $n$  are small enough.*

<sup>22</sup>This is technical condition is satisfied by the matching technologies of Examples 1 2, 3 and 4.

## 4 Competition and Market Coverage with a Continuum of Types

The binary-types model presented above is useful to understand how competition affects quality distortions when firms face asymmetric information regarding consumers types. This model is, however, less appropriate to study the important issue of how competition affects market coverage, and, in particular, on how firms differentiate themselves in equilibrium regarding the breadth of types served by their menus. Indeed, the binary-type model makes the stark prediction that almost all firms serve both consumer types, *regardless* of the degree of competition in the market. This is so even when the monopolist allocation excludes low-type consumers.<sup>23</sup> The aim of this section is to study this issue by extending the binary-types model to a continuum. As we shall see, the model with a continuum of types delivers more nuanced (and interesting) predictions on how competition affects market coverage.

For tractability, we let consumer valuations be uniformly distributed in the unit interval  $[0, 1]$ , and assume that firms costs are quadratic:  $\varphi(q) = \frac{1}{2} \cdot q^2$ . The reason for these assumptions is the following: Characterizing an ordered equilibrium with a continuum of types requires solving a nonlinear partial differential equation with nonstandard boundary conditions (as will be described below). For arbitrary distributions of valuations and cost functions, this equation does not admit a closed-form solution, and the (few) existence results available in the literature on partial differential equations do not apply. While we believe that our results extend to environments other than the uniform-quadratic, computing equilibria in such environments requires numerical techniques which are out of the scope of this work.<sup>24</sup>

The analysis proceeds analogously to that of Section 3. Firms post price-quality menus  $\mathcal{M} \equiv ((q(\theta), x(\theta)) : \theta \in [0, 1])$ , where  $q(\theta)$  is the quality, and  $x(\theta)$  is the price of the contract designed for type  $\theta$ . We let  $u(\theta) \equiv \theta \cdot q(\theta) - x(\theta)$  be the indirect utility of type  $\theta$ . By standard arguments, a menu  $\mathcal{M}$  is incentive compatible (IC) if and only if the indirect utility schedule  $u(\cdot)$  is absolutely continuous (with derivative  $u'(\theta) = q(\theta)$  almost everywhere), and convex. The set of all menus  $\mathcal{M}$  that are incentive compatible and individually rational (i.e.,  $u(\theta) \geq 0$  for all  $\theta \in [0, 1]$ ) is denoted by  $\mathbb{I}$ . For convenience, and in light of incentive compatibility, we write  $\mathcal{M} = u(\cdot)$  to describe the menu  $\mathcal{M} \equiv ((q(\theta), x(\theta)) : \theta \in [0, 1])$ , where  $q(\theta) = u'(\theta)$  and  $x(\theta) = \theta \cdot u'(\theta) - u(\theta)$  for almost every  $\theta$ .

As in the model with binary types, we model the heterogeneity of information possessed by consumers by means of sales functions satisfying Assumption 1. For a given cross-section distribution over menus  $\tilde{F}$  (with support  $\mathbb{S} \subseteq \mathbb{I}$ ), we denote by  $F(u; \theta)$  the marginal distribution over indirect

---

<sup>23</sup>Recall that the reason for this is as follows. If the monopolist's menu excludes the low type, then it offers zero payoff to both types; i.e.,  $(u_l^m, u_h^m) = (0, 0)$ . Under competition, i.e. when the mass of firms  $v$  is positive, firms compete for high types by choosing  $u_h > 0$  with probability 1. For these offers, Proposition 1 predicts that  $0 < u_l < u_h$ , meaning that the low-type quality is positive for almost every menu offered in equilibrium.

<sup>24</sup>In Appendix B, we derive general necessary conditions of equilibria that might be amenable to numerical analysis.

utilities for each type  $\theta$  (with support  $\Upsilon(\theta)$ ). We can therefore write a firm's profit-maximization problem as that of choosing an indirect utility schedule  $u(\cdot)$  to maximize the functional

$$\pi[u] \equiv \int_0^1 \Lambda(F(u(\theta); \theta) | v) \cdot \left( \theta \cdot u'(\theta) - u(\theta) - \frac{1}{2} \cdot [u'(\theta)]^2 \right) d\theta. \quad (18)$$

The expression above computes the total profits of a menu  $u(\cdot)$  by integrating the product of the sales volume,  $\Lambda(F(u(\theta); \theta) | v)$ , with the profits per sale,  $x(\theta) - \frac{1}{2} \cdot q(\theta)^2$ , over all types  $\theta \in [0, 1]$ .

Analogously to the binary type model of the previous sections, we focus on ordered equilibrium, as formally defined below.

**Definition 3 [Ordered Equilibrium]** *An ordered equilibrium is a distribution over menus  $\tilde{F}$  (with marginal distribution over type- $\theta$  indirect utilities  $F(\cdot; \theta)$ ) such that*

1.  $\mathcal{M} = u(\cdot) \in \mathbb{S} \subseteq \mathbb{I}$  implies that  $\mathcal{M} = u(\cdot)$  maximizes (18),
2. if  $u(\cdot), \hat{u}(\cdot) \in \mathbb{S}$ , and  $u(\tilde{\theta}) > \hat{u}(\tilde{\theta})$  for some  $\tilde{\theta} \in [0, 1]$ , then  $u(\theta) \geq \hat{u}(\theta)$  for all  $\theta \in [0, 1]$ , with strict inequality whenever  $u(\theta) > 0$ .

The first condition in the definition above is the usual profit-maximization requirement. The second condition captures the “ordered” feature of our equilibrium: If a menu is “more generous” to one type of consumer, then it is more generous to all consumer types that are served by that menu.

As in the case with binary types, it is convenient to describe the support  $\mathbb{S}$  by indexing each schedule  $u(\cdot) \in \mathbb{S}$  by the indirect utility received by the highest type  $\theta = 1$ . Accordingly, we denote by  $V(\theta, \bar{u})$  the indirect utility received by type  $\theta$  in the menu where the highest type  $\theta = 1$  obtains utility  $\bar{u}$ . For a given ordered equilibrium, we refer to the bivariate function  $V(\cdot, \cdot)$  as its *support schedule*. Note that  $V(\theta, \bar{u})$  is strictly increasing in  $\bar{u}$  at every type  $\theta$  that is not excluded (i.e.,  $V(\theta, \bar{u}) > 0$ ).

We further restrict attention to equilibria exhibiting the following “smoothness” properties. We say that an ordered equilibrium is *smooth* if at every pair  $(\theta, \bar{u})$  such that  $V(\theta, \bar{u}) > 0$  the following conditions hold: (i) the support schedule  $V(\theta, \bar{u})$  is twice continuously differentiable in  $\theta$ , and continuously differentiable in  $\bar{u}$ , (ii) the distribution of type- $\theta$  indirect utilities,  $F(\cdot; \theta)$ , is absolutely continuous with density  $f(\cdot; \theta)$ , and (iii) the mapping  $F(u; \theta)$  is continuously differentiable in  $\theta$  for each  $u$ .

#### 4.1 Equilibrium Characterization

The next proposition describes a smooth ordered equilibrium.

**Proposition 6 [Equilibrium Characterization - Continuum of Types]** *There exists a smooth ordered equilibrium. In this equilibrium, the support of indirect utilities offered by firms is described*

by the support schedule

$$V(\theta, \bar{u}) = \max \left\{ \frac{1}{4 \cdot \bar{u}} \cdot \theta^2 + \left( 1 - \frac{1}{2 \cdot \bar{u}} \right) \cdot \theta + \bar{u} + \frac{1}{4 \cdot \bar{u}} - 1, 0 \right\}, \quad (19)$$

with domain on  $[0, 1] \times [\frac{1}{4}, \frac{1}{2}]$ .

The equilibrium distribution over menus for the highest type solves

$$\frac{\Lambda(F(\bar{u}; 1)|v)}{\Lambda(0|v)} = \frac{\frac{1}{48}}{\int_0^1 \left( \theta \cdot V_1(\theta, \bar{u}) - V(\theta, \bar{u}) - \frac{1}{2} \cdot [V_1(\theta, \bar{u})]^2 \right) d\theta}, \quad (20)$$

where the supremum point of  $\Upsilon(1)$ , denoted  $\bar{u}$ , is determined by  $F(\bar{u}; 1) = 1$ .

**Equilibrium Construction.** The equilibrium construction under a continuum of types closely mirrors that of the binary type model from the previous section. First, the ordered nature of equilibrium, together with the ranking property of kernels, implies that

$$\Lambda(F(\bar{u}; 1)|v) = \Lambda(F(V(\theta, \bar{u}); \theta)|v) \quad (21)$$

at every pair  $(\theta, \bar{u})$  such that  $V(\theta, \bar{u}) > 0$ . Differentiating the expression above with respect to  $\bar{u}$  leads to the continuous analogue of equation (15):

$$V_2(\theta, \bar{u}) = \frac{\Lambda_1(F(\bar{u}; 1)|v) \cdot f(\bar{u}; 1)}{\Lambda(F(\bar{u}; 1)|v)} \cdot \left[ \frac{\Lambda_1(F(V(\theta, \bar{u}); \theta)|v) \cdot f(V(\theta, \bar{u}); \theta)}{\Lambda(F(V(\theta, \bar{u}); \theta)|v)} \right]^{-1}, \quad (22)$$

which states that the partial derivative of the support schedule with respect to  $\bar{u}$  at  $(\theta, \bar{u})$  equals the ratio between the semi-elasticities of sales with respect to indirect utilities between the highest type and type  $\theta$ .

Second, for a smooth equilibrium, the optimality of equilibrium menus implies that the following Euler equation must hold at any  $(\theta, \bar{u})$  where  $V(\theta, \bar{u}) > 0$ :

$$\underbrace{\Lambda_1(F(V(\theta, \bar{u}); \theta)|v) \cdot f(V(\theta, \bar{u}); \theta) \cdot \left( \theta \cdot V_1(\theta, \bar{u}) - V(\theta, \bar{u}) - \frac{1}{2} (V_1(\theta, \bar{u}))^2 \right)}_{\text{sales gains}} \\ = \underbrace{\Lambda(F(V(\theta, \bar{u}); \theta)|v)}_{\text{profit losses}} + \underbrace{\Lambda(F(V(\theta, \bar{u}); \theta)|v) \cdot \frac{d}{d\theta} \{ \theta - V_1(\theta, \bar{u}) \}}_{\text{efficiency effect}}. \quad (23)$$

Analogously to the first-order conditions (7) and (8), the Euler equation above identifies the three effects that determine the firms' optimal choice of menus. The first term captures the effect of generosity on sales, while the second effect accounts for the effect of generosity on profits per sale. More interestingly, the third term captures the effect of increasing the indirect utility of type  $\theta$  on the quality distortions of its "adjacent" types (as implied by incentive constraints). Similarly to the binary type model, the optimality condition alone is not enough to sign the efficiency effect:

While increasing the indirect utility of type  $\theta$  allows the firm to decrease quality distortions to its “lower neighbors”, it also tightens the IC constraints of its “upper neighbors” (which leads to higher distortions).

Combining the ranking condition (22) with the optimality condition (23) leads to the following partial differential equation, that the support schedule has to satisfy in any smooth ordered equilibrium:

$$V_2(\theta, \bar{u}) = \frac{2 - V_{11}(1, \bar{u})}{2 - V_{11}(\theta, \bar{u})} \cdot \frac{\theta \cdot V_1(\theta, \bar{u}) - V(\theta, \bar{u}) - \frac{1}{2} \cdot (V_1(\theta, \bar{u}))^2}{\frac{1}{2} - \bar{u}}. \quad (24)$$

The partial differential equation above is the analogue of the ordinary differential equation (12) from Proposition 1. Guided by the binary type model, we posit that the support schedule has to satisfy the following boundary conditions:

$$V\left(\theta, \frac{1}{4}\right) = \max\left\{\theta^2 - \theta + \frac{1}{4}, 0\right\}, \quad V\left(\theta, \frac{1}{2}\right) = \frac{\theta^2}{2}, \quad V_1(1, \bar{u}) = 1, \quad \text{and} \quad V(1, \bar{u}) = \bar{u}. \quad (25)$$

The first boundary condition in (25) states that the Mussa-Rosen menu is the “lowest” menu in the support  $\mathbb{S}$  (in the sense that it provides the lowest indirect utility to every type). Intuitively, the firm that offers the least generous menu is preferred to any other firm known to the consumer. Therefore, this firm must offer the monopoly menu. The second boundary condition guarantees that equilibrium menus approach the Bertrand (or efficient) menu as firms relinquish the total surplus to consumers. The third boundary condition requires that the highest type is offered the efficient quality in all menus in  $\mathbb{S}$ . The last boundary condition requires that the solution to (24) be consistent with the definition of the support schedule  $V(\theta, \bar{u})$ .

The support schedule  $V(\theta, \bar{u})$  in equation (19) solves the partial differential equation (24) subject to the boundary conditions in (25). In the proof of Proposition 6, contained in the Appendix, we formalize the equilibrium construction sketched above. Most importantly, we establish that the Euler equation (23) is a necessary and sufficient condition that any menu that maximizes (18) has to satisfy, and rule out deviations to menus that offer out-of-equilibrium contracts to any type.

Finally, similarly to Proposition 1, the indifference condition (23) guarantees that all menus offered in equilibrium lead to the same total profits as the Mussa-Rosen menu. As before, the matching technology, captured by the kernel  $\Lambda(y|v)$ , determines the upper limit in the support of indirect utilities to the highest consumer type,  $\bar{u}$ , as well as its cumulative distribution function,  $F(\bar{u}; 1)$ . By virtue of the ordered nature of the equilibrium, the distribution over indirect utilities of any type  $\theta \in [0, 1)$  can be recovered from equation (21).

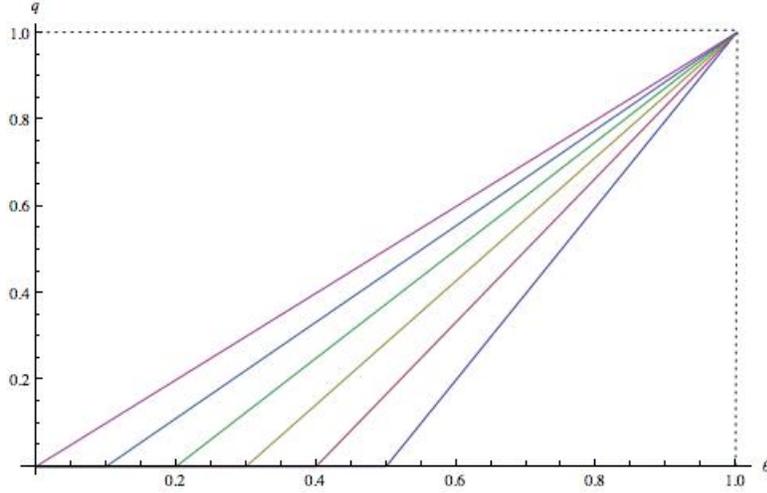


Figure 3: The quality schedules associated with  $\bar{u} = \{0.25, 0.3, 0.35, 0.4, 0.45, 0.5\}$ , from the bottom to the top curves, respectively. The bottom curve is the Mussa-Rosen quality schedule, while the top curve is the Bertrand quality schedule. The schedules offered in equilibrium are those with  $\bar{u} \leq \bar{\bar{u}}$ .

## 4.2 Equilibrium Properties

Let us start our discussion of equilibrium properties with the relationship between generosity and distortions. To do so, let us consider the collection of quality schedules

$$q(\theta, \bar{u}) \equiv V_1(\theta, \bar{u}) = \max \left\{ \frac{1}{2 \cdot \bar{u}} \cdot \theta + \left( 1 - \frac{1}{2 \cdot \bar{u}} \right), 0 \right\}, \quad (26)$$

indexed by the indirect utility offered to the highest type. First, we see that whenever  $q(\theta, \bar{u}) > 0$  we have that

$$\frac{\partial q(\theta, \bar{u})}{\partial \bar{u}} = \frac{1 - \theta}{2\bar{u}^2} > 0.$$

Therefore, as in the binary type model, distortions decrease for all types as firms offer more generous menus. Figure 3 above depicts some quality schedules offered in equilibrium.

We now consider the effects of competition on market coverage. From (26), the range of types served by a menu with highest-type utility  $\bar{u}$  is the interval  $[\alpha(\bar{u}), 1]$ , where

$$\alpha(\bar{u}) = 1 - 2\bar{u}. \quad (27)$$

It follows from (27) that  $\alpha(\bar{u})$  is decreasing in  $\bar{u}$ . Therefore, firms segment themselves according to the range of consumer types served by their menus, which we call *inclusiveness*. As such, more generous firms, as captured by  $\bar{u}$ , are also more inclusive, in the sense that they serve a larger range of types. At one extreme lies the Mussa-Rosen menu, which is the least generous and the least inclusive equilibrium menu. At the other extreme lies the menu associated with highest-type indirect utility  $\bar{\bar{u}}$ , which is the most generous and most inclusive menu offered in equilibrium. This is illustrated in Figure 3 above.

Finally, and analogously to the binary-type model, firms that offer more generous menus make more sales to consumers with low willingness to pay. Formally, for each  $\theta' \in (0, 1)$ , the share of profits obtained from consumers with type  $\theta \in [0, \theta']$  is increasing in  $u$ . We collect these findings in Proposition 7.

**Proposition 7 [Equilibrium Properties]** *The following properties hold in the ordered equilibrium of Proposition 6.*

1. **Efficiency:** *Menus for which consumers earn higher payoffs are more efficient, i.e.,  $q(\theta, \bar{u})$  is increasing in  $\bar{u}$ , strictly whenever  $\theta < 1$  and  $q(\theta, \bar{u}) > 0$ .*
2. **Inclusiveness:** *Firms that offer more generous contracts serve a larger set of consumers, i.e., the range  $[\alpha(\bar{u}), 1]$  of types served expands as  $\bar{u}$  increases.*
3. **Profits:** *Firms that offer more generous menus derive a greater share of profits from consumers with low willingness to pay, i.e., relative to total profits, the ratio of profits derived from consumers with types in any interval of the form  $[0, \theta']$ , where  $\theta' < 1$ , is increasing in  $\bar{u}$ .*

### 4.3 Comparative Statics

The continuous-type model of this section enables us to study how the range of types served in equilibrium,  $[\alpha(\bar{u}), 1]$ , which we call *market coverage*, is affected by competition.

The next proposition shows that the equilibrium market coverage monotonically approaches its competitive level as the mass of firms increases. At one extreme, as  $v \rightarrow 0$ , the equilibrium market coverage approaches its monopolistic level, where only consumers with willingness to pay in the interval  $[\frac{1}{2}, 1]$  are served. At the other extreme, as  $v \rightarrow \infty$ , the equilibrium market coverage approaches its efficient level, i.e., full market coverage.

**Proposition 8 [Competition and Market Coverage: Comparative Statics]** *Consider the smooth ordered equilibrium of Proposition 6, and assume that condition VM holds. Denote by  $F(\cdot; \theta)$  and  $\hat{F}(\cdot; \theta)$  the equilibrium distributions over indirect utilities of type  $\theta$  when the mass of firms is  $v$  and  $\hat{v}$ , respectively.*

1. *If  $v > \hat{v}$ , then  $F(\cdot; \theta)$  first-order stochastically dominates  $\hat{F}(\cdot; \theta)$  for all  $\theta \in [0, 1]$ . In particular, the equilibrium market coverage,  $[\alpha(\bar{u}), 1]$ , expands as  $v$  increases.*
2. *If  $\lim_{v \rightarrow 0} R(1|v) = 1$ , then, as the mass of firms converges to zero,  $v \rightarrow 0$ , the equilibrium distribution over menus converges to a degenerate distribution centered at the monopolistic (Mussa-Rosen) menu. In particular, the equilibrium market coverage monotonically converges to its monopoly level.*

3. If  $\lim_{v \rightarrow \infty} R(y|v) = \infty$  for all  $y \in (0, 1]$ , then, as the mass of firms grows large,  $v \rightarrow \infty$ , the distribution over menus converges to a degenerate distribution centered at the competitive (Bertrand) menu. In particular, the equilibrium market coverage monotonically approaches  $[0, 1]$ , i.e., full market coverage.

Similarly to Section 3, the results above can be recast in terms of the levels of frictions of the matching technology, as discussed in Remarks 4 and 5.

## 5 Extensions

The analysis so far assumed that consumer information is exogenous. This simplification was useful to isolate the effects of competition on the firms' pricing and quality provision. In this section, we endogenize consumer information, and show that the main insights of our analysis naturally extend to this more general environment.

The endogeneity of consumer information stems from two different sources. First, consumers may invest in information acquisition, so as to learn the offers available in the market. Second, firms may invest in advertising, so as to better inform consumers about their offers. The next subsections show that we can easily incorporate information acquisition by consumers, or endogenous advertising by firms, in our model of competitive nonlinear pricing.

### 5.1 Information Acquisition by Consumers

For simplicity, we consider the binary-type model studied in Section 3, and assume that consumer information is generated by the Burdett-Judd random matching model of Example 1.

We model information acquisition by assuming that, after learning their willingness to pay for quality, consumers can make investments that affect the size of their information sets (i.e., the sample of firms they are aware of).<sup>25</sup> If a consumer invests nothing, his sample size distribution is  $\Omega^0 = \{\omega_j^0 : j = 0, 1, 2, \dots\}$ , where  $\omega_j^0 > 0$  is the probability of observing a sample of  $j$  firms. The distribution  $\Omega^0$  captures, in probabilistic terms, the information obtained spontaneously by consumers about the offers available in the market. We assume that  $\omega_0^0 > \frac{1}{2}$ ; i.e., if a consumer makes no investment, then he is more likely to observe no offer than one or more offers.

Investing in information acquisition shifts the consumer sample-size distribution according to first-order stochastic dominance. Specifically, we assume that investing  $z \in [0, 1]$  generates a sample size distribution  $\Omega^z = \{\omega_j^z : j = 0, 1, 2, \dots\}$ , where

$$\omega_j^z \equiv (1 + z) \cdot \omega_j^0 \quad \text{for } j = 1, 2, \dots,$$

---

<sup>25</sup>De los Santos, Hortaçsu and Wildenbeest (2012) provide empirical evidence that sample-size search, as considered here, better explains consumer behavior than other modes of search (for example, sequential).

and  $\omega_0^z \equiv 1 - \sum_{j \geq 1} \omega_j^z$ . Accordingly, investments in information acquisition scale up the probability of sample sizes weakly larger than one.

Let consumers with low and high types invest  $z_l$  and  $z_h$  in information acquisition. Denoting by  $\Lambda^z(\cdot)$  the sales kernel associated with the sample-size distribution  $\Omega^z$ , we can write the firms' profit-maximization problem (in response to the cross-section cdf's over indirect utilities  $\{F_l, F_h\}$ ) as that of choosing a menu  $(u_l, u_h)$  to maximize

$$\sum_{k=l,h} p_k \cdot \Lambda^{z_k}(F_k(u_k)) \cdot (S_k(u_l, u_h) - u_k) = \sum_{k=l,h} p_k \cdot (1 + z_k) \cdot \Lambda^0(F_k(u_k)) \cdot (S_k(u_l, u_h) - u_k),$$

where the equality above follows from the formula for the Burdett-Judd kernel in Example 1. By letting  $\hat{p}_k(z_k) \equiv p_k \cdot (1 + z_k)$ , we can see that the firms' problem when consumers invest  $(z_l, z_h)$  in information acquisition is the same as if the mass of high and low types were  $\hat{p}_l(z_l)$  and  $\hat{p}_h(z_h)$ , and no information acquisition was possible.

It then follows from Proposition 1 that, for any profile of investments  $(z_l, z_h)$ , there exists a unique ordered equilibrium, where the masses of consumer of each type  $(p_l, p_h)$  are replaced by the *adjusted masses*  $(\hat{p}_l(z_l), \hat{p}_h(z_h))$ . The effect of information acquisition by consumers is therefore equivalent to a change in the masses of each consumer type. In particular, the equilibrium behavior of firms satisfies the properties described in Proposition 2.

Of course, when information acquisition is endogenous, the firms' choice of menus and the consumers' investment decisions are jointly determined. To describe the consumers' investment problem, let us assume that investing  $z$  in information acquisition costs  $\psi(z)$  to consumers, where the cost function  $\psi(\cdot)$  is continuously differentiable, strictly increasing and strictly convex, with  $\psi(0) = \psi'(0) = 0$  and  $\lim_{z \rightarrow 1} \psi'(z) = \infty$  (which guarantees an interior solution). It is convenient to denote by  $U_k^{1:j}$  the random variable defined as the highest realization out of  $j$  iid draws from each distribution  $F_k$ , for  $k \in \{l, h\}$ . A consumer with type  $k \in \{l, h\}$  then chooses his investment in information acquisition to maximize his payoff

$$\sum_{j=1}^{\infty} \omega_j^z \cdot \mathbb{E} \left[ U_k^{1:j} \right] - \psi(z) = (1 + z) \cdot \sum_{j=1}^{\infty} \omega_j^0 \cdot \mathbb{E} \left[ U_k^{1:j} \right] - \psi(z). \quad (28)$$

An *equilibrium with information acquisition* is a triple  $(\tilde{F}, z_l, z_h)$  such that the firms' choice of menus and the consumers' investment decision constitute mutual best responses.<sup>26</sup>

Since the marginal benefit of information acquisition is *uniformly* larger for high than for low types, we must have  $z_l < z_h$ . As a consequence, the true and adjusted masses of consumer types satisfy

$$\frac{p_h}{p_l} < \frac{\hat{p}_h(z_h)}{\hat{p}_l(z_l)}.$$

---

<sup>26</sup>The existence of an equilibrium with information acquisition follows from the Kakutani fixed-point theorem. The proof of this claim, which key step establishes that the firms' and consumers' payoffs satisfy the appropriate continuity properties, is available upon request.

This implies that, relative to baseline model of Section 2, consumer information acquisition makes high types “over-represented” in equilibrium, i.e., firms behave as if high-type consumers were more frequent relative to low types than as implied by their actual masses.

## 5.2 Advertising and Entry by Firms

The model analyzed in the previous sections took the mass of firms  $v$  as exogenous. An important and realistic possibility is that the number of firms is endogenously determined. To make things concrete, we will consider in this subsection the model of Butters (1977), described in Example 3. Suppose now that launching an advertising campaign costs  $K$  dollars, and that there is free entry among firms. Firms that do not advertise are not in the informational set of any consumers, and therefore make zero profits. Accordingly, the decision to advertise (and pay the cost  $K$ ) coincides with the firms’ decision of entering the market.

Let  $\pi^m$  be the profit of a monopolist; i.e., the profit generated when the monopoly offer is accepted with probability one. Whenever the entry cost is smaller than the monopolist’s profit,  $K \in (0, \pi^m)$ , the market operates and our model uniquely determines the level of competition in terms of the mass of firms  $v$ . Indeed, the mass of firms  $v(K)$  entering is then given by the zero-profit condition

$$K = \Lambda(0|v(K)) \cdot \pi^m. \quad (29)$$

In words, firms enter until the cost of entry is equal to the benefit. Given that all equilibrium offers yield the same expected profit, this benefit is equal to the expected profit generated by the monopoly offer (which, recall, is itself offered in equilibrium). The existence of a unique  $v(K)$  is guaranteed by the assumption that  $K \in (0, \pi^m)$  and because, in the Butters model,  $\Lambda(0|v)$  is strictly decreasing in  $v$  with  $\Lambda(0|0) = 1$ . The mass of entrants  $v(K)$  is then decreasing in  $K$ , while the distribution over low-type quality is also decreasing in  $K$  in the sense of first-order stochastic dominance (see Proposition 3).

The free-entry version of our model described above is appropriate to address a classic question, raised originally in Butters (1977), regarding the efficient level of advertising/entry. Absent private information on consumer preferences, Butters shows that the level of entry in the decentralized equilibrium is *efficient*. A simple explanation (not given in Butters, but in later work such as Tirole (1988, Section 7.3.2.1) and Stegeman (1991)) is as follows. Absent private information on preferences, all equilibrium offers prescribe efficient qualities, and the social surplus generated by *any* sale is the same. Because in equilibrium firms make the same profits from any offer in the equilibrium support, we can analyze a firm’s entry decision assuming that it makes the monopoly offer (which belongs to the equilibrium support). Recall that the monopoly offer only translates in sales if the offer is received by an otherwise unserved consumer, and that in this offer the firm appropriates the full social surplus. It then follows that the firms’ private gains from advertising coincide with the social

gains from advertising, in which case advertising/entry is efficient.

This observation no longer holds when consumer preferences are private, as studied in this paper. Indeed, a planner able to choose the level of entry, *but not* the menus offered by firms, will choose higher entry than the equilibrium level. To understand why, note that, as in the case of complete information, we can study a firm's entry decision *as if* this firm were to offer the monopolist menu (as implied by the fact that all equilibrium menus lead to the same expected profits). To highlight the novelty of our finding, assume that the monopolist menu is such that only high types are served. Therefore, as under complete information, the monopolist menu gives to the firm the full social surplus generated by this menu. It then follows that, for a firm that offers the monopolist menu, the private gains from advertising coincide with the social gains from advertising, in which case advertising/entry is efficient. However, recall from Proposition 2 that all other equilibrium menus generate strictly more social surplus than the monopolist menu. As a consequence, there is a wedge between the private and the social gains from entry/advertising for all firms *other* than the ones offering the monopolist menu. As result, welfare will increase if the planner subsidizes entry/advertising (even when firms are free to choose their menus once in market).<sup>27</sup>

## 6 Conclusion

This paper studied imperfect competition in price-quality schedules in a market with informational frictions. On the one hand, consumers have private information about their willingness to pay for quality. On the other, consumers are imperfectly informed about the offers in the market, which is the source of firms' market power. As the market becomes more competitive, equilibrium approaches the Bertrand outcome. At the other extreme, all offers are close to the monopolistic menu when competition is sufficiently weak. As we vary the mass of firms in the market, or the degree of informational frictions, we continuously span the entire spectrum of competitive intensity. Equilibrium menus are dispersed and can be ranked in terms of the generosity of their contracts to all consumer types. Firms that offer more generous menus provide more efficient qualities, sell to a more diverse set of consumers, and obtain a higher fraction of their profits from the low-quality goods offered in their menus. These implications are robust to a broad range of possible matching technologies that determine the consumers' information sets.

---

<sup>27</sup>If the monopolist menu serves both low and high types, there is another effect strengthening the inefficiency result described above. In this case, even the firm offering the monopolist menu makes an inefficient advertising/entry decision. The reason is that this firm no longer appropriates the full surplus from the menu it offers, as high-type consumers are endowed to informational rents. This effect is related to the work of Stegeman (1991), who considers a homogeneous good model where consumers have private information as to their reservation values. Stegeman too finds that the competitive level of entry is inefficiently low, which can be attributed to the fact that the least generous (i.e., monopoly) offer in his model leaves positive surplus to the consumer.

Our equilibrium characterization proceeded under the assumption either of two consumer types (with no restrictions on the probability distribution) or a continuum of types with a uniform distribution. The chief difficulty for analyzing other specifications stems from the requirement that implementable quality schedules be monotonic in consumer types. It is difficult to identify conditions on primitives that guarantee monotonicity in more general settings (where a closed-form description of equilibrium is not available). To ascertain the robustness of our findings, we have numerically calculated equilibria for the discrete-type case with various number of types, and different distributions over consumer types. These numerical results indicate that the findings of this paper are robust to other distributions of consumer preferences.

There are several interesting directions for further research. First, for the sake of tractability, we have assumed that firms' capacities are unconstrained, so they are able to fill all orders. Capacity constraints raise the possibility that those firms offering the most generous menus sell out, a possibility that consumers should in turn anticipate. An examination of these "congestion effects" seems difficult but worthwhile.

Second, we assumed that consumers observe the entire menu of qualities offered by each firm. In practice, consumers may fail to consider all of the options that a firm offers; i.e., information imperfections may pertain also to a consumer's ability to observe the entire menu. This possibility has been explicitly recognized in empirical work (e.g., Sovinsky Goeree (2008)). In theoretical work, Villas-Boas (2004) studies a monopolist whose consumers may (randomly) observe only the option designed for the high or low type; extending the analysis to a competitive setting raises additional challenges.

Third, one may hope to introduce additional dimensions to consumer preferences, such as brand preferences, to bring the setting closer to the random utility models popular in empirical work (for instance, suppose that, in addition to the payoffs specified in the model, consumers receive an additional, continuously distributed, "shock"  $\tilde{\varepsilon}$  to their payoffs from purchasing from each seller). We expect that our main qualitative insights are robust to this possibility. More broadly, our model offers a useful theoretical benchmark against which models of imperfect competition with brand preferences can be compared. We hope that the findings of this paper are useful for empirical work trying to assess to what extent the market power enjoyed by firms stems from heterogeneity of information or brand preferences.

Finally, while we focused on sales of goods with variable quality, our results extend to other contexts where information heterogeneity makes sense. A natural application, for instance, is to labor markets where workers have private information about their productivities, and are heterogeneously informed about the job offers available in the market. Contracts might pay wages based on the worker's output. In such settings, our results indicate dispersion over offers, with firms endogenously segmenting themselves relative to (i) the indirect utility left to all worker types, (ii) the efficiency

of effort provision induced by their contracts, (iii) the share of surplus obtained from workers with different productivity levels, and (iv) the composition of the labor force (in what pertains productivity levels) employed by each firm.

## References

- [1] Armstrong, M. and Vickers, J., 2001. "Competitive Price Discrimination." *Rand Journal of Economics*, 32, 579-605.
- [2] Berry, S., J. Levinsohn and A. Pakes, 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63, 841-890.
- [3] Billingsley, P., 1995. *Probability and Measure*. New York: Wiley Series in Probability and Mathematical Statistics.
- [4] Borenstein, S., 1985. "Price discrimination in free-entry markets." *RAND Journal of Economics*, 16, 380-397.
- [5] Borenstein, S. and N. Rose, 1994. "Competition and price dispersion in the U.S. airline industry." *Journal of Political Economy*, 102, 653-683.
- [6] Butters, G., 1977. "Equilibrium Distributions of Sales and Advertising Prices." *The Review of Economic Studies*, 44, 465-491.
- [7] Burdett, K., and Mortensen, D., 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39, 257-273.
- [8] Burdett, K., and Judd, D., 1983. "Equilibrium Price Dispersion." *Econometrica*, 51, 955-970.
- [9] Calzolari, G., and V. Denicolo, 2013. "Competition with Exclusive Contracts and Market-Share Discounts." *American Economic Review* (forthcoming).
- [10] Champsaur, P. and J.-C. Rochet, 1989. "Multi-product duopolists". *Econometrica*, 57, 533-557.
- [11] Clarke, F., 2013. *Functional Analysis, Calculus of Variation and Optimal Control*. London: Springer.
- [12] De los Santos, B., A. Hortaçsu and M. Wildenbeest, 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review*, 102, pp. 2955-2980.
- [13] Draganska, M., and D. Klapper (2011), "Choice Set Heterogeneity and The Role of Advertising: An Analysis with Micro and Macro Data." *Journal of Marketing Research*, 2011, 48, 653-669.
- [14] Eeckhout, J., and P. Kircher, 2010. "Sorting versus Screening: Search Frictions and Competing Mechanisms." *Journal of Economic Theory*, 145, 1354-1385.
- [15] Ellison, G., 2005. "A Model of Add-on Pricing." *Quarterly Journal of Economics*, 120, 585-367.
- [16] Faig, M. and B. Jerez, 2005. "A Theory of Commerce." *Journal of Economic Theory*, 122, 60-99.
- [17] Goldman, M., H. Leland, and D. Sibley. 1984. "Optimal Nonuniform Pricing." *Review of Economic Studies*, 51, 305-320.
- [18] Grossman, G. M. and C. Shapiro, 1984. "Informative Advertising with Differentiated Products." *Review of Economic Studies*, 51, 63-81.
- [19] Guerrieri, V., Shimer, R. and Wright, R., 2010. "Equilibrium Price Dispersion." *Econometrica*, 78, 1823-1862.
- [20] Inderst, R., 2001. "Screening in a Matching Market." *Review of Economic Studies*, 2001, 68, 849-868.
- [21] Inderst, R., 2004. "Contractual Distortions in a (Labor) Market with Frictions." *Journal of Economic Theory*, 116, 155-176.

- [22] Ivaldi, M. and D. Martimort, 1994. "Competition under nonlinear pricing." *Annales d'Economie et de Statistique*, 34, 71–114.
- [23] Katz, M.L., 1984. "Price Discrimination and Monopolist Competition." *Econometrica*, 52, 1453-1471.
- [24] McAfee, P., 1993. "Mechanism Design by Competing Sellers." *Econometrica*, 78, 1823-1862.
- [25] Maskin, E., and Riley, J., 1984. "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15, 171-196.
- [26] Moen, A. and Rosén, A., 2011. "Incentives in Competitive Search Equilibrium." *Review of Economic Studies*, 98, 733-761.
- [27] Mussa, M. and Rosen S., 1978. "Monopoly and Product Quality." *Journal of Economic Theory*, 18, 301-317.
- [28] Pai, M., 2013. "Competing Auctioneers." *Working Paper*.
- [29] Peters, M., 1997. "A Competitive Distribution of Auctions." *Review of Economic Studies*, 64, 97-123.
- [30] Peters, M., 2012. "Non-Contractible Heterogeneity in Directed Search." *Econometrica*, 78, 1173–1200.
- [31] Peters, M., and S. Severinov, 1996. "Competition among Sellers Who Offer Auctions Instead of Prices." *Journal of Economic Theory*, 75, 141-179.
- [32] Petrongolo, B. and C. Pissarides, 2001. "Looking at the Black Box: A Survey of the Matching Function." *Journal of Economic Literature*, 39, 390-431.
- [33] Rochet, J.C., and Stole, L., 1997. "Competitive Nonlinear Pricing." *Working paper*.
- [34] Rochet, J.C., and Stole, L., 2002. "Nonlinear Pricing with Random Participation." *The Review of Economic Studies*, 69, 2002, 277-311.
- [35] Salop, S., and J. Stiglitz, 1977. "Bargains and Rip-Offs: A Model of Monopolistically Competitive Price Discrimination." *Review of Economic Studies*, 44, 493-510.
- [36] Sovinsky Goeree, M., 2008. "Limited Information and Advertising in the U.S. Personal Computer Industry." *Econometrica*, 76, 1017-1074.
- [37] Spulber, D., 1989. "Product variety and competitive discounts." *Journal of Economic Theory*, 48, 510–525.
- [38] Stegeman, M., 1991, "Advertising in Competitive Markets." *American Economic Review*, 81, pp. 210-223.
- [39] Stole, L., 1991. "Mechanism design and common agency". *Working Paper*.
- [40] Stole, L., 1995. "Nonlinear pricing and oligopoly." *Journal of Economics & Management Strategy*, 4, 529– 562.
- [41] Stole, L., 2007. "Price Discrimination and Competition." In: *Handbook of Industrial Organization*, vol. 3 (eds. M. Armstrong and R. Porter) North-Holland.
- [42] Tirole, J., 1988. *The Theory of Industrial Organization*. The MIT Press.
- [43] Varian, H., 1980. "A Model of Sales." *American Economic Review*, 70, 651-659.
- [44] Verboven, F., 1999. "International price discrimination in the European car market." *RAND Journal of Economics*, 27, 240–268.
- [45] Villas-Boas, M., 2004. "Communication Strategies and Product Line Design." *Marketing Science*, 23, 304-316.
- [46] Wilson, R., 1993. *Nonlinear Pricing*. Oxford Univ. Press, Oxford, UK.
- [47] Yang, H., and Ye, L., 2008. "Nonlinear pricing, market coverage and competition." *Theoretical Economics*, 3, 123-153.

# Online Appendix - Not for Publication

This Appendix collects proofs of all results.

## Appendix A: Proofs of results for binary-types model

Throughout, we economize notation by suppressing the dependence of  $\Phi(u_k|F_k, v, p_k)$  on  $(F_k, v, p_k)$  for each  $k \in \{l, h\}$ ; we simply write  $\Phi_k(u_k) = \Phi(u_k|F_k, v, p_k)$ .

**Proof of Lemma 1.** If the low type is offered the quality  $q_l$ , then payoffs must satisfy  $IC_h$ , i.e.,

$$u_h \geq u_l + \Delta\theta q_l. \quad (30)$$

On the other hand,  $IC_l$  requires that

$$u_l \geq u_l - \Delta\theta q_h. \quad (31)$$

The firm would like to make its offer as efficient as possible subject to the payoffs it delivers to the consumer.

If  $u_h - u_l < \bar{\Delta}_l \equiv \Delta\theta q_l^*$ , then offering the efficient quality  $q_l^*$  for the low type is inconsistent with (30), and the firm does best to choose the highest possible value. That is, the firm chooses quality  $q_l(u_l, u_h)$  which satisfies (30) with equality, or

$$q_l(u_l, u_h) \equiv \frac{u_h - u_l}{\Delta\theta}.$$

If  $u_h - u_l \geq \bar{\Delta}_l$ , then the constraint (30) does not bind, and the firm chooses low-type quality efficiently:  $q_l(u_l, u_h) \equiv q_l^*$ . Similarly, let  $\bar{\Delta}_h \equiv \Delta\theta q_h^*$ . If  $u_h - u_l > \bar{\Delta}_h$ , then asking the quality  $q_h^*$  for the high type violates (31), and so the best the firm can do is to choose  $q_h(u_l, u_h)$  defined by

$$q_h(u_l, u_h) \equiv \frac{u_h - u_l}{\Delta\theta}.$$

If  $u_h - u_l < \bar{\Delta}_h$ , the firm offers the high-type an efficient quality:  $q_h(u_l, u_h) \equiv q_h^*$ . Q.E.D.

**Proof of Lemma 2.** To see this claim, note that  $\pi(u_l, u_h^2) - \pi(u_l, u_h^1)$  equals

$$\begin{aligned} & \Phi_l(u_l) (S_l(u_l, u_h^2) - S_l(u_l, u_h^1)) \\ & + (\Phi_h(u_h^2) - \Phi_h(u_h^1)) (S_h(u_l, u_h^2) - u_h^2) \\ & + \Phi_h(u_h^1) \begin{pmatrix} S_h(u_l, u_h^2) - u_h^2 \\ -(S_h(u_l, u_h^1) - u_h^1) \end{pmatrix}. \end{aligned} \quad (32)$$

The cross-partial  $\frac{\partial^2}{\partial u_l \partial u_h} S_l(u_l, u_h)$  is positive if  $u_h - u_l < \Delta\theta q_l^*$  and zero otherwise. Thus the first line of (32) is strictly increasing over  $u_l$  such that  $u_h^1 - u_l \leq \Delta\theta q_l^*$  and constant otherwise. The function

$S_h(\cdot, u_h^2)$  is strictly increasing if  $q_h(u_l, u_h^2) > q_h^*$  and constant otherwise. Thus the second line in (32) is increasing in  $u_l$ . The cross-partial  $\frac{\partial^2}{\partial u_l \partial u_h} S_h(u_l, u_h)$  is positive if  $u_h - u_l > \Delta\theta q_h^*$  and is zero otherwise. Thus the third term is strictly increasing over  $u_l$  such that  $u_h^2 - u_l \geq \Delta\theta q_h^*$  and constant otherwise. These arguments imply the result. Q.E.D.

**Proof of Lemma 3.** We divide the proof in five steps.

**Step 1** *No mass points in the distribution of high-type offers.*

We begin by showing that  $F_h$  has no mass points. Assume towards a contradiction there is an atom of firms offering  $\tilde{u}_h$ .

We first show that, if a firm makes an equilibrium offer  $(\tilde{u}_l, \tilde{u}_h)$ , for some value  $\tilde{u}_l$ , then  $S_h(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_h > 0$ . Suppose not. Then it must be that  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l \leq 0$  (in case  $S_h(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_h \leq 0$  and  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l > 0$ , offering only the option designed for the low type improves the seller's expected profit because high types accept such an offer with positive probability). Hence,  $\pi(\tilde{u}_l, \tilde{u}_h) \leq 0$ . This contradicts seller optimization. Indeed, the seller could offer a menu which yields the Mussa and Rosen utilities  $(u_l^m, u_h^m)$  and obtain a payoff at least as large as  $(S_h^* - u_h^m) \Phi_h(0) > 0$ .

Next, notice that  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l \geq 0$ . If not, the seller can profit by offering the menu  $(q_l, x_l) = (0, 0)$  and  $(q_h, x_h) = (q_h^*, \theta_h q_h^* - \tilde{u}_h)$ . Irrespective of whether the low type finds it incentive compatible to choose the option  $(0, 0)$ , the seller is guaranteed an expected profit at least as high as under the original menu.

These two observations imply that  $\pi(\tilde{u}_l + \varepsilon, \tilde{u}_h + \varepsilon) > \pi(\tilde{u}_l, \tilde{u}_h)$  for  $\varepsilon > 0$  sufficiently small, contradicting the optimality of  $(\tilde{u}_l, \tilde{u}_h)$ . To see this, note that  $\pi(\tilde{u}_l + \varepsilon, \tilde{u}_h + \varepsilon)$  must be bounded below by

$$\begin{aligned} & \pi(\tilde{u}_l, \tilde{u}_h) - \varepsilon [\Phi_h(\tilde{u}_h + \varepsilon) + \Phi_l(\tilde{u}_l + \varepsilon)] \\ & + (S_h(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_h - \varepsilon) [\Phi_h(\tilde{u}_h + \varepsilon) - \Phi_h(\tilde{u}_h)], \end{aligned}$$

Since  $\Phi_h(\tilde{u}_h + \varepsilon) - \Phi_h(\tilde{u}_h)$  is bounded above zero as  $\varepsilon \searrow 0$ , and since  $S_h(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_h > 0$ , the expression above is greater than  $\pi(\tilde{u}_l, \tilde{u}_h)$  whenever  $\varepsilon$  is sufficiently small.

**Step 2** *No mass points in the distribution of low-type offers.*

First, we show that there are no mass points in  $F_l$  at any  $u_l > 0$ . Suppose towards a contradiction that  $F_l$  has a mass point at some  $\tilde{u}_l > 0$ . Take a firm that offers  $(\tilde{u}_l, \tilde{u}_h)$ . Since, as reasoned above,  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l \geq 0$ , we can consider two cases.

**Case 1:**  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l > 0$ .

As noted in Step 1, the expected profit conditional on selling to a high type must also be positive. Notice that in this case  $\pi(\tilde{u}_l + \varepsilon, \tilde{u}_h + \varepsilon)$  is bounded below by

$$\begin{aligned} & \pi(\tilde{u}_l, \tilde{u}_h) - \varepsilon [\Phi_h(\tilde{u}_h + \varepsilon) + \Phi_l(\tilde{u}_l + \varepsilon)] \\ & + (S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l - \varepsilon) (\Phi_l(\tilde{u}_l + \varepsilon) - \Phi_l(\tilde{u}_l)). \end{aligned}$$

Since  $\Phi_l$  has a mass point at  $\tilde{u}_l$ , and since  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l > 0$ , the expression above is strictly greater than  $\pi(\tilde{u}_l, \tilde{u}_h)$  for  $\varepsilon > 0$  sufficiently small.

**Case 2:**  $S_l(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_l = 0$ . Let  $\{(q_l, x_l), (q_h, x_h)\} = \{(q_l(\tilde{u}_l, \tilde{u}_h), x_l(\tilde{u}_l, \tilde{u}_h)), (q_h(\tilde{u}_l, \tilde{u}_h), x_h(\tilde{u}_l, \tilde{u}_h))\}$  be the menu offered by the firm. Consider a deviation to the menu  $\{(q_l, x_l + \varepsilon), (q_h, x_h)\}$  for some  $\varepsilon \in (0, \tilde{u}_l)$ . This menu generates the same expected profits from high types and is accepted with positive probability by low types. Moreover, since  $S_h(\tilde{u}_l, \tilde{u}_h) - \tilde{u}_h > 0$  (see Step 1), the seller makes positive profits whether a low-type buyer chooses the option  $(q_l, x_l + \varepsilon)$  or  $(q_h, x_h)$ . That is, expected profits from low types are strictly positive under the deviating offer.

Finally, we show that there can be no mass point in  $F_l$  at zero. Assume towards a contradiction that  $F_l(0) > 0$ . From Step 1 (i.e., since there are no mass points in the distribution of high-type offers), menus  $(u_l, u_h) \in \{0\} \times [\varepsilon, \infty)$  are then offered with positive probability. It is easy to see that there is  $\chi > 0$  such that  $S_l(0, u_h) > \chi$  for all  $u_h \in [\varepsilon, \infty)$ . Therefore, for small  $\eta > 0$  the difference  $\pi(\eta, u_h) - \pi(0, u_h)$  is

$$\begin{aligned} & (\Phi_l(\eta) - \Phi_l(0)) [S_l(\eta, u_h) - \eta] \\ & - \Phi_l(0) (S_l(0, u_h) - S_l(\eta, u_h) - \eta). \end{aligned} \tag{33}$$

We can take  $\eta^*$  such that  $\eta \in (0, \eta^*)$  implies that the first line of (33) is at least  $(\Phi_l(0_+) - \Phi_l(0)) (\frac{\chi}{2}) > 0$ . Moreover, the second line of (33) converges to 0 as  $\eta \searrow 0$ , which shows a profitable deviation.

**Step 3** *The supports  $\Upsilon_k$  are intervals.*

Suppose for a contradiction that one or both of the supports are disconnected sets. Assume that  $\Upsilon_l$  is disconnected. Then there are  $u'_l$  and  $u''_l$  in  $\Upsilon_l$  with  $u'_l < u''_l$  such that  $(u'_l, u''_l) \cap \Upsilon_l = \emptyset$ . Consider values  $u'_h$  and  $u''_h$  such that  $(u'_l, u'_h)$  and  $(u''_l, u''_h)$  are optimal. From Steps 1 and 2 and Lemma 2 we may assume that  $\Phi_l(u'_l) = \Phi_l(u''_l)$ ,  $\Phi_h(u'_h) = \Phi_h(u''_h)$  and  $u'_h \leq u''_h$ .

If  $u'_h < u''_h$  then there is  $\varepsilon > 0$  for which  $\pi(u''_l - \varepsilon, u''_h - \varepsilon) > \pi(u''_l, u''_h)$ . Thus assume that  $u'_h = u''_h$ . For any  $\varepsilon \in (0, u''_l - u'_l)$ , optimality requires  $\pi(u''_l - \varepsilon, u''_h) \leq \pi(u''_l, u''_h)$ . This implies that  $q_h(u''_l, u''_h) > q_h^*$ , i.e.  $IC_l$  binds. Thus  $\frac{\partial^2 S_h(u_l, u_h)}{\partial u_l^2} < 0$  at  $(u''_l, u''_h)$ , which implies (using  $\Phi_l(u'_l) = \Phi_l(u''_l)$ ) that  $\pi(\lambda u'_l + (1 - \lambda) u''_l, u''_h) > \lambda \pi(u'_l, u''_h) + (1 - \lambda) \pi(u''_l, u''_h)$  for  $\lambda \in (0, 1)$ . Hence,  $(u''_l, u''_h)$  is not optimal. The proof that  $\Upsilon_h$  is connected is analogous and omitted.

**Step 4** *The minimum of the supports  $\Upsilon_l$  and  $\Upsilon_h$  are, respectively,  $u_l^m = 0$  and  $u_h^m$ .*

Let  $\underline{u}_l$  and  $\underline{u}_h$  be the minimum of the supports of  $\Upsilon_l$  and  $\Upsilon_h$  respectively. It follows from Steps 1 and 2 and from Lemma 2 that  $(\underline{u}_l, \underline{u}_h)$  is an optimal menu. IR requires  $\underline{u}_l \geq 0$ , and we next show that  $\underline{u}_l = 0$ . To see this, suppose that  $\underline{u}_l > 0$  and note that  $\underline{u}_h \geq \underline{u}_l$ . Since  $\Phi_l(0) = \Phi_l(\underline{u}_l)$  and  $\Phi_h(\underline{u}_h - \underline{u}_l) = \Phi_h(\underline{u}_h)$ , we have  $\pi(0, \underline{u}_h - \underline{u}_l) > \pi(\underline{u}_l, \underline{u}_h)$ , a contradiction. Hence indeed  $\underline{u}_l = 0$  and so  $\underline{u}_h \geq 0$  maximizes

$$\Phi_l(0) S_l(0, u_h) + \Phi_h(0) (S_h(0, u_h) - u_h). \quad (34)$$

Since  $u_h^m$  is the only maximizer of (34), the claim follows. We have thus established that, for each  $k \in \{l, h\}$ , the support  $\Upsilon_k$  is equal to  $[u_k^m, \bar{u}_k]$ , where  $\bar{u}_k > u_k^m$ .

**Step 5**  $F_l$  and  $F_h$  are absolutely continuous.

We will show that  $F_h$  is Lipschitz continuous (the proof that  $F_l$  is absolutely continuous is analogous and omitted). Notice that from 2. in Assumption 1 it suffices to show that  $\Phi_h$  is Lipschitz continuous. For that, it is enough to show that there are positive values  $K$  and  $\delta$  such that, for all  $u_h \in \Upsilon_h$  and all  $\varepsilon \in (0, \delta)$ ,  $\Phi_h(u_h + \varepsilon) - \Phi_h(u_h) < K\varepsilon$ .

First, we claim that we may find a constant  $\underline{S}_h > 0$  such that we have  $S_h(u'_l, u'_h) - u'_h \geq \underline{S}_h$  for every optimal menu  $(u'_l, u'_h)$ . The claim follows by the same logic as in Step 1. If the claim does not hold, we may find a sequence of optimal menus  $(u_l^n, u_h^n)$  such that  $S_h(u_l^n, u_h^n) - u_h^n \leq \frac{1}{n}$ . Taking a subsequence if necessary, assume that  $(u_l^n, u_h^n) \rightarrow (u_l^*, u_h^*)$ . By the continuity of  $\Phi_k$  (Steps 1 and 2) and the continuity of  $S_k$  (for  $k \in \{l, h\}$ ) we conclude that  $(u_l^*, u_h^*)$  is optimal and that  $S_h(u_l^*, u_h^*) - u_h^* = 0$ . However, we showed in Step 1 that such a menu cannot be optimal.

Next, let  $\delta > 0$  and define  $\xi_h := \sup_{(\bar{u}_l, \bar{u}_h) \in \{[0, \bar{u}_l] \times [u_h^m, \bar{u}_h + \delta] : \bar{u}_h \geq \bar{u}_l\}} \left| \frac{\partial S_h(\bar{u}_l, \bar{u}_h)}{\partial u_h} \right|$ . Take any equilibrium menu  $(u_l, u_h) \in \Upsilon_l \times \Upsilon_h$ . Notice that, for  $\varepsilon \in (0, \delta)$ ,  $\pi(u_l, u_h + \varepsilon)$  is

$$\begin{aligned} & \Phi_l(u_l) [S_l(u_l, u_h + \varepsilon) - u_l] + \Phi_h(u_h + \varepsilon) [S_h(u_l, u_h + \varepsilon) - u_h - \varepsilon] \\ \geq & \left[ \begin{array}{l} \Phi_l(u_l) [S_l(u_l, u_h) - u_l] + \Phi_h(u_h) [S_h(u_l, u_h) - u_h] \\ -\Phi_h(\bar{u}_h) (\xi_h + 1) \varepsilon + [\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)] (\underline{S}_h - (\xi_h + 1) \varepsilon) \end{array} \right]. \end{aligned}$$

Since  $\pi(u_l, u_h + \varepsilon) \leq \pi(u_l, u_h)$  we have:

$$\frac{\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)}{\varepsilon} \leq \frac{\Phi_h(\bar{u}_h) (\xi_h + 1)}{\underline{S}_h - (\xi_h + 1) \varepsilon} < \frac{\Phi_h(\bar{u}_h) (\xi_h + 1)}{\underline{S}_h - (\xi_h + 1) \delta}.$$

Since Part 3 of Assumption 1 implies  $\Phi_h(\bar{u}_h) < +\infty$ , it is then easy to see that our claim holds provided  $K$  is sufficiently large and  $\delta$  sufficiently small. Q.E.D.

**Proof of Proposition 1.** As explained in the proof sketch contained in the text, we divide the proof in three steps.

**Step 1** *Constructing the support function.*

**Necessity of (7) and (8).** We first show that  $\Phi_h(\cdot)$  and  $\Phi_l(\cdot)$  are continuously differentiable. By Assumption 1.2, this implies that each  $F_k(u_k)$  is continuously differentiable as well. Hence, the firm's profits  $\pi(u_l, u_h)$  as defined by (6) are continuously differentiable, with first-order conditions given by (7) and (8).

We focus on the claim that  $\Phi_h(\cdot)$  is continuously differentiable, as the case of  $\Phi_l(\cdot)$  is analogous. Let  $u_h \in \Upsilon_h$  and suppose that  $u_l = \hat{u}_l(u_h)$ , so that  $(u_l, u_h)$  is an optimal menu. Note that for any  $\varepsilon \in \mathbb{R}$ , we have

$$\begin{aligned} & \Phi_l(u_l) [S_l(u_l, u_h + \varepsilon) - u_l] + \Phi_h(u_h + \varepsilon) [S_h(u_l, u_h + \varepsilon) - u_h - \varepsilon] \\ = & \Phi_l(u_l) [S_l(u_l, u_h) - u_l] + \Phi_h(u_h) [S_h(u_l, u_h) - u_h] \\ & + \Phi_l(u_l) [S_l(u_l, u_h + \varepsilon) - S_l(u_l, u_h)] + \Phi_h(u_h) [S_h(u_l, u_h + \varepsilon) - \varepsilon - S_h(u_l, u_h)] \\ & + [\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)] [S_h(u_l, u_h + \varepsilon) - u_h - \varepsilon]. \end{aligned}$$

Since  $\pi(u_l, u_h) \geq \pi(u_l, u_h + \varepsilon)$ , we have

$$\begin{aligned} & [\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)] [S_h(u_l, u_h + \varepsilon) - u_h - \varepsilon] \\ \leq & \Phi_l(u_l) [S_l(u_l, u_h) - S_l(u_l, u_h + \varepsilon)] + \Phi_h(u_h) [S_h(u_l, u_h) - S_h(u_l, u_h + \varepsilon) + \varepsilon]. \end{aligned}$$

Next, for any  $\varepsilon \in \mathbb{R}$  such that  $u_h + \varepsilon \in \Upsilon_h$ , let  $u_{l,\varepsilon} = \hat{u}_l(u_h + \varepsilon)$ . Thus, we have

$$\begin{aligned} & \Phi_l(u_{l,\varepsilon}) [S_l(u_{l,\varepsilon}, u_h + \varepsilon) - u_{l,\varepsilon}] + \Phi_h(u_h + \varepsilon) [S_h(u_{l,\varepsilon}, u_h + \varepsilon) - u_h - \varepsilon] \\ = & \Phi_l(u_{l,\varepsilon}) [S_l(u_{l,\varepsilon}, u_h) - u_{l,\varepsilon}] + \Phi_h(u_h) [S_h(u_{l,\varepsilon}, u_h) - u_h] \\ & + \Phi_l(u_{l,\varepsilon}) [S_l(u_{l,\varepsilon}, u_h + \varepsilon) - S_l(u_{l,\varepsilon}, u_h)] + \Phi_h(u_h) [S_h(u_{l,\varepsilon}, u_h + \varepsilon) - \varepsilon - S_h(u_{l,\varepsilon}, u_h)] \\ & + [\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)] [S_h(u_{l,\varepsilon}, u_h + \varepsilon) - u_h - \varepsilon]. \end{aligned}$$

Since  $\pi(u_{l,\varepsilon}, u_h + \varepsilon) \geq \pi(u_{l,\varepsilon}, u_h)$ , we have

$$\begin{aligned} & [\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)] [S_h(u_{l,\varepsilon}, u_h + \varepsilon) - u_h - \varepsilon] \\ \geq & \Phi_l(u_{l,\varepsilon}) [S_l(u_{l,\varepsilon}, u_h) - S_l(u_{l,\varepsilon}, u_h + \varepsilon)] + \Phi_h(u_h) [S_h(u_{l,\varepsilon}, u_h) - S_h(u_{l,\varepsilon}, u_h + \varepsilon) + \varepsilon]. \end{aligned}$$

For the right derivative we now consider  $\varepsilon > 0$  (the case of the left derivative is analogous). For any  $\varepsilon$  sufficiently small, we have  $S_h(u_l, u_h + \varepsilon) - u_h - \varepsilon > 0$  (to see this, consider the argument in Step 1 of the proof of Lemma 3). For all such  $\varepsilon$ , we have

$$\begin{aligned} & \frac{\left( \begin{array}{l} \Phi_l(u_{l,\varepsilon}) [S_l(u_{l,\varepsilon}, u_h) - S_l(u_{l,\varepsilon}, u_h + \varepsilon)] \\ + \Phi_h(u_h) [S_h(u_{l,\varepsilon}, u_h) - S_h(u_{l,\varepsilon}, u_h + \varepsilon) + \varepsilon] \end{array} \right)}{\varepsilon [S_h(u_{l,\varepsilon}, u_h + \varepsilon) - u_h - \varepsilon]} \leq \frac{\Phi_h(u_h + \varepsilon) - \Phi_h(u_h)}{\varepsilon} \\ & \leq \frac{\left( \begin{array}{l} \Phi_l(u_l) [S_l(u_l, u_h) - S_l(u_l, u_h + \varepsilon)] \\ + \Phi_h(u_h) [S_h(u_l, u_h) - S_h(u_l, u_h + \varepsilon) + \varepsilon] \end{array} \right)}{\varepsilon [S_h(u_l, u_h + \varepsilon) - u_h - \varepsilon]}. \end{aligned}$$

Next, note that  $\hat{u}_l(\cdot)$  must be continuous by Lemma 3, since each  $F_k$  is continuous and  $F_l(\hat{u}_l(u_h)) = F_h(u_h)$  for all  $u_h$ . Hence  $u_{l,\varepsilon} \searrow u_l$  as  $\varepsilon \searrow 0$ , implying that the right derivative of  $\Phi_h(u_h)$  is equal to

$$\frac{-\Phi_l(u_l) \frac{\partial S_l(u_l, u_h)}{\partial u_h} + \Phi_h(u_h) \left(1 - \frac{\partial S_h(u_l, u_h)}{\partial u_h}\right)}{S_h(u_l, u_h) - u_h}$$

The left derivative can similarly be shown to take the same value, i.e.,  $\Phi_h(u_h)$  is differentiable at  $u_h$ . Using our assumption that  $IC_l$  is slack, we can thus conclude that

$$\Phi'_h(u_h) = \frac{-\Phi_l(u_l) \frac{\partial S_l(u_l, u_h)}{\partial u_h} + \Phi_h(u_h)}{S_h^* - u_h} \quad (35)$$

$$\Phi'_l(u_l) = \frac{\Phi_l(u_l) \left(1 - \frac{\partial S_l(u_l, u_h)}{\partial u_l}\right)}{S_l(u_l, u_h) - u_l}. \quad (36)$$

Recall that  $\bar{u}_h < S_h^*$  by Lemma 3. Moreover, we must have  $S_l(u_l, u_h) - u_l > 0$  whenever  $u_h > 0$  (this follows from the argument in Step 2, Case 2 of Lemma 3). Hence, both derivatives are finite over  $u_h \in (u_h^m, \bar{u}_h)$ .

**Verifying (15).** Next we want to verify that  $\hat{u}_l(u_h)$  is differentiable with derivative given by (15). Indeed, note from (36) that  $\Phi'_l(u_l)$  is strictly positive at  $u_l = \hat{u}_l(u_h)$  for any  $u_h \in (u_h^m, \bar{u}_h)$ . Thus, by the implicit function theorem,  $\hat{u}'_l(u_h) = \frac{\Phi'_h(u_h)}{\Phi'_l(\hat{u}_l(u_h))}$ , which is precisely (15).

**Existence and properties of solution to ODE.** As described in the main text, (7), (8) and (15) imply that the support function  $\hat{u}_l$  must satisfy

$$\hat{u}'_l(u_h) = h(\hat{u}_l(u_h), u_h), \quad (37)$$

where

$$h(u_l, u_h) = \frac{S_l(u_l, u_h) - u_l}{S_h^* - u_h} \cdot \frac{1 - \frac{p_l}{p_h} \frac{\partial S_l}{\partial u_h}(u_l, u_h)}{1 - \frac{\partial S_l}{\partial u_l}(u_l, u_h)}, \quad (38)$$

and where we impose the boundary condition  $\hat{u}_l(u_h^m) = 0$ . We now show that there exists a unique solution  $\hat{u}_l(\cdot)$  on  $[u_h^m, S_h^*]$ .

For any  $\varepsilon \in (0, S_h^*)$ , the function  $h(\cdot, \cdot)$  is Lipschitz continuous on

$$\Gamma(\varepsilon) \equiv \{(u_l, u_h) \in [0, S_l^*] \times [u_h^m, S_h^* - \varepsilon] : u_l < u_h\}.$$

Hence, by the Picard-Lindelöf theorem, for any  $\varepsilon \in (0, S_h^*)$ , and for any  $(u_l, u_h)$  in the interior of  $\Gamma(\varepsilon)$ , there is a unique local solution to  $\hat{u}'_l(u_h) = h(\hat{u}_l(u_h), u_h)$ . Local uniqueness will extend to global uniqueness, guaranteeing that the equilibrium we construct is the only ordered equilibrium.

Now consider  $\hat{u}'_l(u_h) = h(\hat{u}_l(u_h), u_h)$  with initial condition  $\hat{u}_l(u_h^m) = 0$  and note the existence of  $\eta > 0$  such that a unique solution exists on  $[u_h^m, u_h^m + \eta]$  where  $(\hat{u}_l(u_h), u_h)$  remains in  $\Gamma(0)$ . We now show that  $h(\hat{u}_l(u_h), u_h)$  remains bounded and that  $(\hat{u}_l(u_h), u_h)$  remains in  $\Gamma(0)$  also as  $u_h$

increases to  $S_h^*$ , implying the existence of a global solution to  $\hat{u}_l'(u_h) = h(\hat{u}_l(u_h), u_h)$  on  $[u_h^m, S_h^*)$ . We further show that  $h(\hat{u}_l(u_h), u_h)$  remains strictly positive on  $[u_h^m, S_h^*)$  (as explained in the main text, this ensures that the equilibrium we construct is ordered). The problem should be considered for two regions of  $u_h$ : we show that there exists a value  $u_h^c \in (u_h^m, S_h^*)$  such that  $u_h^c - \hat{u}_l(u_h^c) = q_l^* \Delta\theta$  and such that  $u_h - \hat{u}_l(u_h) < q_l^* \Delta\theta$  for  $u_h \in [u_h^m, u_h^c]$ . We then show that  $u_h - \hat{u}_l(u_h) > q_l^* \Delta\theta$  for  $u_h > u_h^c$ .

First, we show that, for  $u_h > u_h^m$ , provided  $u_h - \hat{u}_l(u_h)$  remains below  $q_l^* \Delta\theta$ , then  $h(\hat{u}_l(u_h), u_h)$  remains in  $(0, 1)$ . First, note that  $S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)$  remains strictly positive: this follows because  $\frac{d}{du_h} [S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)] > 0$  whenever  $S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)$  is sufficiently close to zero. Second,  $h(\hat{u}_l(u_h), u_h)$  remains below 1 because

$$\begin{aligned}
& S_h^* - u_h - (S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)) \\
&= \theta_h q_h^* - \varphi(q_h^*) - (\theta_l q_l(\hat{u}_l(u_h), u_h) - \varphi(q_l(\hat{u}_l(u_h), u_h))) - q_l(\hat{u}_l(u_h), u_h) \Delta\theta \\
&= \theta_h q_h^* - \varphi(q_h^*) - (\theta_h q_l(\hat{u}_l(u_h), u_h) - \varphi(q_l(\hat{u}_l(u_h), u_h))) \\
&> 0
\end{aligned} \tag{39}$$

and  $\frac{p_l}{p_h} \frac{\partial S_l}{\partial u_h}(\hat{u}_l(u_h), u_h) > \frac{\partial S_l}{\partial u_l}(\hat{u}_l(u_h), u_h)$  whenever  $u_h - \hat{u}_l(u_h) < q_l^* \Delta\theta$ . Finally, to check that  $h(\hat{u}_l(u_h), u_h)$  remains strictly positive, we note that  $\frac{p_l}{p_h} \frac{\partial S_l}{\partial u_h}(\hat{u}_l(u_h), u_h) < 1$  provided  $u_h - \hat{u}_l(u_h) > u_h^m$ , which is guaranteed in turn by the initial condition and that  $h(\hat{u}_l(u_h), u_h)$  remains less than 1.

We now verify the existence of  $u_h^c \in (u_h^m, S_h^*)$  for which  $u_h^c - \hat{u}_l(u_h^c) = q_l^* \Delta\theta$ . Suppose that there is no such value  $u_h^c$ . Then the equalities in (39) must continue to hold for all  $u_h \in (u_h^m, S_h^*)$ . Since these expressions are bounded above zero, we must have  $S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h) < 0$  as  $u_h$  approaches  $S_h^*$ , contradicting the observation in the previous claim.

Next, consider extending the solution to  $u_h \in (u_h^c, S_h^*)$ . It is easily checked that  $\hat{u}_l(u_h) = S_l^* - \alpha(S_h^* - u_h)$  with  $\alpha = \frac{S_l^* - \hat{u}_l(u_h^c)}{S_h^* - u_h^c} \in (0, 1)$  satisfies  $\hat{u}_l'(u_h) = h(\hat{u}_l(u_h), u_h)$  and remains in  $\Gamma(0)$  (that  $\hat{u}_l(u_h)$  remains below  $S_l^*$  follows because  $S_l^* - \hat{u}_l(u_h) = \alpha(S_h^* - u_h) > 0$ ).

The extension of the solution to  $u_h \in (u_h^c, S_h^*)$  completes the construction of the support function. We can then check that the incentive constraint  $IC_l$  (i.e., (31)) is globally satisfied, as we do next.

**The Incentive Constraint  $IC_l$  is globally satisfied.** The above showed that there is  $u_h^c \in (u_h^m, S_h^*)$  for which  $u_h - \hat{u}_l(u_h) < \Delta\theta q_l^*$  for all  $u_h < u_h^c$ . For  $u_h \in [u_h^c, S_h^*)$  we have

$$u_h - \hat{u}_l(u_h) = (u_h - S_l^*) + \alpha(S_h^* - u_h). \tag{40}$$

Notice that the derivative of the RHS of (40) w.r.t.  $u_h$  is  $1 - \alpha > 0$ . Hence (40) achieves its maximum at  $u_h = S_h^*$  and its maximum is given by

$$S_h^* - S_l^* = \Delta\theta q_l^* + \int_{q_l^*}^{q_h^*} (\theta_h - \varphi'(q)) dq < \Delta\theta q_h^*.$$

Thus we conclude that  $u_h - \hat{u}_l(u_h) \in (\Delta\theta q_l^*, \Delta\theta q_h^*)$  for all  $u_h \in (u_h^c, S_h^*)$ . Therefore, the incentive constraint (31) does not bind along the curve  $(\hat{u}_l(u_h), u_h)$ .

**Step 2** *Solving for the distribution  $\tilde{F}$ .*

As noted above (see Step 4 in the proof of Lemma 3), the least generous equilibrium menu must be  $(u_l^m, u_h^m)$ . Moreover, in equilibrium, all offers must yield the same expected profit

$$\pi^* := \sum_{k=l,h} p_k \cdot \Lambda(0|v) \cdot (S_k(u_l^m, u_h^m) - u_k^m).$$

Next, observe that there is a value  $\bar{u}_h > u_h^m$  which solves

$$\Lambda(1|v) \sum_{k=l,h} p_k \cdot (S_k(\hat{u}_l(\bar{u}_h), u_h) - \hat{u}_k(\bar{u}_h)) = \pi^*$$

The existence of such  $\bar{u}_h$  is guaranteed by the intermediate value theorem, since  $\frac{\Lambda(1|v)}{\Lambda(0|v)} \in (1, \infty)$  and since  $\lim_{u_h \uparrow S_h^*} [S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)] = \lim_{u_h \uparrow S_h^*} [S_h^* - u_h] = 0$ .

Condition (13) is then simply the requirement that

$$\Lambda(F_h(u_h)|v) \sum_{k=l,h} p_k \cdot (S_k(\hat{u}_l(u_h), u_h) - \hat{u}_k(u_h)) = \pi^*$$

for  $u_h \in [u_h^m, \bar{u}_h]$  where  $F_h(\bar{u}_h) = 1$ . Note then that  $\frac{d}{d\bar{u}_h} \left[ \sum_{k=l,h} p_k \cdot (S_k(\hat{u}_l(u_h), u_h) - \hat{u}_k(u_h)) \right] < 0$  on  $(u_h^m, \bar{u}_h)$ . This follows because  $\frac{d}{d\bar{u}_h} [u_h - \hat{u}_l(u_h)] > 0$  and  $u_h - \hat{u}_l(u_h) > u_h^m$ , and because  $\hat{u}_l'(u_h) > 0$ . Hence, Assumption 1.1 and 1.2 imply that  $F_h$  is uniquely defined by (13) and is increasing and differentiable on  $(u_h^m, \bar{u}_h)$ . Finally,  $F_l$  is uniquely defined by  $F_l(\hat{u}_l(u_h)) = F_h(u_h)$ .

**Step 3** *Verifying the optimality of equilibrium menus.*

It now remains to check that firms have no incentive to deviate from the putative equilibrium strategies. By construction, all menus  $(u_l, u_h)$  such that  $u_h \in [u_h^m, \bar{u}_h]$  and  $u_l = \hat{u}_l(u_h)$  yield the same profit. Moreover, it is easy to show that we may restrict attention to menus  $(u_l', u_h') \in [u_h^m, \bar{u}_h] \times [u_l^m, \hat{u}_l(\bar{u}_h)]$ . Hence, consider a menu  $(u_l', u_h') \in [u_h^m, \bar{u}_h] \times [u_l^m, \hat{u}_l(\bar{u}_h)]$  such that  $u_l' \neq \hat{u}_l(u_h')$ . We have that

$$\begin{aligned} \pi(\hat{u}_l(u_h'), u_h') - \pi(u_l', u_h') &= \int_{u_l'}^{\hat{u}_l(u_h')} \frac{\partial \pi(\tilde{u}_l, u_h')}{\partial u_l} d\tilde{u}_l \\ &= \int_{u_l'}^{\hat{u}_l(u_h')} \frac{\partial \pi(\tilde{u}_l, u_h')}{\partial u_l} - \frac{\partial \pi(\tilde{u}_l, \hat{u}_l^{-1}(\tilde{u}_l))}{\partial u_l} d\tilde{u}_l \\ &= \int_{u_l'}^{\hat{u}_l(u_h')} \int_{\hat{u}_l^{-1}(\tilde{u}_l)}^{u_h'} \frac{\partial^2 \pi(\tilde{u}_l, \tilde{u}_h)}{\partial u_h \partial u_l} d\tilde{u}_h d\tilde{u}_l \\ &\geq 0. \end{aligned}$$

The second equality follows because  $\frac{\partial \pi(u_l, u_h)}{\partial u_l} = 0$  along the curve  $\{(\hat{u}_l(u_h), u_h) : u_h \in [u_h^m, \bar{u}_h]\}$ . The inequality follows because  $\frac{\partial^2 \pi(\tilde{u}_l, \tilde{u}_h)}{\partial u_h \partial u_l} \geq 0$  for all  $(\tilde{u}_l, \tilde{u}_h)$  by Lemma 2. Thus a deviation to menu  $(u'_l, u'_h)$  is unprofitable. Q.E.D.

**Proof of Proposition 2.** Part (i) follows from the following observations. Consider the differential equation  $\hat{u}'_l(u_h) = h(\hat{u}_l(u_h), u_h)$  with  $h$  given by (38). Step 2 in the proof of Proposition 1 showed that there is  $u_h^c \in (u_h^m, S_h^*)$  such that  $h(\hat{u}_l(u_h), u_h) < 1$  for every  $u_h < u_h^c$ . On the other hand, for  $u_h > u_h^c$  we showed that  $\delta'(u_h) = 1 - \alpha > 0$ .

Now consider Part (ii). For  $u_h < u_h^c$ , we have

$$\begin{aligned} & \frac{d}{du_h} \left[ \frac{p_l \Lambda(F_l(\hat{u}_l(u_h)) | v)(S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h))}{p_h \Lambda(F_h(u_h) | v)(S_h^* - u_h)} \right] \\ = & \left( \frac{p_l}{p_h} \right) \left( \frac{\left( \frac{\partial S_l(\hat{u}_l(u_h), u_h)}{\partial u_l} \hat{u}'_l(u_h) + \frac{\partial S_l(\hat{u}_l(u_h), u_h)}{\partial u_h} \right) (S_h^* - u_h) - \hat{u}'_l(u_h) (S_h^* - u_h) + (S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h))}{(S_h^* - u_h)^2} \right) \\ > & \left( \frac{p_l}{p_h} \right) \left( \frac{-\hat{u}'_l(u_h) (S_h^* - u_h) + (S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h))}{(S_h^* - u_h)^2} \right) \\ > & \left( \frac{p_l}{p_h} \right) \left( \frac{-\left( \frac{S_l(\hat{u}_l(u_h), u_h) - u_l}{S_h^* - u_h} \right) (S_h^* - u_h) + (S_l(\hat{u}_l(u_h), u_h) - u_l)}{(S_h^* - u_h)^2} \right) = 0, \end{aligned}$$

where the first inequality uses Part (i) and the second uses  $\left( \frac{1 - \frac{p_l}{p_h} \frac{\partial S_l(\hat{u}_l(u_h), u_h)}{\partial u_h}}{1 - \frac{\partial S_l(\hat{u}_l(u_h), u_h)}{\partial u_l}} \right) < 1$  and (38). For  $u_h > u_h^c$ ,  $\hat{u}_l(u_h) = S_h^* - \alpha(S_h^* - u_h)$ . Thus, the ratio between the profits from the low and high type is  $\alpha \left( \frac{p_l}{p_h} \right)$  which is locally constant. Q.E.D.

**Proof of Corollary 1.** Because the high-type quality is constant at  $q_h^*$ , it is immediate that  $x_h(\cdot)$  is decreasing in  $u_h$ . The same is true regarding the low-type price  $x_l(\cdot)$  at any  $u_h > u_h^c$  (since the low-type quality is constant at  $q_l^*$ ). So take  $u_h \in [u_h^m, u_h^c]$  and note that

$$x_l(u_h) = \theta_l \frac{u_h - \hat{u}_l(u_h)}{\Delta \theta} - \hat{u}_l(u_h).$$

Consider  $\hat{u}'_l(u_h) = h(\hat{u}_l(u_h), u_h)$  with  $h$  given by (38) and note that  $h(0, u_h^m) = 0$  (which implies that  $\hat{u}'_l(u_h^m) = 0$ ). Therefore, since  $\hat{u}_l(\cdot)$  and  $h(\cdot, \cdot)$  are continuous,

$$x'_l(u_h) = \theta_l \frac{1 - \hat{u}'_l(u_h)}{\Delta \theta} - \hat{u}'_l(u_h) = \theta_l \frac{1 - h(\hat{u}_l(u_h), u_h)}{\Delta \theta} - h(\hat{u}_l(u_h), u_h) > 0$$

for all  $u_h$  sufficiently close to  $u_h^m$ .

We will now show that  $\hat{u}_l(\cdot)$  is convex for  $u_h < u_h^c$ . Note that the convexity of  $\hat{u}_l(\cdot)$  implies that there exists a unique  $u_h^d \in (u_h^m, u_h^c]$  such that  $x'_l(u_h) > 0$  if and only if  $u_h < u_h^d$ . To see why  $\hat{u}_l(\cdot)$  is

convex, let us differentiate (12) to obtain that

$$\hat{u}_l''(u_h) = \frac{\frac{d}{du_h} S_l(\hat{u}_l(u_h), u_h) + \hat{u}_l'(u_h) \cdot \left( \frac{1}{\nabla(u_h)} - 1 \right)}{S_h^* - u_h} + \frac{S_l(\hat{u}_l(u_h), u_h) - \hat{u}_l(u_h)}{S_h^* - u_h} \cdot \nabla'(u_h),$$

where

$$\nabla(u_h) \equiv \frac{1 - \frac{p_l}{p_h} \frac{\partial S_l}{\partial u_h}(\hat{u}_l(u_h), u_h)}{1 - \frac{\partial S_l}{\partial u_l}(\hat{u}_l(u_h), u_h)}.$$

Recall from the proof of Proposition 2 that  $\nabla(u_h) \in (0, 1)$  for  $u_h < u_h^c$  and that  $\frac{d}{du_h} S_l(\hat{u}_l(u_h), u_h) > 0$ . Moreover, straightforward differentiation shows that  $\nabla'(u_h) > 0$ . Coupled together, these facts imply that  $\hat{u}_l''(u_h) > 0$  for  $u_h < u_h^c$ , as claimed. Q.E.D.

**Proof of Proposition 3.** We only prove 1. The proof of 2. is analogous and omitted. Since we are considering ordered equilibrium,<sup>28</sup> it suffices to show that  $F_h(u_h) \leq \hat{F}_h(u_h)$  for all  $u_h$ . Towards a contradiction, take  $\tilde{u}_h$  such that  $F_h(\tilde{u}_h) > \hat{F}_h(\tilde{u}_h)$ . Without loss assume that  $\tilde{u}_h \in \Upsilon_h$  (otherwise, replace  $\tilde{u}_h$  with  $\max \Upsilon_h$ ). Therefore, we have:

$$\begin{aligned} & \Lambda(0 | v) [p_l S_l(0, u_h^m) + p_h (S_h(0, u_h^m) - u_h^m)] \\ &= \Lambda(F_h(\tilde{u}_h) | v) [p_l (S_l(\tilde{u}_h, \hat{u}_l(\tilde{u}_h)) - \hat{u}_l(\tilde{u}_h)) + p_h (S_h(\tilde{u}_h, \hat{u}_l(\tilde{u}_h)) - \tilde{u}_h)] \end{aligned}$$

and

$$\begin{aligned} & \Lambda(0 | \hat{v}) [p_l \cdot S_l(0, u_h^m) + p_h (S_h(0, u_h^m) - u_h^m)] \\ &= \Lambda(\hat{F}_h(\tilde{u}_h) | \hat{v}) [p_l (S_l(\tilde{u}_h, \hat{u}_l(\tilde{u}_h)) - \hat{u}_l(\tilde{u}_h)) + p_h (S_h(\tilde{u}_h, \hat{u}_l(\tilde{u}_h)) - \tilde{u}_h)], \end{aligned}$$

and hence

$$\frac{\Lambda(\hat{F}_h(\tilde{u}_h) | \hat{v})}{\Lambda(0 | \hat{v})} = \frac{\Lambda(F_h(\tilde{u}_h) | v)}{\Lambda(0 | v)}. \quad (41)$$

On the other hand,  $F_h(\tilde{u}_h) > \hat{F}_h(\tilde{u}_h)$  implies  $\frac{\Lambda(F_h(\tilde{u}_h)|v)}{\Lambda(0|v)} > \frac{\Lambda(\hat{F}_h(\tilde{u}_h)|v)}{\Lambda(0|v)}$  and Condition 1 implies  $\frac{\Lambda(\hat{F}_h(\tilde{u}_h)|v)}{\Lambda(0|v)} > \frac{\Lambda(\hat{F}_h(\tilde{u}_h)|\hat{v})}{\Lambda(0|\hat{v})}$  and thus  $\frac{\Lambda(F_h(\tilde{u}_h)|v)}{\Lambda(0|v)} > \frac{\Lambda(\hat{F}_h(\tilde{u}_h)|\hat{v})}{\Lambda(0|\hat{v})}$ , which contradicts (41). Q.E.D.

**Proof of Proposition 4.** We first prove 1. Take  $\bar{u}_h = \max \Upsilon_h$ . We have:

$$\frac{p_l S_l(0, u_h^m) + p_h (S_h(0, u_h^m) - u_h^m)}{p_l (S_l(\bar{u}_h, \hat{u}_l(\bar{u}_h)) - \hat{u}_l(\bar{u}_h)) + p_h (S_h(\bar{u}_h, \hat{u}_l(\bar{u}_h)) - \bar{u}_h)} = \frac{\Lambda(1 | v)}{\Lambda(0 | v)}. \quad (42)$$

Notice that  $\lim_{v \rightarrow 0} R(1|v) = 1$  implies that the RHS of (42) converges to 1 as  $v \rightarrow 0$ . This implies that the LHS of (42) converges to 1 and hence  $(\hat{u}_l(\bar{u}_h), \bar{u}_h) \rightarrow (0, u_h^m)$  as  $v \rightarrow 0$ . The second statement in 1. follows immediately.

Next we prove 2. Take a sequence  $(v_n) \rightarrow \infty$ , and let  $(F_{h,n})$  be the corresponding sequence of distributions over high-type payoffs in the ordered equilibrium. Take  $y \in (0, 1]$  and let  $u_h^n(y) \equiv$

<sup>28</sup>Using Proposition 5, it is easy to see that this result is true for any equilibrium.

$F_{h,n}^{-1}(y)$ . We have:

$$\frac{p_l S_l(0, u_h^m) + p_h (S_h(0, u_h^m) - u_h^m)}{p_l (S_l(u_h^n(y), \hat{u}_l(u_h^n(y))) - \hat{u}_l(u_h^n(y))) + p_h (S_h(u_h^n(y), \hat{u}_h(u_h^n(y))) - u_h^n(y))} = \frac{\Lambda(y | v)}{\Lambda(0 | v)}. \quad (43)$$

Notice that the RHS of (43) diverges to  $\infty$  by assumption. Therefore, the denominator of the LHS of (43) converges to 0, which implies that  $(\hat{u}_l(u_h^n(y)), u_h^n(y))$  converges to the Bertrand menu. The second statement in 2. follows immediately. Q.E.D.

**Proof of Proposition 5.** Proposition 3 implies the following. For the unique ordered distribution described in Proposition 1, there is a value  $v^c$  such that  $v \leq v^c$  implies  $\bar{u}_h \leq u_h^c$ , while  $v > v^c$  implies  $\bar{u}_h > u_h^c$ . What remains to show is that, for  $v \leq v^c$ , the only equilibrium is the ordered equilibrium (i.e., Part 1 of Proposition 3) as well as the uniqueness claims in Part 2 (i.e., regarding menus with payoffs  $u_h \leq u_h^c$  and regarding the marginal distributions  $F_k$ ).

Let  $\tilde{F}$  be any distribution over menus which describes a (not necessarily ordered) equilibrium. Let the marginal distributions over indirect utilities be given by  $F_k$  with supports  $\Upsilon_k$  as given in Lemma 3. We begin with the following lemma.

**Lemma 4** Consider two equilibrium menus  $(u_l, u_h), (u'_l, u'_h) \in \Upsilon_l \times \Upsilon_h$ . If  $u'_h > u_h$ , then either  $u'_l \geq u_l$  or both  $IC_h$  and  $IC_l$  are slack for both menus (i.e.,  $u_h - u_l, u'_h - u'_l \in [q_l^* \Delta\theta, q_h^* \Delta\theta]$ ).

**Proof.** Suppose  $u'_h > u_h$  and  $u'_l < u_l$ , while either  $u_h - u_l \notin [q_l^* \Delta\theta, q_h^* \Delta\theta]$  or  $u'_h - u'_l \notin [q_l^* \Delta\theta, q_h^* \Delta\theta]$ . By Lemma 2, we have

$$\pi(u_l, u'_h) + \pi(u'_l, u_h) > \pi(u'_l, u'_h) + \pi(u_l, u_h),$$

contradicting the optimality of  $(u_l, u_h)$  or  $(u'_l, u'_h)$ . Q.E.D.

An immediate implication of this lemma is that if  $(u_l, u_h)$  is a menu for which  $IC_l$  or  $IC_h$  binds (i.e.,  $u_h - u_l \notin [q_l^* \Delta\theta, q_h^* \Delta\theta]$ ), then there exists no other equilibrium menu  $(u'_l, u'_h)$  for which  $u'_l < u_l$  and  $u'_h > u_h$  or  $u'_l > u_l$  and  $u'_h < u_h$ . Since  $F_l$  and  $F_h$  are absolutely continuous by Lemma 3, we can conclude hence that  $F_l(u_l) = F_h(u_h)$ .

Next, note that there exists  $\varepsilon > 0$  such that  $IC_h$  binds for all  $u_h \leq u_h^m + \varepsilon$ . Thus, for every menu  $(u_l, u_h)$  with  $u_h \leq u_h^m + \varepsilon$ , we have  $F_l(u_l) = F_h(u_h)$ . Define a strictly increasing and continuous function  $\kappa$  by  $\kappa(u_h) := F_l^{-1}(F_h(u_h))$  (here we use Lemma 3 which guarantees both the continuity of  $F_l$  and  $F_h$  and that both are strictly increasing). Using Lemma 4, it is easy to see that there can be no menu  $(u_l, u_h)$  with  $u_l < \kappa(u_h^m + \varepsilon)$  but  $u_l \neq \kappa(u_h)$ . Thus, we have established that, for any equilibrium menu  $(u_l, u_h)$ , with  $u_h < u_h^m + \varepsilon$  or  $u_l < \kappa(u_h^m + \varepsilon)$ ,  $u_l = \kappa(u_h)$ . The arguments in Step 1 of the proof of Proposition 1 then imply that  $\kappa(\cdot) = \hat{u}_l(\cdot)$  on  $[u_h^m, u_h^m + \varepsilon)$ .

We can extend the above argument to show that all menus  $(u_l, u_h)$  with  $u_h < u_h^c$  or  $u_l < \hat{u}_l(u_h^c)$  must also be given by  $(\hat{u}_l(u_h), u_h)$  for some  $u_h < u_h^c$ . To see this, let

$$\check{u}_h := \sup \{u_h : \forall \text{ eqm menus } (u'_l, u'_h), u'_h < u_h \text{ or } u'_l < \hat{u}_l(u_h) \text{ implies } u'_l = \hat{u}_l(u'_h)\}. \quad (44)$$

As argued above,  $\check{u}_h > u_h^m$ . Suppose with a view to contradiction that  $\check{u}_h < u_h^c$ . Since we must have  $u_l \geq \hat{u}_l(\check{u}_h)$  for any equilibrium menu with  $u_h \geq \check{u}_h$ , there must exist  $\eta > 0$  sufficiently small that  $IC_h$  binds for all  $u_h \leq \check{u}_h + \eta$  (indeed, this must follow because  $IC_h$  binds at  $(\hat{u}_l(\check{u}_h), \check{u}_h)$ ). The same arguments as above then imply that, for any equilibrium menu  $(u_l, u_h)$  with  $u_h < \check{u}_h + \eta$  or  $u_l < \hat{u}_l(\check{u}_h + \eta)$ ,  $u_l = \hat{u}_l(u_h)$ . Hence,  $\check{u}_h$  cannot be the supremum in (44), our contradiction.

Thus, we have established that  $\check{u}_h \geq u_h^c$ . This establishes Part 1 of the proposition: In case  $v \leq v^c$ , we have  $u_h \leq u_h^c$  for all equilibrium menus, as implied by the requirement that all menus generate the same expected profits. This also establishes our claim in Part 2 that non-ordered equilibria differ only in menus for which  $u_h > u_h^c$  (the existence of such non-ordered equilibria is straightforward and left to the reader).

To establish our remaining claims, we consider menus for which  $u_h \geq u_h^c$  and  $u_l \geq \hat{u}_l(u_h^c)$ . We show that

$$\Phi'_h(u_h) = \frac{\Phi_h(u_h)}{S_h^* - u_h} \quad (45)$$

$$\Phi'_l(u_l) = \frac{\Phi_l(u_l)}{S_l^* - u_l} \quad (46)$$

for these values of  $u_h$  and  $u_l$ . This implies that  $\Phi_l$  and  $\Phi_h$  are precisely those functions determined in Proposition 1; hence, the marginal distributions  $F_k$  are identical in any equilibrium. As a result, as shown in the proof of Proposition 1, neither incentive constraint can bind for equilibrium menus with  $u_h \geq u_h^c$  and  $u_l \geq \hat{u}_l(u_h^c)$  (a binding incentive constraint at  $(u_l, u_h)$  would imply  $F_l(u_l) = F_h(u_h)$ , but then  $u_l = \hat{u}_l(u_h)$  and neither incentive constraint binds at  $(\hat{u}_l(u_h), u_h)$  as shown in the proof of Proposition 1).

It is easy to see that the equilibrium menu with high-type payoff  $u_h^c$  is unique and equal to  $(\hat{u}_l(u_h^c), u_h^c)$ .<sup>29</sup> Neither of the incentive constraints  $IC_l$  or  $IC_h$  bind at this menu. This allows us to establish that (45) and (46) hold at  $(\hat{u}_l(u_h^c), u_h^c)$ . We consider (45) as the case of (46) is analogous. We use a similar argument to that in Step 1 of the proof of Proposition 1. For any  $\varepsilon \in \mathbb{R}$  such that  $u_h^c + \varepsilon \in \Upsilon_h$ , let  $(u_{l,\varepsilon}, u_h^c + \varepsilon)$  be a corresponding equilibrium menu. The same arguments as in the

---

<sup>29</sup>By the previous argument, any equilibrium menu  $(u_l, u_h^c)$  must satisfy  $u_l \geq \hat{u}_l(u_h^c)$ . If  $u_l > \hat{u}_l(u_h^c)$ , then  $IC_h$  binds at  $(u_l, u_h^c)$  implying that  $F_l(u_l) = F_h(u_h^c)$ , a contradiction (since  $F_h(u_h^c) = F_l(\hat{u}_l(u_h^c))$ ) by the previous argument and continuity of  $F_l$  and  $F_h$ .

proof of Proposition 1 imply

$$\begin{aligned}
& \frac{\left( \begin{array}{l} \Phi_l(u_{l,\varepsilon}) [S_l(u_{l,\varepsilon}, u_h^c) - S_l(u_{l,\varepsilon}, u_h^c + \varepsilon)] \\ + \Phi_h(u_h^c) [S_h(u_{l,\varepsilon}, u_h^c) - S_h(u_{l,\varepsilon}, u_h^c + \varepsilon) + \varepsilon] \end{array} \right)}{\varepsilon [S_h(u_{l,\varepsilon}, u_h^c + \varepsilon) - u_h^c - \varepsilon]} \\
\leq & \frac{\Phi_h(u_h^c + \varepsilon) - \Phi_h(u_h^c)}{\varepsilon} \\
\leq & \frac{\left( \begin{array}{l} \Phi_l(\hat{u}_l(u_h^c)) [S_l(\hat{u}_l(u_h^c), u_h^c) - S_l(\hat{u}_l(u_h^c), u_h^c + \varepsilon)] \\ + \Phi_h(u_h^c) [S_h(\hat{u}_l(u_h^c), u_h^c) - S_h(\hat{u}_l(u_h^c), u_h^c + \varepsilon) + \varepsilon] \end{array} \right)}{\varepsilon [S_h(\hat{u}_l(u_h^c), u_h^c + \varepsilon) - u_h^c - \varepsilon]}.
\end{aligned}$$

We then use that<sup>30</sup>

$$\lim_{\varepsilon \rightarrow 0} \frac{S_l(u_{l,\varepsilon}, u_h^c) - S_l(u_{l,\varepsilon}, u_h^c + \varepsilon)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{S_l(\hat{u}_l(u_h^c), u_h^c) - S_l(\hat{u}_l(u_h^c), u_h^c + \varepsilon)}{\varepsilon} = 0$$

and

$$S_h(u_{l,\varepsilon}, u_h^c) - S_h(u_{l,\varepsilon}, u_h^c + \varepsilon) = S_h(\hat{u}_l(u_h^c), u_h^c) - S_h(\hat{u}_l(u_h^c), u_h^c + \varepsilon) = 0$$

to conclude that

$$\Phi'_h(u_h^c) = \frac{\Phi_h(u_h^c)}{S_h^* - u_h^c}.$$

Next, observe that there exists  $\eta > 0$  such that incentive constraints are slack for any equilibrium menu with  $u_h \in [u_h^c, u_h^c + \eta]$ . This is obtained from (i) the above observation that if  $(u_l, u_h)$  is a menu for which an incentive constraint  $IC_k$  binds, then  $F_l(u_l) = F_h(u_h)$ , and (ii)  $u_h^c - \hat{u}_l(u_h^c) = q_l^* \Delta \theta$  together with  $\Phi'_h(u_h^c) < \Phi'_l(\hat{u}_l(u_h^c))$  (equivalently,  $F'_h(u_h^c) < F'_l(\hat{u}_l(u_h^c))$ ).

As with the derivatives  $\Phi'_h(u_h^c)$  and  $\Phi'_l(\hat{u}_l(u_h^c))$ , one obtains (45) and (46) on  $[u_h^c, u_h^c + \eta]$ . We then use again that  $F_l(u_l) = F_h(u_h)$  for any menu  $(u_l, u_h)$  for which an incentive constraint binds to obtain that the constraints must be slack for any equilibrium menu with  $u_h \geq u_h^c$ . To see this, let

$$u_h^\# = \sup \{ u_h : IC_l \text{ and } IC_h \text{ are slack for all eqm. menus } (u'_l, u'_h) \text{ with } u'_h \in [u_h^c, u_h] \}.$$

The above property, together with continuity of  $F_l$  and  $F_h$ , implies that, if  $u_h^\# < \bar{u}_h$ , then  $u_h^\# - u_l^\# \notin (\Delta \theta q_l^*, \Delta \theta q_h^*)$  for  $u_l^\#$  satisfying  $F_l(u_l^\#) = F_h(u_h^\#)$ . However,  $F_l$  and  $F_h$  must agree with functions derived in Proposition 1 on, respectively,  $[u_l^m, \hat{u}_l(u_h^\#)]$  and  $[u_h^m, u_h^\#]$ . Hence  $u_l^\# = \hat{u}_l(u_h^\#)$ . This contradicts our finding in the proof of Proposition 1 that  $u_h^\# - \hat{u}_l(u_h^\#) \in (\Delta \theta q_l^*, \Delta \theta q_h^*)$ . Q.E.D.

<sup>30</sup>This follows after noticing that, for any  $\nu > 0$ , there exists  $\iota > 0$  such that, for all  $|\varepsilon| < \iota$ ,  $u_h^c - u_{l,\varepsilon} \in (\Delta \theta q_l^* - \nu, \Delta \theta q_h^*]$ . This follows after noticing that either both incentive constraints are slack at  $(u_{l,\varepsilon}, u_h^c + \varepsilon)$ , or one of  $IC_l$  and  $IC_h$  bind, in which case  $u_{l,\varepsilon} = F_l^{-1}(F_h(u_h^c + \varepsilon))$ , which tends to  $\hat{u}_l(u_h^c)$  as  $\varepsilon \rightarrow 0$ .

## 7 Appendix B: Proofs of results for continuum-types model

### 7.1 Outline

The goal of this section is to prove Proposition 6. We start in Section 7.2 deriving necessary conditions for a smooth ordered equilibrium for a general type distribution and cost function. As we will see, these necessary conditions involve a solution to a partial differential equation which relates the quality schedule of a certain menu to its generosity. In Section 7.3 we specialize to the uniform quadratic case. We start by postulating a closed-form solution to this partial differential equation, and we check it is a solution in Section 7.3.1. Given this solution, we are able to propose an equilibrium allocation in Section 7.3.2. The rest of the appendix is then devoted to verifying that we have an equilibrium. In Section 7.3.3 we verify that all menus in the support of the proposed allocation yield the same profit. Next, in Section 7.3.4 we invoke a calculus of variations existence theorem to show that the firms' problem has a solution. Finally, in Section ?? we show that the only menus satisfying calculus of variations necessary conditions for an optimum are those found in the support of the putative equilibrium menus. This shows that all equilibrium menus maximize firm profits, as required.

### 7.2 General Necessary Conditions

Here we will derive the analogue for a continuum of types of the support function obtained in the case of a binary type space. As described in the main text, the firms' problem is to choose an indirect utility schedule  $u(\theta)$  to maximize

$$\int_{\underline{\theta}}^{\bar{\theta}} \Lambda(F(u(\theta); \theta) | v) \cdot (\theta \cdot \dot{u}(\theta) - \varphi(\dot{u}(\theta)) - u(\theta)) \cdot h(\theta) d\theta, \quad (47)$$

where  $\Lambda$  is the kernel of the matching technology,  $\varphi(q)$  is the cost of producing a good of quality  $q$ , and  $h(\theta)$  is the density of type  $\theta$  (with support  $[\underline{\theta}, \bar{\theta}]$ ). In order to ease the notation below we suppress the dependence of  $\Lambda$  on  $v$ , writing  $\Lambda(F(V(\theta, u); \theta))$  for  $\Lambda(F(V(\theta, u); \theta) | v)$ .

We posit that in equilibrium each firm is indifferent between choosing any schedule in some support  $\mathbb{S}$ . Each of the schedules  $u(\theta)$  in  $\mathbb{S}$  are strictly increasing and weakly convex in  $\theta$  (that is, there is an implementable direct-revelation mechanism delivering an indirect utility  $u(\theta)$ ).

We can conveniently describe the support  $\mathbb{S}$  by indexing each schedule by the indirect utility received by type  $\bar{\theta}$ . Denote by  $V(\theta, u)$  the indirect utility received by type  $\theta$  when type  $\bar{\theta}$  obtains utility  $u$ . Note that  $V(\bar{\theta}, u) = u$ . The set of indirect utility schedules is then

$$\mathbb{S} = \{V(\theta, u) : u \in [\bar{u}^m, \bar{S}^*]\},$$

where  $\bar{u}^m$  is the Mussa-Rosen indirect utility of type  $\bar{\theta}$ , and  $\bar{S}^*$  is the Bertrand indirect utility of type  $\bar{\theta}$ :  $\bar{S}^* \equiv \max_q \{\bar{\theta} \cdot q - \varphi(q)\}$ . Given that we wish to characterize a smooth equilibrium, we look

for a support schedule  $V(\theta, u)$  which is twice continuously differentiable at every point such that  $V(\theta, u) > 0$ .

It is a property of the ordered equilibrium that for any two types  $\theta, \bar{\theta} \in [\underline{\theta}, \bar{\theta}]$  such that  $V(\theta, u), V(\bar{\theta}, u) > 0$

$$\Lambda(F(V(\theta, u); \theta)) = \Lambda\left(F\left(V(\bar{\theta}, u); \bar{\theta}\right)\right). \quad (48)$$

It is an implication of (48) that

$$\frac{d}{d\theta} \Lambda(F(V(\theta, u); \theta)) = 0. \quad (49)$$

That  $V(\bar{\theta}, u) = u$  implies

$$\Lambda(F(V(\theta, u); \theta)) = \Lambda(F(u; \bar{\theta})).$$

Differentiating with respect to  $u$  leads to

$$\Lambda'(F(V(\theta, u); \theta)) \cdot f(V(\theta, u); \theta) \cdot V_2(\theta, u) = \Lambda'(F(u; \bar{\theta})) \cdot f(V(\bar{\theta}, u); \theta). \quad (50)$$

Optimality implies that every menu  $V(\theta, u) \in \mathbb{S}$  has to satisfy the following Euler equation at any  $\theta$  where  $V(\theta, u) > 0$ :

$$\begin{aligned} \Lambda'(F(V(\theta, u); \theta)) \cdot f(V(\theta, u); \theta) \cdot (\theta \cdot V_1(\theta, u) - \varphi(V_1(\theta, u)) - V(\theta, u)) \cdot h(\theta) - \Lambda(F(V(\theta, u); \theta)) \cdot h(\theta) \\ = \frac{d}{d\theta} \left\{ \Lambda(F(V(\theta, u); \theta)) \cdot (\theta - \varphi'(V_1(\theta, u))) \cdot h(\theta) \right\}. \end{aligned} \quad (51)$$

Because of (49), it follows that

$$\begin{aligned} \frac{d}{d\theta} \left\{ \Lambda(F(V(\theta, u); \theta)) \cdot (\theta - \varphi'(V_1(\theta, u))) \cdot h(\theta) \right\} \\ = \Lambda(F(V(\theta, u); \theta)) \cdot \frac{d}{d\theta} \left\{ (\theta - \varphi'(V_1(\theta, u))) \cdot h(\theta) \right\}. \end{aligned} \quad (52)$$

Plugging (52) into (51) and manipulating leads to:

$$\frac{\Lambda'(F(V(\theta, u); \theta))}{\Lambda(F(V(\theta, u); \theta))} \cdot f(V(\theta, u); \theta) = \frac{h(\theta) + \frac{d}{d\theta} \left\{ (\theta - \varphi'(V_1(\theta, u))) \cdot h(\theta) \right\}}{(\theta \cdot V_1(\theta, u) - \varphi(V_1(\theta, u)) - V(\theta, u)) \cdot h(\theta)}. \quad (53)$$

Let us choose  $\theta = \bar{\theta}$  in (53) to obtain that:

$$\frac{\Lambda'(F(u; \bar{\theta}))}{\Lambda(F(u; \bar{\theta}))} \cdot f(u; \bar{\theta}) = \frac{h(\bar{\theta}) + \frac{d}{d\theta} \left\{ (\theta - \varphi'(V_1(\theta, u))) \cdot h(\theta) \right\}_{\theta=\bar{\theta}}}{(\bar{\theta} \cdot V_1(\bar{\theta}, u) - \varphi(V_1(\bar{\theta}, u)) - u) \cdot h(\bar{\theta})}. \quad (54)$$

Note that

$$\left( \frac{\Lambda'(F(V(\theta, u); \theta))}{\Lambda(F(V(\theta, u); \theta))} \cdot f(V(\theta, u); \theta) \right)^{-1} \cdot \frac{\Lambda'(F(u; \bar{\theta}))}{\Lambda(F(u; \bar{\theta}))} \cdot f(V(\theta, u); \bar{\theta}) = V_2(\theta, u), \quad (55)$$

where the equality follows from (50).

Dividing (54) by (53), and using the relation (55), we then obtain that

$$V_2(\theta, u) = \frac{(\theta \cdot V_1(\theta, u) - \varphi(V_1(\theta, u)) - V(\theta, u)) \cdot h(\theta)}{(\bar{\theta} \cdot V_1(\bar{\theta}, u) - \varphi(V_1(\bar{\theta}, u)) - V(\bar{\theta}, u)) \cdot h(\bar{\theta})} \cdot \frac{h(\bar{\theta}) + \frac{d}{d\bar{\theta}} \{(V_1(\theta, u) - \varphi'(V_1(\theta, u))) \cdot h(\theta)\}_{\theta=\bar{\theta}}}{h(\theta) + \frac{d}{d\theta} \{(V_1(\theta, u) - \varphi'(V_1(\theta, u))) \cdot h(\theta)\}}.$$

We posit that in the ordered equilibrium the highest type  $\bar{\theta}$  is always assigned the efficient quality level. The PDE is then

$$V_2(\theta, u) = \frac{(\theta \cdot V_1(\theta, u) - \varphi(V_1(\theta, u)) - V(\theta, u)) \cdot h(\theta)}{(\bar{S}^* - u) \cdot h(\bar{\theta})} \cdot \frac{h(\bar{\theta}) + \frac{d}{d\bar{\theta}} \{(V_1(\theta, u) - \varphi'(V_1(\theta, u))) \cdot h(\theta)\}_{\theta=\bar{\theta}}}{h(\theta) + \frac{d}{d\theta} \{(V_1(\theta, u) - \varphi'(V_1(\theta, u))) \cdot h(\theta)\}}. \quad (56)$$

Denote by  $\bar{u}^m$  the indirect utility of type  $\bar{\theta}$  in the Mussa-Rosen schedule. The PDE (56) has to be solved in the range  $u \in [\bar{u}^m, \bar{S}^*]$ ,  $\theta \in [\underline{\theta}, \bar{\theta}]$  with boundary conditions

$$V(\theta, \bar{S}^*) = \max_q \theta \cdot q - \varphi(q), \quad (57)$$

$$V(\theta, \bar{u}^m) = \max_q \left( \theta - \frac{1 - F(\theta)}{h(\theta)} \right) \cdot q - \varphi(q), \quad (58)$$

$$V_1(\bar{\theta}, u) = V_1(\bar{\theta}, \bar{S}^*) \quad (59)$$

and

$$V(\bar{\theta}, u) = u. \quad (60)$$

The boundary condition (57) states that the Bertrand schedule is the ‘‘supremum’’ contract in the support  $\mathbb{S}$ . The boundary condition (58) states that the Mussa-Rosen schedule is the ‘‘infimum’’ contract in the support  $\mathbb{S}$ . The boundary condition (59) requires that the type  $\bar{\theta}$  receives the same quality (which is the efficient one) in all contracts in  $\mathbb{S}$ . The boundary condition (60) requires that the solution to (56) is consistent with the definition of  $V(\theta, u)$ . Conditions (56)-(60) are necessary conditions for a smooth ordered equilibrium when there is a continuum of types. Unfortunately, the well known existence and uniqueness results for partial differential equations do not apply. In order to make progress in this difficult problem, we restrict attention to the linear quadratic case in the next section.

### 7.3 Uniform-Quadratic Case

Assume that production costs are quadratic,  $\varphi(q) = \frac{1}{2} \cdot q^2$ , and types are uniformly distributed,  $\theta \sim U[0, 1]$ .

The PDE (56) becomes

$$V_2(\theta, u) = \frac{2 - V_{11}(\bar{\theta}, u)}{2 - V_{11}(\theta, u)} \cdot \frac{\theta \cdot V_1(\theta, u) - \frac{1}{2} \cdot (V_1(\theta, u))^2 - V(\theta, u)}{\frac{1}{2} - u}, \quad (61)$$

with domain on  $[0, 1] \times [\frac{1}{4}, \frac{1}{2}]$ .

The proposed solution to the PDE above subject to the boundary conditions (57)-(60) is

$$V(\theta, u) = \frac{1}{4 \cdot u} \cdot \theta^2 + \left(1 - \frac{1}{2 \cdot u}\right) \cdot \theta + u + \frac{1}{4 \cdot u} - 1.$$

### 7.3.1 Verification of the Partial Differential Equation

Let us first compute partial derivatives:

$$V_1(\theta, u) = \frac{1}{2 \cdot u} \cdot \theta + \left(1 - \frac{1}{2 \cdot u}\right), \quad (62)$$

and

$$V_2(\theta, u) = -\frac{1}{4 \cdot u^2} \cdot \theta^2 + \frac{1}{2 \cdot u^2} \cdot \theta + 1 - \frac{1}{4 \cdot u^2}.$$

Let us first verify the boundary conditions. To verify (57), note that  $\bar{S}^* = \frac{1}{2}$ . Therefore,

$$V(\theta, \bar{S}^*) = V\left(\theta, \frac{1}{2}\right) = \frac{1}{2} \cdot \theta^2 = \max_q \theta \cdot q - \frac{1}{2} \cdot q^2.$$

To verify (58), note that  $\bar{u}^m = \frac{1}{4}$ . Therefore,

$$V(\theta, \bar{u}^m) = V\left(\theta, \frac{1}{4}\right) = \theta^2 - \theta + \frac{1}{4} = \max_q (2 \cdot \theta - 1) \cdot q - \frac{1}{2} \cdot q^2.$$

To verify (59), note that

$$V_1(1, u) = \frac{1}{2 \cdot u} + \left(1 - \frac{1}{2 \cdot u}\right) = 1.$$

To verify (60), note that

$$V(1, u) = \frac{1}{4 \cdot u} + 1 - \frac{1}{2 \cdot u} + u + \frac{1}{4 \cdot u} - 1 = u.$$

To verify that (61) is satisfied, note that

$$\begin{aligned} & \theta \cdot V_1(\theta, u) - \frac{1}{2} \cdot (V_1(\theta, u))^2 - V(\theta, u) \\ &= \theta \cdot \left(\frac{1}{2 \cdot u} \cdot \theta + \left(1 - \frac{1}{2 \cdot u}\right)\right) - \frac{1}{2} \cdot \left(\frac{1}{2 \cdot u} \cdot \theta + \left(1 - \frac{1}{2 \cdot u}\right)\right)^2 \\ & \quad - \frac{1}{4 \cdot u} \cdot \theta^2 - \left(1 - \frac{1}{2 \cdot u}\right) \cdot \theta - u - \frac{1}{4 \cdot u} + 1, \end{aligned}$$

which, after some algebra, can be shown to be equal to

$$\left(\frac{1}{2} - u\right) \cdot \left(-\frac{1}{4 \cdot u^2} \cdot \theta^2 + \frac{1}{2 \cdot u^2} \cdot \theta + 1 - \frac{1}{4 \cdot u^2}\right) = \left(\frac{1}{2} - u\right) \cdot V_2(\theta, u).$$

Because  $V_{11}(\bar{\theta}, u) = V_{11}(\theta, u)$ , it follows that (61) holds.

### 7.3.2 Ordered Equilibrium

Next, we use the family of curves  $(V(\theta, \bar{u}))_{\theta \in [0,1]}$  for  $\bar{u} \in [\frac{1}{4}, \frac{1}{2}]$  to propose the equilibrium distribution over menus  $\tilde{F}$ . From (54) and the knowledge of  $V$  we obtain a unique  $\bar{u} \in (\frac{1}{4}, \frac{1}{2})$  such that  $\int_{\frac{1}{4}}^{\bar{u}} f(\bar{u}; 1) d\bar{u} = 1$ . Therefore, in the proposed equilibrium, the firms offer menus  $(V(\theta, \bar{u}))_{\theta \in [0,1]}$  for  $\bar{u} \in [\frac{1}{4}, \bar{u}]$  such that a menu less generous than  $(V(\theta, \bar{u}))_{\theta \in [0,1]}$  is offered with probability  $F(\bar{u}; 1)$ . The rest of this appendix verifies that no firm has a profitable deviation. We start verifying that all menus in  $\mathbb{S}$  yield the same profit.

### 7.3.3 Checking Indifference

Here we verify that all menus  $(V(\theta, \bar{u}))_{\theta \in [0,1]}$  for  $\bar{u} \in [\frac{1}{4}, \bar{u}]$  lead to the same profit. We have:

$$\Pi(u) \equiv \int_{\alpha(u)}^{\bar{\theta}} \Lambda(F[V(\theta, u); \theta]) \cdot \left( \theta \cdot V_1(\theta, u) - \frac{1}{2} \cdot (V_1(\theta, u))^2 - V(\theta, u) \right) d\theta,$$

where  $\alpha(u)$  solves

$$V_1(\alpha(u), u) = 0.$$

It is easy to verify that  $V(\alpha(u), u) = 0$ .

To simplify notation, let

$$L(V(\theta, u), V_1(\theta, u), \theta) \equiv \Lambda(F[V(\theta, u); \theta]) \cdot \left( \theta \cdot V_1(\theta, u) - \frac{1}{2} \cdot (V_1(\theta, u))^2 - V(\theta, u) \right).$$

Then

$$\begin{aligned} \Pi'(u) &\equiv \int_{\alpha(u)}^{\bar{\theta}} \{L_1(V(\theta, u), V_1(\theta, u), \theta) \cdot V_2(\theta, u)\} d\theta \\ &+ \int_{\alpha(u)}^{\bar{\theta}} \{L_2(V(\theta, u), V_1(\theta, u), \theta) \cdot V_{12}(\theta, u)\} d\theta \\ &- L(V(\alpha(u), u), V_1(\alpha(u), u), \alpha(u)) \cdot \alpha'(u). \end{aligned} \tag{63}$$

Integration by parts delivers that

$$\begin{aligned} &\int_{\alpha(u)}^{\bar{\theta}} \{L_2(V(\theta, u), V_1(\theta, u), \theta) \cdot V_{12}(\theta, u)\} d\theta \\ &= L_2(V(\theta, u), V_1(\theta, u), \theta) \cdot V_2(\theta, u) \Big|_{\alpha(u)}^{\bar{\theta}} \\ &- \int_{\alpha(u)}^{\bar{\theta}} V_2(\theta, u) \cdot \left\{ \frac{d}{d\theta} [L_2(V(\theta, u), V_1(\theta, u), \theta)] \right\} d\theta. \end{aligned}$$

Plugging the above into (63) and using the fact  $V(\theta, u)$  solves the Euler equation for every  $(\theta, u)$  with  $\theta > \alpha(u)$  leads to

$$\Pi'(u) = L_2(V(\theta, u), V_1(\theta, u), \theta) \cdot V_2(\theta, u)|_{\alpha(u)}^{\bar{\theta}} - L(V(\alpha(u), u), V_1(\alpha(u), u), \alpha(u)) \cdot \alpha'(u).$$

Note that

$$L_2(V(\bar{\theta}, u), V_1(\bar{\theta}, u), \bar{\theta}) = \Lambda(F(V(\bar{\theta}, u); \bar{\theta})) \cdot (\bar{\theta} - V_1(\bar{\theta}, u)) = 0.$$

Recall that  $V(\alpha(u), u) = 0$ . Total differentiation yields

$$V_1(\alpha(u), u) \cdot \alpha'(u) + V_2(\alpha(u), u) = 0.$$

Because by construction  $V_1(\alpha(u), u) = 0$ , it follows that  $V_2(\alpha(u), u) = 0$ . This implies that

$$L_2(V(\theta, u), V_1(\theta, u), \theta) \cdot V_2(\theta, u)|_{\alpha(u)}^{\bar{\theta}} = 0.$$

Finally, because  $V(\alpha(u), u) = 0$  and  $V_1(\alpha(u), u) = 0$ ,

$$L(V(\alpha(u), u), V_1(\alpha(u), u), \alpha(u)) = 0.$$

This establishes that  $\Pi'(u) = 0$  for all  $u \in [\frac{1}{4}, \bar{u}]$ , implying that all proposed menus yield the same profit.

### 7.3.4 Existence of a Solution

We write  $\Psi(u, \theta)$  for the sales function from offering a utility  $u$  to the type  $\theta$ :  $\Psi(u, \theta) \equiv \Lambda(F(u; \theta))$ . For every  $\bar{u} \in [\frac{1}{4}, \bar{u}]$  write  $(u_{\bar{u}}(\theta))$  for the curve  $(V(\theta, \bar{u}))_{\theta \in [0, 1]}$ . We write  $AC[0, 1]$  for the space of absolutely continuous functions from  $[0, 1]$  into  $\mathbb{R}$ . Hence, we can write the firm's problem as:

$$\max_{u \in AC[0, 1]} \int_0^1 \Pi(\theta, u(\theta), \dot{u}(\theta)) d\theta, \quad (64)$$

where

$$\Pi(\theta, u, \dot{u}) \equiv \Psi(u, \theta) \left[ \theta \cdot \dot{u} - \frac{1}{2} \cdot (\dot{u})^2 - u \right].$$

We consider the relaxed problem in which  $\dot{u}(\theta) = q(\theta) \in \mathbb{R}$  and  $q(\theta)$  need not be monotonic. Lemma 5 invokes a calculus of variations existence theorem to show that (64) has a solution.

**Lemma 5** *The problem (64) has a solution.*

**Proof.** Let  $A \equiv \sup_{u,\theta} \Psi(u,\theta)$ . It is straightforward to see that we may restrict attention to allocations such that  $\left[\theta \cdot \dot{u} - \frac{1}{2} \cdot (\dot{u})^2 - u\right] \geq 0$  for almost all  $\theta$ . Therefore

$$\begin{aligned} \Pi(\theta, u, \dot{u}) &\leq A\left[\dot{u} - \frac{1}{2} \cdot (\dot{u})^2\right] = A\dot{u} \left(1 - \frac{1}{4} \cdot \dot{u}\right) - \left(\frac{A}{4}\right) \cdot (\dot{u})^2 \\ &\leq A - \left(\frac{A}{4}\right) \cdot (\dot{u})^2. \end{aligned}$$

Hence, we may assume that  $\Pi$  is coercive of degree 2, that is,  $\Pi(\theta, u, \dot{u}) \leq A - \left(\frac{A}{4}\right) \cdot (\dot{u})^2$  for (almost) every  $(\theta, u, \dot{u})$ . Furthermore, notice that  $\Pi(\theta, u, \dot{u})$  is continuous in  $(\theta, u, \dot{u})$  and it is concave in  $\dot{u}$ . The existence of an absolutely continuous solution follows from Theorem 16.2 in Clarke (2013). Q.E.D.

### 7.3.5 Conclusion

We will show that the solutions to (64) are the curves  $(u_{\bar{u}}(\theta))$  for  $\bar{u} \in \left[\frac{1}{4}, \bar{u}\right]$ . Our approach is to (i) show that we must have  $u^*(1) \in \left[\frac{1}{4}, \bar{u}\right]$  for any optimal menu  $u^*$ , and (ii) show that, for each value of  $u^*(1) \in \left[\frac{1}{4}, \bar{u}\right]$ , the necessary condition for the menu  $u^*(\theta)$  to be optimal in (64) admits a unique solution, which is the menu given by  $u_{u^*(1)}(\theta)$ . Thus, the only candidates for optima in (64) are the curves  $(u_{\bar{u}}(\theta))$ ,  $\bar{u} \in \left[\frac{1}{4}, \bar{u}\right]$ , identified above.

We now outline the proof in more detail. In the proof, we will use Theorem 18.1 in Clarke (2013), which provides necessary conditions for optimal solutions of calculus of variations problems. Hence, we first check that the conditions for Theorem 18.1 hold. Next, we consider an optimal allocation  $(u^*(\theta))$  and divide the analysis into five exhaustive cases.

**Case 1** deals with  $u^*(1) \in \left(u_{\frac{1}{4}}(1), u_{\bar{u}}(1)\right)$ . We show that the Euler equation (51) is necessary. Moreover, this Euler equation leads to an ordinary differential equation which has a unique solution, as desired.

**Case 2** deals with the possibility that  $u^*(1) > u_{\bar{u}}(1)$ . We show that (51) then implies  $u^*(\theta)$  is bounded above  $u_{\bar{u}}(\theta)$  for all  $\theta$ ; hence  $(u^*(\theta))$  cannot be optimal.

**Case 3** deals with the possibility that  $u^*(1) < u_{\frac{1}{4}}(1)$ . We consider two cases. First, assume that the curves  $(u^*(\theta))$  and  $\left(u_{\frac{1}{4}}(\theta)\right)$  never intersect in  $\left(\frac{1}{2}, 1\right)$ . In this case, the firm could profitably deviate by offering the menu  $\left(u_{\frac{1}{4}}(\theta)\right)$ . Second, assume that the curves  $(u^*(\theta))$  and  $\left(u_{\frac{1}{4}}(\theta)\right)$  intersect at some  $\theta^* \in \left(\frac{1}{2}, 1\right)$ . In this case, our argument below implies that the firm could profitably deviate by offering the curve (remember that we are considering a relaxed problem in which monotonicity constraints are ignored):

$$u^{**}(\theta) := \begin{cases} u_{\frac{1}{4}}(\theta) & \text{if } \theta > \theta^* \\ u^*(\theta) & \text{otherwise.} \end{cases}$$

Hence we conclude that this case is not possible.

**Case 4** deals with the case that  $u^*(1) = u_{\bar{u}}(1)$ . This case is complicated by the fact that the function  $\Psi$  is not differentiable along the curve  $(u_{\bar{u}}(\theta))$ , since  $\Psi(u', \theta) = \Psi(u'', \theta)$  whenever  $\min\{u', u''\} \geq u_{\bar{u}}(\theta)$ . Therefore, we have  $0 = \Psi_1(u_{\bar{u}}(\theta)_+, \theta) < \Psi_1(u_{\bar{u}}(\theta)_-, \theta)$ .<sup>31</sup> The proof is thus divided into two cases. First, it is assumed that there is  $\theta' \in (0, 1)$  such that  $u^*(\theta') > u_{\bar{u}}(\theta')$ . The Euler equation (51) then implies that  $u^*(\theta) > u_{\bar{u}}(\theta')$  for all  $\theta < \theta'$  and that  $\dot{u}^*(\theta) < 0$  for sufficiently small  $\theta$ , which contradicts the optimality of  $(u^*(\theta))$ .

Thus we may assume that  $u^*(\theta) \leq u_{\bar{u}}(\theta)$  for every  $\theta$ . We proceed as follows. First, we use calculus of variations necessary conditions which do not require differentiability.<sup>32</sup> Roughly, when the function  $\Psi$  fails to be differentiable along the curve  $u^*(\theta)$  we have a generalized Euler equation in which a subgradient  $\xi(\theta) \in \partial_1 \Psi(u^*(\theta), \theta)$  plays the role of the the derivative of  $\Psi(u^*(\theta), \theta)$ .<sup>33</sup> In this case, we show that the generalized Euler equation picks the left subgradient  $\Psi(u^*(\theta), \theta)$  with respect to  $u$  for almost every point. Hence, the fact that  $\Psi$  is not differentiable along the curve  $(u^*(\theta))$  is immaterial and the conclusion that  $(u^*(\theta)) = (u_{\bar{u}}(\theta))$  follows from the argument given in Case 1. Heuristically, our argument proceeds as follows. Suppose we have an interval  $[\theta_\alpha, \theta_\beta]$  in which the Euler equation picks a subgradient  $\xi(\theta) < \Psi_1(u^*(\theta)_-, \theta)$ . Therefore, since  $\Psi$  is not differentiable only along the curve  $(u_{\bar{u}}(\theta))$  it must be that (for  $\theta_\alpha$  sufficiently close to  $\theta_\beta$ ):

$$u^*(\theta) = u_{\bar{u}}(\theta) \quad \forall \theta \in [\theta_\alpha, \theta_\beta]. \quad (65)$$

Recall that the curve  $(u_{\bar{u}}(\theta))$  is constructed by the Euler equation in which the subgradient  $\Psi(u_{\bar{u}}(\theta)_-, \theta)$  is selected and that  $u^*(\theta_\beta) = u_{\bar{u}}(\theta_\beta)$ . Hence, since  $\xi(\theta) < \Psi_1(u_{\bar{u}}(\theta)_-, \theta)$  for every  $\theta \in [\theta_\alpha, \theta_\beta]$  the curves  $(u_{\bar{u}}(\theta))$  and  $(u^*(\theta))$  drift apart for  $\theta$  sufficiently close to  $\theta_\beta$ , which contradicts (65).<sup>34</sup>

**Case 5** deals with the case that  $u^*(1) = u_{\frac{1}{4}}(1)$ . We show that the function  $\Psi$  is differentiable along  $\left(u_{\frac{1}{4}}(\theta), \theta\right)_{\theta > \frac{1}{2}}$ . Therefore, the difficulties which arose in Case 4 do not appear here and an argument analogous to the one from Case 1 implies that  $(u^*(\theta)) = \left(u_{\frac{1}{4}}(\theta)\right)$ .

**Proof of Proposition 6.** Write  $(u^*(\theta))$  for an optimal menu. We will show that for every  $\bar{u} \in \left[\frac{1}{4}, \bar{u}\right]$  the menu  $(u_{\bar{u}}(\theta))$  is optimal. Our proof will use the necessary conditions from Theorem 18.1 in Clarke (2013). In order to use this Theorem, we must check that our primitives satisfy the following Lipschitz condition.

**Conditions for application of Theorem 18.1 in Clarke (2013).** Let  $K > 0$  be such that  $|\Psi(u_\beta(\theta), \theta) - \Psi(u_\gamma(\theta), \theta)| \leq K |u_\beta(\theta) - u_\gamma(\theta)|$  for all curves  $(u_\beta(\theta)), (u_\gamma(\theta)) \in AC[0, 1]$  and all

<sup>31</sup>For a function  $h : \mathfrak{R}^2 \rightarrow \mathfrak{R}$  we write  $h_1(a_-, b)$  for  $\lim_{x \rightarrow a-} \frac{h(x, b) - h(a, b)}{x - a}$  and  $h_1(a_+, b)$  for  $\lim_{x \rightarrow a+} \frac{h(x, b) - h(a, b)}{x - a}$ .

<sup>32</sup>We follow the analysis of Clarke (2013). Definitions are given in the proof of Proposition 6 below.

<sup>33</sup>For a function  $h : \mathfrak{R}^2 \rightarrow \mathfrak{R}$ , we write  $\partial_1 h(a, b)$  for all subgradients of the function  $x \rightarrow h(x, b)$  at  $h(a, b)$ .

<sup>34</sup>The proof would follow essentially the steps that we just explained if we could assume that the Euler equation picks a subgradient  $\xi(\theta) \in \partial_1 \Psi(u_{\bar{u}}(\theta), \theta) - \varepsilon$  in an interval  $[\theta_\alpha, \theta_\beta]$ . However, assuming that such property holds in an entire interval may be with loss of generality, which makes the formal analysis below a little more complicated.

$\theta$ . Since  $\max_{q,\theta} |\theta \cdot q - \frac{1}{2} \cdot q^2| \leq \frac{1}{2}$  we have, for all  $(u_\beta(\theta), (u_\gamma(\theta)) \in AC[0, 1]$ , all  $\theta$ ,

$$\begin{aligned} & |\Pi(\theta, u_\beta(\theta), \dot{u}_\beta(\theta)) - \Pi(\theta, u_\gamma(\theta), \dot{u}_\gamma(\theta))| \\ &= \left| \Psi(u_\beta(\theta), \theta) \left[ \theta \cdot \dot{u}_\beta(\theta) - \frac{1}{2} \cdot (\dot{u}_\beta(\theta))^2 - u_\beta(\theta) \right] - \Psi(u_\gamma(\theta), \theta) \left[ \theta \cdot \dot{u}_\gamma(\theta) - \frac{1}{2} \cdot (\dot{u}_\gamma(\theta))^2 - u_\gamma(\theta) \right] \right| \\ &\leq K |u_\beta(\theta) - u_\gamma(\theta)| + KA |\dot{u}_\beta(\theta) - \dot{u}_\gamma(\theta)|. \end{aligned}$$

Hence the Lipschitz condition (LH) in page 348 of Clarke (2013) is satisfied, allowing application of Theorem 18.1 in Clarke (2013).

We have five cases.

**Case 1:**  $u^*(1) \in (u_{\frac{1}{4}}(1), u_{\bar{u}}(1))$ .

Notice that, since  $u^*(\theta)$  is absolutely continuous, there is  $\varepsilon_1 > 0$  and a neighborhood of

$$\{(\theta, u^*(\theta), \dot{u}^*(\theta)) : \theta \in [1 - \varepsilon_1, 1]\}$$

for which  $\Pi$  is  $C^2$ . Thus condition (E) for a (locally  $C^1$  function) in Theorem 18.1 assures the existence of an arc  $p^* : [0, 1] \rightarrow \mathbb{R}$  for which, for almost every  $\theta$ ,

$$\dot{p}^*(\theta) = \frac{\partial \Pi(\theta, u^*(\theta), \dot{u}^*(\theta))}{\partial u^*(\theta)} \quad (66)$$

$$p^*(\theta) = \frac{\partial \Pi(\theta, u^*(\theta), \dot{u}^*(\theta))}{\partial \dot{u}^*(\theta)}. \quad (67)$$

From the transversality condition (T) in Theorem 18.1 we conclude that  $p^*(1) = 0$  and thus there is  $\varepsilon_2 \in (0, \varepsilon_1)$  for which  $\dot{u}^*(\theta) \in [\frac{1}{2}, 2]$  for (almost) all  $\theta \in [1 - \varepsilon_2, 1]$ . It follows from (67) that, for (almost) all  $\theta \in [1 - \varepsilon_2, 1]$ , we have

$$\dot{u}(\theta) = \theta - \left( \frac{p^*(\theta)}{\Psi(u(\theta), \theta)} \right),$$

and hence  $\dot{u}(\theta)$  is a Lipschitz function on this interval. Since  $\Pi(\theta, u(\theta), \dot{u}(\theta))$  is strictly concave in  $\dot{u}(\theta)$  on this interval and  $\Psi$  is smooth on this interval, we may apply Theorem 15.7 in Clarke (2013) to conclude that  $u(\theta)$  is a smooth function for this interval. Therefore, the Euler Equation

$$\frac{d}{d\theta} [\Psi(u(\theta), \theta) (\theta - \dot{u}(\theta))] = \frac{\partial}{\partial u(\theta)} \left[ \Psi(u(\theta), \theta) \left[ \theta \cdot \dot{u}(\theta) - \frac{1}{2} \cdot (\dot{u}(\theta))^2 - u(\theta) \right] \right] \quad (68)$$

holds for this interval. Let  $(y(\theta), y'(\theta))$  be a solution of the Euler equation (68) subject to  $y(1) = u^*(1)$ . The Picard-Lindelöf theorem establishes that this solution is unique and  $y(\theta) = u^*(\theta)$  for all  $\theta \in [1 - \varepsilon_2, 1]$ . We can then extend the solution  $(y(\theta), y'(\theta))$  to values  $\theta$  such that  $y(\theta)$  remains strictly positive. Again, the solution is unique for these values. This establishes that  $u^*(\theta) = y(\theta)$  must be the unique function solving (68) for all  $\theta \in [\underline{\theta}(u^*(1)), 1]$ , where  $\underline{\theta}(u^*(1))$  is the largest value of  $\theta$  at which  $y(\theta) = 0$ . One can verify moreover that we must have  $u^*(\theta) = 0$  for  $\theta \leq \underline{\theta}(u^*(1))$ .

**Case 2:**  $u^*(1) > u_{\bar{u}}(1)$ .

To show that this is inconsistent with optimality of  $u^*$ , we consider the cost function  $\frac{1}{2}q^2$  extended over all of  $\mathbb{R}$ . Suboptimality for this cost function implies suboptimality also for the problem when negative effort is not permitted (as assumed in the model).

Using a similar argument to the one from Case 2 one can show that there exists  $\varepsilon > 0$  and a neighborhood of  $\{(\theta, u^*(\theta), \dot{u}^*(\theta)) : \theta \in [1 - \varepsilon, 1]\}$  for which an increase in  $u(\theta)$  does not increase sales, that is  $\Psi(u^*(\theta), \theta) = A$ , and the Euler equation holds. Thus, the Euler equation (68) implies  $\ddot{u}^*(\theta) = 2$ . Since from the transversality condition (T) in Theorem 18.1 we have  $\dot{u}^*(1) = 1$  we conclude that there is an interval for which  $u^*(\theta) = \theta^2 - \theta + c$ . Let  $[\theta_a, 1]$  be the largest interval with this property.

Next, consider the Euler equation (68) evaluated at the curve  $(u_{\bar{u}}(\theta))$ . Since  $\Psi(u_{\bar{u}}(\theta), \theta) = A$  we have:

$$\ddot{u}_{\bar{u}}(\theta) = 2 - \frac{\Psi_{1-(u_{\bar{u}}(\theta)_-, \theta)}}{A} \left[ \theta \cdot \dot{u}_{\bar{u}}(\theta) - \frac{1}{2} \cdot (\dot{u}_{\bar{u}}(\theta))^2 - u_{\bar{u}}(\theta) \right] < \dot{u}^*(\theta) = 2,$$

thus  $\dot{u}_{\bar{u}}(\theta) > \dot{u}^*(\theta)$  for all  $\theta \in (\theta_a, 1)$ . Therefore  $\theta_a < \underline{\theta}(u_{\bar{u}}(1))$ , where  $\underline{\theta}(u_{\bar{u}}(1))$  is the largest value of  $\theta$  such that  $u_{\bar{u}}(\theta) = 0$ . Recalling that  $\alpha(u_{\bar{u}}(1))$  satisfies  $\dot{u}_{\bar{u}}(\alpha(u_{\bar{u}}(1))) = 0$ , we conclude that, for all  $\theta \in (\theta_a, \alpha(u_{\bar{u}}(1)))$  we have  $\dot{u}^*(\theta) < 0$ . It is then easy to see that the seller has a profitable deviation.

**Case 3:**  $u^*(1) < u_{\frac{1}{4}}(1)$ .

Proceeding exactly as in Case 2, we conclude that there is an interval  $(\hat{\theta}, 1)$  for which  $u^*$  is smooth. Furthermore, we have  $\dot{u}^*(1) = 1$  and, for all  $\theta \in (\hat{\theta}, 1)$ ,  $\ddot{u}^*(\theta) = 2$ . We see that:

$$\ddot{u}_{\frac{1}{4}}(\theta) = 2 - \frac{\Psi_1(u_{\frac{1}{4}}(\theta)_+, \theta)}{\Psi(u_{\frac{1}{4}}(\theta)_+, \theta)} \left[ \theta \cdot \dot{u}_{\frac{1}{4}}(\theta) - \frac{1}{2} \cdot (\dot{u}_{\frac{1}{4}}(\theta))^2 - u_{\frac{1}{4}}(\theta) \right],$$

which shows that  $\ddot{u}_{\frac{1}{4}}(\theta) \leq \ddot{u}^*(\theta)$  for every  $\theta \in (\hat{\theta}, 1)$ .

First, assume that  $u_{\frac{1}{4}}(\theta) > u^*(\theta)$  for all  $\theta$  for which  $u^*(\theta) > 0$ . Write  $B \equiv \Psi(u_{\frac{1}{4}}(1)_+, 1)$ . In this case, the firm sells to each consumer for which  $u^*(\theta) > 0$  with probability  $B$ . Therefore, the profit from this contract is weakly lower than the profit from a monopolist who faces a constant sales function equal to  $B$ . Since the unique solution to the later problem is given by the curve  $(u_{\frac{1}{4}}(\theta))$  we conclude that there is a profitable deviation.

Next, assume that there is  $\theta'$  for which  $0 < u_{\frac{1}{4}}(\theta') = u^*(\theta')$  and let  $\theta^*$  be the greatest  $\theta'$  satisfying this condition. Recall that the curve  $\dot{u}_{\frac{1}{4}}(\theta)$  solved:

$$\begin{aligned} & \max_{(u(\theta))} B \int_{\frac{1}{2}}^1 \left[ \theta \cdot \dot{u}(\theta) - \frac{1}{2} \cdot (\dot{u}(\theta))^2 - u(\theta) \right] d\theta \\ & \text{s.t.: } u(\theta) = \int_{\frac{1}{2}}^{\theta} \dot{u}(z) dz \text{ for all } \theta \in \left[ \frac{1}{2}, 1 \right] \end{aligned} \tag{69}$$

which implies that  $\left(\dot{u}_{\frac{1}{4}}(\theta)\right)_{\theta \in [\theta^*, 1]}$  solves:

$$\begin{aligned} & \max_{\left(\dot{u}_{\frac{1}{4}}(\theta)\right)_{\theta \in [\theta^*, 1]}} B \int_{\theta^*}^1 \left[ \theta \cdot \dot{u}(\theta) - \frac{1}{2} \cdot (\dot{u}(\theta))^2 - u(\theta) \right] d\theta \\ & \text{s.t.: } u(\theta) = u_{\frac{1}{4}}(\theta^*) + \int_{\theta^*}^{\theta} \dot{u}(z) dz \text{ for all } \theta \in [\theta^*, 1]. \end{aligned} \quad (70)$$

Since  $\left(\dot{u}_{\frac{1}{4}}(\theta)\right)_{\theta \in [\theta^*, 1]}$  was (a.e.) unique and  $\left(\dot{u}_{\frac{1}{4}}(\theta)\right)_{\theta \in [\theta^*, 1]}$  is different from  $(\dot{u}^*(\theta))_{\theta \in [\theta^*, 1]}$  in a subset of positive measure, we conclude that  $(u^*(\theta))$  is not optimal.

**Case 4:**  $u^*(1) = u_{\bar{u}}(1)$ .

First we claim that  $u^*(\theta) \leq u_{\bar{u}}(\theta)$  for every  $\theta$ . Suppose that there is  $\theta' \in (0, 1)$  such that  $u^*(\theta') > u_{\bar{u}}(\theta')$ . Let  $\theta^* \in (\theta', 1]$  be the smallest  $\theta > \theta'$  such that  $u^*(\theta) = u_{\bar{u}}(\theta)$ . Thus notice that  $\Psi(u^*(\theta), \theta) = A$  for all  $\theta \in (\theta', \theta^*)$ . Therefore, using an argument similar to the one from Case 1, we conclude that  $u^*$  is smooth in  $(\theta', \theta^*)$ . Thus, we have  $\dot{u}^*(\theta^*) \leq \dot{u}_{\bar{u}}(\theta^*)$  and by the same argument as in Case 2,  $\ddot{u}^*(\theta) > \ddot{u}_{\bar{u}}(\theta)$  for all  $\theta \in (\theta', \theta^*)$ . Therefore,  $\dot{u}^*(\theta) < \dot{u}_{\bar{u}}(\theta)$  for all  $\theta \in (\alpha(u_{\bar{u}}(1)), \theta^*)$  and thus there is  $\hat{\theta} \in (\alpha(u_{\bar{u}}(1)), \theta^*)$  such that  $\theta \in (\alpha(u_{\bar{u}}(1)), \hat{\theta})$  implies  $u^*(\theta) > 0$  and  $\dot{u}^*(\theta) < 0$ , a contradiction.

The rest of this proof of this case is complicated because the function  $\Psi(\cdot, \theta)$  is not differentiable along the curve  $(u_{\bar{u}}(\theta), \theta)$ . Indeed, we have  $\partial_1 \Psi(u_{\bar{u}}(\theta), \theta) = [0, \Psi_1(u_{\bar{u}}(\theta)_-, \theta)]$ .<sup>35</sup> In this case, we have to write Condition (E) in Theorem 18.1 in Clarke (2013) in its general form, which implies:<sup>36</sup>

$$p^*(\theta) \in -\Psi(u^*(\theta), \theta) + \partial_1 \Psi(u^*(\theta), \theta) \left[ \theta \cdot \dot{u}^*(\theta) - \frac{1}{2} \cdot (\dot{u}^*(\theta))^2 - u^*(\theta) \right] \quad (71)$$

$$p^*(\theta) = \Psi(u^*(\theta), \theta) [\theta - \dot{u}^*(\theta)]. \quad (72)$$

Condition (T) implies  $p^*(1) = 0$  and hence we can find an interval  $\theta \in [1 - \varepsilon, 1]$  for which  $\dot{u}^*(\theta)$  is Lipschitz. We fix this interval in the analysis below. We claim that

$$p^*(\theta) = -\Psi(u^*(\theta), \theta) + \max_{\xi \in \partial_1 \Psi(u^*(\theta), \theta)} \xi \left[ \theta \cdot \dot{u}^*(\theta) - \frac{1}{2} \cdot (\dot{u}^*(\theta))^2 - u^*(\theta) \right] \quad (73)$$

for every Lebesgue point of this interval. That is, for all  $\theta \in [1 - \varepsilon, 1)$  for which we have a Lebesgue point, i.e.

$$0 = \lim_{\varepsilon \downarrow 0} \frac{\int_{\theta - \varepsilon}^{\theta + \varepsilon} |\dot{p}(\theta) - \dot{p}(x)| dx}{2\varepsilon},$$

<sup>35</sup>Recall that, for a function  $h : \mathfrak{R}^2 \rightarrow \mathfrak{R}$  we write  $\partial_1 h(a, b)$  for all subgradients of the function  $x \rightarrow h(x, b)$  at  $h(a, b)$ .

<sup>36</sup>The condition (E) from Theorem 18.1 in Clarke reads:  $p'(\theta) \in \text{cow} : (\omega, p(\theta)) \in \partial_L \Pi(\theta, u^*(\theta), \dot{u}^*(\theta))$  where  $\partial_L \Pi$  is the limiting subdifferential of  $\Pi$  with respect of  $(u, \dot{u})$  (see Definition 11.10 in Clarke (2013)). The necessary condition above follows from Exercise 18.4 in Clarke (2013) that states that the condition above implies:  $(p'(\theta), p(\theta)) \in \partial_C \Pi(\theta, u^*(\theta), \dot{u}^*(\theta))$ , where  $\partial_C \Pi$  is a generalized subdifferential (see Definition 10.3 in Clarke (2013)) with respect to  $(u, \dot{u})$ . Given the primitives of our model, it is immediate to verify that  $\partial_C \Pi$  is just the set of subgradients of  $\Pi$  with respect to  $(u, \dot{u})$ .

we have (73). Since  $\dot{p}^*$  is integrable, Theorem 7.7 in Rudin (1987) implies that almost every point is a Lebesgue point. Hence (73) holds almost everywhere. Therefore, since non-differentiabilities occur on a zero measure set, and since we have established in the first paragraph of the analysis of this Case that  $u^* \leq u_{\bar{u}}$ , the rest of the analysis is identical to Case 1. Thus we will conclude that proof by showing that (73) holds at every Lebesgue point of  $\dot{p}^*$ .

Towards a contradiction, consider a Lebesgue point for which

$$\dot{p}^*(\theta) < -\Psi(u^*(\theta), \theta) + \max_{\xi \in \partial_1 \Psi(u^*(\theta), \theta)} \xi \left[ \theta \cdot \dot{u}^*(\theta) - \frac{1}{2} \cdot (\dot{u}^*(\theta))^2 - u^*(\theta) \right]. \quad (74)$$

Clearly we have  $u^*(\theta) = u_{\bar{u}}(\theta)$ . Next, we will show that  $\dot{u}^*(\theta) = \dot{u}_{\bar{u}}(\theta)$ . Towards a contradiction assume that  $\dot{u}^*(\theta) \neq \dot{u}_{\bar{u}}(\theta)$ . From (72) we have  $\dot{u}^*(\theta) = \theta - \frac{p^*(\theta)}{\Psi(u^*(\theta), \theta)}$  and hence (since  $p^*(\theta)$  and  $\Psi(u^*(\theta), \theta)$  are continuous) we can find an interval  $(\theta - \epsilon, \theta + \epsilon)$  for which  $|\dot{u}^*(\tilde{\theta}) - \dot{u}_{\bar{u}}(\tilde{\theta})| > 0$  for all  $\tilde{\theta} \in (\theta - \epsilon, \theta + \epsilon)$  which implies that  $\theta$  is the only point that  $u^*(\tilde{\theta}) = u_{\bar{u}}(\tilde{\theta})$  in this interval. Hence, (73) holds (a.e.) in this interval.

Next, assume towards a contradiction that we can find a lebesgue point  $\theta$  of  $\dot{p}^*$  and  $\zeta > 0$  such that  $\dot{u}^*(\theta) = \dot{u}_{\bar{u}}(\theta)$  and

$$\begin{aligned} \dot{p}^*(\theta) &= -\Psi(u^*(\theta), \theta) + \Psi_1(u_{\bar{u}}(\theta)_-, \theta) \left[ \theta \cdot \dot{u}^*(\theta) - \frac{1}{2} \cdot (\dot{u}^*(\theta))^2 - u^*(\theta) \right] - \zeta \\ &= \dot{p}_{\bar{u}}(\theta) - \zeta. \end{aligned} \quad (75)$$

Let  $\dot{p}_{\bar{u}}$  be the (smooth) arc associated with the curve  $(u_{\bar{u}}(\theta))$ . Since  $\theta$  is a Lebesgue point we can find  $\epsilon > 0$  such that for every  $\tilde{\theta} \in (\theta - \epsilon, \theta)$  we have:

$$\frac{\int_{\tilde{\theta}}^{\theta} \dot{p}^*(z) dz}{\theta - \tilde{\theta}} < \frac{\int_{\tilde{\theta}}^{\theta} \dot{p}_{\bar{u}}(z) dz}{\theta - \tilde{\theta}} - \frac{\zeta}{2}. \quad (76)$$

Notice that (76) implies that there is  $\hat{\theta} \in (\theta - \epsilon, \theta)$  such that  $u^*(\hat{\theta}) = u_{\bar{u}}(\hat{\theta})$ . Using (72) we have for all  $\tilde{\theta} \in (\hat{\theta}, \theta)$

$$\begin{aligned} \dot{u}^*(\theta) - \dot{u}^*(\tilde{\theta}) &= \theta - \tilde{\theta} + \left( \frac{p^*(\tilde{\theta})}{\Psi(u^*(\tilde{\theta}), \tilde{\theta})} \right) - \left( \frac{p^*(\theta)}{\Psi(u^*(\theta), \theta)} \right) \\ &\geq \theta - \tilde{\theta} - \left( \frac{p^*(\theta) - p^*(\tilde{\theta})}{\Psi(u^*(\theta), \theta)} \right) \\ &= \theta - \tilde{\theta} - \left( \frac{\int_{\tilde{\theta}}^{\theta} \dot{p}^*(z) dz}{\Psi(u^*(\theta), \theta)} \right) \\ &\geq \theta - \tilde{\theta} - \left( \frac{\int_{\tilde{\theta}}^{\theta} \dot{p}_{\bar{u}}(z) dz}{\Psi(u^*(\theta), \theta)} \right) + \frac{\zeta(\theta - \tilde{\theta})}{A} \\ &= \dot{u}_{\bar{u}}(\theta) - \dot{u}_{\bar{u}}(\tilde{\theta}) + \frac{\zeta(\theta - \tilde{\theta})}{A}, \end{aligned}$$

where the first inequality uses  $\Psi(u^*(\tilde{\theta}), \tilde{\theta}) \leq \Psi(u^*(\theta), \theta) = A$  and the second inequality uses (76). Therefore, since  $\dot{u}^*(\theta) = \dot{u}_{\bar{u}}(\theta)$  we have  $\dot{u}^*(\hat{\theta}) \leq \dot{u}_{\bar{u}}(\hat{\theta}) - \frac{\zeta(\theta - \hat{\theta})}{A}$ . Therefore, we have

$$\begin{aligned}
& u_{\bar{u}}(\theta) - u_{\bar{u}}(\hat{\theta}) \\
&= u^*(\theta) - u^*(\hat{\theta}) \\
&= \int_{\hat{\theta}}^{\theta} \dot{u}^*(z) dz \\
&\leq \int_{\hat{\theta}}^{\theta} \left[ \dot{u}_{\bar{u}}(z) - \frac{\zeta(\theta - z)}{A} \right] dz \\
&= u_{\bar{u}}(\theta) - u_{\bar{u}}(\hat{\theta}) - \left( \frac{\zeta}{A} \right) \int_{\hat{\theta}}^{\theta} (\theta - z) dz \\
&= u_{\bar{u}}(\theta) - u_{\bar{u}}(\hat{\theta}) - \left( \frac{\zeta}{2A} \right) (\theta - \hat{\theta})^2,
\end{aligned}$$

which is absurd.

**Case 5:**  $u^*(1) = u_{\frac{1}{4}}(1)$ .

For all  $\theta \in (\frac{1}{2}, 1)$  the numerator on the RHS of (53) is  $1 + \frac{d}{d\theta} \{\theta - V_1(\theta, u)\}$ . Using (62), the last expression is  $1 + \frac{d}{d\theta} \left\{ \theta - \frac{1}{2 \cdot \frac{1}{4}} \cdot \theta - \left( 1 - \frac{1}{2 \cdot \frac{1}{4}} \right) \right\} = 0$ . Hence, from (53) we conclude that  $\Psi_1(u_{\frac{1}{4}}(\theta)_+, \theta) = \Psi_1(u_{\frac{1}{4}}(\theta)_-, \theta) = 0$ . Thus we can find a neighborhood of  $(\theta, u_{\frac{1}{4}}(\theta), \dot{u}_{\frac{1}{4}}(\theta))_{\theta \in [0, 1]}$  in which  $\Pi$  is (a.e.)  $C^1$ . Therefore, the argument in Case 1 applies, *mutatis mutandis*, to this case. This completes the proof of Proposition 6. Q.E.D.

## 7.4 Equilibrium Properties and Comparative Statics

**Proof of Proposition 7.** We only need to prove (iii). Take  $\theta$  and  $u$  such that the  $q(\theta, u) > 0$ . The profit obtained from this type as a function of  $u$  is  $\Psi(u, \theta) \cdot \hat{\pi}(u, \theta)$ , where

$$\begin{aligned}
& \hat{\pi}(u, \theta) \\
&\equiv \left[ \begin{array}{l} \theta \left( \frac{1}{2 \cdot u} \cdot \theta + \left( 1 - \frac{1}{2 \cdot u} \right) \right) - \frac{1}{2} \left( \frac{1}{2 \cdot u} \cdot \theta + \left( 1 - \frac{1}{2 \cdot u} \right) \right)^2 \\ - \left( \frac{1}{4 \cdot u} \cdot \theta^2 + \left( 1 - \frac{1}{2 \cdot u} \right) \cdot \theta + u + \frac{1}{4 \cdot u} - 1 \right) \end{array} \right].
\end{aligned} \tag{77}$$

Therefore, we have:

$$\begin{aligned}
& \frac{d}{du} \left[ \frac{\int_{\alpha(u)}^{\hat{\theta}} \hat{\pi}(u, \theta) d\theta}{\int_{\hat{\theta}}^1 \hat{\pi}(u, \theta) d\theta} \right] \\
&= \frac{d}{du} \left[ \frac{\int_{1-2u}^{\theta} \left( \begin{array}{l} \theta \left( \frac{1}{2 \cdot u} \cdot \theta + \left( 1 - \frac{1}{2 \cdot u} \right) \right) - \frac{1}{2} \left( \frac{1}{2 \cdot u} \cdot \theta + \left( 1 - \frac{1}{2 \cdot u} \right) \right)^2 \\ - \left( \frac{1}{4 \cdot u} \cdot \theta^2 + \left( 1 - \frac{1}{2 \cdot u} \right) \cdot \theta + u + \frac{1}{4 \cdot u} - 1 \right) \end{array} \right) d\theta}{\int_{\hat{\theta}}^1 \left( \begin{array}{l} \theta \left( \frac{1}{2 \cdot u} \cdot \theta + \left( 1 - \frac{1}{2 \cdot u} \right) \right) - \frac{1}{2} \left( \frac{1}{2 \cdot u} \cdot \theta + \left( 1 - \frac{1}{2 \cdot u} \right) \right)^2 \\ - \left( \frac{1}{4 \cdot u} \cdot \theta^2 + \left( 1 - \frac{1}{2 \cdot u} \right) \cdot \theta + u + \frac{1}{4 \cdot u} - 1 \right) \end{array} \right) d\theta} \right] \\
&= 48 \frac{u^2}{(1 - \theta)} \frac{4u^2 - (\theta - 1)^2}{(12u^2 - \theta^2 + 2\theta - 1)^2}
\end{aligned}$$

First, we claim that  $12u^2 - (\theta - 1)^2 > 0$  for all  $\theta \geq \alpha(u)$ . Indeed, since the expression is decreasing in  $\theta$  we have  $12u^2 - (\theta - 1)^2 \geq 12u^2 - (\alpha(u) - 1)^2 = 8u^2 > 0$ . Therefore, the denominator is always strictly positive. Second notice that  $4u^2 - (\theta - 1)^2$  is strictly decreasing in  $\theta$  and  $q(\theta, u) > 0$  implies  $\theta > \alpha(u) = 1 - 2u$ . Therefore  $4u^2 - (\theta - 1)^2 > 4u^2 - (\alpha(u) - 1)^2 = 0$ , which establishes that the term above is strictly positive. Q.E.D.

**Proof of Proposition 8.** The proof is analogous to the proof of Proposition 4 and is omitted for brevity. Q.E.D.