Nonlinear Pricing, Contract Variety, and Competition*

Huanxing Yang†   Lixin Ye‡

October 25, 2006

Abstract

This paper studies how increased competition affects nonlinear pricing, in particular the variety of contracts offered by firms. We present a model with both horizontally and vertically differentiated products, with the set of consumers served in the market being endogenously determined. Though firms are only able to sort consumers in the vertical dimension, horizontal differentiation affects screening in the vertical dimension. Using a two-phase optimal control technique, we characterize the symmetric equilibrium menu of contracts under different market structures. When the market structure moves from monopoly to duopoly, we show that each firm offers more contracts (serving more types of consumers) and quality distortions decrease. As the market structure becomes more competitive (when the number of firms increases further), the effect of increasing competition exhibits some non-monotonic features: when the initial competition is not too weak, a further increase in the number of firms will lead to more contracts being offered and a reduction in quality distortions; when the initial competition is weak, an increase in the number of firms will lead to fewer contracts being offered, though the effect on quality distortions is not uniform. Our predictions are largely consistent with some empirical studies.

Key words: Nonlinear pricing, product differentiation, contract variety, quality distortions

JEL: D40, D82, L10

*We are grateful to Bill Dupor, Eric Fisher, Howard Marvel, Eugenio Miravete, Massimo Morelli, Frank Page, James Peck, Rafael Rob, Lars Stole, Ching-Jen Sun, Frank Verboven, Fangyang Zheng, and the participants in the Economic Theory Workshop at Ohio State University, the Fall 2005 Midwest Economic Theory Meetings, and the 2006 Far-east Econometric Society Summer Meetings for very helpful comments and suggestions. The remaining errors are our own.

†Department of Economics, The Ohio State University, 405 Arps Hall, 1945 North High Street, Columbus, OH 43210-1172. Phone: (614)292-6523, E-mail: huanxing@econ.ohio-state.edu.

‡Department of Economics, The Ohio State University, 417 Arps Hall, 1945 North High Street, Columbus, OH 43210-1172. Phone: (614)292-6883, E-mail: lixinye@econ.ohio-state.edu.
1 Introduction

As more Japanese car makers enter the US market, will GM or Ford offer more models targeting at different types of consumers? As more competitors enter the cellular phone market, will Verizon or Sprint offer more calling plans? A number of empirical studies suggest that as competition becomes more intense, each firm often offers more variety of goods or services. For example, in the face of increased competition, American Express introduced 12 to 15 new credit cards per year targeted at different customer segments (Forbes, July 1, 1996, “The Battle of Credit Cards”). Similar effects of competition on the variety of contracts/services is observed in the airline industry as well. Borenstein and Rose (1994) find that on routes with more competition, each airline offers more variety of air tickets. In the automotive industry, in response to the increased competition from foreign companies in 1980s, GM and Ford began to offer more variety of car models.

All these cases suggest that increased competition leads to more contracts offered by each individual firm. Since the work of Mussa and Rosen (1978) and Maskin and Riley (1984) on monopolistic nonlinear pricing, there is a growing literature on nonlinear pricing in competitive settings, see, for example, Spulber (1989), Champsaur and Rochet (1989), Wilson (1993), Stole (1995), Villas-Boas and Schmidt-Mohr (1999), Armstrong and Vickers (2001), and Rochet and Stole (1997, 2002). However, much remains to be done in understanding how increased competition affects firms’ nonlinear pricing strategies. In this paper, we focus on the effects of increased (horizontal) competition on the number of (vertical) contracts offered by each firm.

Specifically, we consider a market with both vertically and horizontally differentiated products and consumers’ preferences differ in two dimensions. In the horizontal dimension, consumers have different tastes over different brands; while in the vertical dimension consumers have different marginal utilities over quality. Although neither type is observable to firms, in our model the single crossing property is only satisfied in the vertical dimension. As a result firms can only offer contracts to sort consumers with respect to their vertical types.¹

Such a framework is first analyzed by Rochet and Stole (1997, 2002). Their model is presumably more general since they consider general distributions while for tractability reason we focus on uniform distributions. In Rochet and Stole (2002), horizontal types are also interpreted as outside opportunity costs, which gives rise to consumers’ random participation. By taking random participation into account, they show that in the monopoly case there is either bunching or no quality

¹For this reason our paper does not belong to the multi-dimensional screening literature (e.g., Laffont, Maskin and Rochet, 1987; McAfee and McMillan, 1988; Armstrong, 1996; and Rochet and Chone, 1998).
distortion at the bottom. In the duopoly case, they show that under full market coverage quality distortions disappear and the equilibrium is characterized by the cost-plus-fee pricing feature (a similar result obtained in Armstrong and Vickers (2001)).\(^2\) Both results are in stark contrast to the received wisdom in nonlinear pricing literature (e.g., Mussa and Rosen, 1978).

It is worth noting that Rochet and Stole’s analysis focuses on the case where the lowest (vertical) type of consumers \((\bar{\theta})\) is always served in the market. In particular, all their main results are derived under the condition \(\theta/\bar{\theta} \geq 1/2\), where \(\bar{\theta}\) is the maximal vertical type. It is not clear, from their current analysis, whether their results will continue to hold if the above condition fails.

In this paper, we focus on the case where the lowest vertical type of consumers will typically be excluded from the market. More specifically, we assume that the vertical types of consumers are distributed uniformly over \([0, 1]\). This is a case not analyzed in Rochet and Stole, since the condition \(\theta/\bar{\theta} \geq 1/2\) is clearly violated. A direct consequence is that in our analysis, the minimal (vertical) type of consumers being served in the market is endogenously determined in equilibrium.

Surprisingly, our findings are quite different from those in Rochet and Stole. In all the cases we analyze, the equilibrium exhibits perfect sorting (bunching never occurs), and the quality distortion is maximal for the lowest type (we postpone a detailed discussion on our differences from Rochet and Stole to Section 4). In fact, our results are more in line with those obtained in Mussa and Rosen. More importantly, focusing on the case where the lowest type of consumers being served is endogenously determined allows us to study the effect of varying horizontal differentiation (competition) on the market coverage, and hence the variety of contracts offered, which is the main motivation of this paper. Our analysis in this paper is thus complementary to that in Rochet and Stole.

The key of our analysis comes from the interaction between horizontal differentiation (competition) and screening on the vertical dimension. Although horizontal differentiation does not have direct impact on the incentive compatibility (IC) conditions in the vertical dimension, it affects the IC conditions through the rent provisions to consumers.\(^3\) This interaction in turn affects the menu of contracts offered by each brand (firm). It is through this interaction that we identify the effect of increasing (horizontal) competition on the menu of contracts offered by each firm.

We focus on symmetric equilibria in which each firm (brand) offers the same menu of contracts.

---

\(^2\) Rochet and Stole (2002) focus on either competitive regime or monopoly regime (in terms of consumer coverage in horizontal dimension). The mixed regime with both regimes present is analyzed in Rochet and Stole (1997).

\(^3\) As is standard in the screening literature, any IC contract can be represented by a rent provision schedule, which governs the utilities of consumers in equilibrium.
Applying a two-phase optimal control technique, we characterize the symmetric equilibrium menu of contracts for both monopoly and duopoly. In either case, the equilibrium menu of contracts is unique, and a positive measure of consumers are excluded from the market. Moreover, the equilibrium contracts in both cases exhibit perfect sorting. Thus, the lowest type covered serves as a measure of the number of contracts offered by each firm (brand): the lower the lowest type covered, the more contracts are offered targeting more (vertical) types of consumers. As such in this paper we refer to the variety of the underlying types of consumers being served as the contract variety.\footnote{This is quite intuitive in the case with finite consumer types. In our current model with continuous types, this notion of contract variety can be interpreted as arising from the limiting case when the number of types goes to infinity.}

Compared to the monopoly benchmark, we show that under duopoly more contracts are offered, and quality distortions decrease. This result is due to the interaction between horizontal competition and vertical screening. Intuitively, the competition in duopoly increases the rent provisions for higher type consumers, which relaxes the screening condition in the vertical dimension (informational rent consideration becomes less important, as higher type consumers obtain higher rent anyway due to competition). This leads to additional contracts offered to consumers who were previously excluded, and a reduction in quality distortions.

Our comparison between monopoly and duopoly also has implications about which market structure offers more contracts over any given quality interval. We show that competition has the most effect on the higher end of the quality range: while equally dense contracts are offered over the lower quality range under both market structures, contracts offered over the higher quality range become denser moving from monopoly to duopoly.\footnote{If more types of consumers are served over a quality interval, we say that the contracts offered over that quality interval become denser.} Based on a very different model, Johnson and Myatt (2003) show that an incumbent may respond to entry by either expanding (fighting brand) or contracting (pruning) the product line (the quality range).\footnote{Johnson and Myatt (2003) consider quantity competition between an incumbent and an entrant in a market with vertically differentiated products. In Johnson and Myatt (2005), they extend similar analysis to the case with multiple firms.} In their model, introducing competition only has an effect on the lower end of the quality range, which is quite different from our implication.

We also study how the degree of horizontal differentiation, or the intensity of competition, affects the equilibrium menu of contracts. It turns out that the effects under the two market structures are quite different. Under monopoly, as the two brands become less differentiated, each brand offers fewer contracts, and quality distortions become larger. The effects in the duopoly case are subtle.
When the degree of horizontal differentiation (captured by transportation cost $k$) is smaller than some cutoff value, a decrease in $k$ results in more contracts offered and smaller quality distortions; when $k$ is larger than the cutoff value, a decrease in $k$ results in fewer contracts offered, and the effect on quality distortions is not uniform. Again these results are driven by the interplay between the horizontal differentiation and screening on the vertical dimension.

Finally, we extend our analysis of the duopoly model to any finite $n$-firm case, and demonstrate that the analysis can be translated into that of the duopoly model by proper normalization. We show that an increase in the number of firms is equivalent to a decrease in $k$ in the duopoly model. We thus conclude that when the initial competition level is not too low ($n$ is large), an increase in the number of firms results in more contracts offered by each firm and smaller quality distortions; while when the initial competition level is low ($n$ is small), an increase in the number of firms results in fewer contracts offered by each firm, though the effect on quality distortions is not uniform.

Our results are largely consistent with some existing empirical studies. In particular, the non-monotonic relationship between the number of firms and the number of contracts offered is consistent with a recent empirical study by Seim and Viard (2004), who test the effect of entry on the tariff choices (the number of contracts) of incumbent firms in the US cellular industry. We postpone a discussion of their study to the end of section 6.

As in our approach, a number of papers also study nonlinear pricing in competitive settings with both horizontally and vertically differentiated products (e.g., Gilbert and Matutes, 1993; Stole, 1995; Verboven, 1999; Villas-Boas and Schmidt-Mohr, 1999; and Ellison, 2005). However, these papers assume that all consumer types in the vertical dimension are served in the market. This full market coverage assumption does greatly simplify their analysis, but precludes the effect of competition on the consumer coverage on vertical dimension, which is central to our analysis. Except for Stole (1995), in all the other papers mentioned above firms are only able to produce two exogenously given qualities.\footnote{In Ellison (2005), there are only two types of vertically differentiated consumers. Stole (1995) assumes that either the consumers’ horizontal types or the consumers’ vertical types are observable. Thus the impacts of horizontal differentiation on screening in the vertical dimension in his model are very different from those in our model.}

The paper is organized as follows. Section 2 introduces the base model with two brands. Section 3 lays down necessary preliminaries for our analysis. Section 4 derives the optimal symmetric menu of contracts under monopoly. Section 5 characterizes the symmetric equilibrium in the duopoly model, and investigates how the equilibrium menu of contracts changes as the market structure moves from...
monopoly to duopoly. We extend our analysis to the arbitrary \( n \)-firm case in Section 6. Section 7 concludes.

2 The Model

Following Rochet and Stole, we consider a model in which consumers’ preferences differ both in the vertical and horizontal dimensions. Our basic model studies the two-brand case under both the duopoly and monopoly market structures. Under duopoly, two firms own two distinct brands, brand 1 and brand 2, respectively. Each firm (brand) offers a variety of vertically differentiated products, that is, goods of different qualities, which are indexed by \( q, q \in R_+ \). Quality \( q \) is both observable and contractible.

There are a continuum of consumers in the market, whose preferences differ on two dimensions: the “taste” dimension over the brands and the “quality” dimension. We model the taste dimension as the horizontal “location” of a consumer on a unit circle representing the ideal brand for that consumer.\(^8\) As depicted in Figure 1 below, the locations of brand 1 and brand 2 evenly split the unit circle. Let \( d_i \) be the distance between a consumer’s location and brand \( i \)'s location, then \( d_i \) is this consumer’s horizontal type, \( i = 1, 2 \). Because \( d_1 + d_2 = 1/2 \), either \( d_1 \) or \( d_2 \) alone fully captures a consumer’s preference over two brands.

![Figure 1: A Two-Brand Base Model](image)

\(^8\)For two brands, it would be sufficient to use a unit interval. We work with a unit circle since doing so will make it easier to extend our model to the arbitrary \( n \)-brand or \( n \)-firm case later.
Consumers’ varying preferences over the quality dimension are captured by \( \theta, \theta \in [0, 1] \), which we call a consumer’s vertical type. A consumer is thus characterized by a two-dimensional type \((d_i, \theta)\) (either \( i = 1 \) or \( i = 2 \)). Neither \( \theta \) or \( d_i \) is observable to either firm. We assume that consumers are uniformly located along the unit circle, and the vertical types of consumers at each location are distributed uniformly over the unit interval: \( \theta \sim U[0, 1] \). A consumer’s horizontal location and vertical type are independent.

Each consumer demands at most one unit of a good. If a type-\((d_i, \theta)\) consumer purchases one unit of the brand-\(i\) product with quality \(q\) at price \(t\), her utility is given by

\[
u(q,t,d_i,\theta) = \theta q - t - kd_i \tag{1}\]

where \(k, k > 0\), can be interpreted as the per unit “transportation” cost. Note that the smaller the \(k\), the less horizontally differentiated the two brands are. If a consumer purchases no product, her reservation utility is normalized to be 0.

We assume that the two brands (firms) have the same production technology. Specifically, to produce a unit of quality \(q\) product a firm incurs a cost \(c(q) = q^2/2\).\(^9\) Thus, each firm (brand) has a per-customer profit function given by

\[
\pi(t,q) = t - \frac{q^2}{2}. \tag{2}\]

Each firm offers a menu of contracts, which is a collection of all the quality and price pairs. Given the menus of contracts offered by both firms (brands), consumers decide whether to make a purchase, and if they do, which brand to choose and which contract to accept. It is well known that in the environment of competitive nonlinear pricing, it is no longer without loss of generality to restrict attention to direct contracts.\(^{10}\) To sidestep this problem, as in Rochet and Stole (2002) we restrict

\(^{9}\)The quadratic functional form assumed here is not crucial. What is needed in our analysis is that the cost function should be convex.

\(^{10}\)As demonstrated in a series of examples in Martimort and Stole (1997) and Peck (1997), equilibrium outcomes in indirect mechanisms may not be supported when sellers are restricted to using direct mechanisms where buyers report only their private types. Moreover, as demonstrated by Martimort and Stole (1997), an equilibrium in such direct mechanisms may not be robust to the possibility that sellers might deviate to more complicated mechanisms. The reason for such failures, as pointed out by McAfee (1993) and Katz (1991), is that in competition with nonlinear pricing the offers made by other firms may also be private information of the consumers when they make their purchase decisions, which means that this private information can also potentially be used when firms set up their revelation mechanisms.
attention to deterministic contracts.\textsuperscript{11} Since the preferences of a consumer with vertical type $\theta$ over the available price-quality pairs conditional on purchasing from a firm (brand) are independent of her horizontal type $d_i$, in what follows it is without loss of generality to consider direct contracts of the form $\{q(\theta), t(\theta)\}_{\theta \in [0,1]}$. For brevity of exposition, from now on we will often refer to vertical types as the types, especially when there is no confusion in the context.

Our solution concept is Bertrand-Nash equilibrium: given the other firm’s menu of contracts, each firm maximizes its expected total profit by choosing its menu of contracts.

This basically completes a description of the duopoly model. For the monopoly model, our main goal is to lay down a benchmark with which we can identify the effect of competition on the menu of contracts. As such in the monopoly model we need to control for all but the market structure. We thus assume that in the monopoly case, all the modeling elements are the same as in the duopoly model, except that the two brands are now owned and operated by the same firm, which is the monopolist.\textsuperscript{12} The monopolist’s objective is to maximize the joint profits from the two brands by choosing the menu of contracts for each brand.

3 Preliminaries

As an analytical benchmark, given (1) and (2), the first-best (efficient) quality provision is $q^*(\theta) = \theta$. We can thus define $\theta - q(\theta)$ as the quality distortion for type $\theta$ given quality schedule $q(\cdot)$.

Incentive Compatible Contracts

Let $U_i(\hat{\theta}, \theta, d_i)$ be the utility obtained by a consumer of type $(\theta, d_i)$ who reports $\hat{\theta}$ and purchases a unit of brand $i$ product. Then

$$U_i(\hat{\theta}, \theta, d_i) = \theta q_i(\hat{\theta}) - t_i(\hat{\theta}) - kd_i$$

Incentive compatibility requires

$$\forall(\theta, \hat{\theta}) \in [0, 1]^2, \quad U_i(\theta, \theta, d_i) \geq U_i(\hat{\theta}, \theta, d_i) \text{ for } i = 1, 2$$

\textsuperscript{11}See Rochet and Stole (2002) for a discussion on the restrictions resulted from focusing on deterministic contracts. More general approaches in restoring the “without loss of generality” implication of the revelation principle in the environment of competitive nonlinear pricing have been proposed and developed by, for example, Epstein and Peters (1999), Peters (2001), and Page and Monteiro (2003).

\textsuperscript{12}So our benchmark is a multi-product monopoly, which has an alternative interpretation as being collusive duopoly.
Since (3) satisfies the single crossing property in \((\theta, q_i)\), we can show the following “constraint simplification” lemma.

**Lemma 1**  The IC conditions (4) are satisfied if and only if the following two conditions hold:

1. \( U_i(\theta, \theta, d_i) = \int_{\theta_i^*}^{\theta} q_i(\tau) d\tau - k d_i \) for all \( \theta \geq \theta_i^* \) and \( i = 1, 2 \), where \( \theta_i^* \in [0, 1) \).
2. \( q_i(\theta) \) is increasing in \( \theta \)

where \( \theta_i^* \) is the lowest type that purchases from brand \( i \).

Lemma 1 is a standard result in the one-dimensional screening literature. This also applies to our model because consumers’ utility functions are separable in \( q \) and \( d_i \). Here \( \theta_i^* \) can be regarded as a separate choice variable for brand \( i \): any consumer whose type is below \( \theta_i^* \) is excluded from the market for brand \( i \). Alternatively, one can interpret that brand \( i \) offers a null contract \((q_i = 0 \text{ and } t_i = 0)\) to all consumers whose types are below \( \theta_i^* \). Define

\[
y_i(\theta) = \int_{\theta_i^*}^{\theta} q_i(\tau) d\tau, \quad i = 1, 2.
\]

Then by Lemma 1 \( y_i(\theta) \) is the rent provision to the type \((\theta, 0)\) consumer specified by the menu of IC contracts offered by brand \( i \). The equilibrium utility enjoyed by a consumer of type \((\theta, d_i)\) can now be written as \( y_i(\theta) - k d_i \). Moreover, the quality and the price specified in the original contract can be recovered from \( y_i(\theta) \) as follows:

\[
q_i(\theta) = y'_i(\theta) \quad \text{and} \quad t_i(\theta) = \theta q_i(\theta) - y_i(\theta).
\]

Thus any menu of IC contracts can be characterized by rent provision schedules \( (y_i(\cdot), i = 1, 2) \).
Note that by definition, \( y_i(\theta) \) is continuous in \( \theta \).

**Individual Rationality and Market Shares**

Given rent provision schedules \( \{y_i(\theta)\}, i = 1, 2 \), each consumer decides whether to make a purchase, and if they do, what product (brand and quality) to purchase. If a consumer of type \((\theta, d_i)\) chooses to purchase a product from brand \( i \), then we must have

\[
y_i(\theta) - k d_i \geq \max\{0, y_{-i}(\theta) - k(1/2 - d_i)\}
\]
Alternatively, we have

\[ d_i \leq \min \left\{ \frac{y_i(\theta)}{k}, \frac{1}{4} + \frac{1}{2k}(y_i(\theta) - y_{-i}(\theta)) \right\} \equiv s_i(\theta) \tag{6} \]

\(2s_i(\theta)\) is the total measure of type-\(\theta\) consumers who purchase brand \(i\) products. Figure 2 below illustrates one half of the market share for each brand (the other half not shown is symmetric).

![Figure 2: An Illustration of Market Shares and Market Coverage](image)

From Figure 2, we can see that there is a cutoff type \(\hat{\theta}\) above which the market is fully covered (consumers are served regardless of their horizontal locations), and below which the market is not fully covered. This is because \(y_i(\theta)\) is increasing in \(\theta\) by (5). Under duopoly, the full coverage range \([\hat{\theta}, 1]\) can also be called the competition range since the two firms are competing for customers over this range, and the partial coverage range \([\theta^*_i, \hat{\theta}]\) can also be called the local monopoly range. Note that \(\hat{\theta}\) is endogenously determined by the following condition:

\[ y_1(\hat{\theta}) + y_2(\hat{\theta}) = \frac{k}{2} \tag{7} \]

Given \(y_{-i}(\cdot)\), brand \(i\)'s total expected profit is twice the following:

\[
\int_{\theta^*_i}^{1} [t_i(\theta) - \frac{1}{2}q_i^2(\theta)]s_i(\theta)d\theta = \int_{\theta^*_i}^{1} \left[ \theta q_i(\theta) - y_i(\theta) - \frac{1}{2}q_i^2(\theta) \right] s_i(\theta)d\theta \tag{8}
\]

The maximization of (8) subject to (5) can be formulated as an optimal control problem, with \(q_i\) being the control variable, and \(y_i\) being the state variable. However, the standard optimal control theory requires that the integrand of the objective function be differentiable in the state variable.
(e.g., Kamien and Schwartz, 1992), which fails in our problem: $s_i(\theta)$, and hence the integrand in (8) is not differentiable with respect to $y_i$ at $\hat{\theta}$.

It turns out that we can get around this nondifferentiability problem by using a special optimal control technique. By separating the partial coverage range from the full coverage range, we can rewrite (8) into the sum of two integrations:

\[
\int_{\theta^*_i}^{\hat{\theta}} \left[ \theta q_i(\theta) - y_i(\theta) - \frac{1}{2}q_i^2(\theta) \right] \frac{y_i(\theta)}{k} d\theta \\
+ \int_{\theta}^{1} \left[ \theta q_i(\theta) - y_i(\theta) - \frac{1}{2}q_i^2(\theta) \right] \cdot \left[ \frac{1}{4} + \frac{1}{2k}(y_i(\theta) - y_{-i}(\theta)) \right] d\theta 
\]

(9)

The maximization of (9) subject to the transition equation $y_i'\theta(\theta) = q_i(\theta)$ and the corresponding endpoint conditions can be viewed as an optimal control problem with two potential phases. What makes it different from the ordinary single-phase optimal control is that now we also need to solve for the optimal switching “time” $\hat{\theta}$, at which the first phase switches to the second phase.

A closely related problem is analyzed by Amit (1986), who considers a petroleum recovery process that has two potential phases with different technologies. Unlike in Amit’s model where the state variable is free at the switching time, in our problem $y_i(\hat{\theta})$ is constrained to satisfy (7). As a consequence, the set of necessary conditions for optimality derived in Amit cannot be directly applied to our model. In the following subsection, we formulate a two-phase optimal control problem allowing for constraints on the switching point, and derive necessary conditions that can be directly applied to all our cases.

**A Two-Phase Optimal Control Procedure**

Consider a slightly more general problem below.

\[
\max \int_{\theta_0}^{\theta_1} F_1(\theta, y(\theta), q(\theta)) \ d\theta + \int_{\theta_1}^{\theta_2} F_2(\theta, y(\theta), q(\theta)) \ d\theta \\
subject \ to: \ y'(\theta) = \begin{cases} f_1(\theta, y(\theta), q(\theta)) & : \theta_0 \leq \theta < \theta_1 \\ f_2(\theta, y(\theta), q(\theta)) & : \theta_1 < \theta \leq \theta_2 \end{cases} \\
\theta_0 \ free, \ y(\theta_0) = y_0 \\
\theta_1 \ free, \ y(\theta_1) = R(\theta_1) \\
\theta_2 \ fixed, \ y(\theta_2) \ free. 
\]
where both $F_1(\cdot)$ and $F_2(\cdot)$ are continuously differentiable in $y$, $q$ and $\theta$, and $R(\theta)$ is continuous but may not be differentiable (i.e., $R(\theta)$ is a piecewise continuous function). Basically, the above program is separated into the two phases: phase I ($\theta \in [\theta_0, \theta_1)$) and phase II ($\theta \in (\theta_1, \theta_2]$). The key is to determine the optimal $\theta_1$ at which phase I switches to phase II, which is endogenously determined.

Let $\lambda(\theta)$ be the multiplier associated with the transition equations. Define the Hamiltonian function of the two phases as follows:

$$H = \begin{cases} 
H_1 = F_1 + \lambda f_1 & : \theta_0 \leq \theta < \theta_1 \\
H_2 = F_2 + \lambda f_2 & : \theta_1 < \theta \leq \theta_2
\end{cases}$$

**Lemma 2** The necessary conditions for the above problem are:

$$\frac{\partial H_1}{\partial q} = 0; \quad \frac{\partial H_1}{\partial y} = -\lambda'; \quad \theta_0 \leq \theta < \theta_1$$

$$\frac{\partial H_2}{\partial q} = 0; \quad \frac{\partial H_2}{\partial y} = -\lambda'; \quad \theta_1 < \theta \leq \theta_2$$

$$[H_1 - \lambda R^\prime_0](\theta_1^-) = [H_2 - \lambda R^\prime_0](\theta_1^+) \text{ if } \theta_0 < \theta_1 < \theta_2$$

$$[H_1 - \lambda R^\prime_0](\theta_1^-) \leq [H_2 - \lambda R^\prime_0](\theta_1^+) \text{ if } \theta_0 = \theta_1 < \theta_2$$

$$[H_1 - \lambda R^\prime_0](\theta_1^-) \geq [H_2 - \lambda R^\prime_0](\theta_1^+) \text{ if } \theta_0 < \theta_1 = \theta_2$$

$$H_1(\theta_0) = 0; \quad \lambda(\theta_2) = 0$$

**Proof.** See Appendix. ■

Note that the set of necessary conditions include the conditions governing the optimal switching “time” between the two phases. Let $\hat{H}_i = H_i - \lambda R^\prime_0$ be the pseudo-Hamiltonian modified from the original Hamiltonian (due to the restrictions on the switching “time” $\theta_1$). If there is a type $\theta_1$ at which the pseudo-Hamiltonians of the two phases are equal, then it is optimal to start with phase I and then switch to phase II at $\theta_1$. However, if no such $\theta_1$ can be found, then it is optimal to either skip phase I and start with phase II right away, or stay with phase I and never enter phase II. Lemma 2 can be of some independent interest, since it can be applied to similar two-phase optimal control problems where the state variable at the switching point has to satisfy some restrictions.

## 4 Monopoly

Under monopoly, the two brands are owned by a single firm. The monopolist’s objective is to maximize the joint profits from the two brands. Since consumers are uniformly distributed along the
horizontal dimension and the two brands’ production technologies are symmetric, we focus on the symmetric solution in which each brand offers the same menu of contracts and the resulting market shares are symmetric. We can thus drop the subscripts to write \( y_i(\theta) = y(\theta), i = 1, 2 \). Simplifying (6), the market share becomes \( s_i(\theta) = s(\theta) = \min \{ y(\theta)/k, 1/4 \} \).

The monopolist’s problem can be formulated as the following two-phase optimal control program:

\[
\begin{align*}
\max & \quad \int_{\theta^*}^{\hat{\theta}} \left[ \theta q(\theta) - y(\theta) - \frac{1}{2} q^2(\theta) \right] \frac{y(\theta)}{k} d\theta + \int_{\theta^*}^{1} \left[ \theta q(\theta) - y(\theta) - \frac{1}{2} q^2(\theta) \right] \frac{1}{4} d\theta \\
\text{s.t.} & \quad y'(\theta) = q(\theta) \\
& \quad q'(\theta) \geq 0 \\
& \quad y(\hat{\theta}) = k/4 \\
& \quad y(\theta^*) = 0
\end{align*}
\]

where \( \theta^* \) is the lowest type of consumers served, that is, \( y(\theta^*) = 0 \), and \( \hat{\theta} \) is the unique solution to \( y(\theta)/k = 1/4 \).

As is standard in the literature, we will solve the relaxed program by dropping the constraint \( q'(\theta) \geq 0 \) (the monotonicity of \( q(\theta) \) shall be verified later to justify this approach).

To apply Lemma 2, we define

\[
H = \begin{cases} 
H_1 = [\theta q - y - \frac{1}{2} q^2] \frac{y}{k} + \lambda q & : \theta^* \leq \theta < \hat{\theta} \\
H_2 = [\theta q - y - \frac{1}{2} q^2] \frac{1}{4} + \lambda q & : \hat{\theta} < \theta \leq 1
\end{cases}
\]

It can be verified that Phase I (partial coverage range) is characterized by the following differential equation.

\[
3y - \frac{1}{2} y' - yy'' = 0 \tag{10}
\]

Combining with the lower endpoint condition \( y(\theta^*) = 0 \), it can be verified that the unique solution to (10) is given by:

---

13 We focus on the symmetric solution here for ease of comparison with the duopoly case, where we will focus on symmetric equilibrium in which each firm offers the same menu of contracts. While a formal proof is not attempted here, we conjecture that symmetric solution is optimal for the monopolist.

14 If \( y(\theta^*) > 0 \), then for some sufficiently small \( \epsilon \), it can be verified that some type-(\( \theta^* - \epsilon \)) consumers would prefer accepting contract \( y(\theta^*) \) to staying out of the market, which contradicts the assumption that \( \theta^* \) is the lowest type being served.

15 The uniqueness is implied in Rochet and Stole (2002) (appendix, p. 304): if a convex solution to differential equation (10) exists for a given set of boundary conditions, it is unique.
\[ y(\theta) = \frac{3}{4}(\theta - \theta^*)^2, \quad q(\theta) = \frac{3}{2}(\theta - \theta^*) \]

Similarly, in phase II (full coverage range) we can obtain the differential equation \( y'' = 2 \). Combined with the transversality condition \( \lambda(1) = 0 \), the solution to this system is given by:

\[ y(\theta) = \theta^2 - \theta + \beta, \quad q(\theta) = 2\theta - 1 \]

where \( \beta \) is a parameter yet to be determined. Note that in both phases \( q(\theta) \) is increasing in \( \theta \). Thus the solutions to the relaxed program are also the solutions to the original program. Moreover, since in both phases \( q(\theta) \) is strictly increasing in \( \theta \), the optimal menu of contracts exhibits perfect sorting.

To determine \( \hat{\theta} \), note that

\[ y(\hat{\theta}^-) = \frac{3}{4}(\hat{\theta} - \theta^*)^2 = \frac{k}{4}, \quad \text{and} \quad y(\hat{\theta}^+) = \hat{\theta}^2 - \hat{\theta} + \beta = \frac{k}{4}. \]

In addition, we have a transversality condition about the optimal switching point \( \hat{\theta} \). Since \( R(\theta) = \frac{k}{4}, \quad R' = 0 \). Thus the transversality condition becomes \( H_1(\hat{\theta}) = H_2(\hat{\theta}) \). This translates into the condition that \( q(\cdot) \) is continuous at \( \hat{\theta} \), or equivalently, \( \frac{3}{2}(\hat{\theta} - \theta^*) = 2\hat{\theta} - 1 \). Given all these, we can solve \( \hat{\theta}, \theta^* \) and \( \beta \) as follows:

\[
\begin{align*}
\theta^* &= \frac{1}{2} - \frac{1}{12}\sqrt{3k} \\
\hat{\theta} &= \frac{1}{2} + \frac{1}{4}\sqrt{3k} \\
\beta &= \frac{1}{4} + \frac{1}{16}k
\end{align*}
\]

Proposition 1 In the monopoly model, the optimal symmetric menu of contracts is unique and exhibits perfect sorting. Specifically, for \( k \in (0, \frac{4}{3}) \),

\[
y(\theta) = \begin{cases} 
\frac{3}{4}(\theta - \theta^*)^2 & : \quad \theta^* \leq \theta \leq \hat{\theta} \\
\theta^2 - \theta + \frac{1}{4} + \frac{1}{16}k & : \quad \hat{\theta} < \theta \leq 1 
\end{cases}
\]

where \( \theta^* \) and \( \hat{\theta} \) are given by (11) and (12), respectively. For \( k \geq \frac{4}{3} \),

\[ y(\theta) = \frac{3}{4}\left(\theta - \frac{1}{3}\right)^2, \quad \theta \in \left[\frac{1}{3}, 1\right]. \]
The optimal menu of contracts exhibits several salient features. First, there is always a positive measure of types of consumers (regardless of horizontal location) who are excluded from the market ($\theta^* M > 0$). The underlying reason for the exclusion is the informational rent consideration. Offering contracts to all types may increase the firm’s profit from those types in $[0, \theta^* M)$. However, doing so necessarily increases the informational rent to all types above $\theta^* M$ due to the screening condition (5), which reduces the firm’s profit from those types. The optimal $\theta^* M$, which balances the above two opposing effects, should thus be strictly above zero. Second, there is quality distortion for all but the highest type consumers, i.e., $q(\theta) < \theta$ for all $\theta \in [\theta^* M, 1)$. This is again driven by the informational rent consideration. Finally, the optimal contracts exhibit perfect sorting. That is, different types of consumers choose different contracts. Our results thus imply that bunching does not occur and the quality provision for the lowest type covered is always distorted downwards. These are in stark contrast with the results obtained by Rochet and Stole (1997, 2002), who show that either bunching occurs at a lower interval, or perfect sorting occurs with efficient quality provision for the lowest type.

This sharp difference between our results and theirs first appears puzzling, given that the differential equation (10) is a special case of the Euler equation derived in Rochet and Stole (who allow for more general distributions). The key to solve the puzzle is to observe the difference in boundary conditions. Note that in Rochet and Stole the ratio of the lowest type to the highest type $\gamma = \theta/\Theta$ is assumed to be greater than $1/2$. This implies that all the (vertical) types are covered. As a result, the state variable $y$ is free at the lowest type $\theta$, which gives rise to the boundary condition $\lambda(\theta) = 0$. This boundary condition in turn implies efficient quality provision at $\theta$ if the monotonicity constraint on $q$ is satisfied (perfect sorting case). Note also that sorting can become quite costly for the monopolist given the requirement of no quality distortion at $\theta$, which explains why bunching may occur at a lower interval starting from $\theta$. On the other hand, in our model the lowest possible type $\theta$ is 0 ($\gamma = 0$), thus not all types will be covered and the lowest type covered, $\theta^*$, is endogenously determined. This leads to a different set of boundary conditions: $y(\theta^*) = 0$ and $H(\theta^*) = 0$. Combined with the differential equation (10), these conditions pin down a unique perfect sorting solution in which $q(\theta^*) = 0$.16

Thus in a sense our analysis is complementary to that in Rochet and Stole: while they study the case with full coverage of vertical types ($\gamma$ is big), we analyze the case with endogenously determined

---

16It can be easily verified that the quadratic functional form solution, which works in our case, does not satisfy the differential equation system in Rochet and Stole, simply because it violates their boundary conditions.
coverage of vertical types ($\gamma$ is small). It is worth noting that two cases lead to dramatically different results. To better understand the link between our results and those of Rochet and Stole, let’s fix the upper bound of the vertical type, $\theta$, and gradually raise $\theta$, starting from 0. When $\theta$ is 0, our results apply: there is an endogenously determined lowest type covered, $\theta^*$, with perfect sorting and $q(\theta^*) = 0$. This feature stays the same until $\theta$ is raised just above $\theta^*$. When $\theta$ is just above $\theta^*$, the case of Rochet and Stole applies since all the vertical types are covered. If the monotonicity constraint does not bind, the boundary condition at $\theta$ requires efficient quality provision at $\theta$. But continuity implies that the optimal solution should not change drastically at $\theta = \theta^*$. Thus the monotonicity must fail, leading to bunching at the lower end near $\theta$. Intuitively, when $\theta$ is slightly above $\theta^*$ ($\gamma$ is relatively small), efficient quality provision at $\theta$ is costly since it increases the informational rent for all higher types, the measure of which is big since $\gamma$ is relatively small. The optimality thus requires bunching. As $\theta$ is further raised close to $\theta$ ($\gamma$ becomes big enough), efficient quality provision at $\theta$ becomes less costly since there are fewer higher types. As a result, the monotonicity constraint is more likely to be satisfied even with efficient quality provision at $\theta$. Therefore, perfect sorting is more likely when $\gamma$ is big. This explains why in Rochet and Stole the solution involves perfect sorting when $\gamma$ is sufficiently large.

Given that the optimal contracts exhibit perfect sorting, there is a one-to-one mapping between the number of contracts offered and the (vertical) types of consumers served. Thus the number of contracts offered can be measured in terms of the (Lebesgue) measure of vertical types of consumers covered in the market. As a direct consequence, the lowest type served in the market, $\theta^{*M}$, becomes a measure for the variety of contracts offered. Specifically, as $\theta^{*M}$ decreases, more contracts are offered targeting more types of consumers. On the other hand, as $\theta^{*M}$ increases, fewer contracts are offered targeting fewer types of consumers. We are interested in how the degree of horizontal differentiation, which is parameterized by $k$, affects the variety of contracts offered. Equation (11) shows that for $k \in (0, 4/3)$, $\theta^{*M}$ is decreasing in $k$, and for $k \geq 4/3$, $\theta^{*M} = 1/3$ is independent of $k$. Thus when two brands become more horizontally differentiated (a bigger $k$), the monopolist offers more contracts. From the equilibrium quality schedules it can also be seen that quality distortions become smaller in Phase I but are unaffected in Phase II. We summarize these results in the following

---

17 This would imply that $\lim_{\theta \to \theta^+} q(\theta) = \theta^*$ while $\lim_{\theta \to \theta^-} q(\theta) = 0$.

18 If there is a fixed cost of offering a contract, then the number of contracts offered in equilibrium is positively correlated with the measure of vertical types being served. As this fixed cost goes to zero, the measure of contracts offered in equilibrium coincides with the measure of vertical types being served.
Proposition 2 In the monopoly model, when two brands become more horizontally differentiated, the monopolist offers more contracts, and quality distortions become smaller in the partial coverage range and remain unaffected in the full coverage range.

To understand the intuition of this result, we first need to understand the effects on profit of increasing the rent provision. Raising rent provisions (hence the total rent) to consumers has two effects. The first is to reduce the firm’s profitability per consumer (which can be termed as the marginal effect), and the second is to attract more consumers (which can be termed as the market share effect). Thus profit maximization requires an optimal balance between these two opposing effects. Note that with asymmetric information, the firm cannot freely vary the rent provision for certain types of consumers without affecting the rent provisions to other types. That is, rent provisions can only be adjusted subject to the screening condition, (5), which implies that changing the rent provision for some type will affect the rent provisions for all the types above. Hence the optimal rent provision schedule reflects an optimal trade-off between the marginal effect and market share effect subject to the screening condition.

In view of this insight, it is now straightforward to think through the intuition behind Proposition 2. As $k$ increases, by fixing the previous menu of contracts (holding $y(\cdot)$ fixed), $\hat{\theta}$ increases and $y(\theta)/k$ decreases, which implies that the market shares in both the full and partial coverage ranges shrink. To counter this effect, the monopolist has an incentive to increase $y(\theta)$ in an attempt to partially restore the loss of the market shares. By the screening condition (5), this can be achieved by either moving the schedule $q(\cdot)$ upward or pushing $\theta^* M$ downward, and both occur in equilibrium. Hence Proposition 2 is driven by an interaction between horizontal differentiation and screening in the vertical dimension, which occurs through the rent provision schedule $y(\theta)$.

5 Duopoly

In the duopoly model each firm’s objective is, given the other firm’s menu of contracts, to maximize its own profit by choosing a menu of contracts. Since both firms are symmetric in terms of their production technology and market positions, we focus on symmetric equilibrium, in which each firm offers the same menu of contracts, hence the same rent provision schedule $y^*(\theta)$, $\theta \in [\theta^D, 1]$ ($\theta^D$ is the lowest type that is served in the market). Formally, the pair $(y^*(\theta), y^*(\theta))$ constitutes a Bertrand-
Nash equilibrium if given \( y_{-i}(\theta) = y^*(\theta), \theta \in [\theta^*, 1] \), firm \( i \)'s best response is to choose \( y_i(\theta) = y^*(\theta), \theta \in [\theta^*, 1] \) as well.

Given the two firms’ rent provision schedules \( y_1(\theta) \) and \( y_2(\theta) \), the consumers’ type space is demarcated into two ranges: the competition range \((\theta > \hat{\theta})\), and the local monopoly range \((\theta < \hat{\theta})\). The switching point \( \hat{\theta} \) is determined by \( y_i(\hat{\theta}) = k/2 - y_{-i}(\hat{\theta}) \).

Suppose \( y_{-i}(\theta) = y^*(\theta), \theta \in [\theta^*, 1] \), then firm \( i \)'s relaxed program (by ignoring the monotonicity of \( q_i \)) is as follows:

\[
\max \int_{\theta_i}^{\hat{\theta}} \left[ \theta q_i(\theta) - y_i(\theta) - c(q_i(\theta)) \right] \frac{y_i(\theta)}{k} \, d\theta \\
+ \int_{\theta}^{1} \left[ \theta q_i(\theta) - y_i(\theta) - c(q_i(\theta)) \right] \cdot \left[ \frac{1}{4} + \frac{1}{2k} (y_i(\theta) - y^*(\theta)) \right] \, d\theta \\
s.t. \quad y_i'(\theta) = q_i(\theta) \\
y_i(\theta^*_i) = 0, \quad \theta^*_i \text{ free} \\
y_i(\hat{\theta}) = \frac{k}{2} - y^*(\hat{\theta}), \quad \hat{\theta} \text{ free} \\
y_i(1) \text{ free}
\]  

(13)

To apply Lemma 2, we define

\[
H = \begin{cases} 
H_1 = \left[ \theta q_i - y_i - \frac{1}{2} q_i^2 \right] \frac{y_i}{k} + \lambda q_i : \theta^*_i \leq \theta < \hat{\theta} \\
H_2 = \left[ \theta q_i - y_i - \frac{1}{2} q_i^2 \right] : \left[ \frac{1}{4} + \frac{1}{2k} (y_i(\theta) - y^*(\theta)) \right] + \lambda q_i : \hat{\theta} < \theta \leq 1
\end{cases}
\]

For phase I \((\theta < \hat{\theta})\), we can follow exactly the same steps as in the monopoly model to obtain

\[
y^*(\theta) = \frac{3}{4} (\theta - \theta^*)^2, \quad q^*(\theta) = \frac{3}{2} (\theta - \theta^*)
\]

For phase II \((\theta > \hat{\theta})\), the optimality condition and the co-state equation evaluated at \( y_i = y^* \) are given by

\[
0 = (\theta - q^*) \frac{1}{4} + \lambda \\
\lambda' = \frac{1}{4} - \frac{1}{2k} \left[ \theta q^* - y^* - \frac{1}{2} q^2 \right]
\]

After eliminating \( \lambda \) from the above equations we obtain the following differential equation:

\[
y^{*''} = 2 - \frac{2}{k} \left( \theta y^{*'} - y^* - \frac{1}{2} y^{*2} \right)
\]  

(14)
Letting $y_i = y_{-i} = y^*$, the switching point $\hat{\theta}$ is defined by $y^*(\hat{\theta}) = \frac{k}{4}$. To derive the transversality condition for $\hat{\theta}$, first note that $R(\hat{\theta}) = k/2 - y^*(\hat{\theta})$. Then

$$R'(\hat{\theta}^-) = -q^*(\hat{\theta}^-), \text{ and } R'(\hat{\theta}^+) = -q^*(\hat{\theta}^+)$$

After substituting and simplifying, the transversality condition in Lemma 2, $[H_1 - \lambda R'](\hat{\theta}^-) = [H_2 - \lambda R'](\hat{\theta}^+)$, can be reduced to $q^*(\hat{\theta}^-) = q^*(\hat{\theta}^+)$. By the continuity of $y^*(\cdot)$ at $\hat{\theta}$, we have $\hat{\theta} - \theta^* = \sqrt{k/3}$. From the Phase I solution, we can obtain $y^*(\hat{\theta}) = q^*(\hat{\theta}) = \sqrt{3k}/2$. Finally $\lambda(1) = 0$ implies that $y''(1) = q^*(1) = 1$.

Now the existence of a symmetric equilibrium boils down to the existence of a $\hat{\theta} \in (0, 1]$ and a convex function $y^*(\cdot)$ defined over $[\hat{\theta}, 1]$, which satisfy the following equations (we drop the superscripts to simplify notation):

$$\begin{align*}
    y'' &= 2 - \frac{2}{k} (\theta y' - y - \frac{1}{2} y^2) \\
    y(\hat{\theta}) &= k/4 \\
    y'(\hat{\theta}) &= \sqrt{3k}/2 \\
    y'(1) &= 1
\end{align*}$$

(15)

**Proposition 3** For $k \in (0, 4/3)$, there is a unique symmetric equilibrium in the duopoly model, which exhibits perfect sorting and is given by

$$y(\theta) = \begin{cases} 
    \frac{3}{4}(\theta - \theta^{*D})^2 & : \theta^{*D} \leq \theta \leq \hat{\theta}^D \\
    y^*(\theta) & : \hat{\theta}^D \leq \theta \leq 1 
\end{cases}$$

where $(\hat{\theta}^D, y^*(\theta))$ is the unique solution to the system (15), and $\theta^{*D} = \hat{\theta}^D - \sqrt{k/3}$. For $k \geq \frac{4}{3}$,

$$y(\theta) = \frac{3}{4} \left( \theta - \frac{1}{3} \right)^2, \theta \in \left[ \frac{1}{3}, 1 \right].$$

**Proof.** See Appendix. ■

In the proof we show that given $k \in (0, 4/3)$ the solution to the differential equation system (15) exists and is unique. Moreover, $y^*(\theta)$ is strictly convex. (15) is not a standard ordinary differential

---

19 Our analysis thus justifies the smooth pasting, which is directly applied by Rochet and Stole (1997) in their duopoly analysis of the “mixed regime” (where both the competition and local monopoly regimes are present).

20 We need $y''(\theta) \geq 0$ to ensure $q'(\theta) \geq 0$. 

equation (ODE) system partly due to the fact that the boundary conditions involve an endogenously determined endpoint ($\hat{\theta}$). Thus no existing ODE theorem can be directly applied to show the existence and uniqueness. The proof is somewhat tedious and hence relegated to the appendix. It is clear that system (15) has no closed-form solution. So the schedule of $y^*(\theta)$ can only be obtained from numerical computations.

As in the monopoly model, the equilibrium menu of contracts exhibits perfect sorting. Thus $\theta^{*D}$ also becomes a measure of the variety of contracts offered by each firm. Let $q_D(\cdot)$ and $q_M(\cdot)$ be the equilibrium quality provision schedules in the duopoly model and monopoly model, respectively. Despite the absence of the closed-form solution in the duopoly model, we are able to rank $\theta^{*D}$ and $\theta^{*M}$, and the schedules $q_D(\cdot)$ and $q_M(\cdot)$ unambiguously:

**Proposition 4** Given $k \in (0, \frac{4}{7})$, $\theta^{*D} < \theta^{*M}$, and $q_D(\theta) > q_M(\theta)$ for $\theta \in [\theta^{*D}, 1)$, which implies that compared to the monopoly benchmark, in duopoly equilibrium each firm offers more contracts, and quality distortions are smaller.

**Proof.** See Appendix. ■

Proposition 4 is shown by comparing the two differential equation systems under two market structures. Figure 3 is an illustration for the comparisons.

![Figure 3: Duopoly vs. Monopoly](image)

To see the intuition behind this comparison result, let’s start by assuming that in the duopoly case each firm offers the optimal symmetric menu of contracts as offered in the monopoly case. As a result the partial coverage and full coverage ranges are the same under both market structures.
Note that in the full coverage range (\( \theta \in [\hat{\theta}^M, 1] \)), the market share effect is absent under monopoly since the market is fully covered and the “competition” between the two brands is internalized by the monopolist; however, under duopoly the market share effect is present since each firm (brand) tries to steal the other firm’s market share. Thus the market share effect is stronger under duopoly, and each firm (brand) has an incentive to increase the rent provision. Therefore moving from monopoly to duopoly, \( \theta^*D < \theta^*M \), and \( q_D(\theta) > q_M(\theta) \) (by the screening condition (5)). Another way to see this is that competition under duopoly increases rent provisions to higher-type consumers (served in the full coverage range), which relaxes the screening condition in the vertical dimension: under duopoly firms would worry less about providing additional (informational) rent for the higher-type consumers, as the higher-type consumers are going to enjoy higher rent anyway due to competition. Consequently those consumers not served under monopoly may be served under duopoly, and quality distortions become smaller.

Our result has subtle implications about the distribution of consumers covered over the range of quality provisions. First note that in our analysis, the range of quality provisions is endogenously determined in equilibrium, which is from 0 to 1 in both the monopoly and duopoly cases. Our comparison indicates that moving from monopoly to duopoly denser contracts are offered (covering more consumer types) over the higher quality range. This is illustrated by figure 4 below, where the equilibrium schedules of quality provisions \( (q_M(\cdot) \) and \( q_D(\cdot) \)) are plotted against the consumer types under both market structures:

![Figure 4: Consumer Coverage over Quality Range](image)
By Proposition 4, duopoly induces smaller quality distortions, hence $q_D(\theta) > q_M(\theta)$ except at $\theta_i$. For any given range of quality provisions, we compare the difference in consumer coverages (in the vertical type dimension) under two market structures. This is equivalent to comparing the difference of the intervals on the $\theta$-axis projected from any quality interval. It is easily verified that $q_M(\hat{\theta}^M) = q_D(\hat{\theta}^D) = \sqrt{3k/2}$. Let $\hat{q} = \sqrt{3k/2}$. Thus, as depicted in figure 4, under both market structures the quality range is $[0, \hat{q}]$ in the partial coverage range and it is $[\hat{q}, 1]$ in the full coverage range.

As can be seen from the figure, the effects of changing market structure on the consumer coverage are different in these two ranges. In the range of $[0, \hat{q}]$, the quality provision schedules in both cases are parallel to each other. As a result, the total measure of consumer coverage (the length of the projected interval) remains the same, though the coverage shifts to the lower types as we move from monopoly to duopoly. In the range of $[\hat{q}, 1]$, the slope of the quality schedule is steeper in the monopoly. As a result, the projected interval of consumer coverage from any given quality interval becomes larger as we move from monopoly to duopoly: as illustrated, given any quality interval, say $[q_1, q_2]$, the projected intervals are $[\theta^M_1, \theta^M_2]$ in the monopoly case and $[\theta^D_1, \theta^D_2]$ in the duopoly case. Thus for this range, not only the composition of consumer coverage changes, the total measure of contracts offered also increases: competition leads to denser contracts offered for any quality interval over this range.

In a very different model, Johnson and Myatt (2003) show that an incumbent may respond to entry by either expanding (fighting brand) or contracting (pruning) the product line (the range of qualities). In their model, introducing competition only has an effect on the lower end of quality range, while in our model, moving from monopoly to duopoly unambiguously leads to denser contracts offered in the higher end of the quality range. This difference is obviously due to differences in modeling and assumptions between our approach and theirs. For example, they assume that the entrant cannot produce in some high quality range where the incumbent is able to produce. Thus as entry occurs the incumbent may respond by ceding the market of low quality products to the entrant while focusing on production in the high quality range (pruning). On the other hand, in our model firms are assumed to be technologically equal and the competition for higher type consumers is most intense. As a result, competition has the most effect over the higher end of the quality range.\footnote{This is the case in many market settings. For example, for car models in the relatively higher quality range, more features or options are offered for consumers to choose from, while for car models in the lower quality range, usually}
As in the monopoly case, we are also interested in how changes in \( k \) affect the number of contracts offered by each firm. For convenience of comparison, we show the schedules of both \( \theta^{*D} \) and \( \theta^{*M} \) against \( k \) in Figure 5 below, where the schedule of \( \theta^{*D} \) is plotted from numerical computation.

![Figure 5: Comparison of Participation Thresholds](image)

As can be seen from the figure, \( \theta^{*M} \) is always decreasing as \( k \) increases. But for the duopoly model, there is a cutoff \( k^* \) such that for \( k \in (0,k^*) \), \( \theta^{*D} \) is increasing in \( k \); and for \( k \in (k^*,4/3) \), \( \theta^{*D} \) is decreasing in \( k \) (for \( k \geq 4/3 \), \( \theta^{*D} = \theta^{*M} = 1/3 \) is independent of \( k \)). Our computation shows that the turning point \( k^* \) is approximately .91. Note that the decreasing trend of \( \theta^{*D} \) in the range of \((k^*,4/3)\) is not quantitatively significant; in this range of \( k \), \( \theta^{*D} \) is in the range of \([0.33,0.35]\). On the other hand, the increasing trend of \( \theta^{*D} \) in the range of \((0,k^*)\) is quantitatively significant; when \( k = k^* \), \( \theta^{*D} \) equals to 0.35, while as \( k \) converges to 0, \( \theta^{*D} \) converges to 0 as well. The following comparative statics result is obtained from numerical computations:22

**Proposition 5** In the duopoly case, when \( k \in (0,k^*) \), as \( k \) decreases each firm offers more contracts, and quality distortions become smaller; when \( k \in (k^*,4/3) \), as \( k \) decreases each firm offers fewer contracts, and the effect on quality distortions is not uniform: there is a cutoff type, say \( \tilde{\theta} \), such that when \( \theta \in [0, \tilde{\theta}) \), quality distortions become bigger, while when \( \theta \in (\tilde{\theta},1) \), quality distortions become smaller; when \( k \geq 4/3 \), both firms are local monopolists hence \( k \) does not affect the variety only standard features or options are available.

---

22 The MATLAB code for all the computations in this paper is available upon request.
Thus the effects of changing $k$ on $\theta^*$ and quality distortions in the duopoly case are dramatically different from those in the monopoly benchmark. Again the intuitions spelled out previously continue to help, with the details being a bit more subtle. Under duopoly, a lower $k$ not only implies less horizontal differentiation, but also implies more fierce competition between two firms.

A decrease in $k$ while holding $y(\cdot)$ fixed leads to an increase in the market share in Phase I (the local monopoly range). Following the intuition suggested for Proposition 2, each firm would then have incentive to decrease the rent provision in this range, which can be achieved by raising $\theta^*$ or lowering $q(\cdot)$. However the effect on Phase II (the competition range) is different. As $k$ decreases, the competition becomes more intense. As a result, the impact of the market share effect on firms’ profit becomes relatively more important than that of the marginal effect on firms’ profit (which is further reinforced by a decrease in $\tilde{\theta}$), therefore each firm would have incentive to raise rent provisions, which can be achieved by lowering $\theta^*$ or raising $q(\cdot)$. So the effects on $\theta^*$ and $q(\cdot)$ of decreasing $k$ in two phases work in exactly the opposite directions. The net effect depends on which effect dominates.\textsuperscript{23}

When $k \in (0,k^*)$, i.e., when the initial competition between two firms is not too weak, the competition range is more important relative to the local monopoly range,\textsuperscript{24} thus the effect in the competition range dominates and each firm offers more contracts and quality distortions reduce in equilibrium. On the other hand, when $k \in (k^*,4/3)$, i.e., when the initial competition between two firms is weak, the local monopoly range is relatively more important,\textsuperscript{25} thus the effect in the local monopoly range dominates and each firm offers fewer contracts, though the effect on quality distortions is not uniform: as $k$ decreases, there is a cutoff type, say $\tilde{\theta}$, such that when $\theta \in [0, \tilde{\theta})$, $q(\cdot)$ moves downward, while when $\theta \in (\tilde{\theta}, 1)$, $q(\cdot)$ moves slightly upward. This non-uniform effect actually makes perfect sense. When $k \in (k^*,4/3)$, the competition is weak so the movement of the quality schedule should follow the pattern in the monopoly case. This explains why as $k$ decreases the quality schedule in lower type range moves downward while the schedule in higher type range remains almost unchanged – recall that in the monopoly case, as $k$ decreases the schedule $q(\cdot)$ in the partial coverage range moves downward, while it stays the same in the full coverage range.

\textsuperscript{23}In terms of the rent provision schedule $y(\cdot)$, a decrease in $k$ tends to increase $y(\theta)$ in the competition range and decrease $y(\theta)$ in the local monopoly range. But $y(\cdot)$ has to be continuous at the junction of two ranges to satisfy the IC constraint.

\textsuperscript{24}In the limit as $k \to 0$, the local monopoly range disappears.

\textsuperscript{25}As $k \geq 4/3$, the competition range disappears and both firms behave as if they were local monopolists.
Again our computations show that the effect of changing $k$ on either $\theta^*D$ or quality distortions over the range $k > k^*$ is not quantitatively significant. However, it is qualitatively important as it provides a “continuity” for our intuitions to work when moving from monopoly to duopoly.

From Figure 5, it is apparent that $\theta^*M - \theta^*D$ is decreasing in $k$. Thus one potentially testable empirical implication is that the impact of market structure on the variety of contracts offered depends on the degree of horizontal differentiation: the smaller the $k$, the bigger the impact of introducing competition on the variety of contracts offered.

6 Extension to the $n$-Firm Model

In this section we extend our analysis to any arbitrary finite $n$-firm case. Specifically, in the horizontal dimension there are $n$ brands owned and operated by $n$ distinct firms ($n \geq 2$), the locations of which evenly split the unit circle; and each firm offers vertically differentiated products. Each firm’s objective is to maximize the profit from its own brand, given other firms’ menus of contracts. Again we look for symmetric Bertrand-Nash equilibria in which each firm offers the same menu of contracts.\footnote{As a direct consequence each firm is effectively competing with two adjacent firms.}

An $n$-tuple $(y^*(\theta),...,y^*(\theta))$ constitutes a symmetric equilibrium if, given that all other firms offer $y^*(\theta)$ for $\theta \in [\theta^*,1]$, each firm’s best response is also to choose $y_i(\theta) = y^*(\theta)$, $\theta \in [\theta^*,1]$.

Given that all firms other than $i$ offer the schedule $y^*(\theta)$, $\theta \in [\theta^*,1]$, it can be easily verified that firm $i$’s relaxed program (by ignoring the constraint of the monotonicity of $q_i(\cdot)$) is as follows:

$$\max \quad \int_{\theta^*}^{\hat{\theta}} [\theta q_i(\theta) - y_i(\theta) - c(q_i(\theta))] \frac{y_i(\theta)}{k} d\theta$$

$$+ \int_{\hat{\theta}}^{1} [\theta q_i(\theta) - y_i(\theta) - c(q_i(\theta))] \cdot \left[ \frac{1}{2n} + \frac{1}{2k}(y_i(\theta) - y^*(\theta)) \right] d\theta$$

s.t. $y_i'(\theta) = q_i(\theta)$

$y_i(\theta^*) = 0$, $\theta^*$ free

$y_i(\hat{\theta}) = \frac{k}{n} - y^*(\hat{\theta})$, $\hat{\theta}$ free

$y_i(1)$ free

Applying Lemma 2, firm $i$’s optimal rent provision $y^*(\theta)$ in the local monopoly range ($\theta < \hat{\theta}$) is the same as that in the duopoly model which is independent of $n$. The optimal rent provision in the competition range ($\theta > \hat{\theta}$) and the optimal switching point $\hat{\theta}$ are characterized by the following system:
$$\begin{align*}
y'' &= 2 - \frac{n}{k}(\theta y' - y - \frac{1}{2}y^2) \\
y(\hat{\theta}) &= k/2n \\
y'(\hat{\theta}) &= \sqrt{3k/2n} \\
y'(1) &= 1
\end{align*}$$ (16)

If we define $k' = k/n$ as the normalized degree of horizontal differentiation, then by inspection, in terms of $k'$ the differential equation system (16) is exactly the same as the differential equation system (15) in the duopoly case (where $k' = k/2$). This implies that the analysis of the $n$-firm case can be translated into the analysis of the duopoly case through normalizing $k$ by $n$, and in terms of $k'$ the solution to the $n$-firm model is the same as the solution to the duopoly model. Thus all the results from the duopoly model carry over to the $n$-firm competitive model. In particular, the $n$-firm competitive model has a unique symmetric equilibrium, and such equilibrium exhibits perfect sorting, hence the participation threshold $\theta^*$ becomes a measure for the variety of contracts offered by each firm.\(^{27}\) Moreover, the effect of an increase in $n$ (while holding $k$ fixed) on the equilibrium is exactly the same as the effect of a decrease in $k$ on the duopoly equilibrium. To re-state the results in the duopoly case in terms of $k'$, let’s define $k^* = k/2 \approx .455$. Then as $k'$ increases, for $k' < k^*$, $\theta^*$ increases and $q(\cdot)$ decreases, for $k^* < k' < 2/3$, $\theta^*$ decreases while $q(\cdot)$ increases for lower types but decreases for higher types, and for $k' \geq 2/3$, both $\theta^*$ and $q(\cdot)$ are independent of $k'$. Translating this into $n$-firm case, we have the following result:

**Proposition 6** Fix $k > 0$ and define $n^* = k/k'$. When $n > n^*$, an increase in $n$ leads to more contracts offered by each firm and smaller quality distortions; when $n \in (1.5k, n^*)$, an increase in $n$ leads to fewer contracts offered by each firm, and larger quality distortions for lower types and smaller quality distortions for higher types; when $n \leq 1.5k$, each firm is a local monopolist, hence the contract variety and quality distortions are independent of $n$.

Proposition 6 thus implies that the effect of increasing competition on contract variety or quality distortions depends on the initial state of competition, and that effect is not monotonic. This result is consistent with a recent empirical study of cellular phone markets by Seim and Viard (2004), which we briefly discuss below.

\(^{27}\)In Gal-Or’s (1983) quantity-setting model, symmetric Cournot equilibria may exist when the number of firms is small, but may fail to exist as the number of firms becomes larger. In contrast, in our model the symmetric Bertrand-Nash equilibrium always exists and is unique.
Nonlinear pricing is a standard practice for cellular firms to sort consumers with respect to their different usage. Most wireless service is sold under “three-part tariffs”: monthly fee, peak minutes, and off-peak minutes. Seim and Viard (2004) study the effect of entry on the tariff choices (the number of contracts) of incumbent cellular firms. Before 1996, most geographic cellular market areas (CMA’s) had a duopoly market structure, with two firms operating the wireless service in each CMA. After the FCC auctioned off the PCS spectrum, PCS entrants began to enter the cellular markets. Due to some exogenous reasons, by 1998 there were significant variations in the amount of entry by PCS providers across cellular markets.28

Utilizing this heterogeneity across different cellular markets, Seim and Viard test the relationship between the number of calling plans offered by the incumbent providers and the number of entrants. They found that, generally speaking, incumbents introduce more calling plans in markets with more entrants. Moreover, they found that the relationship is not monotonic: if only one entrant enters, the incumbent duopolists actually reduce the number of calling plans; if more than one entrant enters, the number of calling plans offered by the incumbents increases as the number of entrants increases. They also show that this relationship cannot be explained by demographic heterogeneity or cost differences across markets. This non-monotonic relationship is consistent with the predictions in our Proposition 6.

Our two-brand monopoly can be extended to n-brand multi-product monopoly by a similar normalization. Thus Proposition 2 can be extended to imply that as a monopolist offers more brands, each brand offers less contracts. So for a multi-product monopolist, the horizontal brand variety and the vertical contract variety are substitutes. This can be viewed as another testable implication from our analysis.

7 Conclusion

To our knowledge, this paper is the first to study the effects of the horizontal differentiation (competition) on contract variety. Specifically, using a two-phase optimal control technique, we characterize the unique symmetric optimal menu of contracts in the monopoly benchmark and the unique symmetric equilibrium menu of contracts in the duopoly model. We show that when moving from monopoly to duopoly each brand offers more variety of contracts targeting more types of consumers, and the

---

28 This was due to two factors. First, some licenses were undeveloped because the winning bidders went bankrupt. Second, different cellular markets require different amounts of time to build a sufficiently large network of wireless infrastructure.
quality distortions become smaller. We then extend our analysis to the arbitrary $n$-firm case, and show that the major insights obtained from the base model continue to hold. In particular, we show that as long as the competition among firms is not too weak, further increasing competition (i.e., increasing the number of firms) leads to more variety of contracts offered by each firm and smaller quality distortions.

Our results have empirical relevance regarding how competition affects the variety of goods, services, or contracts offered by firms. The predictions of our model are largely consistent with some existing empirical evidence.

One restriction in this paper is that we assume a uniform distribution of consumers’ types. While we maintain this assumption for ease of analysis, we believe that it is not crucial for our main results to hold. The reason is as follows. If we work with some other distributions instead of the uniform distribution, we may end up with partial pooling in equilibrium. However, in that case Lemma 1 and hence the screening condition (5) still hold.\textsuperscript{29} Thus the same insight regarding the interplay between horizontal differentiation and screening in the vertical dimension continues to apply: for example, in the case that the initial competition is not too weak, as competition increases, the IC constraint relaxes and we conjecture that the range of partial pooling shrinks and the participation threshold moves downward, which in turn implies that more contracts will be offered targeting more consumer types. A rigorous analysis is needed to confirm this insight, which is left for future research.

\textsuperscript{29}Note that the equilibrium rent provision formula (5) holds as long as $q'(\cdot) \geq 0$, which encompasses the case of partial pooling.
Appendix

Proof of Lemma 2:

The idea is to work with a comparison path by constructing the first variation of the functional. First let’s fix the notation:

$q^*$: the optimal control function over $\theta_0 \leq \theta \leq \theta_2$ with $\theta_1$ being the switching point moving from $F_1$ to $F_2$.

$y^*$: the corresponding optimal state variable.

$J^*$: the value achieved when evaluated along $q^*, y^*$.

$\delta \theta_0$ and $\delta \theta_1$: small changes in $\theta_0$ and $\theta_1$, respectively.

$q$: a feasible control over $\theta_0 + \delta \theta_0 \leq \theta \leq \theta_2$ with switching point at $\theta_1 + \delta \theta_1$.

$\delta q$: $q - q^*$.

$y$: the feasible state which corresponds to $q$.

$J$: the value when evaluated along the comparison path.

Let $\lambda(\theta)$ be the multiplier associated with the state variable. Then

$$J - J^* = \int_{\theta_0 + \delta \theta_0}^{\theta_1 + \delta \theta_1} F_1(\theta, y, q) \, d\theta + \int_{\theta_1 + \delta \theta_1}^{\theta_2} F_2(\theta, y, q) \, d\theta - \left[ \int_{\theta_0}^{\theta_1} F_1(\theta, y^*, q^*) \, d\theta + \int_{\theta_1}^{\theta_2} F_2(\theta, y^*, q^*) \, d\theta \right]$$

$$= \int_{\theta_0 + \delta \theta_0}^{\theta_1 + \delta \theta_1} [F_1(\theta, y, q) + \lambda f_1(\theta, y, q) - \lambda y'] \, d\theta + \int_{\theta_1 + \delta \theta_1}^{\theta_2} [F_2(\theta, y, q) + \lambda f_2(\theta, y, q) - \lambda y'] \, d\theta$$

$$- \left[ \int_{\theta_0}^{\theta_1} [F_1(\theta, y^*, q^*) + \lambda f_1(\theta, y^*, q^*) - \lambda y'] \, d\theta + \int_{\theta_1}^{\theta_2} [F_2(\theta, y^*, q^*) + \lambda f_2(\theta, y^*, q^*) - \lambda y'] \, d\theta \right].$$

Let $h(\theta) = y(\theta) - y^*(\theta)$, and define $\delta J$ as the first variation of $J - J^*$. Following the derivations paralleling those in Amit (the proof of Theorem 1), we have

$$\delta J = \lambda(\theta_0) h(\theta_0) - F_1^*(\theta_0) \delta \theta_0 + F_1^*(\theta_1) \delta \theta_1 - F_2^*(\theta_1) \delta \theta_1$$

$$\quad + \int_{\theta_0}^{\theta_1} \left[ \left( \frac{\partial F_1^*}{\partial y} + \lambda \frac{\partial f_1^*}{\partial y} + \lambda' \right) h + \left( \frac{\partial F_1^*}{\partial q} + \lambda \frac{\partial f_1^*}{\partial q} \right) \delta q \right] \, d\theta$$

$$\quad + \int_{\theta_1}^{\theta_2} \left[ \left( \frac{\partial F_2^*}{\partial y} + \lambda \frac{\partial f_2^*}{\partial y} + \lambda' \right) h + \left( \frac{\partial F_2^*}{\partial q} + \lambda \frac{\partial f_2^*}{\partial q} \right) \delta q \right] \, d\theta$$

$$- \lambda(\theta_1^-) h(\theta_1^-) - \lambda(\theta_2) h(\theta_2).$$

(17)

Define $\delta y_0 = y(\theta_0 + \delta \theta_0) - y^*(\theta_0)$, $\delta y_1 = y(\theta_1 + \delta \theta_1) - y^*(\theta_1)$, and $\delta y_2 = y(\theta_2 + \delta \theta_2) - y^*(\theta_2)$. By
approximation, we have
\[ \delta y_0 \approx y^*(\theta_0)\delta \theta_0 + h(\theta_0) \]
\[ \delta y_1 \approx y^*(\theta^+_1)\delta \theta_1 + h(\theta^+_1) \text{ if } \delta \theta_1 < 0 \]
\[ \delta y_1 \approx y^*(\theta^+_1)\delta \theta_1 + h(\theta^+_1) \text{ if } \delta \theta_1 > 0. \]

Therefore,
\[ h(\theta_0) \approx \delta y_0 - y^*(\theta_0)\delta \theta_0 = -y^*(\theta_0)\delta \theta_0 \]
\[ h(\theta^+_1) \approx \delta y_1 - y^*(\theta^+_1)\delta \theta_1 = [R'(\theta^+_1) - y^*(\theta^+_1)]\delta \theta_1 \]
\[ h(\theta^+_1) \approx \delta y_1 - y^*(\theta^+_1)\delta \theta_1 = [R'(\theta^+_1) - y^*(\theta^+_1)]\delta \theta_1 \]
\[ h(\theta_2) = y(\theta_2) - y^*(\theta_2) = \delta y_2. \]

Substituting the above conditions into (17), we have
\[ \delta J = -[F^*_1(\theta_0) + \lambda(\theta_0)y^*(\theta_0)]\delta \theta_0 \]
\[ + \int_{\theta_0}^{\theta_1} \left[ \left( \frac{\partial F^*_1}{\partial y} + \lambda \frac{\partial f^*_1}{\partial y} + \lambda' \right) h + \left( \frac{\partial F^*_1}{\partial q} + \lambda \frac{\partial f^*_1}{\partial q} \right) \delta q \right] d\theta \]
\[ + \int_{\theta_1}^{\theta_2} \left[ \left( \frac{\partial F^*_2}{\partial y} + \lambda \frac{\partial f^*_2}{\partial y} + \lambda' \right) h + \left( \frac{\partial F^*_2}{\partial q} + \lambda \frac{\partial f^*_2}{\partial q} \right) \delta q \right] d\theta \]
\[ + [(F^*_1(\theta_1) + \lambda(\theta_1)y^*(\theta_1) - \lambda(\theta_1)R'(\theta^-_1)) \]
\[ - (F^*_2(\theta_1) + \lambda(\theta^+_1)y^*(\theta^+_1) - \lambda(\theta^+_1)R'(\theta^+_1))]\delta \theta_1 - \lambda(\theta_2)\delta y_2. \tag{18} \]

Let
\[ -\lambda' = \frac{\partial F^*_1}{\partial y} + \lambda \frac{\partial f^*_1}{\partial y} = \frac{\partial H_1}{\partial y} \text{ for } \theta_0 \leq \theta < \theta_1 \]
\[ -\lambda' = \frac{\partial F^*_2}{\partial y} + \lambda \frac{\partial f^*_2}{\partial y} = \frac{\partial H_2}{\partial y} \text{ for } \theta_1 \leq \theta \leq \theta_2. \]

It is possible that \( \delta \theta_0 = \delta \theta_1 = \delta y_2 = 0. \) Then (18) reduces to
\[ \delta J = \int_{\theta_0}^{\theta_1} \left( \frac{\partial F^*_1}{\partial q} + \lambda \frac{\partial f^*_1}{\partial q} \right) \delta q d\theta + \int_{\theta_1}^{\theta_2} \left( \frac{\partial F^*_2}{\partial q} + \lambda \frac{\partial f^*_2}{\partial q} \right) \delta q d\theta. \]

From this we conclude that for optimality, the following conditions must hold:
\[ \frac{\partial F^*_1}{\partial q} + \lambda \frac{\partial f^*_1}{\partial q} = \frac{\partial H_1}{\partial q} = 0, \text{ for } \theta_0 \leq \theta \leq \theta \]
\[ \frac{\partial F^*_2}{\partial q} + \lambda \frac{\partial f^*_2}{\partial q} = \frac{\partial H_2}{\partial q} = 0, \text{ for } \theta_1 \leq \theta \leq \theta_2. \]
Substituting these two equations into (18), we have
\[
\delta J = -[F_1^*(\theta_0) + \lambda(\theta_0)y(\theta_0)]\delta \theta_0 - \lambda(\theta_2)\delta y_2
\]
\[
+[(F_1^*(\theta_1) + \lambda(\theta_1^-)y^-(\theta_1^-) - \lambda(\theta_1^-)R^-(\theta_1^-)) - (F_2^*(\theta_1) + \lambda(\theta_1^+)y^+(\theta_1^+) - \lambda(\theta_1^+)R^+(\theta_1^+))]\delta \theta_1.
\]
Since \(\delta \theta_0, \delta y_2\) are not only independent but also freely variable, we have
\[
F_1^*(\theta_0) + \lambda(\theta_0)y(\theta_0) = H_1(\theta_0) = 0, \text{ and } \lambda(\theta_2) = 0.
\]
(18) now becomes
\[
\delta J = [(F_1^*(\theta_1) + \lambda(\theta_1^-)y^-(\theta_1^-) - \lambda(\theta_1^-)R^-(\theta_1^-))
\]
\[-(F_2^*(\theta_1) + \lambda(\theta_1^+)y^+(\theta_1^+) - \lambda(\theta_1^+)R^+(\theta_1^+))]\delta \theta_1
\]
\[
= [(H_1(\theta_1) - \lambda(\theta_1^-)R^-(\theta_1^-)) - (H_2(\theta_1) - \lambda(\theta_1^+)R^+(\theta_1^+))]\delta \theta_1. \tag{19}
\]
If \(\delta \theta_1\) is freely variable, which occurs when the optimal solution involves \(\theta_0 < \theta_1 < \theta_2\), then nonpositivity of (19) is ensured only if
\[
H_1(\theta_1) - \lambda(\theta_1^-)R^-(\theta_1^-) = H_2(\theta_1) - \lambda(\theta_1^+)R^+(\theta_1^+).
\]
If feasible modifications are \(\delta \theta_1 \geq 0\), which occurs when the optimal solution involves \(\theta_0 = \theta_1 < \theta_2\), then nonpositivity of (19) is ensured only if
\[
H_1(\theta_1) - \lambda(\theta_1^-)R^-(\theta_1^-) \leq H_2(\theta_1) - \lambda(\theta_1^+)R^+(\theta_1^+).
\]
If feasible modifications are \(\delta \theta_1 \leq 0\), which occurs when the optimal solution involves \(\theta_0 < \theta_1 = \theta_2\), then nonpositivity of (19) is ensured only if
\[
H_1(\theta_1) - \lambda(\theta_1^-)R^-(\theta_1^-) \geq H_2(\theta_1) - \lambda(\theta_1^+)R^+(\theta_1^+).
\]
We have thus derived all the necessary conditions in Lemma 2. ■

Proof of Proposition 3:
Following the derivations preceding to the Proposition, the proof will be completed by showing that \(\forall k \in (0, 4/3]\), there is a unique \(\hat{\theta} \in (0, 1]\) and a unique \(y(\theta)\) defined over \([\hat{\theta}, 1]\) satisfying the differential equation system (15). Moreover, the solution of \(y(\theta)\) is strictly convex.
First letting \(z(\theta) = y(\theta) - \frac{1}{2}\theta^2\), we have
\[
z''(\theta) = 1 + \frac{1}{k}(z^2(\theta) + 2z(\theta)). \tag{20}
\]
Let \( z'(\theta) = v(z(\theta)) \), then \( z''(\theta) = v'(z)z'(\theta) = vv'(z) \). (20) thus becomes:

\[
v \frac{dv}{dz} = 1 + \frac{1}{k}(v^2 + 2z). \tag{21}
\]

Substituting \( w(z) = v^2(z) \) into (21), we have \( w' - 2w/k = 2 + 4z/k \), which leads to

\[
w(z) = ce^{2z/k} - 2z - 2k.
\]

where \( c \) is a parameter to be determined by the boundary conditions.

The system (15) can now be written in terms of function \( z(\theta) \) as follows:

\[
(z'(\theta))^2 = ce^{2z(\theta)/k} - 2z(\theta) - 2k
\]

\[
z(\hat{\theta}) = \frac{k}{4} - \frac{1}{2} \hat{\theta}^2 := \hat{z}
\]

\[
z'(\hat{\theta}) = \frac{\sqrt{3k}}{2} - \hat{\theta}
\]

\[
z'(1) = 0.
\]  

(22)

Define \( \alpha \) such that \( c = k\alpha e^{-2z/k} \), and \( \delta \) such that \( \hat{\theta} = \frac{\sqrt{3k}}{2} \delta \) \((\hat{\theta} \in (0, 1] \text{ implies } \delta \in (0, 2/\sqrt{3k}))\).

Also define \( u(\theta) = 2(z(\theta) - \hat{z})/k \).

Then we have

\[
u'^2 = \frac{4}{k^2} z'^2 = \frac{4}{k^2}(k\alpha e^u - 2z - 2k) = \frac{4}{k} \left( \alpha e^u - u - \frac{2}{k} \hat{z} - 2 \right).
\]

Letting \( f(u) = \alpha(e^u - 1) - u + \beta \), where \( \beta = \alpha - \frac{2}{k} \hat{z} - 2 \), then \( u'^2 = 4f(u)/k \).

At \( \hat{\theta} \), \( u(\hat{\theta}) = 0 \), \( u'(\hat{\theta}) = \sqrt{3/k}(1 - \delta) \), hence

\[
\beta = \frac{k}{4} u'^2(\hat{\theta}) = \frac{3}{4}(1 - \delta)^2, \quad \alpha = \beta + 2 \hat{z} + 2 = \frac{13}{4} - \frac{3}{2} \delta.
\]

The system (22) can now be rewritten as follows:

\[
u'^2 = \frac{4}{k} f(u) \tag{23}
\]

\[
u(\hat{\theta}) = 0 := \hat{u} \tag{24}
\]

\[
u'(\hat{\theta}) = \sqrt{3/k}(1 - \delta) := \hat{u}' \tag{25}
\]

\[
u'(1) = 0 := u'_1 \tag{26}
\]

where

\[
f(u) = \alpha(e^u - 1) - u + \beta = \left( \frac{13}{4} - \frac{3}{2} \delta \right) (e^u - 1) - u + \frac{3}{4}(1 - \delta)^2. \tag{27}
\]
For notational convenience let \( u_1 = u(1) \). Then \( u'_1 = 0 \Rightarrow f(u_1) = 0 \).

First, from (25)-(26) it can be verified that \( \hat{\theta} = 1 \Rightarrow k = 4/3 \). So for \( k \in (0, 4/3) \) we must have \( \hat{\theta} < 1 \), or \( \delta < 2/\sqrt{3k} \).

The rest of the proof is completed in 6 steps:

1. Show that (23) implies \( u' = -\frac{2}{\sqrt{k}} \sqrt{f(u)} \) and \( \delta \geq 1 \).

Suppose not, then \( u' = \frac{2}{\sqrt{k}} \sqrt{f(u)} \geq 0 \). By (25) \( \delta \leq 1 \), and \( \alpha \geq 7/4 \), which implies \( f'(u) = \alpha e^u - 1 \geq 7/4 - 1 > 0 \) for all \( u \geq 0 \). But then \( f(u_1) > f(\hat{u}) = f(0) = \beta \geq 0 \), a contradiction. Therefore we must have \( u' = -\frac{2}{\sqrt{k}} \sqrt{f(u)} \leq 0 \) and hence \( \delta \geq 1 \). Since \( u \) is decreasing, we have \( u_1 \leq \hat{u} = 0 \). It can be verified that for \( k \in (0, 4/3) \), \( u_1 = \hat{u} = 0 \) is impossible.\(^{30} \)

Hence \( u_1 < \hat{u} = 0 \) for \( k \in (0, 4/3) \), and \( f(u) \geq 0 \) on \([u_1, 0] \).

2. Show that in the solution to system (23)-(26), \( \alpha > 0 \), which implies that the original solution \( y(\cdot) \) is strictly convex.

Suppose not, i.e., suppose \( \alpha \leq 0 \). Then \( f'(u) = \alpha e^u - 1 < 0 \), which implies that \( f(\hat{u}) < f(u_1) = 0 \). But \( f(\hat{u}) = \alpha(e^{\hat{u}} - 1) - \hat{u} + \beta = \beta \geq 0 \), contradiction. So \( \alpha > 0 \).

Since 
\[
y'' = 1 + z'' = \frac{1}{k} e^{z/k} = \alpha e^{(z-\hat{z})/k},
\]
\( \alpha > 0 \) (or \( \delta < 13/6 \)) implies that the original solution \( y(\cdot) \) must be strictly convex.

3. Show that given \( \delta \) (or \( \hat{\theta} \)), the solution of \( u(\cdot) \) (and hence \( y(\cdot) \)) exists and is unique.

Since \( f''(u) = \alpha e^u > 0 \), \( f \) is strictly convex (with \( f(\pm \infty) = \infty \)). Hence \( f(u) > 0 \) on \((u_1, 0]\).

\( f'(u) = \alpha e^u - 1 = 0 \Rightarrow u_{\min} = -\ln \alpha \).

Let \( A(\delta) =: \min f(u) = f(-\ln \alpha) = \ln \alpha + \frac{3}{2}(\delta^2 - 2) \). Since \( f(u_1) = 0 \), we must have \( A(\delta) \leq 0 \).

We next show that \( A(\delta) < 0 \). Suppose not, then \( u_1 = u_{\min} = -\ln \alpha < 0 \), which implies \( f(u) \approx a(u - u_1)^2 \) near \( u_1 \), where \( a \) is a positive real number. \( \frac{du}{\sqrt{f(u)}} = -\frac{2}{\sqrt{k}} d\theta \) implies that
\[
\int_{u_1}^{0} \frac{du}{\sqrt{f(u)}} = -\frac{2}{\sqrt{k}} \int_{1}^{\hat{\theta}} d\theta = \frac{2}{\sqrt{k}}(1 - \hat{\theta}) < \infty. \tag{28}
\]

\(^{30} u_1 = \hat{u} \Rightarrow u = 0 \), which implies \( z = \hat{z} \) and \( \hat{\theta} = \frac{\sqrt{\pi}}{2} \). Therefore \( y(\theta) = \frac{1}{2} \theta^2 + \hat{z} = \frac{1}{2} \theta^2 + \frac{k}{4} - \frac{1}{2} \hat{\theta}^2 = \frac{1}{2} \theta^2 - \frac{k}{8} \). But then \( y(\theta) \) does not satisfy the differentiation equation in system (15), a contradiction.

33
But on the other hand,
\[
\int_{u_1}^{0} \frac{du}{\sqrt{f(u)}} = \frac{1}{\sqrt{a}} \int_{u_1}^{\hat{u}} \frac{du}{u - u_1} = \infty,
\]
a contradiction.

Therefore \(A(\delta) < 0\) and hence in the neighborhood of \(u_1\), \(f(u) = O(u - u_1)\).

Define
\[
\Phi(u) = \int_{0}^{u} \frac{dv}{\sqrt{f(v)}} = \int_{\hat{\theta}}^{\theta} - \frac{2}{\sqrt{k}} ds = - \frac{2}{\sqrt{k}} (\theta - \hat{\theta}).
\]

Note that \(\Phi(u)\) is well defined for any \(u \in [u_1, 0]\), as \(f(u) = O(u - u_1)\) near \(u_1\) (which implies \(|\int_{0}^{u_1} \frac{dv}{\sqrt{f(v)}}| < \infty\)).

Since \(\Phi(u)\) is a strictly increasing function over \([u_1, 0]\), inverting we have
\[
u(\theta) = \Phi^{-1} \left( \frac{2}{\sqrt{k}} (\theta - \hat{\theta}) \right) \text{ for } \theta \in [\hat{\theta}, 1].
\]

Thus given \(\hat{\theta}, u(\cdot)\) (and hence \(y(\cdot)\)) is uniquely determined by (29). It remains to show that \(\hat{\theta}\) (or \(\delta\)) exists and is unique.

4. Show that in the solution \(\delta \in [1, \min\{\delta_0, 2/\sqrt{3k}\}]\) (where \(\delta_0\) is defined below).

Since \(-\ln \alpha = u_{\min} < u_1 < 0\), we have \(\alpha > 1\) or \(\delta < \frac{3}{2}\). We thus have \(\delta \in [1, \frac{3}{2}]\) (from step 1).

It is straightforward to verify that \(A(\delta)\) is strictly increasing over the interval \([1, \frac{3}{2}]\) and there is a unique \(\delta_0 \in [1, \frac{3}{2}]\) such that \(A(\delta_0) = 0\). Since \(A(\delta) < 0\), we thus have \(\delta \in [1, \delta_0]\). Combining this with \(\delta < \frac{3}{2}\), in the solution to the system (23)-(26) we must have \(\delta \in [1, \min\{\delta_0, 2/\sqrt{3k}\}]\).

By (28) we have \(\int_{u_1}^{0} \frac{du}{\sqrt{f(u)}} = \frac{2}{\sqrt{k}} (1 - \hat{\theta}) = \frac{2}{\sqrt{k}} - \sqrt{3}\delta\).

Define
\[
\xi(\delta) = \sqrt{3}\delta + \int_{u_1}^{0} \frac{du}{\sqrt{f(u)}}.
\]

5. Show that given any \(k \in (0, \frac{4}{3})\), there is a \(\delta \in (1, \min\{\delta_0, 2/\sqrt{3k}\})\) satisfying \(\xi(\delta) = \frac{2}{\sqrt{k}}\).

First \(f' - \beta + u + \alpha = \alpha e^u\) implies \((f - \beta + u + \alpha)' = f - \beta + u + \alpha\). That is, \(f' - f - u = \text{constant} = f'(u_1) - f(u_1) - u_1 = f'(u_1) - u_1\).

Hence \(f' - f - (u - u_1) = f'(u_1) > 0\), which leads to
\[
f'(u_1) \int_{u_1}^{0} \frac{1}{\sqrt{f}} du = \int_{u_1}^{0} \frac{f'}{\sqrt{f}} du - \int_{u_1}^{0} \sqrt{f} du - \int_{u_1}^{0} \frac{u - u_1}{\sqrt{f}} du.
\]
Define \( \xi_1(\delta) = \int_{u_1}^{0} \sqrt{f} \, du \), and \( \xi_2(\delta) = \int_{u_1}^{0} \frac{u - u_1}{\sqrt{f}} \, du \). Note that \( f'(u_1) = \alpha e^{u_1} - 1 > 0 \), and \( \int_{u_1}^{0} \frac{f'}{\sqrt{f}} \, du = 2\sqrt{f(0)} = 2\sqrt{3} = \sqrt{3}(\delta - 1) \). Therefore by (31) we have

\[
\xi(\delta) = \frac{1}{\alpha e^{u_1} - 1} \left[ \sqrt{3}(\delta - 1) - \xi_1(\delta) - \xi_2(\delta) \right] + \sqrt{3}\delta. \tag{32}
\]

Since \( u_1(\delta) \) is continuous in \( \delta \), both \( \xi_1(\delta) \) and \( \xi_2(\delta) \) are also continuous in \( \delta \). Therefore, \( \xi(\delta) \) is continuous in \( \delta \).

First, consider \( \delta \to 1^+ \). It is easily verified that \( \beta \to 0^+, \alpha \to (\frac{7}{4})^- \). Hence \( f(u) \to g(u) = \frac{2}{7}(e^u - 1) - u \), and \( u_1(\delta) \to 0^- \).

By (32), \( \xi(\delta) < \frac{1}{\alpha e^{u_1} - 1} \sqrt{3}(\delta - 1) + \sqrt{3} \to \sqrt{3} \). Since \( \sqrt{3} < 2/\sqrt{k} \), we have \( \xi(\delta) < 2/\sqrt{k} \) for \( \delta \) sufficiently close to \( 1^+ \).

Second, consider \( \delta \to b = \min\{\delta_0, 2/\sqrt{3k}\} \) from the left. We discuss the following two cases:

Case 1: \( \delta_0 > 2/\sqrt{3k} \). Then when \( \delta \to b^- = (2/\sqrt{3k})^- \), \( \xi(\delta) > \sqrt{3}\delta = \sqrt{3}b = 2/\sqrt{k} \) (the inequality is due to (30)).

Case 2: \( \delta_0 \leq 2/\sqrt{3k} \). For \( \delta \to b^- = \delta_0^+ \), \( A(\delta) = f_{\min} \to 0^- \) (since \( A(\delta_0) = 0 \)). So \( u_1 \to (-\ln \alpha)^+ \), and by (30), \( \xi(\delta) \to \infty \). So when \( \delta \to b^- \), \( \xi(\delta) > 2/\sqrt{k} \).

By the mean-value theorem, there exists \( \delta \in (1, \min\{\delta_0, 2/\sqrt{3k}\}) \) such that \( \xi(\delta) = \frac{2}{\sqrt{k}} \).

6. Show that the solution from step 5 is unique.

We have

\[
f(u_1) = \alpha(e^{u_1} - 1) - u_1 + \beta = 0. \tag{33}
\]

Differentiating (33) with respect to \( \delta \), we have

\[
-\frac{3}{2}(e^{u_1} - 1) + \frac{3}{2}(\delta - 1) + (\alpha e^{u_1} - 1)u'_1 = 0.
\]

which gives

\[
u'_1 = \frac{3}{2}(e^{u_1} - \delta) = \frac{1}{\eta} \cdot \frac{3}{2}(e^{u_1} - \delta),
\]

where \( \eta = \alpha e^{u_1} - 1 > 0 \).

By (32),

\[
\xi'_1 = \frac{d\xi_1}{d\delta} = \frac{d\xi_1}{du_1} \frac{du_1}{d\delta} = 0 \cdot \frac{du_1}{d\delta} = 0
\]

\[
\xi'_2 = \frac{d\xi_2}{du_1} u'_1 = \left( 0 + \int_{u_1}^{0} \frac{-1}{\sqrt{f}} \, du \right) u'_1 = -(\xi - \sqrt{3}\delta)u'_1.
\]

35
So
\[
\xi' = \sqrt{3} + \frac{1}{\eta} (\sqrt{3} - \xi_1 - \xi_2) - \frac{1}{\eta} \left( \alpha e^{u_1} u'_1 - \frac{3}{2} e^{u_1} \right) [\sqrt{3}(\delta - 1) - \xi_1 - \xi_2]
\]
\[
= \sqrt{3} + \frac{1}{\eta} \left[ \sqrt{3} + (\xi_1 - \sqrt{3}\delta) u'_1 \right] + \frac{e^{u_1} (\frac{3}{2} - \alpha u'_1)}{\eta} (\xi - \sqrt{3}\delta)
\]
\[
= \sqrt{3} + \frac{1}{\eta} \sqrt{3} + \frac{1}{\eta} (\xi - \sqrt{3}\delta) \left[ u'_1 (1 - \alpha e^{u_1}) + \frac{3}{2} e^{u_1} \right]
\]
\[
= \sqrt{3} + \frac{1}{\eta} \sqrt{3} + \frac{1}{\eta} (\xi - \sqrt{3}\delta) \frac{3}{2} \delta
\]
\[
> 0 \text{ (since } \xi - \sqrt{3}\delta = \int_{u_1}^0 \frac{du}{\sqrt{f(u)}} > 0). \]

Therefore \(\xi(\delta)\) is strictly increasing in \(\delta \in (1, \min\{\delta_0, 2/\sqrt{3}k\})\), which implies that there is a unique \(\delta\) satisfying \(\xi(\delta) = \frac{2}{\sqrt{k}}\). \(\blacksquare\)

**Proof of Proposition 4:**

Suppose \(\theta^*D \leq \theta^*M\). Since \(\hat{\theta}^D - \theta^*D = \hat{\theta}^M - \theta^*M = \sqrt{k/3}\), \(\hat{\theta}^M \leq \hat{\theta}^D\). By the quality provision schedules in the partial coverage range we have

\[
q_M(\hat{\theta}^M) = q_D(\hat{\theta}^D) = \sqrt{3k}/2
\]

From the quality provision schedule in the full coverage range under monopoly,

\[
q'_M(\theta) = 2 > 0 \text{ for } \theta \in [\hat{\theta}^M, \hat{\theta}^D]
\]
\[
\Rightarrow q_M(\hat{\theta}^D) \geq q_D(\hat{\theta}^D)
\]

From (14),

\[
q'_D(\theta) = 2 - \frac{2}{k} \left[ \theta y' - y - \frac{1}{2} y'^2 \right]
\]

In equilibrium, a firm’s profit from a type \(\theta\) consumer is positive for \(\theta > \theta^*D\), i.e. \(\theta y' - y - \frac{1}{2} y'^2 > 0\) hence

\[
q'_M(\theta) = 2 > q'_D(\theta) \text{ for } \theta \in [\hat{\theta}^D, 1]
\]

Note that \(k \in (0, \frac{3}{4})\) implies \(\hat{\theta}^D < 1\). Combining this with (34) and (35), we have \(q_M(1) > q_D(1)\), which contradicts the fact that \(q_M(1) = q_D(1) = 1\). Therefore \(\theta^*M > \theta^*D\) in equilibrium.

To show that \(q_D(\theta) > q_M(\theta)\), we consider the following cases:
For $\theta \in [\theta^*, D]$, by the quality provision schedules in the partial coverage range, we have $q_D(\theta) > q_M(\theta)$ as $\theta^* < \theta^*$. For $\theta \in (\hat{\theta}^D, \check{\theta}^M]$, $q_D(\theta) > q_D(\hat{\theta}^D) = \sqrt{3k}/2$, and $q_M(\theta) \leq q_M(\hat{\theta}^M) = \sqrt{3k}/2$. Hence $q_D(\theta) > q_M(\theta)$.

For $\theta \in (\hat{\theta}^M, 1)$, $q_D'(\theta) < 2 = q_M'(\theta)$ and $q_M(1) = q_D(1)$ implies that $q_D(\theta) > q_M(\theta)$.

To sum up, $q_D(\theta) > q_M(\theta)$, hence $\theta - q_D(\theta) < \theta - q_M(\theta)$ for $\theta \in [\theta^*, 1)$, which implies that quality distortion is smaller in the duopoly case. ■
References


