All-Units Discount, Quantity Forcing, and Capacity Constraint*

Yong Chao[†] and Guofu Tan[‡]

September, 2013

Abstract

An all-units discount (AUD) is a pricing scheme that lowers a buyer's marginal price on *every* unit purchased when the buyer's purchase exceeds or is equal to a pre-specified threshold. The AUD and its variations are commonly used in both final-goods and intermediate-goods markets. The usual antitrust concern about the AUD and its variations is their potential foreclosure effects when adopted by a dominant firm to compete against a small rival. The existing literature has so far focused on interpreting the AUD as a price discrimination tool, investment incentive program, or rent-shifting instrument.

In this article, we investigate strategic effects of volume threshold based pricing schemes used by a dominant firm in the presence of a smaller, capacity-constrained rival. In particular, we consider a three-stage game in which the dominant firm and its rival make price offers to a buyer sequentially before the buyer purchases. We show that the AUD adopted by a dominant firm leads to a partial foreclosure of a capacity-constrained competitor (and full foreclosure is likely, too, if there are fixed costs) in the sense that the small rival is under-supplied strictly below its capacity and its profit is reduced. This result holds even when the rival has a lower marginal cost. When the rival's capacity level is in the range of low values, the buyer is worse off under the AUD as compared to linear pricing. The intuition for our findings is that, due to the limited capacity of the rival, the dominant firm has a "captive" portion of the buyer's demand and is able to use the AUD to leverage its market power on the "captive" portion to the "contestable" portion of the demand, much like the tied-in selling strategy in the context of multiple products.

We compare the AUD with a simple scheme called quantity-forcing (QF), which specifies a single quantity and the corresponding payment. We find that, in equilibrium, when the rival's capacity level is in the range of low values, AUD and QF have the same foreclosure effect; however, when the rival's capacity is in the range of high values, the QF has an additional, softening competition effect.

---

*We thank Luis Cabral, Juan Carrillo, Yongmin Chen, Zhiqi Chen, Harrison Cheng, Dmitry Lubensky, Massimo Motta, and seminar participants at the University of Southern California, the University of Louisville, Koç University, Zhejiang University, Shandong University, 2012 Southern California Symposium on Network Economics and Game Theory, the 9th Workshop on Industrial Organization and Management Strategy, Annual Conference of Mannheim Centre for Competition and Innovation, the 11th Annual International Industrial Organization Conference, 2013 North American Summer Meeting of the Econometric Society, and the Workshop in Industrial Economics, Academia Sinica. The usual caveat applies.

[†]Department of Economics, College of Business, University of Louisville, Louisville, KY 40292.

[‡]Department of Economics, University of Southern California, Los Angeles, CA 90089.

1 Introduction

An all-units discount (AUD) is a pricing scheme that lowers a buyer's marginal price on *every* unit purchased when the buyer's purchase exceeds or is equal to a pre-specified volume threshold. The AUD and its variations are commonly used in both final-goods and intermediate-goods markets, and its adoption by dominant firms has become a prominent antitrust issue. In the recent Intel case, the so-called "first-dollar rebate," in which Intel offered its customers a retroactive rebate if their purchase of microprocessors from Intel exceeded a pre-specified target level, has been challenged.[1] The European Commission has found the AUD adopted by dominant firms to be anticompetitive in several cases, including Hoffmann-La Roche,[2] Michelin I,[3] Michelin II,[4] British Airways,[5] and Tomra.[6]

In all these antitrust cases, the dominant firm holds market power over part of the buyer's demand, which is "captive" by the dominant firm.[7] On the other hand, there is a "contestable" part of the buyer's demand for which the dominant firm faces competition. The major concern about the AUD and its variations is their potential foreclosure effects on the "contestable" portion of the market. Intuitively, a larger firm may take advantage of its "captive" portion of the demand so to induce the buyer to purchase a significant portion of her requirements. This may cause small rivals to be even smaller by limiting their growth possibilities. Such a logic has been pointed out in all the above cases, as well as by the European Commission.[8] However, to the best of our knowledge, it has not been formalized in economic theory yet. In other words, we are still unclear about how the AUD can foreclose small rivals when it is adopted by a dominant firm, although intuition may suggest so.

Here we propose a model to formalize the foreclosure idea and examine the mechanism through which the AUD can impact competition when a dominant firm has a "captive" demand. In reality, the existence of a "captive" market perhaps depends on a variety of factors, such as brand loyalty, product differentiation, switching cost, or capacity constraints faced by small rivals. Regardless of where the "captive" demand comes from, the essence is that the small rival cannot compete for the entire demand of the buyer. To capture this fact, we consider a case where the small rival is capacity-constrained, as this is an intuitive way of giving rise to the "captive" portion of the demand.

In particular, we investigate strategic effects of volume threshold based pricing schemes used by a dom-

[1] *AMD* v. *Intel* (2005); Case COMP/C-3/37.990—–Intel (2009); Docket No. 9341, In the Matter of Intel Corporation (2010). "In general, the rebate schemes operate as follows: quarterly, Intel unilaterally establishes for each of its customers a target level of purchases of Intel microprocessors. If the customer achieves the target, it is entitled to a rebate on all of the quarter's purchases of all microprocessors—–back to the very first one—–generally in the neighborhood of 8-10% of the price paid." (Paragraph 59, *AMD* v. *Intel* Complaint 2005)

[2] See Case 85/76, *Hoffmann-La Roche & Co. AG* v. *Commission of the European Communities*, Judgment of the Court of 13 February 1979.

[3] Case 322/81, *NV Nederlandsche Banden Industrie Michelin* v. *Commission of the European Communities*, Judgment of the Court of 9 November 1983.

[4] Case T-203/01, *Manufacture Française des Pneumatiques Michelin* v. *Commission of the European Communities supported by Bandag Inc.*, Judgment of the Court of First Instance of 30 September 2003. See Motta (2009)[17] for discussions of this case.

[5] *British Airways plc* v. *Commission of the European Communities supported by Virgin Atlantic Airways Ltd.*, C-95/04, Judgment of the European Court of Justice, March 2007.

[6] C-549/10 P, *Tomra Systems and Others* v. *Commission of the European Communities*, Judgment of the Court of 19 April 2012.

[7] For example, "(a)t least in the short run, most if not all of the major OEMs must engage significantly with Intel because AMD is too small to service all their needs." (Paragraph 63, *AMD* v. *Intel* Complaint 2005)

[8] See European Commission (2005[8], 2009[9]).

inant firm in the presence of a smaller, capacity-constrained rival. We show that the dominant firm can use an AUD scheme to limit its rival's supply strictly below its capacity level, and, as a result, the dominant firm gains at the expense of its rival. This result holds even when the rival has a lower marginal cost. Thus, the AUD may lead to a partial foreclosure of a more efficient, capacity-constrained competitor (full foreclosure is likely, too, if there are fixed costs) in the sense that the small firm is under-supplied strictly below its capacity, and its profit is reduced. When the rival's capacity level is in the range of low values, the buyer is worse off under the AUD as compared to linear pricing (LP).

Our analysis suggests that the equilibrium AUD can be reduced to a singleton contract *plus* a per-unit price for incremental demand. Accordingly, we compare the AUD with a simple scheme called quantity-forcing (QF), which specifies a single quantity and the corresponding payment. We find that, in equilibrium, the two pricing schemes are equivalent when the rival's capacity is relatively small. We also find that when the capacity is relatively large, the QF has an additional, softening competition effect. We further explore antitrust implications of the AUD and the QF.

The literature on the AUD and the QF is sparse. Kolay, Shaffer and Ordover (2004)[12] study the price discrimination effect of the AUD offered by a monopolist when the downstream buyer has private information. They show that a menu of AUDs can generate higher profits for the monopolist than a menu of two-part tariffs (2PTs). In a successive, bilateral monopolies setting, O'Brien (2013)[18] shows that the AUD can facilitate non-contractible investments. Feess and Wohlschlegel (2010)[10], in the spirit of Aghion and Bolton (1987)[1], show that the AUD can shift the rent from the entrant to the coalition between the incumbent and the buyer. The crucial element needed for this rent-shifting idea to work is that the adversely affected third party must be absent from the bilateral contracting stage. However, the order of sequential moves in this standard literature of rent-shifting and exclusion might not be consistent with some well-known antitrust cases, where the alleged victims of the exclusionary strategies were already active in the market and could make counteroffers before the buyer could make any purchase.[9]

By contrast, we consider a model in which the competitor is already active in the market and can respond to the dominant firm's pricing scheme with a counteroffer before the buyer makes her purchase decision. In particular, we consider a model with two firms, firms 1 and 2, in the upstream market producing identical products with the same marginal cost. There is a representative buyer in the downstream. We assume complete information, between firms and the buyer, to prevent price discrimination from being a plausible explanation for the AUD. The game is a three-stage sequential-move game in which firms 1 and 2 make offers to the buyer sequentially, and the buyer does not make any binding purchase decision until the last stage. This order of moves automatically excludes the rent shifting possibility between the buyer and any seller, because neither contract is binding unless the buyer purchases from it in the last stage. We provide a new rationale for the AUD in the absence of price discrimination, incentivizing investment or rent shifting motives in the literature. We also find that under some conditions the QF can play a similar role.

A crucial element of our model is the asymmetry between the two firms. The dominant firm (firm 1) has no capacity constraint, whereas its rival (firm 2) is capacity-constrained. It turns out that this capacity

[9] Chao (2013)[4] studies the three-part tariff and allows the rival to respond with a counteroffer before the buyer purchases. But in his setting, the rival has full capacity to serve the whole market, and competing products are differentiated.

constraint plays a key role in the strategic effects of AUD and QF when firms compete. The limited capacity of firm 2 implies that the dominant firm has a "captive" portion of the buyer's demand. The dominant firm is able to use AUD and QF to leverage its market power from the "captive" portion to the "contestable" portion of the demand, much like the tied-in selling strategy in the context of multiple products. Remarkably, although the AUD hurts the capacity-constrained firm all the time, QF may improve the capacity-constrained firm's profit over LP, when the capacity is relatively large.

There is a small body of literature on exclusionary contracts with competition between asymmetric firms. Ordover and Shaffer (2007)[20] consider exclusionary discounts in a two-period model, where one firm is financially constrained, and the buyer incurs switching costs after her first period purchase. They find that the unconstrained firm can exclude the constrained firm by locking in the buyer with a below-cost price for their second period demand. Our model departs from theirs because we consider a one-time purchase from the buyer, and thus there is no switching cost or externality across periods. DeGraba (2013)[6] considers naked exclusive contracts when a dominant firm competes against a small rival with downstream competition. He shows that the large firm can bribe downstream firms for exclusivity, provided that the size difference between the large firm and small firm is sufficiently large. We consider a different model with no downstream competition and do not allow upstream firms to pay the buyer directly for exclusivity. And we find that the AUD can have a partial foreclosure effect for any capacity difference between the large firm and small firm.

Another related literature is the market-share discounts, where discounts are conditional on a seller's percentage share of a buyer's total purchases, instead of an absolute quantity.[10] Majumdar and Shaffer (2009)[13] explain how the market-share discounts can create countervailing incentives for a retailer with private information on demand, when it buys from a dominant firm and competitive fringes. Inderst and Shaffer (2010)[11] point out that the market-share discounts can dampen both intra- and inter-brand competition at the same time. Mills (2010)[16] suggests the market-share discounts can induce non-contractible effort from retailers when their sizes are different, but optimal effort levels are proportional to their sizes. Calzolari and Denicolo (2013)[3] show that the market-share discounts can be anticompetitive when buyers have private information. Chen and Shaffer (2013)[5] study exclusionary contracts with minimum-share requirements. They find that the less than 100% share requirement may be more effective in deterring entry than a 100% naked exclusionary contract. The game in Chen and Shaffer (2013)[5] proceeds as in Rasmusen et al. (1991)[21] and Segal and Whinston (2000)[23], where the incumbent and buyers can sign contracts before the potential entrant enters. Our model differs from theirs in two important respects. First, we abstract away from downstream competition. Second, in our model the small firm is already in the market, and it can make a counteroffer before the buyer makes her purchase decision. As a complement to those mentioned above, our article suggests that we should put a cautious eye on those volume- or share-threshold based contracts when they are adopted by a dominant firm.

The remainder of the article is organized as follows. In Section 2, we set up the model. Section 3 derives two benchmark cases, in which the leading firm can only offer LP or a 2PT. Section 4 offers a

[10] Schwartz and Vincent (2008)[22] provide a survey on QF, bundled discounts and other nonlinear contracts, by reviewing the recent literature and highlighting some open questions.

preliminary analysis showing the similarities and differences between AUD and QF. Sections 5 and 6 present the equilibrium analysis of QF and AUD. In Sections 7 and 8, we compare several pricing schemes and use linear demand examples to illustrate comparative statics analysis and discuss properties of the equilibria. In Section 9, we extend the model and discuss some assumptions of the model. The article closes in Section 10 with some concluding remarks. All proofs are relegated to the Appendix.

2 Model Setting

We consider two types of volume threshold based pricing schemes. The first one is the all-units discount (AUD), which consists of a triple (p_o, Q, p_1) with $p_o > p_1$ and $Q > 0$. Here p_o is the per-unit price when the quantity purchased is less than the quantity threshold Q, and p_1 is the per-unit price for *all* units once the quantity purchased reaches Q. So the AUD is a pricing scheme that rewards a buyer for purchasing some threshold quantity from a firm. In particular, the total payment schedule under AUD is[11]

$$T^{AUD}(q) = \begin{cases} p_o \cdot q & \text{if } q < Q \\ p_1 \cdot q & \text{if } q \geq Q \end{cases} .$$

The second one is the quantity forcing (QF). It is a pair (Q, T) that specifies the quantity to be supplied Q and the corresponding payment T. Any quantity other than Q is not available. In the literature, such single volume threshold QF is also called an "All-or-Nothing" scheme (see Schwartz and Vincent, 2008[22]). Its total payment schedule is

$$T^{QF}(q) = \begin{cases} T & \text{if } q = Q \\ \infty & \text{if } q \neq Q \end{cases} .$$

The two pricing schemes are illustrated in Figure 1.

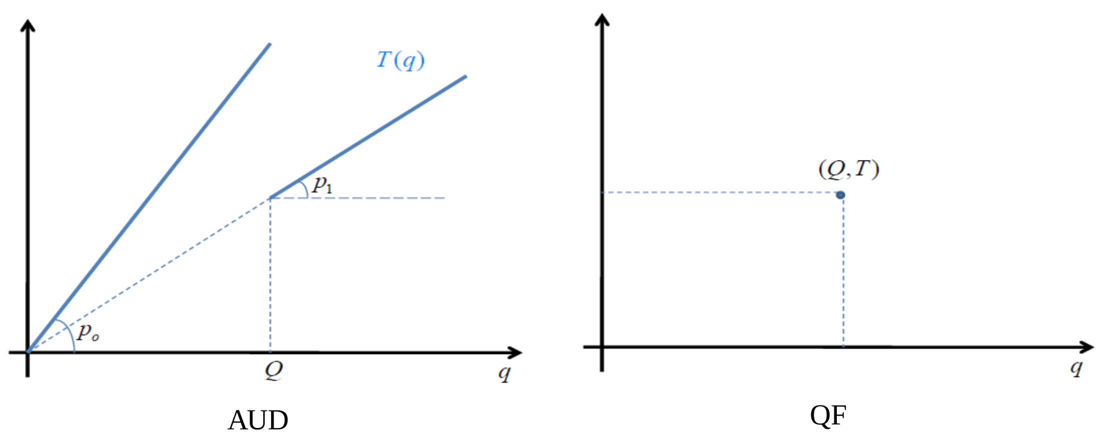


Figure 1: Total Payment Schedules

[11] In practices, multiple volume thresholds are often observed, but we focus on a single volume threshold case. This is because we consider a complete information setting, and it is unnecessary to offer more than one threshold in equilibrium.

There are two firms, say firm 1 and firm 2, in the upstream market that produce identical products with the same marginal cost $c \geq 0$. In order to examine the strategic effects of AUD and QF when a dominant firm competes against a smaller firm, we introduce an asymmetry between two firms—–capacity constraint for the small firm—–into the model.[12] Specifically, firm 1 has full capacity to serve the whole demand of the buyer whereas firm 2 is capacity-constrained in the sense that it can produce at marginal cost c up to its capacity k.[13]

In the downstream, there are a large number of buyers, each of whom is a local monopoly in selling to final consumers, due to local brand names or other attributes of product differentiation. Although each buyer is a local monopoly, none of them has monopsony power. This is because either each of them has only a small share of the whole market, or the number of upstream supplies is quite limited compared with the downstream demand.[14] Moreover, we assume complete information about the demands in every market, and two manufacturers make customized offers to each local monopoly retailer. Therefore, without loss of generality, we can consider a representative buyer with a gross utility function denoted as $u(q)$.

This set up has the following interpretations. As our objective here is to see if an AUD or a QF can have any strategic effects purely coming from upstream competition, we want to rule out any other motives as best as we can. The assumption of one representative buyer helps us to abstract away from strategic interactions resulting from downstream competition. In addition, the complete information assumption in the model prevents price discrimination from being a plausible explanation. As will be illustrated later, even in this simple framework, both AUD and QF have some bite on competition, and their competitive effects can be different depending on the rival firm's capacity level.

We model the interactions between the firms and the buyer as a sequential-move game with three stages. In the first stage, firm 1 offers a pricing scheme to the buyer, which could be LP, a 2PT, an AUD, or a QF. In the second stage, after observing the pricing scheme from firm 1, firm 2 sets its per-unit price for the buyer. In the third stage, the buyer decides where and how many units to purchase. In our setting, we assume firm 2 can only use LP in order to capture the fact that smaller firms in reality usually cannot match the pricing scheme as complicated as offered by a dominant firm. It is worth noting that the buyer here can purchase from both firms. For completeness, we assume that in the event of a tie when the two firms offer the same surplus to the buyer, the buyer will buy from firm 2 with an attempt to fulfill Q (if any) if possible. This tie-breaking rule is used to avoid the need to consider a situation in which the follower charges a price arbitrarily close to, but below the leader's price. The game's timeline is described in Figure 2.

[12] In the Intel case, it is widely known that AMD is capacity constrained, and therefore large computer manufacturers have to carry a significant proportion of their CPU requirements from Intel.

[13] Note that the "capacity constraint" here does not have to be interpreted literally as the physical capacity limit. The small rival can be constrained because of a "must-have" brand from the dominant firm, strong product differentiation, or large switching costs.

[14] Such market structure, where there are a large number of buyers whereas only few sellers, is consistent with many antitrust cases in which contracts offered by the dominant upstream firm give rise to abuse of dominance concern, because otherwise the large buyer power can be a countervailing force to discipline upstream suppliers' abuse of power. In our motivating Intel case, the downstream computer manufacturers only have two major suppliers of CPUs, say Intel and AMD, whereas there are a bunch of computer manufacturers in the downstream. Mathewson and Winter (1987)[15] made such an assumption, too.

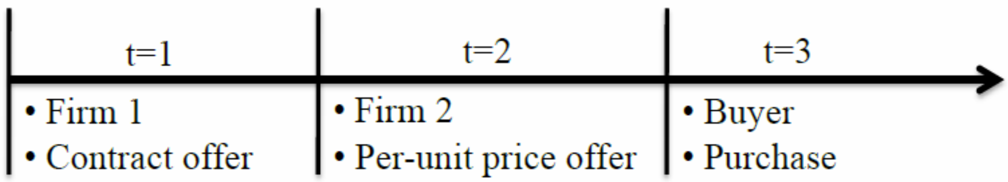


Figure 2: The Timeline of the Game

For the timing of the game, in practice, the nonlinear pricing schemes, such as AUD and QF, become an antitrust concern only when the firm adopting it enjoys a dominant position in the market. When there is a dominant firm, it is the dominant firm that usually moves first, and the number of moves is small. The literature on price leadership shows that the dominant firm will emerge as the price leader.[15] Here we model firm 1 as the dominant firm due to which it moves first and offers a more complicated pricing scheme than the follower firm 2 does. Basically, this sequential-move nature captures the commitment power the dominant firm has in preventing renegotiation.

Moreover, the assumption that the buyer does not make any decision until two competing offers are on the table is to capture the contestable conditions in favor of the buyer. It is worth noting that the nature of the sequential-move game in our model is different from that first introduced by Aghion and Bolton (1987)[1] and then extended by Marx and Shaffer (2004)[14]. In their models, the buyer has to decide whether to accept firm 1's offer or not before seeing firm 2's offer. Once firm 1's offer is accepted, it becomes binding for both firm 1 and the buyer. This is crucial for rent-shifting, which is from firm 2 to firm 1 and the buyer, to occur. Because it is quite possible that the buyer commits to pay firm 1 even if there is no trade between them, such payment is credible when the buyer meets firm 2 after accepting firm 1's offer. So the absence of firm 2 or its inability of making a counteroffer before the buyer accepts firm 1's is where the contracting externality in their models comes from. However, such order of sequential move in this exclusion literature is inconsistent with some well-known antitrust cases, such as in the *FTC v. Intel*, *US v. Microsoft*, *3M v. LePage's*, and *Michelin II* cases, where the alleged victims of the exclusionary strategies were already active in the market and could make counteroffers before the buyer making any purchase.[16] On the contrary, the order of moves in our setting automatically excludes this possibility of rent-shifting between the buyer and any firm, because neither contract is binding for the buyer until the buyer purchases from it in the last stage. And we allow the small firm to respond to the dominant firm's pricing scheme with counteroffers before the buyer makes a purchase decision.

In addition, the equilibrium strategies are renegotiation-proof by nature of the timing because the buyer doesn't commit to any contract before both manufacturers make offers. The nice aspect of this article is that even in this substantially competitive environment at upstream level, both AUD and QF still have some interesting strategic effects.

We make several basic technical assumptions. The first one is on the buyer's utility function, which is assumed to be monotonic and concave.

[15]For literature on price leadership, see Deneckere and Kovenock (1992)[7] and van Damme and Hurkens (2004)[25].

[16]Spector (2011)[24] emphasized this, too, when considering exclusive contracts. In a setting with economies of scale, he showed eviction can occur even if the excluded firm is present and can make counteroffers.

Assumption 1 *$u(q)$ is $\mathbb{C}^2$ on $[0, \overline{q})$, $u'(q) > 0$, $u''(q) < 0, \forall q \in [0, \overline{q})$, $u'(0) > c$, and $u'(\overline{q}) = 0$ where $0 < \overline{q} \leq \infty$.*

Let the optimal quantity demanded by the buyer at the per-unit price p be $q(p) \equiv \arg\max_{x \geq 0}[u(x) - p \cdot x]$. With Assumption 1, $q(p)$ exists and is uniquely determined by $u'(q) = p$ for $c \leq p \leq u'(0)$. Let $v(p) \equiv u(q(p)) - p \cdot q(p)$ be the buyer's surplus when she purchases optimally at per-unit price p.

Assumption 2 $k < q(c)$.

Assumption 2 states that firm 2's capacity level is strictly less than the socially efficient level of quantities, implying that firm 2 cannot serve the whole demand of the buyer when two firms compete à la Bertrand. We can consider $q(c) - k$ as firm 1's "captive" demand. It is the residual demand for firm 1 when firms compete in prices up to the marginal cost. This is also the maximum demand left for firm 1 if firm 2 supplies at its full capacity k. Correspondingly, the "contestable" portion is k, for which both firms compete.

Let the monopoly profit at per-unit price p be $\pi(p) \equiv (p - c) \cdot q(p)$. To facilitate our analysis, we assume the monopoly profit function to be concave. It is clearly satisfied if $q(p)$ is concave.

Assumption 3 *$\pi''(p) < 0, \forall p \in [c, p^{choke}]$, where $q(p^{choke}) = \overline{q}$.*

Denote $p^m \equiv \arg\max_p \pi(p)$ as the monopoly price, and $q^m \equiv q(p^m)$ as the monopoly quantity. In addition, let

$$h(Q) \equiv \max_p (p - c) \cdot [q(p) - Q]$$

for $0 \leq Q \leq q(c)$ be the maximum profit based on the residual demand $q(p) - Q$. Clearly, $h(Q)$ is strictly decreasing and convex in $Q \in [0, q(c)]$.

In the rest of our article, we will determine the subgame perfect equilibrium outcome of the sequential-move game, allowing the dominant firm to choose LP, 2PT, AUD and QF, respectively, and we will compare the equilibrium outcomes.

3 Two Benchmarks

In the first benchmark, the dominant firm can offer LP only.

Proposition 1 (LP vs LP Equilibrium) *(i) The LP equilibrium is uniquely characterized by $p_1^{LP} = p_2^{LP} = \overline{p} \in (c, p^m)$, where $\overline{p}(k)$ is given by*

$$\pi'(\overline{p}) = k. \tag{1}$$

(ii) In the LP equilibrium, firm 1 earns $\pi_1^{LP} = h(k)$ with sales $q_1^{LP} = q(\overline{p}) - k$; firm 2 earns $\pi_2^{LP} = (\overline{p} - c) \cdot k$ with sales $q_2^{LP} = k$; the buyer's surplus $BS^{LP} = v(\overline{p})$.

This proposition indicates that, when firm 1 is restricted to LP, it will have to leave firm 2 its capacity k and only focus on the residual demand $q(\overline{p}) - k$. This is due to the fact that uniform per-unit price from firm 1 is available for the buyer's whole demand forces firm 2 to always undercut it, because otherwise firm

2 would have no sales. Once firm 2 undercuts, the buyer will consider firm 1's supply only after exhausting firm 2's capacity k.

An immediate result following from Proposition 1 is the comparative statics below.

Corollary 1 *For $k \in [0, q(c))$, as k increases, $\overline{p}(k)$ decreases, BS^{LP} increases, and π_1^{LP} decreases.*

As firm 2's capacity k increases, competition becomes more intensive, from which the buyer benefits and firm 1 gets hurt. However, firm 2's profit is not necessarily monotonic in k, because there are two opposing effects on its price and sales respectively: $\overline{p}$ falls while k rises. Indeed, firm 2's profit increases with k when k is small, whereas decreases with k when k is large.

Next we consider the second benchmark in which the dominant firm offers a 2PT, say a pair (T_1, p_1).

Proposition 2 (2PT vs LP Equilibrium) *(i) The 2PT equilibrium is uniquely characterized by*

$$p_1^{2PT} = c,\ T_1^{2PT} = v(c) - [u(k) - c \cdot k];\ p_2^{2PT} = c.$$

(ii) In the 2PT equilibrium, firm 1 earns $\pi_1^{2PT} = v(c) - [u(k) - c \cdot k]$ with sales $q_1^{2PT} = q(c) - k$; firm 2 earns $\pi_2^{2PT} = 0$ with sales $q_2^{2PT} = k$; the buyer's surplus $BS^{2PT} = u(k) - c \cdot k$.

This proposition says that when firm 1 can use a 2PT, it will leave firm 2 its full capacity k again, as in the LP equilibrium. The difference is that firm 1 now can extract all the surplus from the residual demand through the fixed fee. Therefore, firm 1 has an incentive to ensure that the total surplus is maximized so that the incremental surplus for it to extract is maximized, too. It is easy to see that firm 1 earns more profit whereas firm 2 gets hurt under the 2PT equilibrium than under the LP equilibrium.

In the following analysis, we will see how an AUD or a QF can further increase firm 1's profit over a 2PT, given that firm 1 has already extracted the full surplus from its captive portion $q(c) - k$.

4 Preliminary Analysis of AUD and QF

We now study two volume-threshold based pricing schemes, AUD and QF. A common feature between the two schemes is the volume threshold. As such, the buyer needs to decide whether to meet the volume threshold from firm 1 or not. As the first mover of the game, firm 1 will have incentives to design such a volume threshold together with payment structure to induce the buyer to reach the threshold in equilibrium. Thus, firm 1 has incentives to set a sufficiently high initial price p_o in order to make not meeting the threshold option unattractive to the buyer. In that sense, the volume target under AUD becomes a quantity requirement under QF de facto.

There are also differences between AUD and QF schemes. The marginal price p_1 for incremental demand is present under AUD whereas absent under QF. Such marginal price for incremental demand needs to be restricted, which in turn limits firm 2's choice of p_2.

To understand the common features and differences between AUD and QF, we begin with analyzing the buyer's purchase decisions in the last stage of the game.

Given an AUD (p_o, Q, p_1) offered by firm 1, and a uniform price p_2 from firm 2, the buyer's maximization problem

$$\max_{\substack{q_1 \\ q_2 \le k}} [u(q_1 + q_2) - T^{AUD}(q_1) - p_2 \cdot q_2]$$

can be decomposed into the following two maximization problems. The first one is given by

$$\max_{\substack{q_1 < Q \\ q_2 \le k}} [u(q_1 + q_2) - p_o \cdot q_1 - p_2 \cdot q_2], \tag{2}$$

which represents the case when the buyer does not meet firm 1's volume threshold Q. The second one is given by

$$\max_{\substack{\Delta \ge 0 \\ q_2 \le k}} [u(Q + \Delta + q_2) - p_1 \cdot (Q + \Delta) - p_2 \cdot q_2], \tag{3}$$

which represents the case when the buyer meets firm 1's volume threshold Q. The buyer chooses one of the two options that gives her higher surplus.

■ **Single Sourcing from Firm 2.** In order for the AUD to improve firm 1's profit over LP, the buyer must meet firm 1's volume threshold Q in the AUD equilibrium. This is because the outcome of (2) can always be achieved by LP (p_o) vs LP (p_2). Therefore, firm 1 does not want the buyer to choose (2) in equilibrium, and it is without loss of generality to restrict our attention to $p_o = \infty$.[17] So from the buyer's point of view, *the equilibrium AUD* (p_o, Q, p_1) *can be reduced to a QF scheme* (Q, T) *with* $T = p_1 Q$ *plus a per-unit price* p_1 *for incremental demand*.

As a result of sufficiently high p_o, (2) is reduced to

$$\max_{q_2 \le k} [u(q_2) - p_2 \cdot q_2], \tag{SS}$$

which represents single-sourcing (SS) when the buyer does not meet firm 1's volume threshold and thus purchases from firm 2 only.[18] That is, under both AUD and QF contracts, if the buyer decides not to meet Q, she essentially chooses SS from firm 2.

The solution to the (SS) problem serves as an outside option for firm 2 as well as for the buyer. Such an outside option applies whether firm 1 uses an AUD or QF scheme. Denote the buyer's demand under SS as $\overline{q}(k, p_2) \equiv \min\{k, q(p_2)\}$. We can write the buyer's surplus under SS as

$$BS_S(p_2) = u(\overline{q}(k, p_2)) - p_2 \cdot \overline{q}(k, p_2). \tag{4}$$

Two firms' profits under SS are $\pi_1 = 0$ and

$$\pi_2 = (p_2 - c) \cdot \overline{q}(k, p_2). \tag{5}$$

[17] Here p_o does not have to be ∞, literally. In fact, we only need p_o to be above a certain level in equilibrium, ensuring that any amount below Q from firm 1 is never optimal for the buyer.

[18] Note that there is another kind of SS in which the buyer only purchases from firm 1. However, as shown in the proof of Lemma 1, introducing buyer SS from firm 1 only can *at most* give firm 1 the 2PT equilibrium profit.

Apparently, the SS problem under AUD is exactly the same as under QF.

■ **Dual Sourcing.** Now we study (3) carefully, as this is the case that will emerge in equilibrium. Moreover, we will see the differences between AUD and QF from (3).

Under (3), when the buyer meets firm 1's volume threshold, she will continue to buy from the cheaper source, as long as her marginal utility is above the corresponding price. Thus, in order to have positive sales, firm 2 as a follower must always set $p_2 \le w \equiv \min\{p_1, u'(Q)\}$ as long as $c < w$. As a result, the buyer buys exactly Q units from firm 1 and her residual demand from firm 2. Therefore, if we denote $T = p_1 Q$, then with $p_2 \le w$, (3) will be reduced to

$$\max_{q_2 \le k}[u(Q + q_2) - T - p_2 \cdot q_2], \tag{DS}$$

which represents dual-sourcing (DS) when the buyer meets firm 1's volume threshold and may purchase her remaining demand from firm 2.

Under an AUD (p_o, Q, p_1) with $p_o = \infty$ and $T = p_1 Q$, the buyer's surplus in (3) is

$$BS_D^{AUD}(p_2) = \begin{cases} u(\overline{q}(Q+k, p_2)) - p_2 \cdot \overline{q}(Q+k, p_2) + p_2 \cdot Q - T & \text{if } p_2 \le w \\ u(q(w)) - p_1 \cdot q(w) & \text{if } w < p_2 \end{cases}. \tag{6}$$

The two firms' profits from (3) are

$$\pi_1^{AUD} = \begin{cases} T - c \cdot Q & \text{if } p_2 \le w \\ (p_1 - c) \cdot q(w) & \text{if } w < p_2 \end{cases}, \tag{7}$$

and

$$\pi_2^{AUD} = (p_2 - c) \cdot [\overline{q}(Q+k, p_2) - Q] \tag{8}$$

for $p_2 \le w$, and 0 otherwise.

By contrast, under a QF (Q, T), the buyer's surplus from (DS) is

$$BS_D^{QF}(p_2) = \begin{cases} u(\overline{q}(Q+k, p_2)) - p_2 \cdot \overline{q}(Q+k, p_2) + p_2 \cdot Q - T & \text{if } p_2 < u'(Q) \\ u(Q) - T & \text{if } u'(Q) \le p_2 \end{cases}. \tag{9}$$

Accordingly, the two firms' profits under DS are $\pi_1^{QF} = T - c \cdot Q$, and

$$\pi_2^{QF} = (p_2 - c) \cdot [\overline{q}(Q+k, p_2) - Q] \tag{10}$$

for $p_2 < u'(Q)$, and 0 otherwise. Note that the buyer's surplus and both firms' profits in (3) under AUD when $p_2 \le w$ are exactly the same as those under QF when $p_2 < u'(Q)$.

As firm 1 would have no sales under SS, in order for firm 1 to earn possible positive profit, it must ensure the buyer to choose DS under both AUD and QF. The following lemma shows that the buyer will meet firm 1's quantity threshold Q in the AUD and QF equilibria, and firm 2 will supply too, but at a level strictly below its capacity k.

Lemma 1 (Firm 1 must induce DS and firm 2 undersupplies) *In both AUD and QF equilibria, (i)* $q_1 = Q \in (0, q(c)]$*; (ii)* $0 < q(p_2) - Q < k$*.*

Lemma 1 tells us that, in the AUD and QF equilibria, the buyer will buy from both firms——Q from firm 1 and $q(p_2) - Q$ from firm 2. So firm 2 becomes a residual demand supplier after Q. Note that after the buyer fulfills firm 1's threshold Q, firm 2 will always set $p_2 \leq u'(Q)$, because otherwise the buyer would never buy anything from firm 2 in DS. So $Q < q(p_2)$ indicates that firm 1 will leave some demand for firm 2 under both AUD and QF. But at the same time firm 1 contains firm 2. $q(p_2) - Q < k$ implies that in the AUD and QF equilibria, firm 2 strictly undersupplies as a residual demand supplier. This contrasts remarkably with the case of LP or a 2PT, where firm 2 always supplies its full capacity.

■ **Differences between AUD and QF.** The above discussions illustrate the common features of AUD and QF schemes. We now discuss the major differences between these two pricing mechanisms.

It is worth noting that, by their definitions, the marginal price p_1 is absent in QF, whereas is available in AUD. As such, AUD entails two more constraints compared with QF. First, due to the availability of p_1 for incremental demand, firm 2 faces one more constraint $p_2 \leq p_1$ under AUD. Second, in the AUD equilibrium, p_1 cannot be set too high, i.e., $p_1 < u'(k)$, because otherwise the buyer always chooses SS when $p_2 \leq p_1$.

We now summarize our comparison of QF and AUD. The equilibrium AUD (p_o, Q, p_1) is equivalent to a QF (Q, T) *plus* a per-unit price p_1 for incremental demand, where $T = p_1 Q$ and $p_o = \infty$. It is the very marginal price p_1 under AUD only that gives rise to the differences between AUD and QF, which are highlighted in the lemma below.

Lemma 2 (Price Constraints Under AUD) *The equilibrium AUD* (p_o, Q, p_1) *with* $p_o = \infty$ *needs to satisfy the following two constraints:*

$$p_1 < u'(k), \tag{C1}$$

and

$$p_2 \leq p_1. \tag{C2}$$

Compared an AUD (p_o, Q, p_1) with $T = p_1 Q$ and $p_o = \infty$, a QF (Q, T) does not entail constraints (C1) and (C2), simply because the marginal price p_1 for the incremental demand is absent under QF. In our setting, such p_1 restricts firm 2's choice of p_2. As we will see next, such restriction on firm 2 turns out to backfire on firm 1.

Consequently, it is instructive to characterize the QF equilibrium first, before determining the AUD equilibrium.

5 QF Equilibrium

In this section, we characterize the QF equilibrium, which provides a basis for our analysis of the AUD equilibrium later.

We can solve our sequential-move game by backward induction. It turns out that the determination of the leader's optimal QF can be reduced to a mechanism design problem. In particular, by judiciously choosing the quantity threshold along with the corresponding fixed fee, the leading firm induces the buyer to reach the threshold and firm 2 to be indifferent between supplying the residual demand at a higher price and being a sole supplier by undercutting. Through this way, the leading firm can leverage its market power in its captive market to the contestable part, which the smaller firm would otherwise be interested in competing for.

Below we will first present several lemmas, which offer a set of necessary conditions for equilibrium. The logic is supported by iterated elimination of dominated strategies using firm 1 and firm 2's forward thinking. We will then formulate firm 1's maximization problem, and characterize the equilibrium.

5.1 Dual-Sourcing vs. Single-Sourcing, and the Implied Threat Price

From (4) and (9), the buyer's surplus curves under both SS and DS weakly decrease with p_2, and BS_S curve as a function of p_2 is everywhere no flatter than BS_D^{QF} curve, as illustrated in Figure 3. Intuitively, the impact of p_2 on BS_S is larger than that on BS_D^{QF}, because firm 2 is the sole supplier under SS whereas firm 1, as a substitute supplier, becomes available under DS. If BS_D^{QF} is everywhere below BS_S, then the buyer would never choose DS. But if BS_D^{QF} is everywhere above BS_S, it is not optimal for firm 1, either. Note that BS_D^{QF} decreases with T. Whenever BS_D^{QF} is everywhere above BS_S, although the buyer will choose DS, firm 1 can always increase its profit by increasing T. Hence, BS_D^{QF} and BS_S must cross once, as shown in Figure 3. Such a unique crossing point is firm 2's threat price to undercut and induce SS.

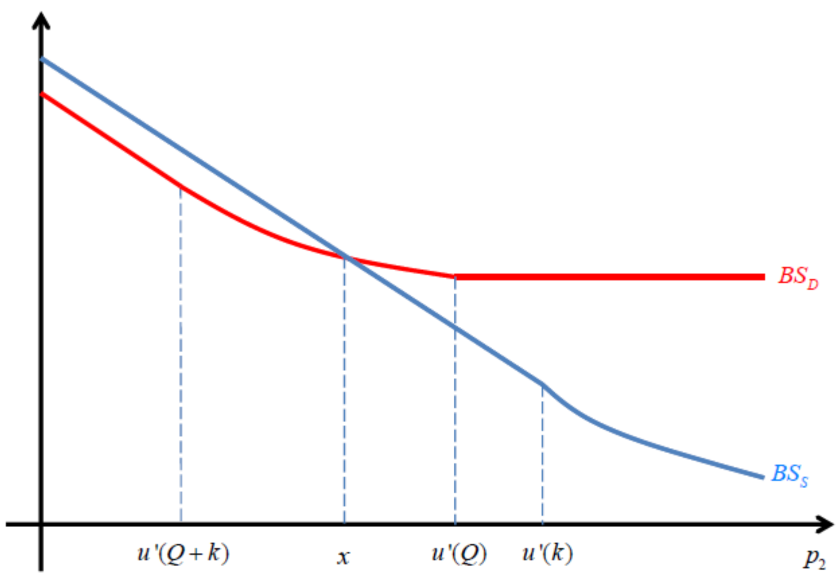


Figure 3: Buyer's Surpluses

Lemma 3 (Firm 2's equilibrium threat price) *In the QF equilibrium, there exists a unique* $x \in (u'(Q+k), u'(Q))$ *determined by*

$$u(\overline{q}(k,x)) - x \cdot \overline{q}(k,x) = v(x) + x \cdot Q - T, \tag{11}$$

such that $BS_S(p_2) \gtreqless BS_D^{QF}(p_2), \forall p \lesseqgtr x$.

The left-hand side (LHS) of (11) is BS_S at $p_2 = x$ when buying $\overline{q}(k,x)$ from firm 2 only. The right-hand side (RHS) of (11) is BS_D^{QF} at $p_2 = x$ when buying Q from firm 1 and residual demand $q(x) - Q$ from firm 2. The condition (11) uniquely determines such x at which the buyer is indifferent between SS and DS, given (Q,T).

Given a QF (Q,T) from firm 1, firm 2 can always induce the buyer to choose SS by undercutting sufficiently. The upper bound of such an undercutting threshold for SS is threat price x. That is, if firm 2 charges p_2 below x, the buyer will choose SS from firm 2 only for $\overline{q}(k,p_2)$. If p_2 is above x, the buyer will choose DS: buys Q from firm 1 and $q(p_2) - Q$ from firm 2. So the most firm 1 can extract using its fixed fee T is the incremental surplus the buyer and firm 1 as a coalition can gain over the buyer's outside option of SS from firm 2 only, when firm 2 undercuts at x. Hence, the total payment T to firm 1 is determined as

$$T = v(x) + x \cdot Q - [u(\overline{q}(k,x)) - x \cdot \overline{q}(k,x)]. \tag{12}$$

That is, firm 1 will charge a fixed fee such that the buyer is just indifferent between SS from firm 2 and DS from both firms at firm 2's undercutting threat price x.

Now we can see firm 2's trade-offs introduced by a QF. Such trade-offs are absent under LP or a 2PT. Under LP or a 2PT, firm 2's only viable option is to undercut or match firm 1's per-unit price p_1, as p_1 is uniformly applied to all units supplied by firm 1. Nonetheless, with the quantity requirement Q, firm 1 commits to supply only Q units with a fixed fee T, and thus creates trade-offs for firm 2: undercuts below x to be a monopoly supplier, or instead charges a price above x to be a residual demand supplier after Q.

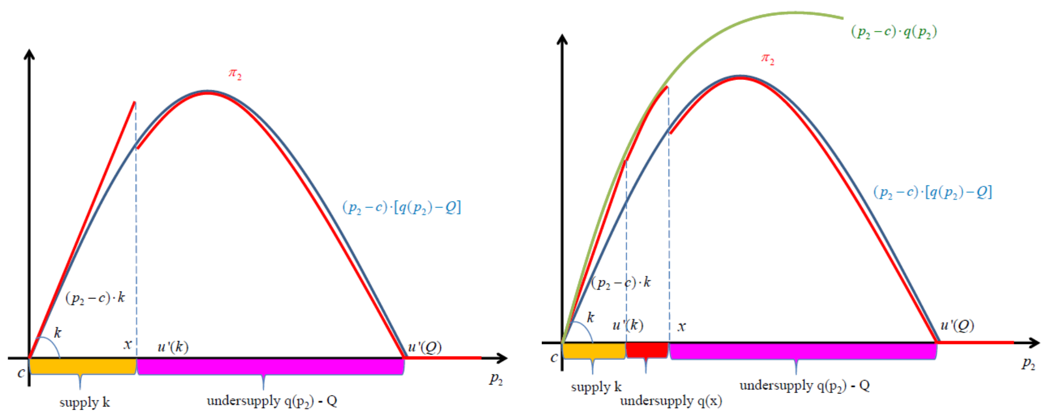
5.2 Firm 2's Pricing Decision

Lemma 3 tells us that, if firm 2 sets its p_2 below the cutoff x, then it will be a monopoly supplier for $\overline{q}(k,p_2)$; if it sets its p_2 above x but below $u'(Q)$, then it will supply the residual demand $q(p_2) - Q$. As a result, firm 2's profit can be written as

,

$$\pi_2(p_2) = \begin{cases} (p_2 - c) \cdot \overline{q}(k,p_2) & \text{if } p_2 < x \\ (p_2 - c) \cdot [q(p_2) - Q] & \text{if } x \leq p_2 < u'(Q) \\ 0 & \text{if } u'(Q) \leq p_2 \end{cases} .$$

Note that there is a discontinuous drop at x in firm 2's profit curve. And there are two possible cases, depending on whether $q(x) < k$ holds or not. Firm 2's profit curves for the two cases are shown as the red curves in Figure 4.



Case (i): when $q(x) > k$ Case (ii): when $q(x) \leq k$

Figure 4: Firm 2's Profit

From its profit curve, we can clearly see the trade-offs firm 2 faces: undercutting below x with its limited capacity k and making itself a monopoly supplier, or giving up part of the contestable market by leaving Q units to firm 1 but charging a higher price between x and $u'(Q)$. Accordingly, firm 1's profit is

$$\pi_1 = T - c \cdot Q$$

for $x \leq p_2$, and 0 otherwise.

Note that firm 2 would never choose p_2 equal to or greater than $u'(Q)$, because it would earn zero in that case. But setting $p_2 < x$ would leave zero profit for firm 1. Thus, for a profitable improvement, firm 1 must ensure $x \leq p_2 < u'(Q)$, instead of $p_2 < x$. That is,

$$\max_{p_2 < x}(p_2 - c) \cdot \overline{q}(k, p_2) \leq \max_{x \leq p_2 < u'(Q)}(p_2 - c) \cdot [q(p_2) - Q], \tag{13}$$

which says being a residual demand supplier is at least as profitable as being an undercutting monopoly. Because there is a discontinuous drop at x in firm 2's profit curve, firm 2 would prefer $p_2 < x$ if $p_2 = x$ is the optimal solution to the RHS problem in (13). Thus, firm 2's optimal price p_2 must be an interior solution. We can further show that (13) must be binding in equilibrium.

Lemma 4 (Firm 2's Choices) *In the QF equilibrium,*

$$(x - c) \cdot \overline{q}(k, x) = h(Q), \tag{14}$$

and

$$\pi'(p_2) = Q, \tag{15}$$

with $x < p_2 < u'(Q)$.

The LHS of (14) is firm 2's profit when it supplies $\overline{q}(k,x)$ as an undercutting monopoly. The RHS of (14) is firm 2's maximum profit when it supplies the residual demand and undersupplies. Recall from (12) that T increases with x, as $u'(Q+k) < x$. So whenever the LHS of (14) is smaller than the RHS of (14), firm 1 can always increase its profit by increasing fixed fee T, thereby increasing threat price x. Lemma 4 demonstrates that in equilibrium, firm 1 will design its QF to induce firm 2 to be just satisfied as a residual demand supplier, rather than an undercutting sole supplier. In the QF equilibrium, firm 2 undersupplies and sets its price p_2 above threat price x to maximize the residual profit.

5.3 Firm 1's Optimal QF

Note that firm 1's choice of a QF scheme can be reduced to an incentive contract design problem in which firm 1 chooses (Q,T) to maximize its profit such that (i) the buyer prefers DS to SS, and (ii) firm 2 chooses its uniform price p_2 optimally and yet is indifferent between choosing p_2 and threat price x. From the discussion in Section 5.1 and 5.2, firm 1's optimization problem is

$$\max_{(Q,T)} \pi_1^{QF} = T - c \cdot Q \tag{OP-QF}$$

$$s.t. \text{ (11), (14), (15)}$$

$$u'(Q+k) < x < p_2 < u'(Q) \tag{16}$$

To better understand strategic roles of the quantity threshold, we now denote all variables in terms of Q. For $0 \leq Q \leq q(c)$, let $x(Q)$ and $p_2(Q)$ satisfy (14) and (15) respectively. Using (12), the profit function of firm 1 can be expressed as

$$\pi_1^{QF}(Q) = \underbrace{v(x) + (x-c)\cdot Q}_{\text{Sum of surpluses for firm 1 and the buyer under DS at } x} - \underbrace{[u(\overline{q}(k,x)) - x\cdot \overline{q}(k,x)]}_{BS \text{ under SS at } x},$$

where $x = x(Q)$ is determined by (14). From such profit function expression, in the QF equilibrium, firm 1 extracts all the incremental surplus over the buyer's outside option at threat price x. Note that when $x = c$, the profit above is $v(c) - [u(k) - c\cdot k]$, which is firm 1's profit in the 2PT equilibrium. Moreover, it is easy to see that $x = c$ satisfies all constraints. So QF can at least reach the 2PT equilibrium profit by choosing $Q = q(c)$.

Define $\widehat{Q}_k$ such that $h(\widehat{Q}_k) = \pi(u'(k))$ if $k > q^m$. Lemma 5 below summarizes properties of $x(Q)$, $p_2(Q)$ and $\pi_1^{QF}(Q)$.

Lemma 5 *(i) $p_2(Q)$ strictly decreases with Q for $Q \in [0, q(c)]$.*

(ii) $x(Q)$ strictly decreases with Q for $Q \in [0, q(c)]$, and has a kink at $Q = \widehat{Q}_k$ if $k > q^m$.

(iii) $\pi_1^{QF}(Q)$ is continuously differentiable in Q for $Q \in [0, q(c)]$, except that it has a kink at $Q = \widehat{Q}_k$ if $k > q^m$.

When the quantity requirement Q increases, the competitive pressure on firm 2 becomes larger. In particular, the residual demand $q(p_2) - Q$ becomes more elastic as Q increases. So from (15), firm 2's equilibrium price p_2 falls with Q. By (14), the equilibrium threat price x will also be lower when Q becomes larger. When $k \leq q^m$, we always have $k < q(x)$ because $x < p^m$. Thus, in (14) $\overline{q}(k,x) = k$ all the time. However, when $k > q^m$, both $k < q(x)$ and $k \geq q(x)$ are possible. The presence of the kink $\widehat{Q}_k$ in this case is the result of two possibilities of $\overline{q}(k,x)$ in (14).

Note that

$$\begin{aligned} \frac{d\pi_1^{QF}}{dQ} &= \frac{\partial \pi_1^{QF}}{\partial Q} + \frac{\partial \pi_1^{QF}}{\partial x} \cdot x'(Q) \\ &= \underbrace{x - c}_{\text{Direct Effect}} + \underbrace{\{\overline{q}(k,x) - [q(x) - Q]\} \cdot x'(Q)}_{\text{Indirect Effect}}. \end{aligned} \tag{17}$$

Clearly, when Q increases by one unit, firm 1 has to incur an extra per-unit production cost c while it saves x, because x is the amount of per-unit payment to firm 2 for a coalition of firm 1 and the buyer. The difference $x-c$ is thus the direct effect of setting a higher Q. There is an indirect effect of increasing Q. It is through its impact on the most profitable undercutting price $x(Q)$. By the Envelope theorem, an increase in x reduces BS under SS by $\overline{q}(k,x)$. This helps firm 1, as it needs to compensate less to the buyer when inducing DS. Meanwhile, the higher x means the sum of surpluses for firm 1 and the buyer under DS is reduced, thanks to the greater payment to firm 2. By the Envelope Theorem, the magnitude of such reduction in surplus (or the increased payment to firm 2) is the residual demand purchased from firm 2 under DS at x, i.e., $q(x) - Q$. This hurts firm 1's profit. Consequently, the overall impact from x is $\overline{q}(k,x) - [q(x) - Q]$. So the indirect effect of Q through x is $\{\overline{q}(k,x) - [q(x) - Q]\} \cdot x'(Q)$. To maximize its profit, firm 1 will balance these two effects.

We now consider two cases of k. If $k \leq q^m$, we always have $k < q(x)$. Hence, (17) becomes[19]

$$q(p_2) - Q = k + Q - q(x). \tag{FOC-R}$$

That is, firm 1 sets its volume threshold such that the direct effect measured by the residual demand $q(p_2)-Q$ is equal to the indirect effect measured by the difference $k - [q(x) - Q]$.

To ensure the sufficiency and the uniqueness of (FOC-R) for the optimum and facilitate our comparative statics analysis, we make Assumption 4 below.

Assumption 4 (Concavity of Demand) $q''(p) \leq 0, \forall p \in [c, u'(0)]$.

Assumption 4, which is stronger than Assumption 3, guarantees that $\pi_1^{QF}(Q)$ is single-peaked in Q, and thus (FOC-R) characterizes the optimal solution.

If $k > q^m$, we need to consider the possibilities of $k < q(x)$ and $k \geq q(x)$. Accordingly, the objective function $\pi_1^{QF}(Q)$ has a kink at $\widehat{Q}_k$, where $k = q(x)$. Hence, $\pi_1^{QF}(Q)$ may have two local maximum points as shown in Figure 5. When $Q > \widehat{Q}_k$, we have $k < q(x)$, and the local maximum R is characterized by

[19] From (14), we get $x - c = h(Q)/k$ and $x'(Q) = h'(Q)/k = -(p_2 - c)/k$. Substituting them into (17) yields

$$\frac{d\pi_1^{QF}}{dQ} = \frac{p_2 - c}{k} \cdot \{[q(p_2) - Q] - [k - (q(x) - Q)]\}.$$

(FOC-R). When $Q \leq \widehat{Q}_k$, we have $k \geq q(x)$, and the local maximum L is characterized by[20]

$$(x - c)\pi'(x) = (p_2 - c)\pi'(p_2). \tag{FOC-L}$$

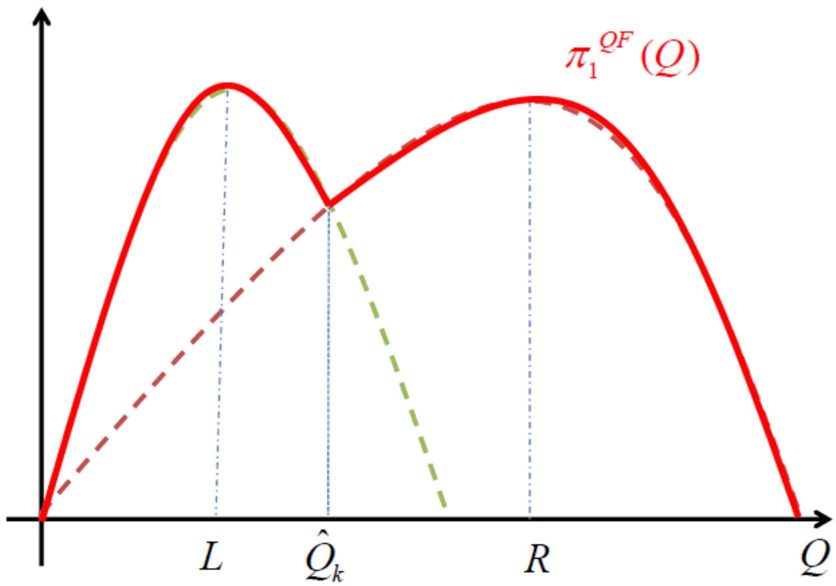


Figure 5: Kinky $\pi_1^{QF}(Q)$ if $k > q^m$

Similarly, for the sufficiency of (FOC-L) and the uniqueness of the solution to it, we make Assumption 5 below.

Assumption 5 (Single-Peakedness) *$(p - c)\pi'(p)$ is single-peaked in $[c, p^m]$.*

Both Assumptions 4 and 5 are satisfied by linear demand and generalized linear demand such as $q(p) = 1 - p^r$ $(r \geq 1)$. However, they do not generally imply each other. Assumption 4 is equivalent to $u'''(q) \leq 0$. Assumption 5 is implied by $q''(p) \leq 0$ and $q'''(p) \leq 0$.

When $k \leq q^m$, $\pi_1^{QF}(Q)$ only has one peak characterized by (FOC-R). When $k > q^m$, with the two local maximums, we identify a unique cutoff in k below which the right peak R dominates, and above which the left peak L becomes the global maximum. The following proposition characterizes the QF equilibrium.

Proposition 3 (QF Equilibrium) *There exists a unique QF equilibrium, which is characterized as follows. There exists a unique $\widehat{k} \in (q^m, q(c))$ such that*

- *when $k \in [0, \widehat{k})$, the equilibrium outcome (Q, T, p_2) along with threat price x is jointly determined by (11), (14), (15), and (FOC-R);*

[20] (14) leads to $x'(Q) = h'(Q)/\pi'(x) = -(p_2 - c)/\pi'(x)$ and hence

$$\frac{d\pi_1^{QF}}{dQ} = \frac{1}{\pi'(x)}[(x - c)\pi'(x) - (p_2 - c)Q]\}.$$

- *when $k \in [\widehat{k}, q(c))$, the equilibrium outcome (Q, T, p_2) along with threat price x is jointly determined by (11), (14), (15), and (FOC-L).*

Under both LP and a 2PT, firm 2 always undercuts and sells at its full capacity. So the contestable portion k becomes firm 2's turf. Accordingly, the best firm 1 can do is to use a fixed fee to extract the incremental surplus from its captive demand. Such incremental surplus is maximized at the efficient outcome under a 2PT, and thus firm 1 extracts its marginal contribution to the efficiency $v(c) - [u(k) - c \cdot k]$.

How can a QF further increase firm 1's profit over a 2PT, given that the 2PT equilibrium outcome is efficient and firm 1 has already extracted the full surplus from its captive portion $q(c) - k$? The crux is to leverage its market power from the captive portion to the contestable portion, and at the same time prevent firm 2 from undercutting.

The unique component of a QF, compared with LP or a 2PT, is its quantity requirement Q. Under QF, firm 1 now can take the initiative to dictate a quantity target beyond its captive portion, and commit not to supply any amount other than that. By doing so, the buyer faces trade-offs between SS and DS——if she buys from firm 2 at p_2 for $\overline{q}(k, p_2)$, she would not be able to meet firm 1's quantity requirement, and thus is forced to rely on firm 2's limited supply only; instead, if she meets firm 1's quantity target, her residual demand does not allow her to enjoy firm 2's lower price up to firm 2's full capacity. So with the quantity target instrument, firm 1 acts more aggressively and encroaches on the contestable portion. It induces the buyer to treat firm 2, instead of firm 1 as under LP or a 2PT, as a residual demand supplier.

Correspondingly, under QF, firm 2 now faces trade-offs that are missing under LP or a 2PT. Recall that under LP or a 2PT, firm 2's only option to survive is to undercut and hence sell its full capacity. Facing a QF, firm 2 has two options——undercut low enough to be a sole supplier, or set a high price serving the residual demand only. Hence, the quantity target creates another option other than undercutting for firm 2, so that preventing undercutting that is implausible under LP or a 2PT becomes possible now.

In the QF equilibrium, firm 1 judiciously designs the quantity requirement subject to two incentive constraints. One is from the buyer. Firm 1 has to ensure that the buyer will meet the quantity target rather than miss it and rely on firm 2 only. It is guaranteed by inducing firm 2 to set $x < p_2$, where x is given by (11). The other incentive constraint is from firm 2. Firm 1 has to induce firm 2 to be satisfied as a residual demand supplier instead of undercutting to be a sole supplier, as stated by (14).

For such leverage to work, firm 1's stake is its captive demand due to firm 2's limited capacity, and the quantity requirement is the instrument. What makes the QF interesting is that such quantity target plays two roles of "carrot and stick" at the same time. On the one hand, firm 1 designs its quantity requirement not too high so that it leaves some room for firm 2 to supply, i.e., $Q < q(p_2)$. It thus creates a new option for firm 2——be a residual demand supplier by setting a high price, like a "carrot" to induce firm 2 not to compete too harshly. On the other hand, firm 1 intentionally sets the quantity requirement beyond the captive portion at p_2, i.e., $Q > q(p_2) - k$, and use its captive portion as a threat. Firm 1's refusal to supply any amount other than Q makes firm 2's undercutting and selling at its full capacity more costly, because it then has to compensate the buyer's foregone purchase from firm 1 using its limited capacity. So the quantity requirement acts as a "stick" to prevent firm 2 from undercutting.

The corollary below illustrates such "stick"——the quantity expansion effect of the QF.

Corollary 2 (Quantity Expansion of QF) *In the QF equilibrium, $Q > q(c) - k$ for any $k > 0$.*

Under QF, firm 1 will expand its quantity requirement so large that the buyer would not be able to absorb firm 2's full capacity, even if firm 2 undercuts towards marginal cost c. Note that $Q > q(c) - k > q(p_2) - k$ for any $p_2 > c$. So Corollary 2 is stronger than Part (ii) of Lemma 1. Such significant quantity expansion squeezes the buyer's demand for firm 2's product to a level that it is strictly below its full capacity for any above-cost price it can charge.

Define the total surplus TS as the sum of both firms' profits and the buyer's surplus. The following corollary summarizes how the equilibrium outcomes change as k varies.

Corollary 3 (The Impacts of Limited Capacity) *For $0 < k < q(c)$, as k increases, the following hold:*

(i) The equilibrium quantity threshold Q and the total output weakly decrease.

(ii) The equilibrium p_2 (and also x) weakly increases.

(iii) The equilibrium profit π_1^{QF} weakly decreases, and π_2^{QF} weakly increases.

(iv) TS^{QF} weakly decreases.

As k increases, firm 2's competitive position becomes stronger. Therefore, firm 1, when designing its quantity target, has to leave more room for firm 2, in order to prevent firm 2 from undercutting. So the equilibrium Q decreases as k increases. But when k is above $\widehat{k}$, Q becomes independent of k.[21] Recall from Lemma 3 that x is the threat price firm 2 would charge when undercutting to induce SS. At threat price x, the buyer would buy $\overline{q}(k, x)$ from firm 2 only. For sufficiently large k, it is in firm 2's interest not to flood the market with its full capacity when undercutting. When this is the case, firm 2's deviation profit will be independent of k, i.e., $\pi(x)$ instead of $(x - c) \cdot k$. From the optimization program (OP-QF), it is easy to see that the whole problem becomes independent of k then. So is the case for Q when k is large.

Other comparative statics follow from the pattern of Q and Lemma 5. The results that π_1^{QF} decreases whereas π_2^{QF} increases when k increases are easy to understand. The fact that both total surplus and total output decrease with k suggests that as firm 2 becomes more competitive, the QF equilibrium deviates from the efficient outcome further. That is, QF behaves more as a collusive device for firm 1 and firm 2 to soften competition. However, the buyer's surplus is not monotonic in k, as will be shown in our illustrative examples later.

6 AUD Equilibrium

We now turn to the AUD equilibrium. Given our discussion on the relationships between AUD and QF schemes in Section 4 and our characterization of the QF equilibrium in Section 5, the determination of the AUD equilibrium can be simplified. The logic is still the backward induction.

As discussed in Section 4, the equilibrium AUD (p_o, Q, p_1) is equivalent to a QF (Q, T) *plus* a per-unit price p_1 for incremental demand, where $T = p_1 \cdot Q$ and $p_o = \infty$. From Lemma 2, compared with a QF, AUD involves two more constraints, (C1) and (C2), due to the presence of the marginal price p_1. Thus, from

[21]Due to the kinky π_1^{QF} and the presence of two local maximums as shown in Figure 5, Q drops at $\widehat{k}$.

firm 1's point of view, it can at most achieve the profit level that it would get under the optimal QF. This only occurs when neither (C1) nor (C2) is binding. So the key question here is when and which of the constraints (C1) and (C2) will be binding.

Similar to the procedure of characterizing the QF equilibrium, Lemmas 3 and 4 hold by a slight modification. In particular, similar to Lemma 3, there exists a unique threat price x at which the two buyer's surplus curves cross. That is, firm 2 will be a monopoly supplier if it sets p_2 below x whereas it will be a residual demand supplier if it sets p_2 above x.

It is worth noting that under AUD, $x < p_2$, (C2) and (C1) together yield $x < p_2 \leq p_1 < u'(k)$. Because now $x < u'(k)$ (or $Q > \widehat{Q}_k$ when $k > q^m$), $\overline{q}(k, x) = k$ all the time. Consequently, the buyer's indifference condition (11) in Lemma 3 now becomes

$$u(k) - x \cdot k = v(x) + x \cdot Q - p_1 \cdot Q. \tag{11'}$$

Parallel to Lemma 4, firm 2 will be induced to be satisfied as a residual demand supplier, instead of an undercutting monopoly. Thus, firm 2's indifference condition (14) can be written as

$$(x - c) \cdot k = h(Q), \tag{14'}$$

and firm 2's optimality condition (15) remains the same. Accordingly, firm 1's profit function under AUD is

$$\pi_1^{AUD}(Q) = (p_1 - c) \cdot Q = (x - c) \cdot Q + v(x) - [u(k) - x \cdot k]$$

for $p_2 \leq p_1$. So firm 1's optimization problem for AUD can be written as

$$\max_Q \pi_1^{AUD}(Q) \tag{OP-AUD}$$

$$s.t. \; (x - c) \cdot k = h(Q) \tag{14'}$$

$$\pi'(p_2) = Q \tag{15}$$

$$u'(Q + k) < x < p_2 < u'(Q) \tag{16}$$

$$p_1 < u'(k) \tag{C1}$$

$$p_2 \leq p_1 \tag{C2}$$

Clearly, due to (C1) and (C2), firm 1 weakly prefers QF to AUD. It can be shown that given that other constraints hold, (C1) is redundant.[22] So the only possible binding constraint is (C2). Thus, under AUD we have (i) one more constraint (C2), which could be binding, and (ii) one less case ($Q \leq \widehat{Q}_k$) to be considered due to (C1), as compared to the QF.

If we ignore (C2), then our analysis on the QF can be applied to the AUD here. When (C2) is not binding, then the equilibrium outcomes of AUD and QF are equivalent. When (C2) is binding, the equilibrium outcomes of AUD and QF differ. Thus, the crux is when (C2) will be binding in the equilibrium.

[22]This is formally shown in Lemma 6.

The following presents when (C1) is binding or not in the AUD equilibrium.

Lemma 6 (When (C2) is binding) *Given (Q, T, p_2, x) jointly determined by (11'), (14'), (15), and (FOC-R), (i) $u'(Q+k) < x < u'(k)$ implies (C1); (ii) there exists a unique $\overline{k} \in (0, q(c))$ such that (C2) is binding if and only if $k \geq \overline{k}$.*

This lemma states that (C2) is not binding for k below $\overline{k}$, and is for k above $\overline{k}$. Therefore, the AUD equilibrium condition for the determination of Q is (FOC-R) for $k < \overline{k}$, whereas (C2) for $k \geq \overline{k}$. The only thing that remains to be checked is the existence of the equilibrium when $k \geq \overline{k}$. The following proposition confirms this.

Proposition 4 (AUD Equilibrium) *A unique AUD equilibrium exists with $p_o = \infty$ and is characterized as follows. There exists a unique $\overline{k} \in (0, q(c))$ such that*

- *when $k \in [0, \overline{k})$, the equilibrium outcome (Q, p_1, p_2) along with threat price x is jointly determined by (11'), (14'), (15), and (FOC-R);*
- *when $k \in [\overline{k}, q(c))$, the equilibrium outcome (Q, p_1, p_2) along with threat price x is jointly determined by (11'), (14'), (15), and (C2).*

As discussed in Section 4, we can view the equilibrium AUD as a minimum quantity requirement Q with a quasi-fixed fee $T = p_1 \cdot Q$ plus a per-unit price p_1 for incremental demand. The quasi-fixed fee T serves as a surplus extraction tool, whereas the per-unit price p_1 may become a nuisance for firm 1. Because p_1 is an upper bound restriction on firm 2's choice p_2, e.g., (C2), it may force firm 2 to price aggressively low and backfire on firm 1. To avoid the possibly aggressive pricing from firm 2, firm 1 would like to set p_1 sufficiently high as under QF. Nevertheless, by AUD's definition, the per-unit price for both before and after the threshold $p_1 = T/Q$ is automatically implied by Q and T, instead of being freely chosen.

When k is small, firm 1 can extract surplus without worrying too much about competition, given firm 2's rather limited capacity. It will set a large requirement Q, and its average price for the Q units T/Q will be high, too. From (15), the large Q squeezes firm 2's residual demand and forces its optimal price p_2 to be low. So (C2) is not binding in this case. On the contrary, when k is large, the market becomes more competitive as firm 2's capacity grows. The competitive pressure forces firm 1 to set a small Q as well as a low average price for the Q units. The small Q results in a high p_2 from (15). That is, as k increases, p_1 is forced to fall whereas firm 2's optimal price rises. Then the constraint (C2) becomes binding and, in equilibrium, firm 2 will just match p_1 by setting $p_2 = p_1 = T/Q$.

An immediate result following the proposition is that the result in Corollary 2, i.e., $Q > q(c) - k$ for any k, remains valid in the AUD equilibrium, as shown in the corollary below.

Corollary 4 (Quantity Expansion of AUD) *In the AUD equilibrium, $Q > q(c) - k$ for any $k > 0$.*

Similar to the effect of the QF, the AUD has a significant quantity expansion effect. Such a quantity expansion effect illustrates how the dominant firm can leverage its market power from its captive portion to the contestable portion of the demand. By setting p_o to prohibitively high and the quantity threshold above its captive demand, the leverage is realized through a refusal-to-deal threat if the buyer's purchase is less than the threshold.

7 Comparisons

To further understand the similarities and differences between AUD and QF as well as their strategic effects compared with LP and 2PT, here we compare the equilibrium outcomes when firm 1 chooses LP, 2PT AUD and QF, respectively.

An immediate corollary follows from Propositions 3 and 4 as below.

Corollary 5 *When $k \leq \min\{\widehat{k}, \overline{k}\}$, the equilibrium outcomes between AUD and QF are equivalent; when $\min\{\widehat{k}, \overline{k}\} < k$, the QF yields higher profits for firm 1 than the AUD.*

Because QF involves less constraints, it weakly dominates AUD from firm 1's point of view. The divergence between AUD and QF when k is large is due to the nature of the marginal price p_1 in AUD.

As discussed in Section 4, the equilibrium AUD can be reduced to a QF (Q, T) with $T = p_1 \cdot Q$ *plus* a per-unit price p_1 for incremental demand. By definition of the AUD, once the buyer meets Q, firm 1's average prices for the Q units and beyond have to be the same. So once Q and the corresponding payment T are set, $p_1 = T/Q$ is automatically implied under AUD, instead of being freely chosen. Recall that in our sequential-move price-setting game, firm 2 has a second-mover advantage, although it is capacity constrained. Moreover, firm 2's second-mover advantage changes as k varies. This can be seen from the LP vs. LP case. Corollary 1 there demonstrates that higher k forces firm 1 to charge a lower price $\overline{p}$ and get lower profits. Under AUD, firm 1 would like to adjust its marginal price p_1 according to firm 2's second-mover advantage, because such p_1 may interfere with firm 2's choice of p_2, e.g., the constraint (C2), as k changes. But it cannot set p_1 freely.

When k is small, firm 2's second-mover advantage is small. So firm 1 mainly uses (Q, T) in AUD or QF to extract surplus without worrying much about the competitive pressure from firm 2. Hence, the implied $p_1 = T/Q$ is higher than p_2 and hence does not bother firm 1 or firm 2.

When k is large, firm 2's competitive pressure becomes firm 1's major concern. Now firm 1 as a first mover would like to use Q as its commitment to restrict its own supply only to that level, because firm 2's undercutting threat is significant. The QF can do the job because any amount other than Q, especially beyond that, is unavailable. However, the implied p_1 for incremental demand under AUD can be a nuisance for firm 1. The competitive pressure from firm 2 pushes firm 1's average price T/Q to be low. So the implied p_1 is forced to be low. However, in this case, firm 1 would like to set p_1 sufficiently high in order to avoid aggressive undercutting from firm 2. So under AUD, the presence of p_1 conflicts with firm 1's intention of committing not to supply beyond Q when k is large. The QF is immune to this because it involves no p_1.

Thus, the equivalence and divergence between AUD and QF for different values of k demonstrate how firm 2's capacity constraint along with its implied second-mover advantage affects firm 1's competitive concern.

When k is in the range of high values, we only know that firm 1 strictly prefers the QF to the AUD. We are unable to determine whether firm 2 or the buyer becomes better off or worse off under AUD than under QF, and we are unable to rank $\widehat{k}$ and $\overline{k}$ generally.[23] In Section 8, we will use numerical examples to illustrate the comparisons between AUD and QF for a full range of values of k.

In the corollary below, we provide a comparison of LP and AUD equilibria. Note that the LP equilibrium price $\overline{p}$ increases with k whereas the AUD equilibrium price p_2^{AUD} decreases with k. Because $p_2^{AUD}(0) = c < \overline{p}(0) = p^m$, there must be a cutoff $k_0 > 0$ such that $p_2^{AUD}(k_0) = \overline{p}(k_0)$.

Corollary 6 (Comparison between AUD and LP) *(i)* $p_2^{AUD} < p_2^{LP}, \forall 0 < k < \min\{k_0, \overline{k}\}$;

(ii) $q_1^{LP} < q_1^{AUD}, q_2^{AUD} < q_2^{LP} = k, \forall 0 < k < q(c)$;

(iii) $\pi_1^{LP} < \pi_1^{AUD}, \forall 0 < k < q(c); \pi_2^{AUD} < \pi_2^{LP}, \forall 0 < k < \min\{k_0, \overline{k}\}$.

(iv) Buyer's Surpluses: There exists a $k_1 \in (0, q(c))$ *s.t.* $BS^{AUD} < BS^{LP}$ *for* $0 < k < k_1$.

(v) Total Surpluses: There exists a $k_2 \in (0, q(c)]$ *s.t.* $TS^{AUD} > TS^{LP}$ *for* $0 < k < k_2$.

Hence, when k is relatively small, firm 1 gains from the AUD, firm 2 gets hurt in terms of both profit and volume sales, and the buyer gets hurt, compared with LP equilibrium. In the next section, we provide examples to illustrate that k does not have to be really small in order for the results in Corollary 6 to hold. So under AUD, we have a partial foreclosure in the sense that firm 2 is under-supplied strictly below its capacity and its profit is reduced. If firm 2 has certain fixed cost, then the AUD adopted by a dominant firm can partially deny firm 2's profitable access to the otherwise contestable market, and it may induce firm 2 to exit.

Compared with LP, the AUD has a fixed fee effect and a quantity-forcing effect. With an AUD, firm 1 can extract incremental surplus using its quasi-fixed fee $T = p_1 Q$. So it has an incentive to push the equilibrium outcome towards a more efficient one. That's why the total surplus can be higher and at the same time the buyer's surplus can be lower under AUD. Meanwhile, another instrument from the AUD—–the quantity target—–can squeeze firm 2's space by creating trade-offs between SS and DS. It turns out that under AUD, firm 1 will exploit its dominant position, and the quantity target acts more as a "stick" to push firm 2 into a corner. For relatively small k, the quasi-fixed fee under AUD extracts most of the buyer's surplus. Our results support the antitrust concern on AUD when k is relatively small.

The following corollary summarizes the comparison between QF and LP equilibria.

Corollary 7 (Comparison between QF and LP) *For* $0 < k < q(c)$,

(i) Prices: There exists a $k_3 \in (0, \widehat{k})$ *s.t.* $p_2^{QF} \lesseqgtr p_2^{LP}$ *for* $k \lesseqgtr k_3$.

(ii) Quantities: $q_1^{LP} < q_1^{QF}$*, and* $q_2^{QF} < q_2^{LP} = k$.

(iii) Profits: $\pi_1^{LP} < \pi_1^{QF}$*. There exists a* $k_4 \in (k_3, \widehat{k})$ *s.t.* $\pi_2^{QF} \lesseqgtr \pi_2^{LP}$ *for* $k \lesseqgtr k_4$.

(iv) Buyer's Surpluses: There exists a $k_5 \in (0, q(c))$ *s.t.* $BS^{AUD} < BS^{LP}$ *for* $k < k_5$.

(v) Total Surpluses: There exists a $k_6 \in (0, q(c))$ *s.t.* $TS^{QF} \gtreqless TS^{LP}$ *for* $k \lesseqgtr k_6$.

[23] It is true that $\overline{k} < \widehat{k}$ in various numerical computations we perform for generalized linear demand such as $q(p) = 1 - p^r$ $(r \geq 1)$. This may imply that AUD will be equivalent to QF as long as the constraint (C2) is not binding, and they differ only when (C2) becomes binding.

Similar to AUD, QF also has both the fixed fee effect and the quantity-forcing effect. The buyer gets hurt when k is small, too. However, because of the absence of the marginal price p_1, firm 1 under QF faces less constraints. As a result, the quantity target now can set the tone of competition, depending on firm 2's capacity level, without worrying about its possible implication on p_1. When firm 2's capacity is small, it plays the role of a "stick" as it does under AUD. In this case, the total surplus is higher due to the intensified competition. However, when firm 2's capacity becomes comparable to firm 1's, firm 1 will use the quantity target to restrict its own supply and thus leave some room for firm 2. In this case, the quantity target acts as a "carrot" to prevent firm 2 from aggressive undercutting, thereby softening competition. Therefore firm 2 benefits from a QF and total surplus is reduced.

In the corollary below, we provide a comparison between AUD (or QF) and 2PT equilibria.

Corollary 8 (Comparison with a 2PT) *Compared with a 2PT, both AUD and QF adopted by firm 1 increase firm 1's profits, reduce firm 2's profits, and decrease the buyer's surplus and total surplus.*

Under AUD (or QF), firm 1 always gains more in profits as well as volume sales than that under a 2PT, whereas firm 2 is under-supplied all the time. The buyer's surplus under AUD (or QF) is always below that under 2PT. This is because the 2PT equilibrium outcome is efficient, and the buyer enjoys its outside option—–buying k from firm 2 at marginal cost c. Nevertheless, the AUD (or QF) equilibrium outcome is inefficient, and the realized equilibrium price p_2 firm 2 charges is above marginal cost c.

In summary, both AUD and QF have two effects in general: the surplus extraction effect via fixed fee T, and the quantity-forcing effect via the quantity target Q. First, the AUD and QF can extract surplus from the buyer through its fixed fee. Second, the quantity-forcing effect always softens the over-fierce competition under a 2PT. But compared with LP, it always intensifies competition under AUD, whereas it can either intensify competition and hurt firm 2 or soften competition and benefit firm 2 under QF, depending on the magnitude of firm 2's capacity level. Under AUD, firm 1 ideally would like to set the tone of competition as it does using the QF. However, the per-unit price $p_1 = T/Q$ is automatically implied once T and Q are determined under AUD. When k is large, it is impossible for AUD to achieve both objectives of surplus extraction and controlling competition. The QF does not suffer from this because the marginal price p_1 is absent under QF.

8 Illustrative Examples

To illustrate our analyses above and gain more insights on how the limited capacity affects the equilibrium, in this section, we use numerical examples to investigate competitive effects of capacity constraint.

We consider a linear demand function $q(p) = 1 - p$, which is generated by the gross utility function $u(q) = q(1 - q/2)$. For simplicity, we assume $c = 0$. Accordingly, k is normalized to be in the range of $[0, 1)$. It is easy to verify that such a linear demand function satisfies Assumptions 1 and 3~5.

Let's take a quick look at an example of partial foreclosure effect of AUD (or QF). Table 1 shows the LP and AUD (or QF) equilibrium outcomes when $k = 0.1$.

Table 1: A Linear Demand When $k = 0.1$

	q_1	q_2	π_1	π_2	BS	TS
LP	0.45	0.1	0.2025	0.045	0.1513	0.3987
AUD (or QF)	0.9236	0.0382	0.4055	0.0015	0.0924	0.4993

In this case, firm 1's captive demand is $1 - k = 0.9$. Under AUD (or QF), firm 1 expands its volume sales beyond its captive portion, i.e., $q_1^{AUD} = 0.9236 > 0.9$. Compared with LP, firm 1 gains fairly large in both profit and volume sales, whereas firm 2 loses significantly in both dimensions, i.e., $q_2^{AUD} = 0.0382 < k = 0.1$ and $\pi_2^{AUD} = 0.0015 < \pi_2^{LP} = 0.045$. So under AUD (or QF), firm 2 incurs a 62% loss in volume sales, and a 97% loss in profit, compared to those under LP. The buyer's surplus is lowered by 39%, too.

Indeed, this partial foreclosure under AUD is true for all $k \in [0, 1)$. Moreover, the buyer is worse off under AUD than under LP for all $k < 0.23$.

Now we perform our comparative statics analyses for the full range of $k \in [0, 1)$, by directly applying Propositions 1~4. The computed results are listed in Tables A1 and A2 in the Appendix. It is easy to compute $\overline{k} \simeq 0.5354$, and $\widehat{k} \simeq 0.8642$. So according to Corollary 5, AUD and QF are equivalent for $k \leq 0.5354$.

■ **Firm 2's Volume Sales and Profits.** The equilibrium volume sales for firm 2 under four pricing schemes are shown in Figure 6. Firm 2's volume sales are severely hurt by the AUD. As firm 2 will supply to its full capacity k under LP, the difference between the blue line and red line tells us the idle capacity of firm 2 $k - [q(p_2) - Q]$.

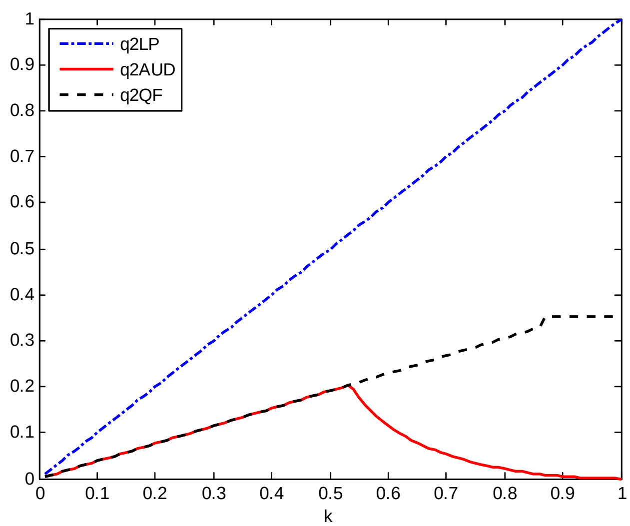


Figure 6: Firm 2's Volume Sales

Firm 2's volumes under AUD and QF are identical for low values of k and start to diverge starting from $k = \overline{k}$, where the constraint (C2) becomes binding. This fact demonstrates the difference between the AUD and the QF is thanks to the constraint (C2). Although firm 1 would like to free firm 2 in setting p_2 for residual demand, it is unfortunately restricted to do so when k is above $\overline{k}$ due to the nature of the AUD. Moreover, firms' divergent quantity paths tell us that as k increases firm 1 would like to leave more room for firm 2 in order to induce favorable response from it, because firm 2's competitive power becomes stronger. That's why we see firm 1's quantity sales keep falling whereas firm 2's sales keep rising under the QF.

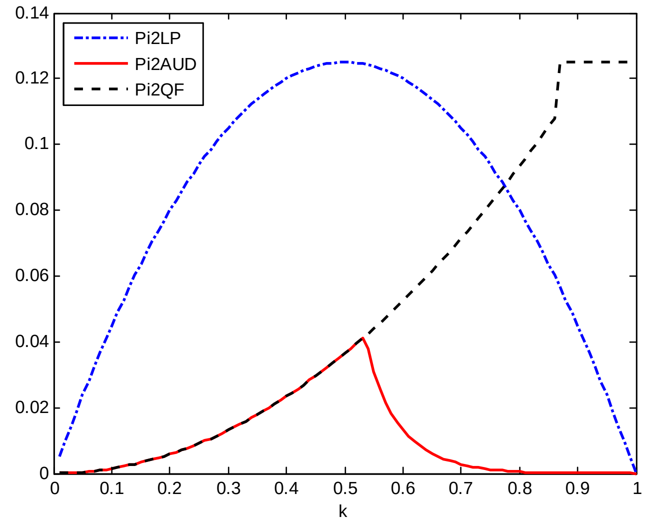


Figure 7: Firm 2's Profits

As shown in Figure 7, firm 2's profit is reduced dramatically when firm 1 adopts the AUD, and this is true for the full range of k. So firm 2 gets partially foreclosed by the dominant firm's AUD for all levels of k. This result may raise antitrust concerns when a dominant firm competes against a capacity-constrained competitor and the dominant firm uses the AUD.

However, firm 2's profit curve under the QF crosses its profit curve under LP from below when $k \simeq 0.77$. This implies that the competitive effects of the QF will change as k varies. When k is in the range of low values, the QF will intensify competition and hurt firm 2 as the AUD does; but when k is in the range of high values, the QF will soften competition and firm 2 earns more profit under the QF than under LP. This is in stark contrast with the competitive impacts of the AUD, where it always hurts firm 2 for the whole range of k.

■ **Buyer's Surpluses.** The equilibrium buyer's surpluses under LP, 2PT, AUD, and QF equilibria are shown in Figure 8. Note that BS^{AUD} crosses BS^{LP} from below at $k \simeq 0.23$. So when $k < 0.23$, $BS^{AUD} < BS^{LP}$; when $k \geq 0.23$, $BS^{AUD} \geq BS^{LP}$. This shows two effects of the AUD on the buyer. First, the AUD is a more efficient surplus extraction tool than LP, which in principle hurts the buyer. Second, the adoption of the AUD intensifies competition by pushing firm 2 to set a lower price. As shown in Figure 8, when k is relatively small, the former effect dominates the latter because the competitive pressure from firm 2 is limited due to its small capacity; when k is relatively large, the latter effect dominates the former for more intensified competition becomes significant when firm 2's capacity is large.

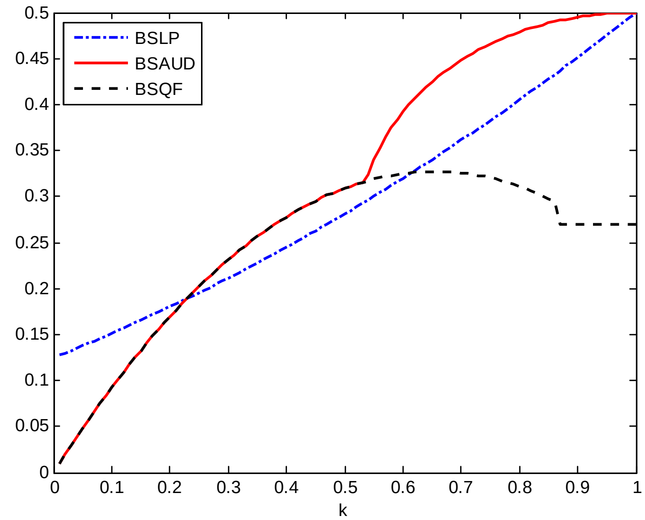


Figure 8: Buyer's Surpluses

The case for the QF is more complicated. BS^{QF} crosses BS^{LP} twice——one from below at $k \simeq 0.23$ as BS^{AUD} does, the other from above at $k \simeq 0.62$. This is because of the two different impacts of the QF. First, it has the feature of the AUD in extracting surplus more efficiently through the fixed fee along with the minimum quantity requirement, which may provoke more aggressive response from the follower. Second, the QF's commitment power of restricting its supply level by setting the quantity target strategically low helps it to soften the rival's second-mover advantage. When k is in the range of low values, restricting supply to induce favorable response becomes secondary because the second-mover advantage is diluted given its limited capacity. The QF will work more as a surplus extraction instrument. In this way, the buyer gets worse off than LP equilibrium. Besides, when k is in the range of high values, the competitive pressure from the follower becomes a major concern of firm 1. The QF will lessen competition by credibly committing to a limited supply without worrying about the constraint (C2) as under the AUD. This limited supply induces the follower to accommodate rather than compete against the leader. Hence, the buyer is worse off, too. Note that the buyer gets hurt in these two end cases, but for different reasons——in the former case, the buyer's surplus is extracted more by the fixed fee; in the latter case, the QF harms the buyer by softening competition and preventing the follower from competing aggressively.

From numerical examples above, we find that when k is relatively small, both the competitor and the buyer are hurt by the dominant firm's adoption of the AUD. This observation appears to be consistent with antitrust concerns put forward in a number of recent cases. Moreover, when k is relatively large, the buyer may not be hurt by the adoption of the AUD in the short run, but the competitor is always partially foreclosed as when k is small. So if there is any fixed costs, such limited profit level as well as not enough growth opportunity under AUD may induce the competitor to exit the market. Hence, the buyer's welfare may get hurt due to the adoption of the AUD by the dominant firm in the long run.

9 Discussions

We now extend our analysis and discuss some assumptions of the model. In the first subsection, we illustrate that our results are robust when there is a cost differential between firms. This suggests that the AUD can be an effective tool to squeeze firm 2's profit even when firm 2 is more efficient up to its capacity level. In the second subsection, we consider a game in which two firms make offers simultaneously. In the third subsection, we offer some thoughts on whether our results would be affected if firm 2's feasible contract set is expanded.

■ **Differential Marginal Costs.** We generalize our analysis of AUD and QF to allow firm 2 to be more efficient up to its capacity level. The major findings still hold when the difference of the two marginal costs is not too large.

Suppose firm 2's marginal cost c_2 is no higher than firm 1's marginal cost c_1, i.e., $c_2 \leq c_1$. We adapt Assumption 2 to $k < q(c_1)$. This means that firm 2 cannot serve the whole demand of the buyer when firm 1 undercuts to its marginal cost c_1. Denote firm 2's profit as $\pi(p; c_2) \equiv (p - c_2) \cdot q(p)$ and its monopoly price as $p^m(c_2) \equiv \max_p \pi(p; c_2)$. Corresponding to Assumption 5, here we assume $(p - c_2) \cdot \pi'(p; c_2)$ to be single-peaked in $(c_2, p^m(c_2))$. The following proposition shows that our analysis of QF and AUD works with differential marginal costs when the difference is not too large.

Proposition 5 (With More Efficient Rival) *Suppose*

$$c_2 \leq c_1 < p^m(c_2) \text{ and } k + (c_1 - c_2) \cdot q'(c_1) > 0, \tag{18}$$

the QF and AUD equilibrium outcomes are, respectively, characterized by Propositions 3 and 4, with adaptations of $c = c_2$ in (14), (14'), (15), (FOC-R) being replaced by

$$q(p_2) - Q = k + Q - q(x) + \frac{c_1 - c_2}{p_2 - c_2} \cdot k, \tag{19}$$

and (FOC-L) being replaced by

$$(x - c_1) \cdot \pi'(x; c_2) = (p_2 - c_2) \cdot \pi'(p_2; c_2). \tag{20}$$

Even when facing a more efficient rival up to its capacity limit, as long as the rival's cost advantage is within a certain range, both QF and AUD are effective competition instruments to improve firm 1's profit over LP and a 2PT. Other comparative statics also hold. Particularly, the buyer's surplus under AUD or QF is lower than the one under LP when k is relatively small. Proposition 5 implies that, even when facing a more efficient, capacity-constrained competitor, the AUD may still lead to a partial foreclosure of the competitor (and full foreclosure is likely if there are fixed costs).

■ **Simultaneous Move.** In our model, the major concern of firm 1 as a first-mover is possible price undercutting from firm 2, whereas firm 2 is immune from the undercutting threat once firm 1 has committed its offer in the first stage.

With a simultaneous move, given firm 1's offer, firm 2's best response will be exactly the same as in our previous analysis. However, firm 1 with no capacity constraint will behave as a full-capacity firm 2 in our previous setting. More importantly, firm 1 can use nonlinear contracts such as a 2PT, QF or AUD to extract any incremental surplus. From firm 1's perspective, inducing SS now weakly dominates inducing DS from the buyer. This is because the most firm 1 can extract under DS is its incremental surplus given firm 2's price p_2, but it can get at least the same amount by supplying what firm 2 supplies under DS simply through undercutting p_2 a bit and extracting the buyer's surplus $v(p_2) - [u(\overline{q}(k, p_2)) - p_2 \cdot \overline{q}(k, p_2)]$ using a 2PT, QF or AUD. Such undercutting reasoning drives p_2 towards marginal cost c, and hence firm 1 earns $v(c) - [u(k) - c \cdot k]$, whereas firm 2 gets zero. Therefore, equipped with a 2PT, QF or an AUD, firm 1 will always undercut and maximize the joint profit between it and the buyer. The equilibrium outcome will be efficient, as in the settings of common agency when there is complete information and nonlinear contracts are allowed (see O'Brien and Shaffer (1997)[19], Bernheim and Whinston (1998)[2]).

In the simultaneous move game, firm 1's equilibrium profit $v(c) - [u(k) - c \cdot k]$ when using a 2PT, QF or an AUD is the same as its profit when it moves first and uses a 2PT in Proposition 2. With a sequential move, we have shown that firm 1 can improve its profit over a 2PT by adopting an AUD or a QF, and firm 2 can also earn positive profits. This implies that firm 1 is better off by moving first and hence has incentives to make such a commitment. In addition, in this setting, firm 2 also prefers to be a second mover than moving simultaneously.

■ **Nonlinear Counteroffer.** In the analysis up to this point, we have restricted our attention to the equilibrium when firm 2 can use LP only. This is to capture the fact that small firms in practice usually cannot offer contracts as complicated as offered by a dominant firm. One reason could be that the buyer considers the dominant firm's product as a must-carry one and thus she is reluctant to sign another AUD or QF with a small supplier. Moreover, due to the lack of experience, small firms often don't have sufficient information on market demand compared to the dominant firm. Even if allowing them to offer an AUD or a QF scheme, setting proper threshold requirements and the corresponding payments would be hard for small firms, not to mention their insufficient ability in monitoring and enforcing those complicated nonlinear contracts.

With complete information, when both firms can use nonlinear contracts such as a 2PT, QF or AUD, it is well known in the common agency literature that the equilibrium outcomes are efficient, when two principals can both offer complicated enough nonlinear contracts (see O'Brien and Shaffer (1997)[19], Bernheim and Whinston (1998)[2], and Marx and Shaffer (2004)[14].).[24] This is in stark contrast with our equilibrium outcome when firm 2 can only offer LP——our QF or AUD equilibrium is not efficient. Because the AUD is often observed in practice, our analysis provides a theoretic explanation for its prevalence in the strategic context, complementary to the existing common agency literature.

[24]A formal proof of this result in our setting is available upon request. Besides, when both firms use complex contracts, the surplus division between firm 2 and the buyer is not uniquely determined, with firm 2's profit falls in a range between 0 and $u(q(c)) - u(q(c) - k) - c \cdot k$. Such a multiplicity of surplus division between firm 2 and the buyer could cause mis-coordination or uncertainty for both of them.

10 Conclusion

The use of AUD pricing schemes by a dominant firm has become a hotly debated topic in antitrust economics and competition policy enforcement. A key feature in some of the antitrust cases involving AUD pricing schemes is that a dominant firm's competitors often have limited capacity of production and thus cannot economically match the dominant firm's AUD offer to serve a customer's whole demand requirement. Although the existing literature has thus far focused on interpreting AUD schemes as a price discrimination tool, investment incentive program, or rent-shifting tool, the antitrust concerns on the AUD are often on its plausible exclusionary effects.

In absence of asymmetric information, downstream competition, or contract externality, we establish strategic effects of AUD and its variations, such as the QF scheme, when a dominant firm competes against an equally efficient (or more efficient) but capacity-constrained competitor. In our setting, we find that the dominant firm is able to use AUD or QF to partially foreclose its competitor's access to the otherwise would-be contestable portion of the market, when the competitor's capacity is limited. Essentially, the dominant firm can use volume-threshold based pricing schemes, such as AUD and QF, to leverage its market power from its captive portion of the market to the contestable portion. Our finding appears to be consistent with the following logic by the European Commission:

> Intel is an unavoidable trading partner. The rebate therefore enables Intel to use the inelastic or "non-contestable" share of the demand of each customer, that is to say the amount that would anyhow be purchased by the customer from the dominant undertaking, as leverage to decrease the price of the elastic or "contestable" share of demand market to lower the price in the contestable market, that is to say the amount for which the customer may prefer and be able to find substitutes.
>
> ——Intel (Case COMP/C-3/37.990), Commission Decision of 13 May 2009 D(2009) 3726 Final

Moreover, we find that when the rival's capacity level is in the range of low values, AUD and QF have the same foreclosure effect; however, when the rival's capacity is in the range of high values, the QF has an additional, softening competition effect.

Appendix

Proof of Proposition 1. First, firm 2 must set $p_2 \leq p_1$ unless $p_1 < c$, because otherwise firm 2 would have no sales and thus zero profit. But $p_1 < c$ is ruled out as it gives firm 1 negative profit.

Second, firm 1 must set $p_1 < u'(k)$. This is because the buyer always buys k units from firm 2 at p_2 first due to $p_2 \leq p_1$, and $u'(k) \leq p_1$ would result in no sale for firm 1.

Hence, with $p_2 \leq p_1 < u'(k)$, the buyer buys k from firm 2 at p_2 and $q(p_1) - k$ from firm 1 at p_1. Firm 2's profit is $(p_2 - c) \cdot k$ and firm 1's profit is $(p_1 - c) \cdot [q(p_1) - k]$. It is easy to see that firm 2 must set $p_2 = p_1$ and firm 1 will set $p_1 = \overline{p}$ with $\pi'(\overline{p}) = k$. The existence and uniqueness of the equilibrium follows from the concavity of $\pi(p)$ and the fact that $\pi'(c) = q(c) > k$ and $\pi'(p^m) = 0 \leq k$. ■

Proof of Proposition 2. First, in the 2PT equilibrium, firm 1 must ensure that the buyer accepts the 2PT. Then firm 2 must set $p_2 \leq p_1$ unless $p_1 < c$, because otherwise firm 2 would have no sales and thus zero profit. But $p_1 < c$ is ruled out, because if so then $\pi_1 = T_1 + (p_1 - c) \cdot q(p_1)$ such that the buyer is better off by buying from firm 1 only, i.e., $v(p_1) - T_1 \geq u(k) - c \cdot k$. But then $\pi_1 \leq v(p_1) + (p_1 - c) \cdot q(p_1) - [u(k) - c \cdot k]$, which is maximized at $p_1 = c$.

By the same argument in the proof of Proposition 1, we must have $p_2 \leq p_1 < u'(k)$, and firm 2 must set $p_2 = p_1$. Then $\pi_1 = T_1 + (p_1 - c) \cdot [q(p_1) - k]$ subject to $v(p_1) - T_1 \geq u(k) - p_1 \cdot k$. So $\pi_1 \leq v(p_1) + (p_1 - c) \cdot q(p_1) - u(k) + c \cdot k$, which is maximized at $p_1 = c$. The claim follows. ■

Proof of Lemma 1. (i) Under QF, if $q_1 \neq Q$, then $\pi_1 = 0$. Thus, we must have $q_1 = Q$. Under AUD, if $q_1 < Q$, then $q_1 = 0$ and $\pi_1 = 0$ because $p_o = \infty$; if $q_1 > Q$, then it is equivalent to LP (p_1) vs. LP (p_2), and the AUD cannot improve firm 1's profit.

In the following, we use $T = p_1 \cdot Q$ under AUD.

We now show $Q < q(c)$. Suppose not, i.e., $u'(Q) \leq c$. Then under DS, firm 2 would have no sales, and it would try its best to undercut until c in order to induce SS, if possible. To ensure the buyer meets Q, firm 1 must make $u(Q) - T \geq u(k) - ck$, i.e., $T \leq u(Q) - [u(k) - ck]$. Thus,

$$
\begin{aligned}
\pi_1 &= T - cQ \\
&\leq u(Q) - cQ - [u(k) - ck] \\
&\leq v(c) - [u(k) - ck] = \pi_1^{2PT}.
\end{aligned}
$$

So in order to have a strictly profitable improvement over a 2PT, we must have $c < u'(Q)$.

(ii) $Q < q(p_2)$ follows from the fact that $c < u'(Q)$ and $u'(Q) \leq p_2$ would result in no sales for firm 2.

We now show $q(p_2) < Q+k$. Suppose not, i.e., $p_2 \leq u'(Q+k)$. It follows that $\pi_2 = (p_2 - c)k \leq [u'(Q+k) - c]k$. Then firm 2 can always increase its profit without losing any sales, as long as $p_2 < u'(Q+k)$. Next, we rule out the case of $p_2 = u'(Q+k)$. Suppose $BS_D^{AUD}(u'(Q+k)) = BS_D^{QF}(u'(Q+k)) \geq$

$BS_S(u'(Q+k))$, i.e., $u(Q+k) - T - u'(Q+k)k \geq u(k) - u'(Q+k)k$. So $T \leq u(Q+k) - u(k)$. Then

$$\begin{aligned}\pi_1 &= T - cQ \\ &\leq u(Q+k) - u(k) - cQ \\ &\leq v(c) - [u(k) - ck] = \pi_1^{2PT}.\end{aligned}$$

For $\pi_1 > \pi_1^{2PT}$, we must have $BS_D^{AUD}(u'(Q+k)) = BS_D^{QF}(u'(Q+k)) < BS_S(u'(Q+k))$, but then the buyer would choose SS. Thus, in order to induce the buyer to choose DS, firm 1 has to ensure $u'(Q+k) < p_2$. ■

Proof of Lemma 2. It is easy to see $p_2 \leq p_1$ under AUD, because otherwise firm 2 would have no sales under DS. Here we only show $p_1 < u'(k)$, based on the idea that otherwise the buyer would choose SS when $p_2 \leq w$.

From (4) and (6), both BS_S and BS_D^{AUD} weakly decrease with p_2, and BS_S as a function of p_2 is everywhere no flatter than BS_D^{AUD} curve. Next, we show that if $u'(\dot{k}) \leq p_1$, then $BS_S(w) \geq BS_D^{AUD}(w)$, which implies $BS_S(p_2) \geq BS_D^{AUD}(p_2)$ for $p_2 \leq w$. That is, if $u'(\dot{k}) \leq p_1$, then the buyer always chooses SS when $p_2 \leq w$. Therefore, we must have $p_1 < u'(k)$ under AUD.

Suppose $u'(\dot{k}) \leq p_1$. When $q(w) \leq k$, $BS_S(w) = v(w)$, $BS_D^{AUD}(w) = v(w) + (w - p_1) \cdot Q$. Because $w \leq p_1$, we have $BS_S(w) \geq BS_D^{AUD}(w)$. When $k < q(w)$, we must have $w = u'(Q)$ because $u'(\dot{k}) \leq p_1$, thereby $k < Q$.

$$\begin{aligned}BS_S(w) &= u(k) - u'(Q) \cdot k \\ &> u(Q) + u'(k) \cdot (k - Q) - u'(Q) \cdot k \\ &= u(Q) + [u'(k) - u'(Q)] \cdot k - u'(k) \cdot Q \\ &> u(Q) - p_1 Q \\ &= BS_D^{AUD}(w),\end{aligned}$$

where the first inequality follows from $u''(\cdot) < 0$ and the second inequality follows from $k < Q$ and $u'(\dot{k}) \leq p_1$. ■

Proof of Lemma 3. First, we show that $BS_D^{QF}(u'(Q)) > BS_S(u'(Q))$. Suppose not. Then $BS_D^{QF}(u'(Q)) \leq BS_S(u'(Q))$ implies $BS_D^{QF}(p_2) \leq BS_S(p_2), \forall p_2 \leq u'(Q)$, because $\partial BS_S/\partial p_2 \leq \partial BS_D^{QF}/\partial p_2 \leq 0$. By Lemma 1, $p_2 < u'(Q)$, and hence the buyer would choose SS from firm 2. Thus, in order to induce DS, we must have $BS_D^{QF}(u'(Q)) > BS_S(u'(Q))$.

Recall from the proof of Lemma 1 that $BS_D^{QF}(u'(Q+k)) < BS_S(u'(Q+k))$. Combining it and $BS_D^{QF}(u'(Q)) > BS_S(u'(Q))$ with $\partial BS_S/\partial p_2 \leq \partial BS_D^{QF}/\partial p_2 \leq 0$, the unique intersection, and its determination (11), follow. ■

Proof of Lemma 4. First, in QF equilibrium, $x \leq p_2 < u'(Q)$. The first inequality holds because otherwise the buyer would SS and firm 1 would have no sale. The second inequality follows from Lemma 1.

To ensure firm 2 chooses p_2 s.t. $x \leq p_2 < u'(Q)$, we must have (13). Note that firm 2's profit has a drop at $p_2 = x$, i.e., $(x - c) \cdot \overline{q}(k, x) > (x - c) \cdot [q(x) - Q]$. In order to have (13), we must have the optimal p_2 to

$\max_{x \le p_2 < u'(Q)} (p_2 - c)[q(p_2) - Q]$ as an interior solution, i.e., $x < p_2 < u'(Q)$. The first-order condition for an interior solution satisfies (15).

Meanwhile, $x < p_2$ and (15) imply that $\pi'(x) > Q$. Hence, $\frac{d[(p_2-c)\cdot\overline{q}(k,p_2)]}{dp_2}|_{p_2=x} > 0$ and thus $\max_{p_2<x}(p_2 - c)\cdot\overline{q}(k, p_2) = (x - c)\cdot\overline{q}(k, x)$.

Next, we show that (14) holds in equilibrium. Using (12),

$$\pi_1 = T - c \cdot Q = (x - c)Q + v(x) - [u(\overline{q}(k, x)) - x \cdot \overline{q}(k, x)],$$

$$\frac{\partial \pi_1}{\partial x} = Q + \overline{q}(k, x) - q(x) > 0,$$

where the inequality follows from $u'(Q + k) < x$ and $Q > 0$. Consequently, as long as (13) is not binding, π_1 can always be increased by increasing x. ■

Proof of Lemma 5. (i) From Assumption 3, $p_2'(Q) < 0$ follows from (15).

(ii) From (14),

$$x(Q) = \begin{cases} \frac{h(Q)}{k} + c & \text{if } \frac{h(Q)}{k} + c < u'(k) \\ \pi^{-1}(h(Q)) & \text{if } p^m \ge \pi^{-1}(h(Q)) \ge u'(k) \end{cases}.$$

Note that $x'(Q) < 0$ follows directly from $h'(Q) < 0$.

When $k \le q^m$, $u'(k) \ge p^m = \pi^{-1}(h(0)) \ge \pi^{-1}(h(Q))$, and hence $x(Q) = \frac{h(Q)}{k} + c$ for any $Q \in [0, q(c)]$.

When $k > q^m$, $\pi^{-1}(h(0)) = p^m > \frac{\pi(p^m)}{k} + c = \frac{h(0)}{k} + c$. Note that $\frac{h(\widehat{Q}_k)}{k} + c = u'(k) = \pi^{-1}(h(\widehat{Q}_k))$, and

$$\frac{d[\pi^{-1}(h(Q))]}{dQ}\Big|_{Q=\widehat{Q}_k} = \frac{h'(\widehat{Q}_k)}{\pi'(u'(k))} < \frac{h'(\widehat{Q}_k)}{k} = \frac{d[\frac{h(Q)}{k} + c]}{dQ}\Big|_{Q=\widehat{Q}_k},$$

where the inequality follows from $h'(Q) < 0$ and $\pi'(u'(k)) < k$. These imply that $\pi^{-1}(h(Q))$ crosses $\frac{h(Q)}{k} + c$ from above only once at $\widehat{Q}_k$. That is, $\pi^{-1}(h(Q)) > \frac{h(Q)}{k} + c > u'(k)$ for $Q \in [0, \widehat{Q}_k)$ and $\pi^{-1}(h(Q)) < \frac{h(Q)}{k} + c < u'(k)$ for $Q \in (\widehat{Q}_k, q(c))$. So $x(Q) = \pi^{-1}(h(Q))$ for $Q \in [0, \widehat{Q}_k)$, $x(Q) = \frac{h(Q)}{k} + c$ for $Q \in (\widehat{Q}_k, q(c))$, and there is a kink at $\widehat{Q}_k$.

(iii) When $k > q^m$, it is easy to see $\pi_1^{QF}(Q)$ is continuous at $\widehat{Q}_k$. When $Q > \widehat{Q}_k$, $x'(Q) = \frac{h'(Q)}{k}$. So $\lim_{Q\searrow\widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ} = u'(k) - c + \widehat{Q}_k \cdot \frac{h'(\widehat{Q}_k)}{k}$. When $Q \le \widehat{Q}_k$, $x'(Q) = \frac{h'(Q)}{\pi'(\pi^{-1}(h(Q)))}$. So $\lim_{Q\nearrow\widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ} = u'(k) - c + \widehat{Q}_k \cdot \frac{h'(\widehat{Q}_k)}{\pi'(u'(k))}$. As $\pi'(u'(k)) < k$, we have $\lim_{Q\searrow\widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ} > \lim_{Q\nearrow\widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ}$. ■

Proof of Proposition 3. Here we characterize the equilibrium for two cases of k. In each case, we first show the existence, uniqueness and sufficiency of (FOC-R) and (FOC-L), and then we prove the solution to them satisfy the corresponding constraints. Last, we show the existence and uniqueness of the cutoff $\widehat{k}$.

Part (A): $k \le q^m$**,**

Step 1: Existence, Uniqueness and Sufficiency of (FOC-R)

In this case, $k < q(x)$ all the time. Thus, $x(Q) = \frac{h(Q)}{k} + c$ for all $Q \in [0, q(c)]$. Note that

$$\frac{d\pi_1^{QF}}{dQ} = \frac{p_2 - c}{k} \cdot \varphi_R(Q), \tag{21}$$

where $\varphi_R(Q) \equiv q(x) + q(p_2) - 2Q - k$.

Let $Q_k \equiv \pi'(p_k)$ with p_k satisfying $k + (p_k - c) \cdot q'(p_k) = 0$. Next we show that $\varphi'_R(Q) < 0$ for $Q \in [Q_k, q(c)]$ and there exists a unique $Q_R(k) \in [Q_k, q(c)]$ s.t. $\varphi_R(Q_R(k)) = 0$. These together imply that π_1^{QF} is single-peaked in Q, and $Q_R(k)$ is the unique optimal solution.

$$\begin{aligned}\varphi'_R(Q) &= q'(x) \cdot x'(Q) + q'(p_2) \cdot p'_2(Q) - 2 \\ &= q'(x) \cdot (-\frac{p_2 - c}{k}) + q'(p_2) \cdot \frac{1}{\pi''(p_2)} - 2 \\ &= [\frac{q'(p_2)}{\pi''(p_2)} - 1] - \frac{k + (p_2 - c) \cdot q'(x)}{k},\end{aligned}$$

where the second equality follows from (14) and (15).

From Assumption 4, $\pi''(p_2) = 2q'(p_2) + (p_2 - c)q''(p_2) < q'(p_2) < 0$. Hence, $0 < \frac{q'(p_2)}{\pi''(p_2)} < 1$. For $Q > Q_k$, we have $p_2 = p_2(Q) < p_2(Q_k) = p_k$, thereby $k + (p_2 - c) \cdot q'(p_2) > 0$. From (14), it follows that $x < p_2$ for $Q > Q_k$. Hence, $k + (p_2 - c) \cdot q'(x) \geq k + (p_2 - c) \cdot q'(p_2) > 0$ for $Q > Q_k$. As a result, we have $\varphi'_R(Q) < 0$ for $Q \in [Q_k, q(c)]$.

Last, we show that $\varphi_R(Q)$ does cross zero from above. At $Q = q(c)$, $p_2(q(c)) = x(q(c)) = c$, thus $\varphi_R(q(c)) = -k < 0$. At $Q = Q_k$, $x = p_2 = p_k$, thus $\varphi_R(Q_k) = 2 \cdot [q(p_k) - Q_k] - k = k \geq 0$ with "=" only if $k = 0$. So there exists a unique $Q_R(k) \in [Q_k, q(c))$ s.t. $\varphi_R(Q_R(k)) = 0$.

Step 2: Check Constraints $u'(Q + k) < x < p_2 < u'(Q)$

Note that $x < p_2$ has been shown in Step 1 for $Q > Q_k$, and that $p_2 < u'(Q)$ follows from (15). Moreover, $u'(Q_R + k) < x$ follows from (FOC-R), because $Q_R(k) + k - q(x) = q(p_2) - Q_R(k) > 0$ due to $p_2 < u'(Q_R)$.

Part (B): $k > q^m$,

In this case, we need to consider both $Q > \widehat{Q}_k$ (or $k < q(x)$) and $Q \leq \widehat{Q}_k$ (or $k \geq q(x)$). From Lemma 5, $x(Q) = \pi^{-1}(h(Q))$ for $Q \leq \widehat{Q}_k$, $x(Q) = \frac{h(Q)}{k} + c$ for $Q > \widehat{Q}_k$.

Step 1: Existence, Uniqueness and Sufficiency of (FOC-R) and (FOC-L)

Case (i): $Q > \widehat{Q}_k$

By the same argument in Part (A), we can show that $\varphi'_R(Q) < 0$ for $Q \in [0, q(c)]$. Note that Q_k defined when $k \leq q^m$ does not exist when $k > q^m$, because now for all $Q \in [0, q(c)]$, $k + (p_2 - c) \cdot q'(p_2) \geq k + (p^m - c) \cdot q'(p^m) > \pi'(p^m) = 0$. Assumption 4 and $p_2(Q) \leq p^m$ yield the first inequality, and the second inequality follows from $k > q^m$.

At $Q = 0$, $p_2(0) = p^m$ and $x(0) > c$, thus $\varphi_R(0) = q(x) + q(p^m) - k > 0$, which follows from $x < p^m$ and $2q(p^m) \geq q(c) > k$. Here $2q(p^m) \geq q(c)$ is the result of $q''(p) \leq 0$. So combining with the fact that $\varphi_R(q(c)) = -k < 0$, the existence of $Q_R(k)$ follows.

Case (ii): $Q \leq \widehat{Q}_k$

Let $g(p) \equiv (p - c)\pi'(p)$. Assumption 5 says $g(p)$ is single-peaked. Let the peak is reached at $\widetilde{p}$, where $\widetilde{p} < p^m$. Note that

$$\frac{d\pi_1^{QF}}{dQ} = \frac{1}{\pi'(x)} \cdot \varphi_L(Q),$$

where $\varphi_L(Q) \equiv g(x) - g(p_2)$.

Next we show that $\varphi'_L(Q)\,|_{\varphi_L(Q)=0} < 0$ and there exists a unique Q_L s.t. $\varphi_L(Q_L) = 0$ in the relevant range. This implies that $\varphi_L(Q)$ crosses zero from above only once, and Q_L is the unique optimal solution.

Note that (14): $\pi(x) = \pi(p_2) - (p_2 - c) \cdot \pi'(p_2)$. Here $\pi(x)$ is increasing in x for $x < p^m$. Moreover, $\pi(c) = 0 < \pi(p_2) - (p_2 - c) \cdot \pi'(p_2)$ for $c < p_2$, and $\pi(p_2) > \pi(p_2) - (p_2 - c) \cdot \pi'(p_2)$ for $c < p_2 < p^m$. Hence, for any $c < p_2 < p^m$, there exists a unique $x(p_2) \in (c, p_2)$ s.t. (14) holds.

At $\varphi_L(Q) = 0$, we have $g(x) = g(p_2)$. From (14), we have $c < x < p_2 < p^m$. Assumption 5 implies $x < \widetilde{p} < p_2$, thereby $g'(x) > 0 > g'(p_2)$. Note that

$$\varphi'_L(Q) = [g'(x) \cdot x'(p_2) - g'(p_2)] \cdot p'_2(Q).$$

Thus, $\varphi'_L(Q)\,|_{\varphi_L(Q)=0} < 0$ follows from $x'(p_2) = \frac{-(p_2-c)\pi''(p_2)}{\pi'(x)} > 0$ and $p'_2(Q) = \frac{1}{\pi''(p_2)} < 0$.

Last, we show that $\varphi_L(Q)$ does cross zero from above.

Let $\widetilde{Q} \equiv \pi'(\widetilde{p}) \in (0, q(c))$. At $\widetilde{Q}$, $p_2(\widetilde{Q}) = \widetilde{p} > x_L(\widetilde{Q})$, thus $\varphi_L(\widetilde{Q}) = g(x) - g(\widetilde{p}) < 0$. As $Q \searrow 0$, $x_L(Q) \nearrow p^m$, $p_2(Q) \nearrow p^m$, and $\widetilde{p} < x(Q) < p_2(Q) < p^m$, thus $\varphi_L(0) = g(x) - g(p_2) > 0$. So there exists a unique $Q_L \in (0, \widetilde{Q})$ s.t. $\varphi_L(Q_L) = 0$.

Step 2: When $Q_L \leq \widehat{Q}_k$ and When $\widehat{Q}_k < Q_R(k)$

In this part, we show that there exists a α s.t. $Q_L \leq \widehat{Q}_k$ when $\alpha \leq k$; there exists a β s.t. $\widehat{Q}_k < Q_R(k)$ when $k < \beta$. Moreover, $q^m < \alpha < \beta$.

Recall that for $k > q^m$, $\widehat{Q}_k$ is given by $h(\widehat{Q}_k) = \pi(u'(k))$, and $\widehat{Q}_k$ is increasing in k. At $k = q^m$, $\widehat{Q}_k = 0$. At $k = q(c)$, $\widehat{Q}_k = q(c)$. Note that Q_L is independent of k, and $Q_L \in (0, \widetilde{Q})$, where $\widetilde{Q} \in (0, q(c))$. Thus, there must exist a $\alpha \in (q^m, q(c))$ s.t. $Q_L \gtreqless \widehat{Q}_k$ for $k \lesseqgtr \alpha$, and $q^m < \alpha$.

Note that

$$\varphi'_R(Q) \cdot Q'_R(k) = 1 - q'(x) \cdot \frac{\partial x}{\partial k}$$
$$\therefore \varphi'_R(Q) \cdot Q'_R(k) = \frac{k + (x - c)q'(x)}{k},$$

where the second equality follows from (14). Recall that in Step 1, we have shown that $\varphi'_R(Q) < 0$ and $k + (p_2 - c)q'(x) > 0$ for all $Q \in [0, q(c)]$. From the fact that $k + (x - c)q'(x) > k + (p_2 - c)q'(x) > 0$, we know that $Q'_R(k) < 0$. At $k = q^m$, $Q_R(k) = \pi'(p_2(k)) > 0 = \widehat{Q}_k$ as $p_2(k) < p^m$. At $k = q(c)$, $Q_R(k) < q(c) = \widehat{Q}_k$. Thus, there must exist a $\beta \in (q^m, q(c))$ s.t. $Q_R(k) \gtreqless \widehat{Q}_k$ for $k \lesseqgtr \beta$.

Evaluating $\varphi_R(Q)$ evaluating at Q_L :

$$\begin{aligned}
\varphi_R(Q_L) &= q(x) + q(p_2) - 2Q_L - k \\
&= q(p_2) - 2\pi'(p_2) \text{ } (\because x(Q_L) = u'(\alpha) \text{ and } Q_L = \pi'(p_2)) \\
&= \pi'(p_2) \cdot [\frac{q(p_2)}{\pi'(p_2)} - 2] \\
&= \pi'(p_2) \cdot [\frac{q(x)}{\pi'(x)} - 1] \\
&> 0,
\end{aligned}$$

where the fourth equality follows from (14) and (FOC-L). Recall that $\varphi'_R(Q) < 0$ for $Q \in [0, q(c)]$ when $k > q(p^m)$. Thus, at $k = \alpha$, $Q_R(\alpha) > Q_L = \widehat{Q}_k$. By definition of β, we have $\alpha < \beta$.

Step 3: Check Constraints $u'(Q+k) < x < p_2 < u'(Q)$

$x < p_2$ has been shown in Step 1. $p_2 < u'(Q)$ follows from (15). The only constraint that remains to be checked is $u'(Q+k) < x$.

When $\alpha < k$, $0 < Q_L < \widehat{Q}_k$ and $u'(k) < x$. So $u'(Q_L + k) < u'(k) < x$.

When $k < \beta$, (FOC-R) yields $Q_R(k) + k - q(x) = q(p_2) - Q_R(k) > 0$ because $p_2 < u'(Q_R)$. Thus, $u'(Q_R + k) < x$.

Part (C): Existence and Uniqueness of a Cutoff $\widehat{k}$

When $k \leq q^m$, $Q_R(k)$ is the unique equilibrium solution.

When $k > q^m$, recall that in Case (i), $\pi_1^{QF} = v(x) - [u(k) - xk] + (x-c) \cdot Q$, and $x < u'(k)$.

$$\begin{aligned}
\frac{d\pi_1^{QF}}{dk} &= [k + Q - q(x)] \cdot \frac{\partial x}{\partial k} - [u'(k) - x] \\
&= [k + Q - q(x)] \cdot (-\frac{x-c}{k}) - [u'(k) - x] \text{ (By (14))} \\
&= [q(p_2) - Q] \cdot (-\frac{x-c}{k}) - [u'(k) - x] \text{ (By (FOC-R))} \\
&< 0.
\end{aligned}$$

Thus, when $k < \beta$, π_1^{QF} in Case (i) is decreasing in k. Note that, when $k > \alpha$, π_1^{QF} in Case (ii) is independent of k.

Recall that at $k = \alpha$, $Q_R(\alpha) > \widehat{Q}_k = Q_L$, we have $\lim_{Q \nearrow \widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ} = 0 < \lim_{Q \searrow \widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ}$. Thus, $\pi_1^{QF}(Q_R(k)) > \pi_1^{QF}(\widehat{Q}_k) = \pi_1^{QF}(Q_L)$. When $k = \beta$, $Q(\beta) = \widehat{Q}_k > Q_L$, we have $\lim_{Q \nearrow \widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ} < 0 = \lim_{Q \searrow \widehat{Q}_k} \frac{d\pi_1^{QF}}{dQ}$. Thus, $\pi_1^{QF}(Q_L) > \pi_1^{QF}(\widehat{Q}_k) = \pi_1^{QF}(Q_R(k))$.

So we can conclude that there exists a $\widehat{k} \in (\alpha, \beta)$ s.t. $\pi_1^{QF}(Q_R(k)) \gtreqless \pi_1^{QF}(Q_L)$ for $k \lesseqgtr \widehat{k}$. $\widehat{k} > q^m$ follows from $\alpha > q^m$. ■

Proof of Corollary 2. When $k < \widehat{k}$,

$$\begin{aligned}
Q + k &= q(x) + q(p_2) - Q \text{ (By (FOC-R))} \\
&= q(x) - (p_2 - c) q'(p_2) \text{ (By (15))} \\
&\geq q(x) - (x - c) q'(x) \\
&\geq q(c),
\end{aligned}$$

where the first inequality is from $(p_2 - c)q'(p_2)$ decreases with p_2 for $c \leq p_2$ and $x \leq p_2$, and the second inequality follows from $q(x) - (x-c)q'(x)$ is weakly increasing in x for $c \leq x$. Note that "=" occurs only when $x = p_2 = c$, that is, only when $k = 0$.

When $k \geq \widehat{k}$, we have $x \geq u'(k)$, i.e., $k \leq q(x)$. Thus if we can show $Q + q(x) > q(c)$, then

$Q + k > q(c)$ follows.

$$\begin{aligned} Q + q(x) - q(c) &= \pi'(p_2) + q(x) - q(c) \text{ (By (15))} \\ &= q(p_2) - q(c) + q(x) + (p_2 - c)q'(p_2) \\ &\geq q(x) + 2(p_2 - c)q'(p_2) \text{ } (\because q''(p_2) \leq 0 \text{ and } p_2 \geq c). \end{aligned}$$

Hence, to show $Q + q(x) > q(c)$, it suffices to show that $q(x) + 2(p_2 - c)q'(p_2) > 0$.

Multiplying $q(x) + 2(p_2 - c)q'(p_2)$ by $x - c$,

$$\begin{aligned} & \pi(x) + 2(x - c)(p_2 - c)q'(p_2) \\ = & -(p_2 - c)q'(p_2)[(p_2 - c) - 2(x - c)] \text{ (By (14))}. \end{aligned}$$

So if we can show $\frac{p_2 - c}{x - c} > 2$, then we are done.

From (FOC-L), $\frac{p_2 - c}{x - c} = \frac{\pi'(x)}{\pi'(p_2)}$. Recall that (14) $\pi(x) = \pi(p_2) - (p_2 - c) \cdot \pi'(p_2)$. So

$$\begin{aligned} (p_2 - c)\pi'(p_2) &= \pi(p_2) - \pi(x) \\ &< \pi'(x)(p_2 - x) \text{ } (\because \pi''(x) < 0 \text{ and } p_2 > x) \\ &= \pi'(x)(p_2 - c) - \pi'(x)(x - c) \\ &= \pi'(x)(p_2 - c) - \pi'(p_2)(p_2 - c) \text{ (By (FOC-L))} \\ &\therefore \text{ } 2(p_2 - c)\pi'(p_2) < (p_2 - c)\pi'(x). \end{aligned}$$

Thus, we have $\frac{p_2 - c}{x - c} = \frac{\pi'(x)}{\pi'(p_2)} > 2$, thereby $Q + k \geq Q + q(x) > q(c)$. ■

Proof of Corollary 3. (i) From Step 2 of the Proof of Proposition 3, we have $Q_R'(k) < 0$ for $k < \beta$. For $k \geq \widehat{k}$, Q_L is independent of k. Step 4 of it shows that at $k = \beta$, $Q_R(k) = \widehat{Q}_k > Q_L$. Therefore, at $k = \widehat{k} < \beta$, $Q_L < Q_R(k)$.

Total output is $q(p_2)$. $p_2(Q)$ decreases with Q from (15). Because $Q^* = Q_R(k)$ first decreases with k, drops at $\widehat{k}$ and then $Q^* = Q_L$ becomes flat for $k > \widehat{k}$, $p_2(k)$ must increase with k, jumps at $\widehat{k}$ and then becomes flat for $k > \widehat{k}$. So total output $q(p_2)$ first decreases with k, drops at $\widehat{k}$ and then becomes flat for $k > \widehat{k}$.

(ii) The comparative statics on $p_2(k)$ is shown in Part (i). Similarly, we can derive the same result for $x(k)$ from (14).

(iii) From Step 4 of the Proof of Proposition 3 π_1^{QF} is decreasing for $k < \widehat{k}$, smooth at $\widehat{k}$ and becomes flat for $k > \widehat{k}$. Because $\pi_2^{QF} = h(Q)$ decreases with Q, this part follows from Part (i).

(iv) $TS^{QF} = u(q(p_2)) - c \cdot q(p_2)$. $\frac{dTS^{QF}}{dp_2} = [u'(q(p_2)) - c] \cdot q'(p_2) < 0$. Then this part follows from the result on $p_2(k)$ of Part (ii). ■

Proof of Lemma 6. (i) When $x < u'(k)$, (11') yields $p_1 \cdot Q = x \cdot Q + v(x) - [u(k) - x \cdot k]$. So

$$\begin{aligned}
[p_1 - u'(k)] \cdot Q &= [x - u'(k)] \cdot Q + v(x) - [u(k) - x \cdot k] \\
&< [x - u'(k)] \cdot Q + [u'(k) - x] \cdot [q(x) - k] \\
&= [u'(k) - x] \cdot [q(x) - k - Q] \\
&< 0,
\end{aligned}$$

where the concavity of $u(q)$ leads to the first inequality, and the second inequality follows from $u'(Q+k) < x < u'(k)$ in the AUD equilibrium.

(ii) (C2) is $p_2 \leq T/Q$, which can be rewritten as $T - c \cdot Q - (p_2 - c) \cdot Q \geq 0$. Let

$$\begin{aligned}
D(k) &\equiv \pi_1^{AUD}(Q) - (p_2 - c) \cdot Q \\
&= v(x) - [u(k) - x \cdot k] - (p_2 - x) \cdot Q,
\end{aligned}$$

where $(Q(k), p_2(k), x(k))$ is jointly determined by (14'), (15), and (FOC-R). Then, (C2) becomes $D(k) \geq 0$.

Here we first show that $D(0) > 0$ and $D(k) < 0$ for $\gamma \leq k$, then we prove $D(k)$ decreases with k for $k \leq \gamma$. Hence, we conclude with the existence of $\overline{k} \in (0, \gamma)$ s.t. $D(k) \gtreqless 0$ for $k \lesseqgtr \overline{k}$.

Step 1: $D(0) > 0$ *and* $D(k) < 0$ *for* $\gamma \leq k$.

When $k = 0$, $x = p_2 = c$. $D(0) = v(c) > 0$.

By Corollary 3, $p_2(k)$ increases with k. From the concavity of $u(q)$, $u'(k)$ decreases with k. $p_2(0) = c < u'(0)$, $p_2(\beta) > x(\beta) = u'(\beta)$. Thus, there exists a unique $\gamma \in (0, \beta)$ s.t. $p_2(\gamma) = u'(\gamma)$.

$$\begin{aligned}
D(k) &= v(x) - [u(k) - x \cdot k] - (p_2 - x) \cdot Q \\
&< [u'(k) - x] \cdot [q(x) - k] - (p_2 - x) \cdot \pi'(p_2) \text{ ($\because u''(q) < 0$ and $x < u'(k)$)} \\
&= [u'(k) - x] \cdot [\pi'(p_2) + (p_2 - c) \cdot q'(p_2)] - (p_2 - x) \cdot \pi'(p_2) \text{ (by (FOC-R))} \\
&= [u'(k) - p_2] \cdot \pi'(p_2) + [u'(k) - x] \cdot (p_2 - c) \cdot q'(p_2) \\
&< 0
\end{aligned}$$

By definition of β and γ, for $\gamma \leq k \leq \beta$, $p_2 \geq u'(k) \geq x$. So the last inequality follows. That is, $D(k) < 0$ for $\gamma \leq k \leq \beta$.

For $\beta < k$, $p_2 > x > u'(k)$.

$$\begin{aligned}
D(k) &< [u'(k) - p_2] \cdot \pi'(p_2) + [u'(k) - x] \cdot (p_2 - c) \cdot q'(p_2) \\
&= -\{[p_2 - u'(k)] \cdot \pi'(p_2) - [x - u'(k)] \cdot [-(p_2 - c) \cdot q'(p_2)]\}.
\end{aligned}$$

If we can show that $[p_2 - u'(k)]/[x - u'(k)] > [-(p_2 - c) \cdot q'(p_2)]/\pi'(p_2)$, then $D(k) < 0$.

Because $u'(k) > c$ and $p_2 > x$, $[p_2 - u'(k)]/[x - u'(k)] > (p_2 - c)/(x - c)$. Thus, to show $D(k) > 0$, it suffices to show

$$\frac{p_2 - c}{x - c} \geq \frac{-(p_2 - c) \cdot q'(p_2)}{\pi'(p_2)},$$

which is reduced to

$$\pi'(p_2) + (x - c) \cdot q'(p_2) \geq 0.$$

$$\begin{aligned} \pi'(p_2) + (x - c) \cdot q'(p_2) &= q(x) - k - (p_2 - x) \cdot q'(p_2) \text{ (by (FOC-R))} \\ &\geq (p_2 - x) \cdot [-q'(p_2)] - [x - u'(k)] \cdot [-q'(x)] \text{ } (\because q''(p) \leq 0) \end{aligned}$$

Because $q''(p) \leq 0$ and $p_2 > x$, $-q'(p_2) \geq -q'(x)$. If we can show that $p_2 - x > x - u'(k)$, then $\pi'(p_2) + (x - c) \cdot q'(p_2) \geq 0$.

Because $u'(k) > c$,

$$p_2 - x > x - c \tag{22}$$

would imply $p_2 - x > x - u'(k)$.

(14') and (FOC-R) together imply (22) as follows. (14') gives

$$\begin{aligned} \frac{p_2 - c}{x - c} &= \frac{k}{-(p_2 - c) \cdot q'(p_2)} \\ &= \frac{q(x) - q(p_2) - 2 \cdot (p_2 - c) \cdot q'(p_2)}{-(p_2 - c) \cdot q'(p_2)} \text{ (by (FOC-R))} \\ &= 2 + \frac{q(x) - q(p_2)}{-(p_2 - c) \cdot q'(p_2)} \\ &> 2 \text{ } (\because x < p_2) \end{aligned}$$

Hence, $D(k) < 0$ for $\beta < k$. This completes Step 1.

Step 2: $D'(k) < 0$ for $k \leq \gamma$.

Using (FOC-R),

$$\begin{aligned} D'(k) &= \frac{\partial D}{\partial k} + \frac{\partial D}{\partial x} \cdot \frac{\partial x}{\partial k} + \frac{\partial D}{\partial p_2} \cdot p_2'(k) \\ &= x - u'(k) + [k + \pi'(p_2) - q(x)] \cdot (-\frac{x - c}{k}) - [\pi'(p_2) + p_2 \pi''(p_2)] \cdot p_2'(k) \\ &= x - u'(k) + [k - q(x)] \cdot p_2'(k) + (p_2 - c) q'(p_2) \cdot \frac{x - c}{k} \\ &\quad + (p_2 - c) \cdot [-\frac{p_2 + c}{p_2 - c} q'(p_2) - p_2 q''(p_2)] \cdot p_2'(k) \text{ (by (FOC-R))} \\ &\leq [k - q(x)] \cdot p_2'(k) + \underbrace{(x - p_2)}_{\langle 1 \rangle} + \underbrace{(p_2 - c) q'(p_2) \cdot \frac{x - c}{k}}_{\langle 2 \rangle} \\ &\quad \underbrace{+ (p_2 - c) \cdot [-\frac{p_2 + c}{p_2 - c} q'(p_2) - p_2 q''(p_2)] \cdot p_2'(k)}_{\langle 3 \rangle}, \end{aligned}$$

where the inequality follows from $p_2 \leq u'(k)$ for $k \leq \gamma$. Note that $[k - q(x)] \cdot p_2'(k)$ is negative, because $x < u'(k)$ for $k \leq \gamma$ and $p_2'(k) > 0$ from Corollary 3. So if we can show $\langle 1 \rangle + \langle 2 \rangle + \langle 3 \rangle$ is negative, then this part is complete.

From (FOC-R),

$$\begin{aligned} p_2'(k) &= \frac{k + (x-c) \cdot q'(x)}{-\pi''(p_2) \cdot [k + (p_2 - c)q'(x)] + k \cdot [-q'(p_2) - (p_2 - c)q''(p_2)]} \\ &< \frac{k + (x-c)q'(x)}{2k + (p_2 - c)q'(x)} \cdot \frac{1}{-q'(p_2) - (p_2 - c)q''(p_2)}, \end{aligned} \tag{23}$$

where the inequality follows from $-\pi''(p) > -q'(p) - (p-c)q''(p)$. It is easy to verify that

$$-\frac{p_2 + c}{p_2 - c} q'(p_2) - p_2 q''(p_2) < -q'(p_2) - (p_2 - c)q''(p_2), \tag{24}$$

as it is equivalent to $\pi''(p) < 0$. Hence,

$$\begin{aligned} & \langle 2 \rangle + \langle 3 \rangle \\ < \; & (p_2 - c)q'(p_2) \cdot \frac{x-c}{k} + (p_2 - c) \cdot [-q'(p_2) - (p_2 - c)q''(p_2)] \cdot p_2'(k) \\ < \; & (p_2 - c)q'(p_2) \cdot \frac{x-c}{k} + (p_2 - c) \cdot \frac{k + (x-c)q'(x)}{2k + (p_2 - c)q'(x)} \\ = \; & \frac{(p_2 - c)q'(p_2) \cdot (x-c)2k - (x-c)q'(x) \cdot (x-c)k + k \cdot (p_2 - c) \cdot [k + (x-c) \cdot q'(x)]}{k \cdot [2k + (p_2 - c)q'(x)]} \\ = \; & \frac{(p_2 - c) \cdot [k + 2(x-c)q'(p_2)] + (x-c)q'(x) \cdot (p_2 - x)}{2k + (p_2 - c)q'(x)}, \end{aligned}$$

where the first inequality follows from (24), the second follows by (23), and the last equality is from (11'). Therefore,

$$\begin{aligned} & \langle 1 \rangle + \langle 2 \rangle + \langle 3 \rangle \\ < \; & (x - p_2) + \frac{(p_2 - c) \cdot [k + 2(x-c)q'(p_2)] + (x-c)q'(x) \cdot (p_2 - x)}{2k + (p_2 - c)q'(x)} \\ = \; & \frac{(p_2 + c - 2x)[q(p_2) - q(x)] + 2(p_2 - x) \cdot (p_2 - c)q'(p_2) - (p_2 - x)^2 \cdot q'(x)}{2k + (p_2 - c)q'(x)} \\ \leq \; & \frac{(p_2 + c - 2x)q'(x)(p_2 - x) + 2(p_2 - x) \cdot (p_2 - c)q'(p_2) - (p_2 - x)^2 \cdot q'(x)}{2k + (p_2 - c)q'(x)} \\ = \; & \frac{p_2 - x}{2k + (p_2 - c)q'(x)} \cdot [(p_2 - c)q'(p_2) + (p_2 - c)q'(p_2) - (x-c)q'(x)], \end{aligned}$$

where the first equality follows from (FOC-R), and the second inequality is due to $q''(p) \leq 0$ and $x < p_2$. Indeed, $2k + (p_2 - c)q'(x) > k + (p_2 - c)q'(x) > k + (p_2 - c) \cdot q'(p_2) > 0$, where the second inequality follows from $q''(p) \leq 0$ and $x < p_2$. Because $(p - c)q'(p)$ is decreasing in p, $\forall p > c$, $(p_2 - c)q'(p_2) - (x-c)q'(x) < 0$ as $x < p_2$. Thus, $(p_2 - c)q'(p_2) + (p_2 - c)q'(p_2) - (x-c)q'(x) < 0$, thereby Step 2 is completed.

Step 3: There exists a unique $\overline{k} \in (0, \gamma)$ s.t. $D(k) \gtreqless 0$ for $k \lesseqgtr \overline{k}$.

This follows directly from Steps 1 and 2. ■

Proof of Proposition 4. With Lemmas 2 and 6, we know that when $k < \overline{k}$, the equilibrium outcome (Q, T, p_2) along with threat price x is jointly determined by (11'), (14'), (15), and (FOC-R), with (FOC-R) being replaced by (C2) when $\overline{k} \leq k$. The sufficiency of (FOC-R) has already been shown in the proof of Proposition 3. In the proof of Lemma 6, $u'(Q+k) < x < u'(k)$ ensures that $p_1 = T/Q < u'(k)$. And we know that $u'(Q+k) < x < u'(k)$ is true under (FOC-R) for $k < \overline{k} < \gamma$. So here we only need to show the existence of equilibrium when $\overline{k} \leq k$, and check the constraint $u'(Q+k) < x < u'(k)$.

Step 1: Existence of the Solution to $p_2 = p_1$ when $\overline{k} \leq k$.

Similar to $D(k)$ but without using the equilibrium $Q_R(k)$ from (FOC-R), we can define

$$
\begin{aligned}
d(Q,k) &\equiv \pi_1^{AUD}(Q) - (p_2 - c) \cdot Q \\
&= v(x) - [u(k) - x \cdot k] - (p_2 - x) \cdot Q,
\end{aligned}
$$

where $(p_2(Q), x(Q))$ is jointly determined by (14') and (15). So $D(k) = d(Q_R(k), k)$, and the constraint (C2) $p_2 \leq p_1$ is $d(Q,k) \geq 0$.

When $\overline{k} \leq k$, $d(Q_R(k), k) = D(k) < 0$. At $Q = q(c)$, (14') and (15) lead to $x = p_2 = c$. So $d(q(c), k) = v(c) - [u(k) - c \cdot k] > 0$. From the continuity of $d(Q,k)$, there must exist a $\overline{Q}(k) \in (Q_R(k), q(c))$ s.t. $d(\overline{Q}(k), k) = 0$, i.e., $p_2(\overline{Q}(k)) = p_1(\overline{Q}(k))$.

Step 2: Check Constraints $u'(Q+k) < x < p_2 < u'(Q)$ and $p_1 = T/Q < u'(k)$

Because $\widehat{Q}_k < Q_R(k) < \overline{Q}(k)$ and $x'(Q) < 0$ from Lemma 5, we have $x(\overline{Q}(k)) < x(\widehat{Q}_k) = u'(k)$.

In Step 1 of the proof of Proposition 3, when showing the sufficiency of (FOC-R), we proved $\varphi'_R(Q) < 0$ in the relevant range. $\widehat{Q}(k) < \overline{Q}(k)$ yields $\varphi'_R(\overline{Q}(k)) < 0$, which is

$$
q(p_2) - \overline{Q} < k + \overline{Q} - q(x).
$$

So $k + \overline{Q} - q(x) > q(p_2) - \overline{Q} = -(p_2 - c) \cdot q'(p_2) \geq 0$. $u'(\overline{Q} + k) < x$ is satisfied.

It is easy to see $p_2 < u'(\overline{Q})$ from (15) $\pi'(p_2) = \overline{Q}$.

From the proof of Lemma 6, $u'(\overline{Q} + k) < x < u'(k)$ ensures that $p_1 = T/\overline{Q} < u'(k)$. ■

Proof of Corollary 4. When $k < \overline{k}$, (FOC-R) characterizes the equilibrium solution of Q. So the same argument in Corollary 2 when $k < \widehat{k}$ can be applied.

When $k \geq \widehat{k}$, (FOC-R) is replaced by (C2). So the equilibrium solution $\overline{Q}(k) > Q_R(k)$, where $Q_R(k)$ is characterized by (FOC-R).

$$
\begin{aligned}
\overline{Q} + k &> q(x) + q(p_2) - \overline{Q} \text{ (By } \varphi'_R(Q) < 0 \text{ and } \overline{Q}(k) > Q_R(k)) \\
&= q(x) - (p_2 - c)q'(p_2) \text{ (By (15))} \\
&\geq q(x) - (x - c)q'(x) \\
&\geq q(c),
\end{aligned}
$$

where the first inequality is from $(p_2 - c)q'(p_2)$ decreases with p_2 for $c \leq p_2$ and $x \leq p_2$, and the second inequality follows from $q(x) - (x - c)q'(x)$ is weakly increasing in x for $c \leq x$. ■

Proof of Corollary 6. (i) Note that $p_2(0) = c < p^m = \overline{p}(0)$, and $p_2(\alpha) = u'(\alpha) > \overline{p}(\alpha)$ ($\because \pi'(p) < q(p)$). Moreover, $p_2'(k) > 0$ and $\overline{p}'(k) < 0$. Hence, $\exists$ a unique $k_0 \in (0, \alpha)$ s.t. $p_2(k) \lesseqqgtr \overline{p}(k), \forall k \lesseqqgtr k_0$.

(ii) $q_1^{AUD} = Q > q(c) - k > q(\overline{p}) - k$ follows from Corollary 2. $q_2^{AUD} = q(p_2) - Q < k = q_2^{LP}$ follows from the fact that $u'(Q + k) < p_2$.

(iii) It is obvious that $\pi_1^{AUD} > \pi_1^{LP}$. $\pi_2^{AUD} = (p_2 - c) \cdot [q(p_2) - q_o] < (p_2 - c) \cdot k < (p_2^{LP} - c) \cdot k = \pi_2^{LP}$, where the first inequality follows from the fact that $u'(Q + k) < p_2$ and the second one follows from $p_2^{AUD} < p_2^{LP}, \forall k < \min\{k_0, \overline{k}\}$.

(iv) When $k \to 0$, $BS^{AUD} \to 0$, and $BS^{LP} \to v(p^m) > 0$ as $\overline{p} \to p^m$. From continuity of BS^{AUD} and BS^{LP}, this part follows.

(v) $TS^{LP} = u(q(\overline{p})) - c \cdot q(\overline{p})$. Because $\overline{p}'(k) < 0$ and $q(\overline{p}) < c$, $\frac{dTS^{LP}}{dk} = [u'(q(\overline{p})) - c] \cdot q'(\overline{p}) \cdot \overline{p}'(k) > 0$.

$TS^{AUD} = u(q(p_2)) - c \cdot q(p_2)$. Because $p_2'(k) > 0$ for $k < \overline{k}$, $\frac{dTS^{QF}}{dk} = [u'(q(p_2)) - c] \cdot q'(p_2) \cdot p_2'(k) < 0$ for $k < \overline{k}$.

At $k = 0$, $\overline{p} = p^m > c = p_2$. So $TS^{AUD} = v(c) > u(q^m) - c \cdot q^m = TS^{LP}$. At $k = q(c)$, $\overline{p} = c = p_2$. So $TS^{AUD} = v(c) = TS^{LP}$. Consequently, there exists a $k_2 \in (0, q(c)]$ s.t. $TS^{AUD} > TS^{LP}$ for $k < k_2$. ∎

Proof of Corollary 7. (i) Part (ii) of Corollary 3 says $p_2(k)$ increases with k. Corollary 1 tells us that $\overline{p}(k)$ is decreasing in k. At $k = 0$, $\overline{p} = p^m > c = p_2$. At $k = \widehat{k}$, $\overline{p} < u'(k) < x(Q_L) < p_2(Q_L)$. Thus, there exists a $k_3 \in (0, \widehat{k})$ s.t. $p_2^{QF} \lesseqqgtr p_2^{LP}$ for $k \lesseqqgtr k_3$.

(ii) $q_1^{LP} = q(\overline{p}) - k < q(c) - k$. Corollary 2 gives $q(c) - k < Q$ and $q_2^{QF} = q(p_2) - Q \leq k = q_2^{LP}$.

(iii) As LP is a special case of a 2PT, $\pi_1^{LP} < \pi_1^{2PT}$. Recall that $\pi_1^{QF}(Q) = v(x) + (x - c) \cdot Q - [u(\overline{q}(k, x)) - x \cdot \overline{q}(k, x)]$. Note that when $x = c$, it becomes $v(c) - [u(k) - ck]$, which is π_1^{2PT}. Because $x = c$ does not violate any constraints, $\pi_1^{QF} > \pi_1^{2PT} > \pi_1^{LP}$.

p_2^{LP} is given by $\pi'(p_2^{LP}) = k$. So $p_2^{LP} < u'(k)$ for any k. $\pi_2^{LP} = (p_2^{LP} - c)k$. When $0 < k < \widehat{k}$, $\pi_2^{QF} = (x(k) - c) \cdot k$. Corollary 1 says p_2^{LP} decreases with k. Part (ii) of Corollary 3 states x weakly increases with k. Note that at $k = 0$, $p_2^{LP} = p^m > c = x$. At $k = \widehat{k}$, $x > u'(\widehat{k}) > p_2^{LP}$, as $\widehat{k} > \alpha$. So there must exist a $k_2 < \widehat{k}$ s.t. p_2^{LP} crosses x from above once at k_2, where $x(k_2) = p_2^{LP}(k_2)$. It follows that π_2^{LP} must cross π_2^{QF} from above once at k_4. Recall that k_3 is defined by $p_2(k_1) = p_2^{LP}(k_1)$ and $p_2(k) > x(k)$ for any k. These imply that $k_4 > k_3$. Moreover, for $\widehat{k} < k$, $\pi_2^{QF} = (x - c) \cdot q(x) > (x(k) - c) \cdot k$, because setting $x(k)$ and selling k is an option for firm 2 when undercutting. Hence, $\pi_2^{QF} > (x(k) - c) \cdot k > (p_2^{LP} - c)k = \pi_2^{LP}$, for $\widehat{k} < k$.

(iv) When $k \to 0$, $BS^{QF} \to 0$, and $BS^{LP} \to v(p^m) > 0$ as $\overline{p} \to p^m$. From continuity of BS^{AUD} and BS^{LP}, this part follows.

(v) $TS^{LP} = u(q(\overline{p})) - c \cdot q(\overline{p})$. Because $\overline{p}'(k) < 0$ and $q(\overline{p}) < c$, $\frac{dTS^{LP}}{dk} = [u'(q(\overline{p})) - c] \cdot q'(\overline{p}) \cdot \overline{p}'(k) > 0$.

$TS^{QF} = u(q(p_2)) - c \cdot q(p_2)$. Because $p_2'(k) > 0$ for $k < \widehat{k}$, and $p_2'(k) = 0$ for $\widehat{k} \leq k$, $\frac{dTS^{QF}}{dk} = [u'(q(p_2)) - c] \cdot q'(p_2) \cdot p_2'(k) \leq 0$ with "=" only when $\widehat{k} \leq k$.

At $k = 0$, $\overline{p} = p^m > c = p_2$. So $TS^{QF} = v(c) > u(q^m) - c \cdot q^m = TS^{LP}$. At $k = q(c)$, $\overline{p} = c < p_2$. So $TS^{QF} = u(q(p_2)) - c \cdot q(p_2) < v(c) = TS^{LP}$. Consequently, there exists a $k_6 \in (0, q(c))$ s.t. $TS^{QF} \gtreqqless TS^{LP}$ for $k \lesseqqgtr k_6$. ∎

Proof of Corollary 8. Given the equilibrium characterizations in Propositions 2~4, the comparisons on profits and total surpluses are straightforward. Here we only compare the buyer's surpluses.

Note that $BS^{2PT} = u(k) - ck$. We express the buyer's surplus under QF in two cases.

When $k < \widehat{k}$,

$$\begin{aligned} BS^{QF} &= u(q(p_2)) - T - p_2 \cdot [q(p_2) - Q] \\ &= [v(p_2) + p_2 \cdot Q] - [v(x) + x \cdot Q] + [u(k) - x \cdot k] \\ &\leq u(k) - x \cdot k \\ &\leq u(k) - c \cdot k = BS^{2PT}, \end{aligned}$$

where the first inequality is from the fact that $v(p) + p \cdot Q$ decreases with p for any $q(p) \geq Q$. Note that "=" occurs only when $k = 0$.

When $\widehat{k} \leq k$,

$$\begin{aligned} BS^{QF} &= u(q(p_2)) - T - p_2 \cdot [q(p_2) - Q] \\ &= v(p_2) + p_2 \cdot Q - x \cdot Q \\ &< v(x) \\ &< v(u'(k)) \\ &< u(k) - c \cdot k = BS^{2PT}, \end{aligned}$$

where the first inequality is from the fact that $v(p) + p \cdot Q$ decreases with p for any $q(p) > Q$, the second inequality follows from $x > u'(k)$ when $\widehat{k} \leq k$, and the last one is from $u'(k) > c$. $BS^{AUD} < BS^{2PT}$ follows similarly. ■

Proof of Proposition 5. First, the buyer's problem in the last stage is the same as when $c_1 = c_2 = c$, because the buyer does not face those costs.

Lemmas 1~4 can be applied with adaptations of c replaced by c_2 in (14), (14'), (15), because those conditions are for firm 2. Firm 1's profit is the same as before with replacement of c by c_1, so is for its first-order derivative (17), which becomes

$$\begin{aligned} \frac{d\pi_1}{dQ} &= x - c_1 + \{\overline{q}(k, x) - [q(x) - Q]\} \cdot x'(Q) \\ &= \begin{cases} x - c_1 - [k + Q - q(x)] \cdot \frac{p_2 - c_2}{k} & \text{when } k < q(x) \\ x - c_1 - Q \cdot \frac{p_2 - c_2}{\pi'(x; c_2)} & \text{when } q(x) \leq k \end{cases} \quad \text{(from (14) with } c = c_2\text{)}. \end{aligned}$$

Thus, (19) and (20) follow.

When $k < q(x)$, $\pi_1 = (x - c_1) \cdot Q + v(x) - [u(k) - x \cdot k]$. Note that when $x = c_1$, $\pi_1 = v(c_1) - [u(k) - c_1 \cdot k]$, which is firm 1's profit under a 2PT. It is easy to verify that $x = p_2 = c_1$, $Q = q(c_1) + (c_1 - c_2) \cdot q'(c_1)$ and $k = -(c_1 - c_2) \cdot q'(c_1)$ satisfy (19) and $k < q(x) = q(c_1)$. (18) ensures that $x > c_1$ so that both QF and AUD improve firm 1's profit over a 2PT. Moreover, (18) implies that $\varphi_R(0) > 0$. Thus the sufficiency of (19) follows.

When $k \geq q(x)$, (18) assures that $x > c_1$. The assumption that $(p - c_2) \cdot \pi'(p; c_2)$ is single-peaked in

$(c_2, p^m(c_2))$ guarantees the sufficiency of (20) and the uniqueness of the solution to it.

The existence and uniqueness of the equilibrium follows from the same analysis in Propositions 3~4. ■

Table A1: Equillibrium Tariffs for Linear Demand

	LP	2PT	AUD	QF (or a 3PT)
Fixed Fee	N/A	$\frac{(1-k)^2}{2}$	N/A	$\frac{1}{2} + \frac{(5\sqrt{5}-9)k-4}{4} \cdot k$ when $k < \widehat{k}$ $\frac{3-2\sqrt{2}}{4}$ when $k \geq \widehat{k}$
Quantity Threshold	N/A	N/A	$1-(3-\sqrt{5})k$ when $k < \overline{k}$ $1-2a$ when $k \geq \overline{k}$	$1-(3-\sqrt{5})k$ when $k < \widehat{k}$ $1-\frac{\sqrt{2}}{2}$ when $k \geq \widehat{k}$
Firm 1's Per-Unit Price	$\frac{1-k}{2}$	0	$\frac{\frac{1}{2}+\frac{(5\sqrt{5}-9)k-4}{4}\cdot k}{1-(3-\sqrt{5})\cdot k}$ when $k < \overline{k}$ a when $k \geq \overline{k}$	N/A (or ∞ for 3PT)
Firm 2's Per-Unit Price	$\frac{1-k}{2}$	0	$\frac{3-\sqrt{5}}{2} \cdot k$ when $k < \overline{k}$ a when $k \geq \overline{k}$	$\frac{3-\sqrt{5}}{2} \cdot k$ when $k < \widehat{k}$ $\frac{\sqrt{2}}{4}$ when $k \geq \widehat{k}$

Table A2: Equillibrium Surpluses for Linear Demand

	LP	2PT	AUD	QF (or a 3PT)
Firm 1's Profit	$\frac{(1-k)^2}{4}$	$\frac{(1-k)^2}{2}$	$\frac{1}{2} + \frac{(5\sqrt{5}-9)k-4}{4} \cdot k$ when $k < \overline{k}$ $a(1-2a)$ when $k \geq \overline{k}$	$\frac{1}{2} + \frac{(5\sqrt{5}-9)k-4}{4} \cdot k$ when $k < \widehat{k}$ $\frac{3-2\sqrt{2}}{4}$ when $k \geq \widehat{k}$
Firm 2's Profit	$\frac{k(1-k)}{2}$	0	$\frac{7-3\sqrt{5}}{2} \cdot k^2$ when $k < \overline{k}$ a^2 when $k \geq \overline{k}$	$\frac{7-3\sqrt{5}}{2} \cdot k^2$ when $k < \widehat{k}$ $\frac{1}{8}$ when $k \geq \widehat{k}$
Buyer's Surplus	$\frac{(1+k)^2}{8}$	$\frac{k(2-k)}{2}$	$k \cdot [1-(3-\sqrt{5})k]$ when $k < \overline{k}$ $\frac{(1-a)^2}{2}$ when $k \geq \overline{k}$	$k \cdot [1-(3-\sqrt{5})k]$ when $k < \widehat{k}$ $\frac{8\sqrt{2}-7}{16}$ when $k \geq \widehat{k}$
Total Surplus	$\frac{(1+k)(3-k)}{8}$	$\frac{1}{2}$	$\frac{1}{2} + \frac{3\sqrt{5}-7}{4} \cdot k^2$ when $k < \overline{k}$ $\frac{1-a^2}{2}$ when $k \geq \overline{k}$	$\frac{1}{2} + \frac{3\sqrt{5}-7}{4} \cdot k^2$ when $k < \widehat{k}$ $\frac{7}{16}$ when $k \geq \widehat{k}$

Note: In Tables A1 and A2, $\overline{k} = \frac{5-\sqrt{5}-2\sqrt{\sqrt{5}-2}}{19-7\sqrt{5}} \simeq 0.5354$, $\widehat{k} = \frac{2-\sqrt{18\sqrt{2}+5\sqrt{5}-10\sqrt{10}-5}}{5\sqrt{5}-9} \simeq 0.8642$.
a is determined by $a(a^3 - 4a^2k + 6ak^2 - 2k^2) + k^2(1-k)^2 = 0$ $(a < k)$.

References

[1] P. Aghion and P. Bolton. Contracts as a barrier to entry. *The American Economic Review*, 77(3):pp. 388–401, 1987.

[2] B. D. Bernheim and M. D. Whinston. Exclusive dealing. *Journal of Political Economy*, 106(1):pp. 64–103, 1998.

[3] G. Calzolari and V. Denicolo. Competition with exclusive contracts and market-share discounts. *American Economic Review*, forthcoming.

[4] Y. Chao. Strategic effects of three-part tariffs under oligopoly. *International Economic Review*, 54(3):977–1015, 2013.

[5] Z. Chen and G. Shaffer. Naked exclusion with minimum-share requirements. *RAND Journal of Economics*, forthcoming.

[6] P. DeGraba. Naked exclusion by an input supplier: Exclusive contracting loyalty discounts. *International Journal of Industrial Organization*, forthcoming.

[7] R. J. Deneckere and D. Kovenock. Price leadership. *The Review of Economic Studies*, 59(1):pp. 143–162, 1992.

[8] European Commission. DG competition discussion paper on the application of article 82 of the treaty to exclusionary abuses. 2005.

[9] European Commission. Guidance on the commission's enforcement priorities in applying article 82 ec treaty to abusive exclusionary conduct by dominant undertakings. *Communication from the Commission*, 2009.

[10] E. Feess and A. Wohlschlegel. All-unit discounts and the problem of surplus division. *Review of Industrial Organization*, 37:161–178, 2010. 10.1007/s11151-010-9266-4.

[11] R. Inderst and G. Shaffer. Market-share contracts as facilitating practices. *The RAND Journal of Economics*, 41(4):709–729, 2010.

[12] S. Kolay, G. Shaffer, and J. A. Ordover. All-units discounts in retail contracts. *Journal of Economics & Management Strategy*, 13(3):429–459, 2004.

[13] A. Majumdar and G. Shaffer. Market-share contracts with asymmetric information. *Journal of Economics & Management Strategy*, 18(2):393–421, 2009.

[14] L. Marx and G. Shaffer. Rent-shifting, exclusion, and market-share discounts. 2004.

[15] G. F. Mathewson and R. A. Winter. The competitive effects of vertical agreements: Comment. *The American Economic Review*, 77(5):1057–1062, 1987.

[16] D. E. Mills. Inducing downstream selling effort with market share discounts. *International Journal of the Economics of Business*, 17(2):129–146, 2010.

[17] M. Motta. Michelin ii: The treatment of rebates. *Cases in European competition policy*, pages 29–49, 2009.

[18] D. O'Brien. All-units discounts and double moral hazard. 2013.

[19] D. O'Brien and G. Shaffer. Nonlinear supply contracts, exclusive dealing, and equilibrium market foreclosure. *Journal of Economics & Management Strategy*, 6(4):755–785, 1997.

[20] J. Ordover and G. Shaffer. Exclusionary discounts. 2007.

[21] E. B. Rasmusen, J. M. Ramseyer, and J. Wiley, John S. Naked exclusion. *The American Economic Review*, 81(5):pp. 1137–1145, 1991.

[22] M. Schwartz and D. Vincent. Quantity "forcing" and exclusion: Bundled discounts and nonlinear pricing. *Issues in Competition Law and Policy*, II, 2008.

[23] I. R. Segal and M. D. Whinston. Naked exclusion: Comment. *The American Economic Review*, 90(1):pp. 296–309, 2000.

[24] D. Spector. Exclusive contracts and demand foreclosure. *The RAND Journal of Economics*, 42(4):619–638, 2011.

[25] E. van Damme and S. Hurkens. Endogenous price leadership. *Games and Economic Behavior*, 47(2):404 – 420, 2004.