Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Does Pooling Purchases Lead to Higher Profits?

Bin Hu, Izak Duenyas, Damian R. Beil,

To cite this article:
Bin Hu, Izak Duenyas, Damian R. Beil, (2013) Does Pooling Purchases Lead to Higher Profits?. Management Science 59(7):1576-1593. http://dx.doi.org/10.1287/mnsc.1120.1651

MANAGEMENT SCIENCE
Vol. 59, No. 7, July 2013, pp. 1576–1593
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1120.1651

Does Pooling Purchases Lead to Higher Profits?

Bin Hu
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599,
bin_hu@unc.edu

Izak Duenyas, Damian R. Beil
Ross School of Business, University of Michigan, Ann Arbor, Michigan 48109
{duenyas@umich.edu, dbeil@umich.edu}

Consider two buyers facing uncertain demands who need to purchase a common critical component from a powerful sole-source supplier. If the two buyers pool their demands and purchase from the supplier as a single entity, will they necessarily earn higher profits than purchasing separately? We show that when a powerful supplier extracts profits from the buyers through optimal contract design, the demand variability reduction achieved by pooling can harm the buyers because it makes extracting profits easier for the supplier. We characterize cases when pooling is disadvantageous and also provide insights into when it is still advantageous. Our result is in contrast to the case where the price of the component is exogenous (typically assumed in most of the pooling literature), in which case pooling demands is always beneficial for the buyers.

Key words: games; inventory; production; policies; pricing; scale; diseconomies; principal–agent; demand variability
History: Received April 1, 2010; accepted May 24, 2012, by Yossi Aviv, operations management. Published online in *Articles in Advance* January 28, 2013.

1. Introduction

There exists a large research literature on inventory pooling, dating back to the seminal paper by Eppen (1979). In Eppen's canonical example, a firm with several locations considers satisfying different locations' demands from a central stock. More generally, the inventory pooling concept applies to areas such as component commonality, whereby the same component is used for multiple products, rather than a specific, noninterchangeable component being used for each product. In this literature it has been generally found that pooling enables demand variability reduction, which results in higher profits. This powerful intuition is widely taught in operations management courses when examining pooling. The pooling literature, however, has virtually always assumed that the buyers can purchase any quantity of their choosing at a fixed unit price. In this paper, we ask if and how this intuition will carry over if the goods are instead purchased from a powerful supplier who can tailor—to her own benefit—the price and quantity terms she offers to the buyers.

We became interested in this problem after interacting with a large manufacturer. This manufacturer makes several products in its different divisions that all require an essentially common critical component, and there is only one supplier who has the capability to make this component in a way that meets the manufacturer's engineering requirements. The component is slightly customized for each product (e.g., in tolerance levels), but it is possible to revise the specifications of their products so all divisions can pool their purchases and share the exact same component. However, we noticed that different divisions currently enter into individual purchase contracts with the supplier. Although our initial reaction was that the manufacturer could improve its profit by revising the specifications and pooling its purchases of the common component, we wondered whether it mattered that these components were bought from a sole supplier.

A similar situation in a different industry involves the iPhone. In the United States, most iPhones sold are bundled with wireless service contracts. For the package "iPhone+service," Apple is the sole supplier of the iPhone, and a carrier buys the iPhone, adds service, and sells the package to customers.[1] As a result, Apple enjoys a near 60% gross margin on the iPhone (Elmer-DeWitt 2010). The wireless service carriers that sell iPhones, on the other hand, have been less fortunate. AT&T struck a 24-month contract with Apple in 2008 and did not break even until 17 months later, a situation believed to be due to a contracted heavy subsidy to Apple on every iPhone sold (Whitney 2009). In addition to price, Apple can also dictate quantity terms. For example, in 2011, Sprint became a

[1] Apple also sells iPhones from its retail stores, but a large majority of those iPhones are also bundled with wireless service contracts as well; thus, Apple's retail stores merely serve as an outlet for the carriers. Apple does sell some unlocked iPhones, but these constitute a small minority in the United States.

carrier for the iPhone; its contract with Apple includes a commitment to purchase 30.5 million units in the next four years, and Sprint expected not to make any profit on the iPhone until 2015 (Lublin and Ante 2011). These examples indicate that Apple is a powerful sole-source supplier for the iPhone that can dictate price and quantity to extract the majority of the profits in its supply chain. Clearly, Apple tailors its contracts to the particular provider. For example, recently Apple struck an agreement with C Spire, a small carrier in the southern United States that has only 900,000 customers. It is clear that Apple cannot ask C Spire to purchase 30 million units for the next four years, so the quantities and other contract terms that Apple offers to its carriers are customized to that carrier in a way that benefits Apple but leaves the carrier just enough expected profits to make it worthwhile to offer the iPhone. Now consider the hypothetical scenario that Sprint acquired C Spire and pooled the purchasing of the iPhone. Given that Apple can set prices and quantities that benefit it, would pooling purchases necessarily increase the total profits that Sprint and C Spire make from the iPhone?

Another example is Monsanto, an agricultural biotechnology conglomerate. Monsanto has developed and patented seed strains that have superior yields and pesticide resistance and is therefore able to dictate contract terms for the strains to buyers. Monsanto uses what it calls "value-based pricing" whereby it adjusts its prices according to the "value" that it expects the buyer will obtain from using its products. Monsanto genes are inserted into roughly 95% of soybeans and 80% of corn grown in the United States, both extremely important crops in the United States (Leonard 2009). Given that Monsanto is able to customize contract terms for its buyers, an interesting question is if two independent seed companies (buyers) merge and approach Monsanto for seed strains as one entity rather than purchasing separately, will they necessarily improve their profits? These examples clearly illustrate that powerful suppliers who can dictate contract terms like we model in our paper exist in very important industries ranging from seeds to smartphones.

Although a sole-source supplier has the power to extract profits from the buyers, this power is mitigated because the buyers often hold more accurate information on the products' demands. The buyers' informational privilege may arise, for example, because the buyer conducts a market study or receives informal purchase intentions from downstream customers. Buyers typically guard such valuable information very carefully and rely on the informational privilege in their negotiations with the sole-source supplier. For example, the seed companies have a better sense of their demands—or "value"—from the seed strains, which Monsanto seeks to extract through contract terms. Similarly, Apple will not have as good information as C Spire will about how many iPhones C Spire expects to sell, especially because this will also depend on what other smartphones C Spire intends to include in its portfolio and at what prices, a piece of information that Apple does not necessarily have. The research question we aim to answer is when several buyers with private demand information purchase a common critical component from a powerful sole-source supplier, does pooling purchases necessarily lead to higher profits for the buyers?

To answer this question, we compare two parallel scenarios. In the *nonpooling scenario*, two firms, A and B, sell two products, a and b, respectively. Each firm possesses private information on its respective products' demands. Both products a and b require a common component from a sole-source supplier, and firms A and B independently approach the supplier and enter into separate purchase contracts for the component. In the *pooling scenario*, instead of firms A and B, a single firm, D, sells both products a and b and enters into one purchase contract with the supplier for the component. Firm D has the same private demand information on products a and b as firms A and B. Assume that products a's and b's demand distributions and the costs of processing the component into products a and b in both scenarios are identical. Would firm D always earn higher profits than firms A and B combined?

In contrast to the usual benefits of pooling, we find that the presence of a powerful sole supplier can actually result in situations where firms A and B (with unpooled demands for the component) have higher combined profits than firm does D (with pooled demands for the component). One of our key insights is that in the presence of a powerful supplier, the demand variability reduction achieved by pooling may become a disadvantage. A powerful supplier can design contracts adapted to a buyer's demand to extract profit; when demand variability is reduced, the supplier may more easily adapt the contracts to the demand, extracting more profit from the buyer. In this paper, we characterize cases when pooling is disadvantageous and also provide insights into when it is still advantageous.

The remainder of this paper is organized as follows: Section 2 provides a literature review. Section 3 introduces the model. Section 4 provides an analysis of whether the pooling or the nonpooling scenario yields higher profits for the buyers. Section 5 concludes the paper.

2. Literature Review

Our research is closely related to two streams of existing literature. The first stream is component commonality and inventory pooling. Stemming from the seminal paper by Eppen (1979), a very large literature explores statistical economies of scale (the reduction of uncertainty upon merging multiple stochastic demand streams) and shows that under various settings, component commonality, storage centralization, and inventory sharing can reduce operations costs. Examples from this literature include Eppen and Schrage (1981), Gerchak and He (2003), Benjaafar et al. (2005), and, more recently, Hanany and Gerchak (2008). However, one tacit assumption in this literature is that the purchase price of the good discussed is exogenous, and usually the supplier is not modeled. This would be a reasonable assumption if the good is a commodity, but less so when one supplier is the only source of supply (e.g., because the supplier owns a patented technology). In this paper, we assume that the component can only be purchased from a sole-source supplier who strategically takes into account the operational structure of the buyer(s), which leads to the second related stream of literature—procurement contract design. Based on the principal–agent model (e.g., Laffont and Martimort 2002), this literature analyzes how a powerful, profit-maximizing member of the supply chain (the principal, who may be a buyer or a supplier) should optimally design procurement contracts for the other members (agents) who possess private information. The principal–agent model captures the general practice of tailoring a contract to a specific buyer or supplier (instead of relying on one-size-fits-all contracting) and is a canonical modeling construct. In the operations management literature, several papers take the perspective of a buyer and find the optimal supplier selection and contracting mechanism when the buyer faces operational issues like the need for fast delivery (Cachon and Zhang 2006), random demand (Chen 2007), or uncertain supplier qualifications (Wan and Beil 2009). In our paper, we analyze a different operational issue, namely, component pooling, and study a supplier who designs contracts. Other operations papers have examined the supplier's perspective (although absent the component pooling issue that is the crux of our paper), with early references including Corbett and de Groote (2000) and Ha (2001).

Relatively distantly related is a literature on the analysis of group purchasing, particularly coalition forming and stability issues, such as Hartman and Dror (2003) and Nagarajan et al. (2009). This literature usually assumes that either each buyer faces uncertain demand and group purchasing benefits the buyers because of statistical economies of scale or that the supplier announces a price schedule that offers greater discounts for larger purchase quantities and then analyzes how to allocate the benefit from group purchasing to form a stable coalition. In both cases the supplier is not acting strategically, and the presumed benefit of group purchasing is a premise for the analysis. In contrast, we model a supplier who is strategic (namely, seeks to maximize her own profit) and ask whether there is always benefit in pooling purchases. Thus, the core research problems are very different.

3. Model and Preliminaries

We first describe and analyze a stylized model for the purchase contract negotiation process before formally defining our research question. Suppose a firm (buyer) needs to purchase a critical component from a powerful sole-source supplier that is needed for a product i (e.g., product a or b that we introduced in §1) that faces uncertain future demand. Without loss of generality, assume that each unit of product i requires a single unit of the component and processing the component into the product incurs no cost. We assume the product's demand to have the structure $\mu^i + e^i$, where μ^i is the *mean demand* that can be forecasted through a market study and e^i is a random *forecast error*. For tractability, we make the simplifying assumption that product i's mean demand μ^i can only take one of n *known* values μ^i_θ, where θ denotes one of n possible *demand types* and the prior probability of having demand type θ is p^i_θ (of course, this implies that $\sum_\theta p^i_\theta = 1$). The possible mean demands and the prior probabilities are public information. For example, a publicly available industry report may estimate that the demand type θ could be "high" or "low" with equal likelihood ($p^i_{\text{high}} = p^i_{\text{low}} = 0.5$), and the corresponding mean demands are $\mu^i_{\text{high}} = 20{,}000$ and $\mu^i_{\text{low}} = 10{,}000$, in which case $n = 2$. Prior to approaching the supplier for the component, the buyer learns his demand type (e.g., whether demand is going to be high or low) through a market study, and the learned demand type is his private information. However, the realized demand is unlikely to exactly match the forecast; to model this aspect, we assume that product i of demand type θ will actually experience a realized demand $\mu^i_\theta + e^i$ where e^i is a random forecast error that is assumed to have mean zero, probability density function (pdf) f^i, and cumulative density function (cdf) F^i. The forecast error is an abstraction of all the uncertainties that neither the supplier nor the buyer is capable of predicting: for example, economic cycle, competitor movement, and weather. Therefore, we assume e^i's distribution is public information, and neither the supplier nor the buyer can learn e^i's exact value before the demand realization. We assume e^i is independent of the product's demand type θ. To summarize, the buyer has strictly more accurate demand information than the supplier because the buyer knows the demand type whereas the supplier only knows the prior distribution of the demand types; however, even though the buyer knows the demand type with certainty, he still faces some demand uncertainty due to the forecast error.

The supply chain is decentralized (the supplier and the buyer are independent decision makers), and we assume the supplier and the buyer each seek to maximize their respective expected profits. Because the supplier is the sole source of the critical component required by the buyer, the supplier can offer the buyer take-it-or-leave-it contracts. In such a setting, if the supplier knew exactly the demand type (thus the mean demand) of the product sold by the buyer, she could extract all expected supply chain profits, leaving only a reservation profit (which we assume to be zero without loss of generality) to induce the buyer to participate. However, because only the buyer is privileged with the forecasted demand type information and the supplier only knows its prior distribution, the buyer can possibly earn a higher-than-reservation profit. The order of events is as follows:

1. The buyer learns his demand type.
2. The supplier offers the buyer a menu of contracts consisting of several quantity-payment pairs (Q_θ, t_θ, each meant for a possible type θ); that is, the buyer can choose to buy Q_θ units at total cost t_θ. The buyer decides which (if any) quantity-payment pair to choose.
3. Once the buyer has chosen a contract, the supplier produces the agreed-upon quantity at per unit cost c, delivers the units to the buyer, and receives the corresponding payment.
4. The buyer processes the components into finished goods. The demand for the finished goods is realized. The buyer receives revenue r for each unit of demand he satisfies. We assume $r > c$ and define $q \doteq c/r$ ($c = qr$). Unsatisfied demand is lost and excess inventory has no salvage value.

Notice that in stage 2 the supplier offers a menu of contracts to the buyer, with each contract intended for a possible demand type. Because the supplier does not know the buyer's demand type, she needs to design the menu of contracts that provides the incentives for the buyer of each type to pick the contract intended for him. Assuming that the supplier uses this particular type of mechanism is not restrictive because by the standard *revelation-principle* argument in mechanism design theory (e.g., Laffont and Martimort 2002), the outcome of *any* mechanism that concludes the transaction between the supplier and the buyer before the buyer's demand realizes can be replicated by an equivalent menu of contracts. Therefore, by finding the optimal menu of contracts, we are actually maximizing the supplier's expected profit among a wide range of general mechanisms, including but not limited to wholesale price, quantity discount, and price negotiation.

Formally, the supplier's problem is

$$\max_{Q_\theta,\, t_\theta} \; E_\Theta[t_\Theta - cQ_\Theta] \tag{1a}$$

subject to

$$E_e[r\min\{\mu_\theta + e, Q_\theta\}] - t_\theta \geq 0, \quad \forall \theta, \tag{1b}$$

$$E_e[r\min\{\mu_\theta + e, Q_\theta\}] - t_\theta \geq E_e[r\min\{\mu_\theta + e, Q_{\theta'}\}] - t_{\theta'}, \quad \forall \theta' \neq \theta, \tag{1c}$$

where random variable Θ reflects the supplier's prior on the buyer's types. In the above formulation (and those to follow where it does not cause confusion), we suppress superscript "i" for readability. Constraint (1b) is the *participation constraint* (PC) with reservation profit set to zero, and (1c) is the *incentive compatibility constraint* (IC), which ensures that the buyer chooses the contract designed for his demand type. Following convention, we define *information rent* π_θ as the expected profit of a buyer having demand type θ who chooses the contract designed for his demand type: $\pi_\theta \doteq E_e[r\min\{\mu_\theta + e, Q_\theta\}] - t_\theta$. For convenience and without loss of generality, henceforth we denote a contract by a quantity-information rent pair (Q_θ, π_θ) rather than a quantity-payment pair. Using this notation and explicitly writing out the expectations, (1a)–(1c) can be shown to be equivalent to

$$\max_{Q_\theta,\, \pi_\theta} \; \sum_\theta p_\theta \left[r\left((1-q)Q_\theta - \int_{-\infty}^{Q_\theta - \mu_\theta} F(x)\,dx \right) - \pi_\theta \right] \tag{2a}$$

subject to

$$\pi_\theta \geq 0, \quad \forall \theta, \tag{2b}$$

$$\pi_\theta \geq \pi_{\theta'} + r\int_{Q_{\theta'} - \mu_\theta}^{Q_{\theta'} - \mu_{\theta'}} F(x)\,dx, \quad \forall \theta' \neq \theta. \tag{2c}$$

The formulation can be further simplified once one notices that the buyer's expected profit function satisfies the *Spence-Mirrlees property* (Laffont and Martimort 2002, p. 53); namely, the *marginal rate of substitution* $(\partial \pi_\theta / \partial Q_\theta)/(\partial \pi_\theta / \partial t_\theta)$ is monotonic in type θ. With this property, all the constraints can be substituted with the lowest type's participation constraint, local downward incentive constraints, and *monotonicity constraints* (MC); (contract quantity nondecreasing in type). Rank the demand types by mean demand: $\mu_{\theta_1} \leq \mu_{\theta_2} \leq \cdots \leq \mu_{\theta_n}$, and (2a)–(2c) are equivalent to

$$\max_{Q_\theta,\, \pi_\theta} \; \sum_{\theta = \theta_j} p_\theta \left[r\left((1-q)Q_\theta - \int_{-\infty}^{Q_\theta - \mu_\theta} F(x)\,dx \right) - \pi_\theta \right] \tag{3a}$$

subject to

$$\pi_{\theta_1} \geq 0 \tag{3b}$$

$$\pi_{\theta_{j+1}} \geq \pi_{\theta_j} + r\int_{Q_{\theta_j} - \mu_{\theta_{j+1}}}^{Q_{\theta_j} - \mu_{\theta_j}} F(x)\,dx, \quad j = 1, \ldots, n-1, \tag{3c}$$

$$Q_{\theta_{j+1}} \geq Q_{\theta_j}, \quad j = 1, \ldots, n-1. \tag{3d}$$

Equations (3a)–(3d) describe how the supplier derives an optimal menu of quantity-payment pairs (represented by equivalent quantity-information rent pairs) to offer to the buyer. Thanks to the buyer's private information on the demand type, the supplier cannot extract all supply chain profit but has to leave some information rent to the buyer, which constitutes the buyer's profit.

Now we can state our research question. Consider two parallel scenarios. In the nonpooling scenario, a firm A makes product a and a firm B makes product b. The firms have learned their respective demand types and separately purchase the component from the sole-source supplier. In the pooling scenario, a firm D makes both products a and b. Firm D has learned both products' demand types and purchases the component from the sole-source supplier. Does the pooled firm D always earn higher expected profit (information rent) than firms A and B combined? Because this research question involves comparing the firms' information rents, we next obtain results that characterize the information rent structure in preparation for answering this question.

3.1. Solving for Information Rent

By the analysis of Laffont and Martimort (2002, p. 43), we know that participation and incentive compatibility constraints (3b) and (3c) are binding at optimality. Applying this insight, the objective function (3a) can be recast as a function solely of Q_θ. Relaxing the MCs (3d) for now (we will test these constraints ex post and, in the case of them being violated, revise our solution), the first-order condition (FOC) for (3a) as a function of Q_{θ_j} can be written as

$$\begin{aligned}
&\Pr(\Theta=\theta_j)(1-q)\\
&\quad=\Pr(\Theta\ge\theta_j)F(Q_{\theta_j}-\mu_{\theta_j})-\Pr(\Theta\ge\theta_{j+1})F(Q_{\theta_j}-\mu_{\theta_{j+1}})\\
&\quad\Longleftrightarrow\quad \lambda(\theta_j)(1-q)=F(Q_{\theta_j}-\mu_{\theta_j})-(1-\lambda(\theta_j))\\
&\qquad\qquad\cdot F(Q_{\theta_j}-\mu_{\theta_{j+1}}),\quad j=1,\ldots,n-1, \qquad (4)
\end{aligned}$$

where $\lambda(\theta_j)\doteq\Pr(\Theta=\theta_j)/\Pr(\Theta\ge\theta_j)$, and the information rents can be derived recursively as

$$\pi_1=0,\qquad \pi_{\theta_{j+1}}=\pi_{\theta_j}+r\int_{Q_{\theta_j}-\mu_{\theta_{j+1}}}^{Q_{\theta_j}-\mu_{\theta_j}}F(x)\,dx. \qquad (5)$$

For concision, we use $\Theta\ge\theta_j$ to denote $\Theta\in\{\bar{\theta}\mid\mu_{\bar{\theta}}\ge\mu_{\theta_j}\}$. As an example, if Θ can take two values, high or low, then $\Pr(\Theta\ge\text{low})=\Pr(\Theta=\text{low or high})$ and $\Pr(\Theta\ge\text{high})=\Pr(\Theta=\text{high})$.

Equation (3a), which represents the supplier's objective function, is generally not concave in its decision variable Q_{θ_j}. Consequently, solving FOC (4) cannot always guarantee a global maximizer. In the following proposition we provide a set of *sufficient* conditions that guarantee (4) has a unique solution and it is the global maximizer.

Proposition 1. *Equation* (4) *has a unique solution and the solution is a global maximizer of the supplier's expected profit, if the forecast error e's pdf f satisfies the following conditions*:

1. $f(-x)=f(x)$;
2. $f(x)$ *is continuous in* x;
3. $f(x_2)\le f(x_1)$ *for all* $x_2>x_1\ge 0$;
4. *For all* $\delta>0$ *and* $x\ge 0$, $f(x+\delta)/f(x)$ *is nonincreasing in* x.

The conditions in Proposition 1 ensure that e is reasonably well behaved. The first three conditions require that e has a symmetric, continuous, and unimodal pdf. The fourth condition has the intuitive meaning that the pdf must be sufficiently smooth. This condition will be violated, for example, if the pdf is piecewise linear and alternates between being flat and steep. Considering that e is a demand forecast error, the conditions in Proposition 1 are quite natural and mild. In fact, it is trivial to test that many common distributions including uniform, triangular, and normal satisfy all four conditions. For the rest of this paper, we assume these conditions are satisfied; therefore, (4) determines the unique global maximizer.

Recall that we want to answer the following question: Does firm D selling the two products a and b always earn higher profit than the combined profits of firms A and B, who sell products a and b, respectively? To gain insights into this question using the simplest possible setting, we assume that the mean demand for products a and b takes only two possible types h and l: product i $(i=a,b)$ has either high mean demand μ_h^i or low mean demand μ_l^i. Because each unit of the product requires one unit of the component, these demands are also the firms' demands for the component. Thus, in the nonpooling scenario, when firm A approaches the supplier, he is offered a menu of two contracts, (Q_h^a,π_h^a) and (Q_l^a,π_l^a). Similarly, when firm B approaches the supplier he is offered (Q_h^b,π_h^b) and (Q_l^b,π_l^b). We define $\delta^i\doteq\mu_h^i-\mu_l^i$ and assume $\delta^i\ge 0$. On the other hand, in the pooling scenario, firm D sells both products a and b, each of which has two demand types h and l. As a result, firm D has four possible demand types hh, hl, lh, or ll, where the type $\theta^a\theta^b$ means product a is type θ^a and product b is type θ^b. The mean demands for the component are denoted by $\mu_{hh}^D,\mu_{hl}^D,\mu_{lh}^D$, or μ_{ll}^D. Naturally, $\mu_{\theta^a\theta^b}^D=\mu_{\theta^a}^a+\mu_{\theta^b}^b$. For example, firm D having type hl means product a is type h and product b is type l, and so the total mean demand for the component equals $\mu_{hl}^D=\mu_h^a+\mu_l^b$. Accordingly, the supplier will offer firm D a menu of four contracts: (Q_{hh}^D,π_{hh}^D), (Q_{hl}^D,π_{hl}^D), (Q_{lh}^D,π_{lh}^D), and (Q_{ll}^D,π_{ll}^D). Without loss of generality, we assume $\delta^a\ge\delta^b$, and as a result $\mu_{hh}^D\ge\mu_{hl}^D\ge\mu_{lh}^D\ge\mu_{ll}^D$. In the rest of the paper, we calculate the total information rents obtained by firms A

and B each selling a single product and compare it to the information rent obtained by firm D selling both products to understand if and when firm D is better off than firms A and B combined.

In the nonpooling scenario, when firms A and B separately approach the supplier, she designs an optimal menu of contracts for each firm. For each product i, the supplier's problem is a two-type version of (3a)–(3d):

$$\max_{Q^i_\theta,\pi^i_\theta} \sum_{\theta=h,l} p^i_\theta\left[r\left((1-q)Q^i_\theta-\int_{-\infty}^{Q^i_\theta-\mu^i_\theta}F^i(x)\,\mathrm{d}x\right)-\pi^i_\theta\right]$$

$$\text{subject to}\quad \pi^i_l\ge 0,\quad \pi^i_h\ge\pi^i_l+r\int_{Q^i_l-\mu^i_h}^{Q^i_l-\mu^i_l}F^i(x)\,\mathrm{d}x$$

$$Q^i_h\ge Q^i_l. \tag{6}$$

Ignoring the MC (6), the FOC solution to the above problem is $p^i_l(1-q)=F^i(Q^i_l-\mu^i_l)-p^i_hF^i(Q^i_l-\mu^i_h)$, $Q^i_h=\mu^i_h+(F^i)^{-1}(1-q)$. It is trivial to test that this solution satisfies (6); thus, this is the optimal solution. At this solution, the information rents are $\pi^i_l=0$, $\pi^i_h=r\int_{Q^i_l-\mu^i_h}^{Q^i_l-\mu^i_l}F^i(x)\,\mathrm{d}x$.

In the parallel scenario, when firm D approaches the supplier, the supplier's problem is a four-type version of (3a)–(3d). Because firm D's total demand for the component equals the sum of products a and b's demands and the two products' demand types are assumed independent, D's demand type $\theta^a\theta^b$ has mean demand $\mu^D_{\theta^a\theta^b}=\mu^a_{\theta^a}+\mu^b_{\theta^b}$ and prior probability $p^D_{\theta^a\theta^b}=p^a_{\theta^a}p^b_{\theta^b}$. Similarly, because firm D depends on the demand forecasts for products a and b to forecast the total demand of the component, the component demand's forecast error will be the sum of the forecast errors for the two products: $e^D=e^a+e^b$. Assume e^D has pdf f^D and cdf F^D. The supplier's problem is

$$\max_{Q^D_\theta,\pi^D_\theta}\sum_{\theta=hh,hl,lh,ll}p^D_\theta\left[r\left((1-q)Q^D_\theta-\int_{-\infty}^{Q^D_\theta-\mu^D_\theta}F^D(x)\,\mathrm{d}x\right)-\pi^D_\theta\right] \tag{7a}$$

$$\text{subject to}\quad \pi^D_{ll}\ge 0, \tag{7b}$$

$$\pi^D_{lh}\ge\pi^D_{ll}+r\int_{Q^D_{ll}-\mu^D_{lh}}^{Q^D_{ll}-\mu^D_{ll}}F^D(x)\,\mathrm{d}x, \tag{7c}$$

$$\pi^D_{hl}\ge\pi^D_{lh}+r\int_{Q^D_{lh}-\mu^D_{hl}}^{Q^D_{lh}-\mu^D_{lh}}F^D(x)\,\mathrm{d}x, \tag{7d}$$

$$\pi^D_{hh}\ge\pi^D_{hl}+r\int_{Q^D_{hl}-\mu^D_{hh}}^{Q^D_{hl}-\mu^D_{hl}}F^D(x)\,\mathrm{d}x, \tag{7e}$$

$$Q^D_{hh}\ge Q^D_{hl}\ge Q^D_{lh}\ge Q^D_{ll}. \tag{7f}$$

Ignoring the MCs (7f), the FOC solution of (7a)–(7e) is

$$p^D_{ll}(1-q)=F^D(Q^D_{ll}-\mu^D_{ll})-(p^D_{lh}+p^D_{hl}+p^D_{hh})F^D(Q^D_{ll}-\mu^D_{lh}), \tag{8}$$

$$p^D_{lh}(1-q)=(p^D_{lh}+p^D_{hl}+p^D_{hh})F^D(Q^D_{lh}-\mu^D_{lh})-(p^D_{hl}+p^D_{hh})F^D(Q^D_{lh}-\mu^D_{hl}), \tag{9}$$

$$p^D_{hl}(1-q)=(p^D_{hl}+p^D_{hh})F^D(Q^D_{hl}-\mu^D_{hl})-p^D_{hh}F^D(Q^D_{hl}-\mu^D_{hh}), \tag{10}$$

$$Q^D_{hh}=\mu^D_{hh}+(F^D)^{-1}(1-q). \tag{11}$$

At this solution, the information rents are

$$\pi^D_{ll}=0,$$

$$\pi^D_{lh}=r\int_{Q^D_{ll}-\mu^D_{lh}}^{Q^D_{ll}-\mu^D_{ll}}F^D(x)\,\mathrm{d}x, \tag{12}$$

$$\pi^D_{hl}=\pi^D_{lh}+r\int_{Q^D_{lh}-\mu^D_{hl}}^{Q^D_{lh}-\mu^D_{lh}}F^D(x)\,\mathrm{d}x,$$

$$\pi^D_{hh}=\pi^D_{hl}+r\int_{Q^D_{hl}-\mu^D_{hh}}^{Q^D_{hl}-\mu^D_{hl}}F^D(x)\,\mathrm{d}x. \tag{13}$$

If the solution satisfies (7f), then it is indeed an optimal solution. Otherwise, if the MCs (7f) are not satisfied, it is necessary to revise the solution.

Proposition 2. *The contract quantities determined by* (8)–(11) *can only violate $Q^D_{ll}\le Q^D_{lh}$ or $Q^D_{lh}\le Q^D_{hl}$ and not both. When the solution of* (8)–(11) *violates $Q^D_{ll}\le Q^D_{lh}$, replace* (8) *and* (9) *with*

$$(p^D_{ll}+p^D_{lh})(1-q)=F^D(\underline{Q}-\mu^D_{ll})-(p^D_{hl}+p^D_{hh})F^D(\underline{Q}-\mu^D_{hl}), \tag{14}$$
$$Q^D_{ll}=Q^D_{lh}=\underline{Q}$$

and keep (10) *and* (11) *unchanged, and the revised FOCs determine the optimal solution. When the solution of* (8)–(11) *violates $Q^D_{lh}\le Q^D_{hl}$, replace* (9)–(10) *with*

$$(p^D_{lh}+p^D_{hl})(1-q)=(p^D_{lh}+p^D_{hl}+p^D_{hh})F^D(\bar{Q}-\mu^D_{lh})-p^D_{hh}F^D(\bar{Q}-\mu^D_{hh}), \tag{15}$$
$$Q^D_{lh}=Q^D_{hl}=\bar{Q}$$

and keep (8) *and* (11) *unchanged, and the revised FOCs determine the optimal solution.*

In the above revised solutions, two types are offered the same contract (e.g., $Q^D_{ll}=Q^D_{lh}=\underline{Q}$); this situation is referred to as *bunching*. If the solution of (8)–(11) violates a monotonicity constraint, an optimal solution can be obtained by forcing equality of the violated constraint and ignoring the rest of the MCs (they will always be satisfied), then rederiving the FOC. The revised FOC (14) is actually the sum of (8) and (9) with $\underline{Q}\doteq Q^D_{ll}=Q^D_{lh}$, and (15) is the sum of (9) and (10) with $\bar{Q}\doteq Q^D_{lh}=Q^D_{hl}$. Note that (14) and (15) can still be represented by (4) (with $\lambda(\theta_j)$ replaced by $\underline{\lambda}\doteq(p^D_{ll}+p^D_{lh})/(p^D_{ll}+p^D_{lh}+p^D_{hl}+p^D_{hh})$ corresponding to $\underline{Q}$, and $\lambda(\theta_j)$ replaced by $\bar{\lambda} \doteq (p^D_{lh} + p^D_{hl})/(p^D_{lh} + p^D_{hl} + p^D_{hh})$ corresponding to $\bar{Q}$); thus, properties derived from (4) will remain true in both cases of bunching.

In conclusion, to solve for information rents, we ignore the MCs and assume all PCs and ICs are binding, then solve the FOCs. In the nonpooling scenario for individual firms A and B, this solution is always optimal. In the pooling scenario for firm D, we must test whether the monotonicity constraints are satisfied. If they are satisfied, the solution is optimal. Otherwise, we need to solve the revised FOCs to obtain the optimal solution.

3.2. Three Key Drivers of Information Rent

The above analysis reveals that the information rent for any demand type can be expressed as the sum of several incremental information rents, where each incremental information rent is characterized by (4) and (5). Therefore, understanding the properties of the incremental information rents as governed by (4) and (5) is crucial for comparing pooling versus nonpooling profits. The three lemmas below identify three key drivers of incremental information rents. In what follows, θ and θ' always refer to a type and its adjacent higher type (assuming θ is not the highest type).

Lemma 1 (Type Rareness). *Incremental information rent $\pi_{\theta'} - \pi_\theta$ is increasing in $\lambda(\theta)$.*[2]

Recall that $\lambda(\theta)$ was defined as $\Pr(\Theta = \theta)/\Pr(\Theta \geq \theta)$. The intuition behind Lemma 1 is easily seen for an unpooled firm, say A, having high-type mean demand. In this case, $\lambda(l)$ equals p^a_l, and Lemma 1 states that π^a_h is increasing in p^a_l, which means it is decreasing in p^a_h. To understand this, note that the larger p^a_h is, the more the supplier anticipates that firm A has high-type mean demand, and consequently the lower information rent the high-type firm A will receive. In the extreme case, if the supplier knows firm A's mean demand is of high type with certainty ($p^a_h = 1$), then the firm A would get no information rent at all. Similarly, the rarer the high-type firm A is, the more information rent he receives. For firm D, who can have four different demand types, $\lambda(\theta)$ reflects how likely type θ is, compared only within the set of types θ and higher. (The lower types do not matter because we are only considering the *incremental* information rent.) In summary, Lemma 1 characterizes the impact of type rareness on incremental information rents.

We define $\delta = \mu_{\theta'} - \mu_\theta$ to be the "gap" between the type-θ' and type-θ mean demands (recall that θ is the lower type of the two). Lemma 2 shows that the incremental information rent depends on the mean demands only through their gap.

[2] When bunching occurs, $\lambda(\theta)$ should be understood as $\underline{\lambda}$ or $\bar{\lambda}$.

Lemma 2 (Gap Between Types). *Incremental information rent $\pi_{\theta'} - \pi_\theta$ as a function of μ_θ and $\mu_{\theta'}$ is determined only by their difference: $\delta = \mu_{\theta'} - \mu_\theta$. Furthermore, $\pi_{\theta'} - \pi_\theta < r\delta$.*

The supplier's contracts provide the buyer firm with information rent to ensure that the firm picks the contract meant for his demand type and not the contract meant for a lower demand type. Lemma 2 shows that the incremental information rent does not depend on individual mean demands but only on the gap between them. Additionally, the incremental information rent is bounded by the largest possible revenue difference of the two types, i.e., $r\delta$. This means that when there is almost no gap between the mean demands of two adjacent types ($\delta \to 0$), the two adjacent types are almost identical and so there is little incremental information rent.

Another important element in the model is the demand forecast error. To be able to quantify the effect of forecast error variability, we first define variability in a family of distributions.

Definition 1 (Rescaling). Suppose F is the cdf of a zero-mean random variable. For any constant $\gamma > 0$, define cdf $F_{(\gamma)}$ as $F_{(\gamma)}(x) \doteq F(x/\gamma)$.

The defined cdf $F_{(\gamma)}$ is a γ rescaling of F. $\{F_{(\gamma)}, \gamma > 0\}$ could be seen as a family of random variables stemming from F. One could easily verify that the variance of $F_{(\gamma)}$ is γ^2 times that of F. Thus, when $\gamma > 1$ the variance of $F_{(\gamma)}$ is greater than that of F.

Lemma 3 (Demand Variability). *Suppose $\gamma > 1$, and replace F by $F_{(\gamma)}$ in (4) and (5). Then the incremental information rent $\pi_{\theta'} - \pi_\theta$ is increasing in γ.*

Lemma 3 implies that the incremental information rent is actually *increasing* in demand variability, within the same family of forecast error distributions. Notice that information rent stems from the supplier's uncertainty about the firm's demand. Therefore, it is understandable that the firm's profit increases in his demand variability. Recall that firm D pools the demand for products a and b and therefore faces reduced variability for his component demand. The interesting question then is whether the reduced variability could actually lead to firm D receiving lower information rent than firms A and B combined.

4. Pooling vs. Nonpooling

In this section, we compare firm D's information rent to the total information rents of firms A and B. We have seen that incremental information rents are determined by contract quantities (see (5)), which are in turn determined implicitly by FOC (4). For firm D, whose demand for the component can have four different types, calculating information rent involves adding multiple layers of incremental information rents. The complex multilayer structure of firm D's information rent makes it difficult to directly compare the information rents in closed form. Thus, we use the insights of Lemmas 1–3 to facilitate comparisons.

To ensure that the forecast error for firm D's component demand (i.e., the forecast errors for products a and b combined) is tractable, we assume that products a and b's forecast errors e^a and e^b have normal distributions $N(0, \sigma^a)$ and $N(0, \sigma^b)$, respectively.[3] We assume that forecast errors are small relative to mean demands, so that the probability of having negative demand is negligible. As we discussed earlier, the forecast error represents the uncertain factors that neither the supplier nor the buyer firm is able to predict. Some uncertain factors that are common to both firms A and B, e.g., material price and economic cycle, may lead to positive correlation between e^a and e^b. Some other uncertain factors, such as competition between products a and b, may lead to negative correlation between e^a and e^b. For simplicity, we assume e^a and e^b are independently distributed, and thus firm D's component demand forecast error $e^D = e^a + e^b$ has normal distribution $N(0, \sqrt{(\sigma^a)^2 + (\sigma^b)^2})$.[4]

We will now focus on when firm D receives lower information rent than firms A and B. We will study this analytically for demand types lh, hl, and hh faced by firm D. (For demand type ll, both products a and b have low-type demands; thus, firms A and B in the nonpooling scenario and firm D in the pooling scenario will all earn zero information rent, making the pooling versus nonpooling comparison trivial.) We will also numerically illustrate what parameters lead to firm D or firms A and B receiving higher information rent. Our primary method of presenting numerical results will be showing the "*pooling*" and "*nonpooling*" regions—denoted with P and N, respectively—where firm D's information rent is *greater* or *smaller* than that of firms A and B combined. We plot these regions in a two-dimensional box of δ^b versus $\delta^a - \delta^b$ (recall that we assumed $\delta^a \geq \delta^b$), and in different plots we vary either σ^i or p_h^i. If not indicated otherwise, the default parameters are $p_h^i = 0.5$, $q(=c/r) = 0.2$, and $\sigma^i = 1$, $i = a, b$. Because of Lemma 2, the values of μ_θ^i are irrelevant except for $\delta^i = \mu_h^i - \mu_l^i$; thus, we do not assume any value for μ_θ^i (δ^i are indicated at the axes).

We first study the comparison for demand type lh.

Theorem 1. *Assume demand type is lh.*

1. *With sufficiently small σ^a, firm D receives lower information rent than A and B combined.*

2. *Suppose firm D receives lower information rent than A and B combined. Then as σ^a decreases, or p_h^a increases, firm D receives even lower information rent and is still outperformed by A and B.*

To understand our result, first notice that when firm A has low demand and firm B has high demand, the combined information rent of A and B is just the information rent of B because the low-type firm A earns no information rent. The theorem's first part establishes the existence of cases where firm D earns less profit than firms A and B combined. Compared to firm B, what firm D loses by having pooled demand is his position in the type hierarchy: firm D of type lh has the second-lowest demand type, so the supplier does not have to grant him significant information rent, whereas firm B's type is h and therefore he will be granted significant information rent. (Recall that calculating information rent involves adding multiple layers of incremental information rents, so higher types earn more information rent.) On the other hand, firm D has higher demand variability $\sqrt{(\sigma^a)^2 + (\sigma^b)^2}$ than does firm B's σ^b, and we know (from Lemma 3) that higher demand variability improves information rent. However, when σ^a is small, firm D's gain from the increased demand variability is minimal because $\sqrt{(\sigma^a)^2 + (\sigma^b)^2}$ is not much higher than σ^b. Therefore, when σ^a is small enough, firm D will make less profit compared to firms A and B combined. This is the intuition behind the first part of Theorem 1.

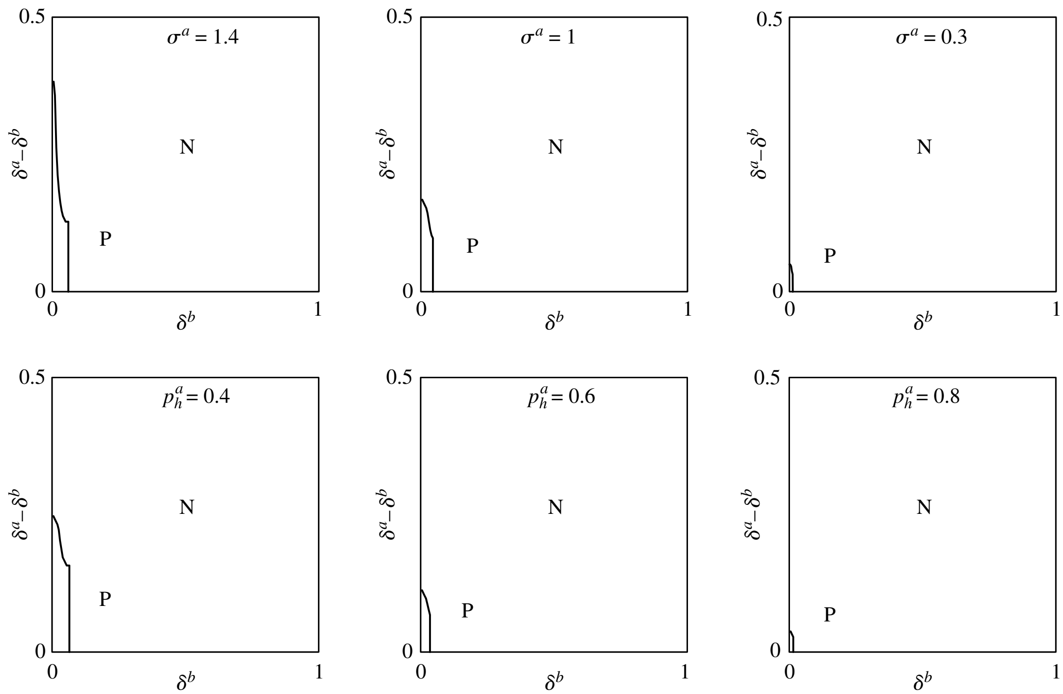
Now we turn to the second part of the theorem. Because firm A always gets zero information rent, lowering σ^a or increasing p_h^a has no impact on the combined information rent of firms A and B. However, lowering σ^a or increasing p_h^a does reduce the information rent of firm D, who has pooled demand for products a and b. The reason is that information rent depends on demand variability (Lemma 3) and type rareness (Lemma 1), and firm D—through pooling demand for products a and b—is negatively affected by either lowering σ^a or increasing p_h^a.

Figure 1 illustrates our result. Comparing the panels from left to right, one can see that the region in which A and B outperform D (the nonpooled region, denoted with N) grows as the demand variability for a, σ^a, decreases and as the probability of product a having high-type demand, p_h^a, increases. Furthermore, Figure 1 shows that the regions in which nonpooling is optimal are fairly large and not limited to very low values of σ^a.

[3] We have numerically tested error distributions other than normal (e.g., uniform and triangular) and verified that the qualitative insights we obtain in this section using normally distributed forecast errors remain valid.

[4] When e^a and e^b have correlation coefficient $\rho \neq 0$, e^D will have distribution $N(0, \sqrt{(\sigma^a)^2 + (\sigma^b)^2 + 2\rho\sigma^a\sigma^b})$. Theorems 1 and 2 will not be affected. Theorem 3's structure will not be affected, but the threshold number will change. When $\rho > 0$, Theorem 4 will not be affected. When $\rho < 0$, Theorem 4 may fail to hold, which means pooling is even less favorable than nonpooling when there is negative correlation between the forecast errors. When $\rho \neq 1$, Theorem 5 will not be affected.

Figure 1 **Pooling and Nonpooling Regions of Demand Type lh, Varying σ^a and p_h^a**



It is interesting to compare the intuition behind Theorem 1 to the traditional pooling intuition. In traditional inventory theory, pooling yields higher profits because it reduces demand variability. However, the implicit assumption there is that the firm's price is exogenous and fixed regardless of its purchase quantity. In our setting, because of information asymmetry, the supplier offers the buyer a menu of price-quantity pairs and has to offer the buyer information rent. Variability is in fact one of the drivers of this information rent, and therefore reducing it can result in a reduction of the firms' profits. Furthermore, in a setting with information asymmetry, all types except the highest type receive quantities that are distorted from the centrally optimal newsvendor quantities. The supplier has to do this to ensure that every type picks a contract that is meant for him and not for anyone else. In fact, this *downward distortion* can result in situations where pooling is not beneficial for anyone in the supply chain. For example, in Figure 1, when $\sigma^a = 0.3$, $\delta^a = 0.7$, and $\delta^b = 0.2$, we have numerically observed that both the supplier and the buyers' profits are reduced when pooling (and thus the whole supply chain's profit is also reduced). In this region the buyers prefer not to pool because of the above intuition. On the other hand, the supply chain also suffers from the reduced efficiency because of the downward distortion of purchase quantities. Thus, it is interesting to note that in our setting, pooling could be disadvantageous to all parties in the supply chain.

However, similar to the traditional setting, in our setting pooling can also benefit buyers, although not for the usual reason. In our setting pooling can be beneficial for buyers because not all unpooled firms can make good use of their demand variability to generate information rents. (For example, in designing the contracts to offer to buyers, the supplier has to offer information rent to firms that have higher mean demands to keep them from choosing the contract meant for the lower mean demand firms. However, firms with lower mean demands do not receive much if any information rent—they would not choose the contracts with higher quantities meant for the other firms in any case—with the lowest mean demand firm receiving no information rent.) Thus, if a high mean demand firm pools its demand with a low mean but very high variability firm, then pooling may benefit both firms by making use of the low mean demand firm's variability that would be "wasted" had that firm purchased on its own. This is a very interesting reason for buyers wanting to pool that again does not appear in traditional settings with exogenous pricing.

Thus far we have shown that in the presence of a powerful strategic supplier, a firm D having demand type lh can earn less profit than firms A and B with unpooled demands. In fact, this is not unique to demand type lh but can also occur for type hl. The intuition behind these results is fairly similar and is explained below.

Theorem 2. *Assume demand type is hl.*

1. *When σ^b is sufficiently small and δ^b is sufficiently close to δ^a, firms D receives lower information rent than A and B combined.*

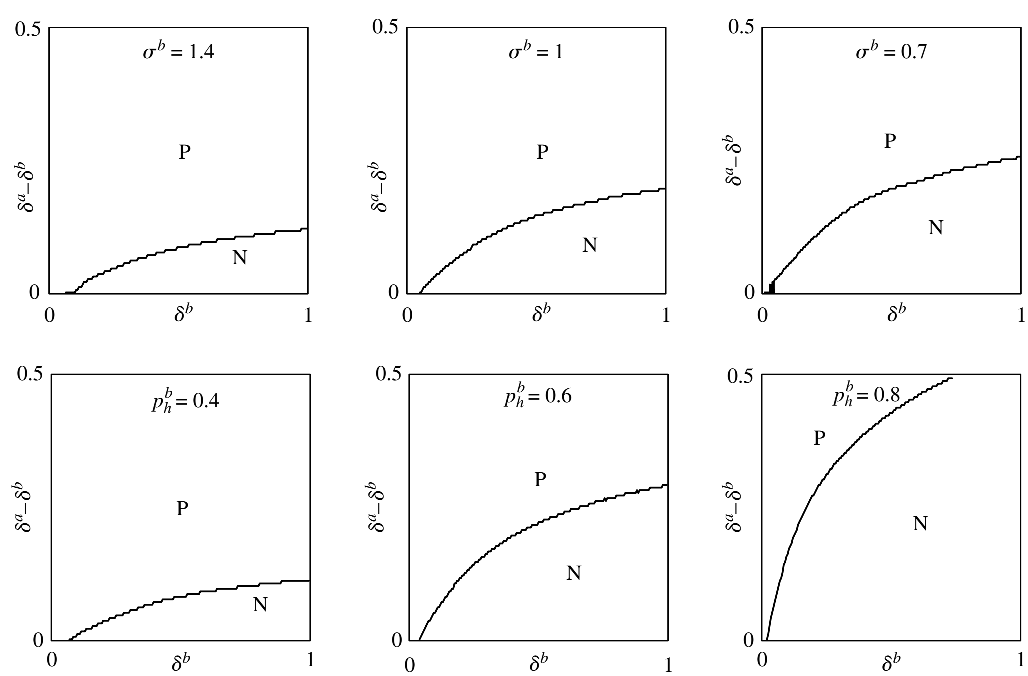
2. *Suppose firm D receives lower information rent than A and B combined. Then as σ^b decreases, firm D receives even lower information rent and is still outperformed by A and B. Similarly, increasing p_h^b results in firm D still receiving lower information rent than A and B combined, provided δ^a and δ^b are sufficiently close.*

The intuition behind this theorem is similar to that of Theorem 1s. The first part establishes the existence of cases where firm D can be outperformed by firms A and B for demand type hl. With type hl, firm A receives information rents, whereas firm B does not. For firm D, the following trade-off arises. On one hand, firm A is of the highest type whereas firm D is not, so firm D is lower in the type hierarchy, which reduces his information rent. (In fact, when δ^b is sufficiently close to δ^a, firm D of type hl is offered the same contract as type lh; therefore, firm D's loss of information rent because of lower type hierarchy is significant.) On the other hand, firm D can make use of the demand variability of product b, which firm B could not make use of. (Again, recall by Lemma 3 that higher demand variability results in more information rent.) However, when product b's demand variability is low, firm D's gain in information rent due to product b's demand variability is negligible, and as a result, firm D ends up with lower profit than firms A and B combined. For the second part, notice that when firm B has low-type demand and thus earns no information rent, lowering σ^b or increasing p_h^b has no impact on the combined information rent of A and B. Therefore, the result follows by noting that firm D's information rent decreases when σ^b is decreased (by Lemma 3) or p_h^b is increased (by Lemma 1).

Figure 2 illustrates this result. Comparing the panels from left to right, the nonpooling region in which firms A and B outperform firm D grows as the demand variability for b, σ^b, decreases, or the probability that product b has high demand, p_h^b, increases. Figure 2 again shows that the regions in which nonpooling is optimal are fairly large and not limited to very low values of σ^b. Notice that for any given σ^b and δ^b, as $\delta^a - \delta^b$ approaches zero, nonpooling is more likely to be preferred. However, the first two panels reveal that even when $\delta^a - \delta^b$ is zero, there are cases where firm D earns higher profits. This occurs when σ^b is high, thus revealing that gaining sufficiently high demand variability from product b can indeed compensate for firm D's decrease in type hierarchy when pooling. Once again, notice that buyers benefit from pooling when doing so utilizes *more* demand variability for generating information rent.

Concluding the results for types lh and hl, we make the following observation. Facing a powerful strategic supplier, the only source of profit for the buyer firms is information rent. When not pooling purchases, only the high-type buyer firm gets information rent. When pooling purchases, firm D has a hierarchy disadvantage compared with a high-type firm A or B because D does not have the highest type. This can potentially be compensated for because firm D can make use of demand variability for both products, whereas for types lh or hl, the low-type firm A or B cannot. Once again, we would like to draw the reader's attention

Figure 2 Pooling and Nonpooling Regions of Demand Type hl, Varying σ^b and p_h^b



to our finding that the presence of a powerful strategic supplier can reverse the common wisdom about the benefit of variability reduction through pooling. In our setting, reduced variability harms the buyer firm, but pooling that utilizes more demand variability for generating information rent can actually be beneficial for the buyer. This is counter to the received intuition about pooling.

But what happens when both products a and b have high mean demands? In this case, firms A, B, and D all have the highest type, so the trade-off of hierarchy disadvantage versus increased variability that we set up above does not apply. Interestingly, situations where pooling results in lower profits can be easily found in this case as well.

Theorem 3. *Suppose demands for products a and b are symmetric:* $\delta^a = \delta^b = \delta$, $p_h^a = p_h^b = p_h$, $\sigma^a = \sigma^b = \sigma$, $p_h > 0.15$, *and* σ *is sufficiently large. If firm D of type hh receives lower information rent than A and B combined, then as* p_h *increases, firm D receives even lower information rent and is still outperformed by A and B.*

With demand type hh, both products' demand variability will generate information rent when firms A and B purchase separately, but for firm D the variability is reduced upon pooling because of statistical economies of scale. This gives firms A and B a variability advantage over firm D, which gets greater as σ gets larger. On the other hand, firm D's type-hh demand has the highest rank of four possible types in the demand type hierarchy, whereas A and B's type-h demands are only the higher of two possible demand types. As a result, D gains an advantage in type hierarchy. This however is only a significant advantage when p_h is small (e.g., $p_h = 0.05$) because then p_h^2 would be *very* small and, by Lemma 1, the rareness of type-hh demand leads to a large information rent for firm D of type hh. As p_h increases, the hh type becomes less rare, and this coupled with the decrease in variability from pooling ends up making firm D worse off compared to firms A and B. Figure 3 illustrates this trend. Reading the panels from left to right, as p_h increases, the nonpooling region in which firms A and B outperform D grows. (Notice that although the sufficient conditions of Theorem 3 require $p_h^a = p_h^b$, it is easy to generate examples where firms A and B dominate firm D when $p_h^a = p_h^b$ is violated, as can be seen on the last panel of Figure 3.)

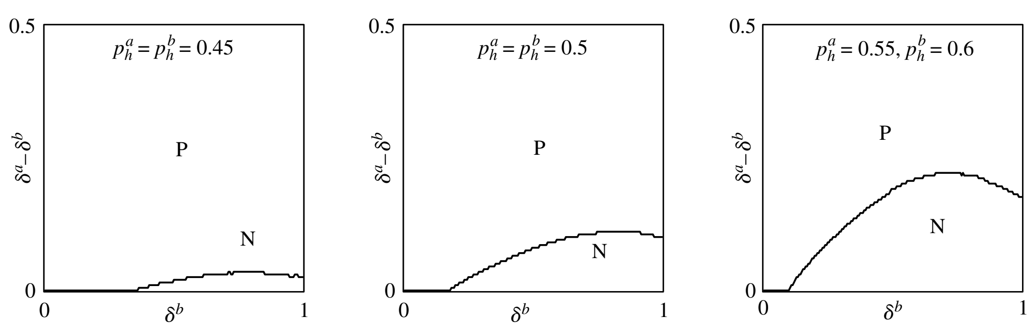
Although in the above results we focused on cases when firms A and B outperform firm D, the intuition we gained can also predict when the opposite happens. For example, consider the case where product b's demand has small gap but huge variability. On his own, firm B would earn little information rent (e.g., *zero* information rent if the gap is zero, per Lemma 2). In this case firm B's demand variability is "wasted." However, firm D can combine the gap from product a with the variability from product b to generate significant information rent. This is captured in Theorem 4.

Theorem 4. *Fixing all other parameters, as* δ^b *becomes sufficiently small, firm D of types hl and hh receive higher information rent than A and B combined.*

Notice that Theorem 4 does not consider type lh. In this case, firm A receives no information rent because of having low mean demand, and when δ^b is small, firms B and D both receive only negligible information rents. Therefore, the profit comparisons for type lh are trivial. Theorem 4 is clearly demonstrated in Figures 2 and 3 because the pooling regions occur near the left edge (where $\delta^b \approx 0$). We again point out that, although Theorem 4 describes a case of our model where pooling is indeed beneficial for buyers, the reason is completely different from the standard pooling logic: firm D receives higher profit because compared to firm B, he can better utilize product b's demand variability to generate information rent.

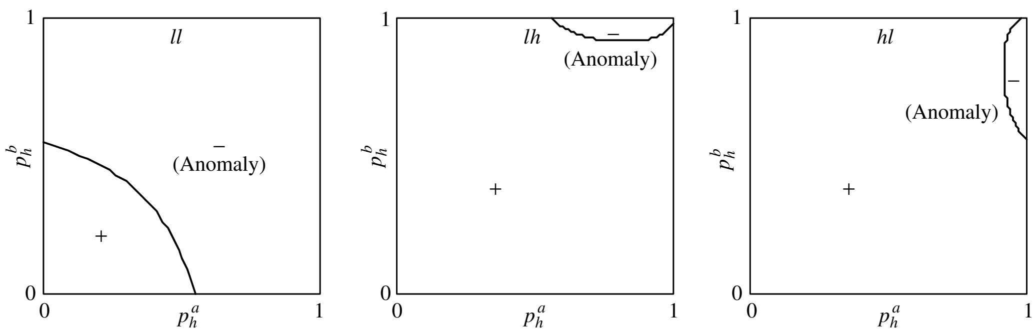
It is also interesting to note that unlike the lh type we discussed earlier, where pooling caused the entire supply chain to be worse off, when both firms have high demands, the supply chain always benefits from the two firms pooling demands. This is because when a firm's demand has the highest type, its purchase quantity suffers no downward distortion and

Figure 3 Pooling and Nonpooling Regions of Demand Type hh, Varying p_h^i



Note. In this figure, $\sigma^i = 3$.

Figure 4 $Q^D_{\theta^a\theta^b} - Q^a_{\theta^a} - Q^b_{\theta^b}$ **of Demand Types** ll, lh, **and** hl, **and** $q = 0.6$



always equals the supply chain-optimal newsvendor quantity. This raises the interesting question of how purchase quantities are affected by pooling in our setting. Below we explore the similarities and differences between the effect of pooling in traditional settings with exogenous pricing and in our setting with a powerful supplier offering price-quantity menus.

In a traditional pooling model (where the good is purchased from a commodity market at a fixed price c), how inventories change upon pooling is primarily driven by the *critical ratio* $(r - c)/r$ $(= 1 - q$ in our model), where r is the buyer's revenue from selling one unit of the good. When the demand pdf is symmetric about its mean (as it is in the seminal paper Eppen 1979), total inventory decreases upon pooling if $(r - c)/r > 0.5$ and increases if $(r - c)/r < 0.5$. The intuition is very simple: When the critical ratio is high (low), lost sales are more (less) expensive than leftover inventory; thus, it is optimal to overstock (understock). Pooling reduces demand uncertainty, so the level of overstock (understock) to achieve the critical ratio is also reduced. This translates into decreased (increased) inventory when $(r-c)/r > (<)\ 0.5$. Note that the assumption of symmetric demand pdf is crucial; for example, Yang and Schrage (2009) show that with a right-skewed demand pdf, "inventory anomaly" can occur; namely, inventory can increase upon pooling even when $(r - c)/r > 0.5$. In our model, interestingly, even when assuming symmetric (normal) demand pdfs, "inventory anomaly" can still occur, thanks to the presence of information asymmetry. The next figure provides examples. In these examples, we set $q = 0.6$, $\sigma^a = \sigma^b = 1$, $\delta^a = \delta^b = 1$, and plot $Q^D_{\theta^a\theta^b} - Q^a_{\theta^a} - Q^b_{\theta^b}$ for varying p_h^a and p_h^b and types ll, lh, and hl. (There is no anomaly for type hh, as is explained in the next theorem.) In the traditional pooling model, setting $q = 0.6$ (corresponding to a critical ratio of 0.4) would result in positive values of $Q^D_{\theta^a\theta^b} - Q^a_{\theta^a} - Q^b_{\theta^b}$ (increased inventory upon pooling), but this is not always the case in our model.

The reason for the "anomaly" in our model is that information asymmetry causes downward distortion of purchase quantities. In fact, one can show (see Lemma 1's proof in the appendix) that the rarer the type, the greater the quantity distortion for that type. Pooling changes the type distribution and affects the level of downward distortion, which can lead to the anomaly. For example, in the first panel (type ll), when p_l^a and p_l^b are small (say 0.3), firm D of type ll is much rarer than firms A and B of type l $(p^D_{ll} = 0.09)$. Consequently, firm D experiences much stronger downward distortion in purchase quantity than firms A and B, leading to the anomaly. Similar observations can be made in the other two figures.

On the other hand, when p_l^a and p_l^b are sufficiently large, downward distortion in purchase quantity is weak. We also know that the highest type never experiences downward distortion. In these cases the purchase quantities behave as in the traditional pooling literature. The theorem below states conditions for the purchase quantities to behave as in traditional pooling models despite the presence of the strategic supplier and information asymmetry when the critical ratio $(r - c)/r > 0.5$. (The result for the case where $(r - c)/r < 0.5$ is similar and omitted for brevity.)

Theorem 5. *Assume the critical ratio* $(r - c)/r > 0.5$. *Then for any firm A's type* θ^a *and firm B's type* θ^b, *when* p_l^a *and* p_l^b *are sufficiently close to* 1, $Q^a_{\theta^a} + Q^b_{\theta^b} > Q^D_{\theta^a\theta^b}$. *In particular, for demand type hh*, $Q^a_h + Q^b_h > Q^D_{hh}$ *if* $(r - c)/r > 0.5$.

The above discussion shows that in our model the traditional pooling intuition and information asymmetry both influence the behavior of purchase quantities upon pooling. When downward distortion is weak, the purchase quantities behave much like in a traditional model. On the other hand, information asymmetry and downward distortion can lead to "inventory anomaly"; namely, inventory moves in the opposite direction of what traditional pooling intuition suggests. This means we have identified information asymmetry as another possible cause of inventory anomaly, along with skewed demand distributions as previously identified in Yang and Schrage (2009).

5. Concluding Discussion

In this paper, we consider several buyers with private demand information purchasing a common critical component from a powerful sole-source supplier and ask whether pooling purchases necessarily improves the buyers' profits. Our setting and the traditional pooling literature have a fundamental difference: We consider a powerful supplier who maximizes her profit by strategically designing purchase contracts based on her anticipation about her customers' demands, whereas the traditional literature usually does not model a strategic supplier and the buyers effectively purchase from a commodity market at a fixed price. In our setting, pooling that significantly reduces demand variability can actually result in reduced profits for the buyers because of reduced information rent. We additionally show that the analysis is complex and subtle because the comparison does not only depend on the variability but also the "rank order" of the buyers' demands. Besides comparing pooling versus nonpooling profits, we also show that information asymmetry and the resulting downward distortion can be a cause of "inventory anomaly," which is different from the cause previously identified in the literature.

In our paper, we assumed that a powerful supplier screens the buyers of their private demand forecast information using two instruments combined: quantity and price. A powerful supplier dictating both quantity and price in a purchase contract is evidenced by real-life business cases (e.g., see the Apple and wireless carriers example in §1), and under this assumption we showed that pooling purchases may reduce the buyers' profits. In contrast, an essential assumption in the traditional pooling literature is that the price is fixed regardless of whether or not the buyers are pooling purchases. An interesting question is, then, if the supplier still dictates quantities but can no longer dictate prices (i.e., provides a menu of quantities for a buyer to choose from but always charges a fixed unit purchase price), can pooling purchases ever reduce the buyers' profits?

With fixed purchase price, the model—though seemingly simpler—lacks the nice solution structure of the more "complex" original model, and as a result, analytically solving this model is much more difficult. For example, unlike in the original model where the optimal contract quantities are interior solutions determined by FOCs and are thus continuous in their parameters, the optimal contract quantities in this model are found at boundaries and are discontinuous in their parameters. One feature of the "quantity-only" model is that the supplier may find it optimal to provide quantity levels so large that some of the low-demand type buyers would not find it profitable to purchase. (This is akin to companies imposing large minimum order quantities. However, in our setting, the supplier is imposing these large quantities in an attempt to improve profits from higher types.) Such a situation where multiple types find it unprofitable to purchase does not arise when the supplier can set both quantity and price because she can achieve a finer control and extract high-type buyers' profits without having to give up low-type buyers.

Interestingly, pooling can still reduce the buyers' profits even with the purchase price fixed. Below we present one numerical comparison of the purchase quantities Q, prices p, and profits π of pooling and nonpooling buyers, for the cases where the price is fixed and where the supplier sets both price and quantity. In this example, $p_h^a = p_h^b = 0.7$, $\sigma^a = \sigma^b = 1$, $\mu_l^a = \mu_l^b = 6$, $\mu_h^a = 8.5$, and $\mu_h^b = 8$ ($\delta^a = 2.5$, $\delta^b = 2$). The retail price is $r = 1$, and the supplier's manufacturing cost is $c = 0.4$. In the fixed price scenario, we set the buyers' purchase price to be $p = 0.8$.

In the example illustrated in Table 1, when purchase price is fixed, pooling hurts buyers[5] of demand types lh, hl, and hh (buyers of type ll are indifferent as in the rest of the paper), whereas when the supplier strategically sets the purchase price, pooling only hurts buyers of type lh. This is because when the supplier is forced to sell the component to everyone at the same price, facing the pooled buyers, she offers the product in only one quantity level (17.5). Buyers of type ll would lose money purchasing this quantity and so choose not to purchase at all (and type lh buyers just break even purchasing this quantity). In other words, the supplier requires a large order quantity to extract the most profit from higher-type firms but has to give up on lower-type firms who cannot afford the quantity. Interestingly, when the supplier can set both price and quantity, she can extract *more* profit from higher types *and* induce lower types to also purchase by adjusting prices; but when she can only set quantity, her instrument is much more blunt and her effort to maximize her profits results in rejecting the lowest type buyers, which significantly reduces profits for the pooled buyers. The takeaway here is that our paper's central message—pooling purchases may reduce the buyers' profits when the supplier strategically tailors contracts to the buyers—continues to hold even when the supplier is less powerful and can only set quantities but not prices.

In our model we assume that either unpooled firm A or B or pooled firm D approaches the supplier to purchase the component, and the supplier designs a menu of contracts accordingly. This implies that the pooling decision has been made before the buyers

[5] For example, firm A of type h and firm B of type l would make a total profit of 1.41 when purchasing separately, but a profit of only 0.5 when purchasing together as firm D of type hl.

Table 1 Purchase Quantities, Prices, and Profits of Nonpooling and Pooling Buyers, with Fixed and Strategic Price

	Q_l^a	p_l^a	π_l^a	Q_h^a	p_h^a	π_h^a	Q_l^b	p_l^b	π_l^b	Q_h^b	p_h^b	π_h^b
Nonpooling, fixed price	7.46	0.80	0.00	7.85	0.80	1.41	7.46	0.80	0.00	7.46	0.80	1.31
Nonpooling, strategic price	5.09	0.98	0.00	8.75	0.93	0.10	5.09	0.98	0.00	8.25	0.92	0.10
	Q_{ll}^D	p_{ll}^D	π_{ll}^D	Q_{lh}^D	p_{lh}^D	π_{lh}^D	Q_{hl}^D	p_{hl}^D	π_{hl}^D	Q_{hh}^D	p_{hh}^D	π_{hh}^D
Pooling, fixed price	—	0.80	—	17.50	0.80	0.00	17.50	0.80	0.50	17.50	0.80	2.30
Pooling, strategic price	9.74	1.00	0.00	13.10	0.98	0.03	13.25	0.98	0.14	16.86	0.94	0.28

approach the supplier, and the pooling decision is observable and verifiable by the supplier. This implication is supported by the observation that suppliers often customize goods to individual customers (through specifications, line-side delivery agreements, etc.), in which case the buyers' pooling decisions would be transparent to the supplier. For example, Apple implements a baseband encryption mechanism that locks every iPhone to the purchasing carrier, so if Sprint and C Spire separately sign purchase contracts with Apple, Apple knows that they cannot share their iPhone inventories; and if Sprint acquired C Spire, it will ask Apple to offer iPhones locked to the new combined carrier, which signals to Apple that they are pooling their purchases.

Our paper explores when and why pooling purchases leads to higher/lower buyer profits. We did so for demand types lh, hl, and hh (the comparison for type ll is trivial). An interesting follow-up question is whether two buyers *should* pool their purchases, i.e., the buyers' pooling decision. For example, if Sprint acquired C Spire, what would be the effect on the profit that the firm makes from purchasing iPhones from Apple? Our results can be used to evaluate the long-term impact of pooling: One only needs to consider the prior probabilities of having each demand type and then calculate the ex ante expectation of pooling and nonpooling profits (information rents). Of course, other factors also need consideration, such as the implementation costs of pooling (e.g., reconfiguring the carriers' networks), details that we leave for future research.

Acknowledgments
The authors acknowledge the department editor, associate editor, and referees for constructive suggestions that improved the paper.

Appendix

Proof of Proposition 1
Let the first derivative of the supplier's objective function with respect to Q_{θ_j} be denoted by $FD_{\theta_j}(Q_{\theta_j}) \doteq r\Pr(\Theta \geq \theta_j) \cdot [\lambda(\theta_j)(1-q) - F(Q_{\theta_j} - \mu_{\theta_j}) + (1-\lambda(\theta_j))F(Q_{\theta_j} - \mu_{\theta_{j+1}})]$ and the second derivative be denoted by $SD_{\theta_j}(Q_{\theta_j}) \doteq r\Pr(\Theta \geq \theta_j) \cdot [(1-\lambda(\theta_j))f(Q_{\theta_j} - \mu_{\theta_{j+1}}) - f(Q_{\theta_j} - \mu_{\theta_j})]$. It is straightforward to observe that $FD_{\theta_j}(Q_{\theta_j})$ is positive when Q_{θ_j} is sufficiently small (so that $F(Q_{\theta_j} - \mu_{\theta_j}) = 0$) and negative when Q_{θ_j} is sufficiently large (so that $F(Q_{\theta_j} - \mu_{\theta_{j+1}}) = 1$). Assume f has support over $(-s, s)$ and satisfies all four conditions in Proposition 1. Next, for each of three possible cases, we show that $FD_{\theta_j}(Q_{\theta_j})$ is first positive and then negative; thus, the supplier's objective function is unimodal and the unique solution to (4) characterizes the global optimal solution.

Case 1. $\mu_{\theta_{j+1}} - \mu_{\theta_j} \geq 2s$ ($\mu_{\theta_j} - s < \mu_{\theta_j} + s \leq \mu_{\theta_{j+1}} - s < \mu_{\theta_{j+1}} + s$).

In this case, it is clear that $FD_{\theta_j}(Q_{\theta_j})$ is decreasing over $(\mu_{\theta_j} - s, \mu_{\theta_j} + s)$, constant over $[\mu_{\theta_j} + s, \mu_{\theta_{j+1}} - s]$, and increasing over $(\mu_{\theta_{j+1}} - s, \mu_{\theta_{j+1}} + s)$. Yet we already know $FD_{\theta_j}(Q_{\theta_j}) > 0$ when $Q_{\theta_j} = \mu_{\theta_j} - s$ and $FD_{\theta_j}(Q_{\theta_j}) < 0$ when $Q_{\theta_j} = \mu_{\theta_{j+1}} + s$. Thus, it is clear that $FD_{\theta_j}(Q_{\theta_j})$ is first positive and then negative.

Case 2. $s < \mu_{\theta_{j+1}} - \mu_{\theta_j} < 2s$ ($\mu_{\theta_j} - s < \mu_{\theta_j} < \mu_{\theta_{j+1}} - s < \mu_{\theta_j} + s < \mu_{\theta_{j+1}} < \mu_{\theta_{j+1}} + s$).

In this case, it is clear that $FD_{\theta_j}(Q_{\theta_j})$ is decreasing over $(\mu_{\theta_j} - s, \mu_{\theta_{j+1}} - s)$ and increasing over $(\mu_{\theta_j} + s, \mu_{\theta_{j+1}} + s)$. Between $\mu_{\theta_{j+1}} - s$ and $\mu_{\theta_j} + s$, the second derivative $SD_{\theta_j}(Q_{\theta_j})$ equals $r\Pr(\Theta \geq \theta_j)[(1-\lambda(\theta_j))f(Q_{\theta_j} - \mu_{\theta_{j+1}}) - f(Q_{\theta_j} - \mu_{\theta_j})]$. By Condition 3, $f(Q_{\theta_j} - \mu_{\theta_{j+1}})$ is nondecreasing and $f(Q_{\theta_j} - \mu_{\theta_j})$ is nonincreasing; therefore, $SD_{\theta_j}(Q_{\theta_j})$ is nondecreasing. Because of this, we know $FD_{\theta_j}(Q_{\theta_j})$ is convex over $(\mu_{\theta_{j+1}} - s, \mu_{\theta_j} + s)$. If $FD_{\theta_j}(Q_{\theta_j}) < 0$ at $\mu_{\theta_{j+1}} - s$, then by convexity of $FD_{\theta_j}(Q_{\theta_j})$, $FD_{\theta_j}(Q_{\theta_j})$ must be negative over $(\mu_{\theta_{j+1}} - s, \mu_{\theta_j} + s)$. If $FD_{\theta_j}(Q_{\theta_j}) > 0$ at $\mu_{\theta_{j+1}} - s$, then by convexity of $FD_{\theta_j}(Q_{\theta_j})$, $FD_{\theta_j}(Q_{\theta_j})$ crosses zero only once. In both cases, it is clear that $FD_{\theta_j}(Q_{\theta_j})$ is first positive and then negative.

Case 3. $\mu_{\theta_{j+1}} - \mu_{\theta_j} \leq s$ ($\mu_{\theta_j} - s < \mu_{\theta_{j+1}} - s \leq \mu_{\theta_j} < \mu_{\theta_{j+1}} \leq \mu_{\theta_j} + s < \mu_{\theta_{j+1}} + s$).

In this case, it is clear that $FD_{\theta_j}(Q_{\theta_j})$ is decreasing over $(\mu_{\theta_j} - s, \mu_{\theta_j})$, convex over $(\mu_{\theta_j}, \mu_{\theta_{j+1}})$ (similar to Case 2), and increasing over $(\mu_{\theta_j} + s, \mu_{\theta_{j+1}} + s)$. We now characterize $FD_{\theta_j}(Q_{\theta_j})$ over $(\mu_{\theta_{j+1}}, \mu_{\theta_j} + s)$. Recall that the second derivative $SD_{\theta_j}(Q_{\theta_j})$ equals $r\Pr(\Theta \geq \theta_j)[(1-\lambda(\theta_j))f(Q_{\theta_j} - \mu_{\theta_{j+1}}) - f(Q_{\theta_j} - \mu_{\theta_j})]$. By Condition 4, $f(Q_{\theta_j} - \mu_{\theta_j})/f(Q_{\theta_j} - \mu_{\theta_{j+1}})$ is nonincreasing in Q_{θ_j}. Thus, we know over $(\mu_{\theta_{j+1}}, \mu_{\theta_j} + s)$, $SD_{\theta_j}(Q_{\theta_j})$ cannot first be positive and then negative; it may only be always positive, always negative, or first negative and then positive.

When $SD_{\theta_j}(Q_{\theta_j})$ is always positive over $(\mu_{\theta_{j+1}}, \mu_{\theta_j} + s)$, $FD_{\theta_j}(Q_{\theta_j})$ is increasing over $(\mu_{\theta_{j+1}}, \mu_{\theta_{j+1}} + s)$. Following the same proof as in Case 2, we know $FD_{\theta_j}(Q_{\theta_j})$ must be first positive and then negative.

When $SD_{\theta_j}(Q_{\theta_j})$ is negative over $(\mu_{\theta_{j+1}}, t)$ and positive over $[t, \mu_{\theta_j} + s)$ where $\mu_{\theta_{j+1}} < t \leq \mu_{\theta_j} + s$, we know the right-derivative of $FD_{\theta_j}(Q_{\theta_j})$ at $\mu_{\theta_{j+1}}$ is negative. Condition 2 then guarantees that the left-derivative of $FD_{\theta_j}(Q_{\theta_j})$ at $\mu_{\theta_{j+1}}$ is also negative. Combining this and the fact that $FD_{\theta_j}(Q_{\theta_j})$ is convex over $(\mu_{\theta_j}, \mu_{\theta_{j+1}})$, we know $FD_{\theta_j}(Q_{\theta_j})$ must also be decreasing over $(\mu_{\theta_j}, \mu_{\theta_{j+1}})$. As a result, $FD_{\theta_j}(Q_{\theta_j})$ is decreasing over $(\mu_{\theta_j}-s, t)$ and increasing over $(t, \mu_{\theta_{j+1}}+s)$, so it is first positive and then negative. Q.E.D.

Proof of Proposition 2

To prove that the violation of the MCs can only occur in the two cases mentioned in the proposition, we need to show the solution to (8)–(11) (ignoring MCs (7f)) always satisfies $Q^D_{ll} < Q^D_{hl} < Q^D_{hh}$.

As discussed in §3.1, the participation and incentive compatibility constraints are always binding at optimality. First plug the binding constraints (7b)–(7e) into (7a); then take the first derivative of the objective function with respect to Q^D_{hl}. Doing so yields

$$FD_{hl}(Q) \doteq \left.\frac{\partial[\text{Equation (7a)}]}{\partial Q^D_{hl}}\right|_{Q^D_{hl}=Q}$$
$$= p^D_{hl}(1-q)-(p^D_{hl}+p^D_{hh})F^D(Q-\mu^D_{hl})+p^D_{hh}F^D(Q-\mu^D_{hh}).$$

Because of Proposition 1, we know that $FD_{hl}(Q)=0$ always has an interior solution. Recall that from (11) we have $Q^D_{hh}=\mu^D_{hh}+(F^D)^{-1}(1-q)$. If we evaluate $FD_{hl}(Q)$ at $Q=Q^D_{hh}=\mu^D_{hh}+(F^D)^{-1}(1-q)$, we can easily see that $FD_{hl}(Q^D_{hh})<0$. This means the solution to $FD_{hl}(Q)=0$, Q^D_{hl}, is smaller than Q^D_{hh}.

We next show that $Q^D_{ll} < Q^D_{hl}$. Notice that $\mu^D_{lh}-\mu^D_{ll} = (\mu^a_l+\mu^b_h)-(\mu^a_l+\mu^b_l) = (\mu^a_h+\mu^b_h)-(\mu^a_h+\mu^b_l) = \mu^D_{hh}-\mu^D_{hl}$. Define $\delta \doteq \mu^D_{lh}-\mu^D_{ll} = \mu^D_{hh}-\mu^D_{hl}$. Using (8), (10), (12), and (13), and applying the definitions of δ and γ, (8) and (10) can be rewritten as

$$\lambda(\theta)(1-q) = F(s)-(1-\lambda(\theta))F(s-\delta), \tag{16}$$

and (12) and (13) can be rewritten as

$$\pi_{\theta'}-\pi_\theta = r\int_{s-\delta}^{s} F(x)\,\mathrm{d}x \tag{17}$$

for types $(\theta,\theta') = (ll, lh)$ and (hl, hh), where $s = Q_\theta-\mu_\theta$, $\lambda(ll)=p^D_{ll}$ for (8), and $\lambda(hl)=p^D_{hl}/(p^D_{hl}+p^D_{hh})$ for (10). Taking the total derivative of (16) with respect to $\lambda(\theta)$ yields

$$1-q = \frac{\mathrm{d}s}{\mathrm{d}\lambda(\theta)}[f(s)-(1-\lambda(\theta))f(s-\delta)]+F(s-\delta)$$
$$\Longrightarrow \quad \frac{\mathrm{d}s}{\mathrm{d}\lambda(\theta)} = \frac{1-q-F(s-\delta)}{f(s)-(1-\lambda(\theta))f(s-\delta)}. \tag{18}$$

Because the objective function is concave at the global optimal solution (by Proposition 1), we know the denominator of (18) is positive. The numerator is also positive, because (16) and the fact that $F(s) > F(s-\delta)$ together imply $\lambda(\theta)\cdot(1-q) = F(s)-(1-\lambda(\theta))F(s-\delta) > \lambda(\theta)F(s-\delta)$. Therefore, we know s is increasing in $\lambda(\theta)$. Notice that $\lambda(ll) = p^D_{ll} = p^a_l p^b_l < \lambda(hl) = p^D_{hl}/(p^D_{hl}+p^D_{hh}) = p^a_h p^b_l/(p^a_h(p^b_l+p^b_h)) = p^b_l$. As a result, $Q^D_{ll}-\mu^D_{ll} = Q^D_{ll}-\mu^a_l-\mu^b_l < Q^D_{hl}-\mu^a_h-\mu^b_l = Q^D_{hl}-\mu^D_{hl}$, which immediately yields $Q^D_{ll} < Q^D_{hl}$.

Ultimately we wish to solve optimization problem (7a)–(7f). Thus far we have examined the solution obtained with MCs (7f) ignored. If the solution satisfies (7f), then it is the optimal solution to (7a)–(7f). If it violates (7f), which as argued above can only be true if $Q^D_{ll} > Q^D_{lh}$ or $Q^D_{lh} > Q^D_{hl}$, then we must revise our solution. We now show how to revise (8)–(10) to obtain the true optimal solution. We show this for the case when $Q^D_{ll} \le Q^D_{lh}$ is violated; the other case is similar.

Suppose when ignoring the MCs, the FOC solution is such that $Q^D_{ll} > Q^D_{lh}$. Consider a function $I(x,y) = G(x)+H(y)$ such that both G and H are unimodal, and suppose the x_0 that maximizes G and the y_0 that maximizes H are such that $x_0 > y_0$. If we impose the requirement that $x \le y$ and find the optimal $(x^*, y^*) \doteq \{(x,y) \mid x \le y,\ G(x)+H(y) \ge G(x')+H(y'),\ \forall x' \le y'\}$, it is easy to see (by a contradiction argument) that $y_0 \le x^* = y^* \le x_0$. Because by Proposition 1 our maximization problem is decomposable, and unimodal in both Q^D_{ll} and Q^D_{lh}, this result applies to our problem. We thus know that adding back the violated constraint $Q^D_{ll} \le Q^D_{lh}$ will result in a new optimal solution for which $Q^D_{ll} = Q^D_{lh} = \underline{Q}$ for some $\underline{Q}$. Differentiating the objective function with this equality enforced and setting the derivative to zero yields FOC (14). Furthermore, because (14) is a specific instance of (4), we can apply Proposition 1 to conclude that the FOC solution $\underline{Q}$ will be the unique global maximizer of the supplier's objective function. Q.E.D.

Proof of Lemma 1

By (5), $\pi_{\theta'}-\pi_\theta = r\int_{Q_\theta-\mu_{\theta'}}^{Q_\theta-\mu_\theta} F(x)\,\mathrm{d}x$, so to show that $\pi_{\theta'}-\pi_\theta$ increases in $\lambda(\theta)$, it is sufficient to show that Q_θ increases in $\lambda(\theta)$. We focus on type ll; arguments for the other types are similar. Depending on whether the MC for type ll in (7f) is slack or binding, the equation that determines Q_{ll} is either (8) or (14), respectively. We denote the solution to FOC (8) by $Q^*(8)$, and the solution to FOC (14) by $Q^*(14)$. Similarly denote the solution to FOC (9) by $Q^*(9)$.

Suppose when $\lambda(ll)=\lambda_1$, $Q^*(8) < Q^*(9)$ (so the MC for type ll in (7f) is slack and $Q^D_{ll} = Q^*(8)$). To prove Lemma 1 for this case, we need to show three things: (i) While the MC for type ll remains slack, when $\lambda(ll)$ increases, $Q^*(8)$ increases. (ii) If increasing $\lambda(ll)$ to some point $\lambda(ll)=\lambda_2$ causes the MC for type ll to become binding, then we switch our consideration to the parameter $\underline{\lambda}$ and need to show that when $\underline{\lambda}$ increases, $Q^*(14)$ increases. (iii) At the "switch point" $\lambda(ll)=\lambda_2$ between the nonbunching and bunching cases, we have continuity; namely, $Q^D_{ll} = Q^*(8) = Q^*(14)$.

Arguments (i) and (ii) follow because both (8) and (14) fit the structure of (4), so as argued following (16), we know $Q^*(8)$ and $Q^*(14)$ are increasing in $\lambda(ll)$ and $\underline{\lambda}$, respectively. We next prove (iii). $Q^*(8)$ increases in $\lambda(ll)$. Because $Q^*(9)$ remains constant (because we are increasing $\lambda(ll)$ in a way that keeps $\lambda(lh)$ fixed), if type ll's MC becomes binding at $\lambda(ll)=\lambda_2$, then $Q^*(8)=Q^*(9)$ at $\lambda(ll)=\lambda_2$. Furthermore, as argued following (16), we know that $Q^*(8) \le Q^*(14) \le Q^*(9)$, which in turn implies that $Q^D_{ll} = Q^*(8) = Q^*(14)$ at $\lambda(ll)=\lambda_2$.

The above arguments proved the lemma when increasing $\lambda(ll)$ took us from nonbunching to bunching. The case where increasing $\underline{\lambda}$ takes us from bunching to nonbunching can be treated analogously. Iteratively applying these results yields the lemma over any interval of $\lambda(ll)$ or $\underline{\lambda}$. Q.E.D.

Proof of Lemma 2
Regardless of whether there is bunching, the equations characterizing the optimal solution and information rent always fit the structure of (4) and (5), respectively. Therefore, we only need to prove Lemma 2 for the general Equations (4) and (5).

Consider two pairs of mean demands $\mu_\theta, \mu_{\theta'}$ and $\mu'_\theta, \mu'_{\theta'}$ such that $\mu_{\theta'} - \mu_\theta = \mu'_{\theta'} - \mu'_\theta = \delta$. Notice that (4) and (5), with either pair of mean demands, can be rewritten as (16) and (17) where $s = Q_\theta - \mu_\theta$ or $s = Q'_\theta - \mu'_\theta$. Therefore, $\mu_\theta, \mu_{\theta'}$ and $\mu'_\theta, \mu'_{\theta'}$ yield the same information rent $\pi_{\theta'} - \pi_\theta$. This shows that μ_θ and $\mu_{\theta'}$ determine information rent only through $\delta = \mu_{\theta'} - \mu_\theta$. The inequality $\pi_{\theta'} - \pi_\theta \le r\delta$ follows because $F(x) \le 1$ for all x. Q.E.D.

Proof of Lemma 3
We prove the result for type $\theta = ll$ only; the proofs for other types are analogous. The proof's structure is similar to that of Lemma 1's proof. Suppose when $\gamma = \gamma_1$, $Q^*(8) < Q^*(9)$ (so the MC for type ll in (7f) is slack, meaning $Q^D_{ll} = Q^*(8)$). To prove Lemma 3 for this case, we need to show three things. (i) While the MC for type ll remains slack, when γ increases $\pi_{lh} - \pi_{ll}$ also increases. (ii) If increasing γ to some point $\gamma = \gamma_1$ causes the MC for type ll to become binding, then we switch our consideration to the bunching case, meaning $Q^D_{ll} = Q^*(9)$, and we need to show that when γ continues to increase, $\pi_{lh} - \pi_{ll}$ increases. (iii) At the "switch point" $\gamma = \gamma_1$ between the nonbunching and bunching cases, we have continuity, namely $Q^D_{ll} = Q^*(8) = Q^*(14)$ (which implies that the information rent gap $\pi_{lh} - \pi_{ll}$ is also continuous at the switch point).

We begin by addressing (iii). The arguments are precisely the same as that used to prove point (iii) in the proof of Lemma 1, except with γ's in the role of λ's. Therefore, it suffices to show (i) and (ii).

Notice that the incremental information rent $\pi_{lh} - \pi_{ll}$ is determined by (12) together with (8) or (14) (corresponding to nonbunching or bunching, respectively). In both cases the incremental information rent can be expressed in the following general way, where θ and θ' play the roles of ll and lh, respectively:

$$\lambda(\theta)(1-q) = F_{(\gamma)}(\tilde{s}) - (1-\lambda(\theta))F_{(\gamma)}(\tilde{s}-\delta), \tag{19}$$

$$\tilde{\pi}_{\theta'} - \tilde{\pi}_\theta = r\int_{\tilde{s}-\delta}^{\tilde{s}} F_{(\gamma)}(x)\,\mathrm{d}x, \tag{20}$$

where $\delta \doteq \mu_{\theta'} - \mu_\theta$. Suppose $\gamma > 1$. To show that the incremental information rent increases in γ, it suffices to show that $\tilde{\pi}_{\theta'} - \tilde{\pi}_\theta$ is greater than $\pi_{\theta'} - \pi_\theta$, which we will define to be the incremental information rent when $\gamma = 1$; namely, $\pi_{\theta'} - \pi_\theta = r\int_{s-\delta}^{s} F_{(1)}(x)\,\mathrm{d}x$ where $\lambda(\theta)(1-q) = F_{(1)}(s) - (1-\lambda(\theta))F_{(1)}(s-\delta)$. To this end, let s' be the root of the following equation:

$$\lambda(\theta)(1-q) = F_{(\gamma)}(s') - (1-\lambda(\theta))F_{(\gamma)}(s'-\gamma\delta), \tag{21}$$

and notice that $s' = \gamma s$. We now show that $s' \le \tilde{s}$. Because s' is the solution to (21), we have $F_{(\gamma)}(s') - (1-\lambda(\theta))F_{(\gamma)}(s'-\delta) \le F_{(\gamma)}(s') - (1-\lambda(\theta))F_{(\gamma)}(s'-\gamma\delta) = \lambda(\theta)(1-q)$. By the proof of Proposition 1, we know that when $\tilde{s}$ increases, the RHS of (19) is first smaller then greater than $\lambda(\theta)(1-q)$. Therefore, we know the solution to (19), $\tilde{s}$, must be greater than s'.

Now we show that $\tilde{\pi}_{\theta'} - \tilde{\pi}_\theta \ge \pi_{\theta'} - \pi_\theta$. In fact, since $\tilde{s} \ge s'$, we have

$$\begin{aligned}
\tilde{\pi}_{\theta'} - \tilde{\pi}_\theta &= r\int_{\tilde{s}-\delta}^{\tilde{s}} F_{(\gamma)}(x)\,\mathrm{d}x \ge r\int_{s'-\delta}^{s'} F_{(\gamma)}(x)\,\mathrm{d}x\\
&= r\int_0^{\delta} F_{(\gamma)}(s'-t)\,\mathrm{d}t = \frac{1}{\gamma} r\int_0^{\gamma\delta} F_{(\gamma)}(s'-k/\gamma)\,\mathrm{d}k\\
&\ge \frac{1}{\gamma} r\int_0^{\gamma\delta} F_{(\gamma)}(s'-k)\,\mathrm{d}k = \frac{1}{\gamma} r\int_{s'-\gamma\delta}^{s'} F_{(\gamma)}(y)\,\mathrm{d}y\\
&= \frac{1}{\gamma}\gamma(\pi_{\theta'} - \pi_\theta) = \pi_{\theta'} - \pi_\theta,
\end{aligned}$$

where $k = \gamma t$ and $y = s' - k$. This proves (i) and (ii).

The above arguments prove the lemma when increasing γ took us from nonbunching to bunching. The opposite case where increasing γ takes us from bunching to nonbunching can be treated analogously. Iteratively applying these results yields the lemma over any interval of γ. Q.E.D.

Proof of Theorem 1
Before proving the theorem we introduce a technique that will appear multiple times in this and later proofs. We first note that the incremental information rent $\pi_{\theta'} - \pi_\theta$ is continuous in $\lambda(\theta)$, σ, and δ^θ. (Continuity in $\lambda(\theta)$ and γ was shown in the course of proving Lemmas 1 and 3. Continuity in δ^θ comes from similar arguments.)

By the definition of continuity, when we establish a property about the size of the incremental information rent at a certain point in the parameter space, it will also hold in a sufficiently small neighborhood around that point. With this technique, we can show that our desired properties hold in a neighborhood of a particular point by showing they hold *at* the particular point.

We begin with the first part of the theorem. When $\sigma^a = 0$, $\sqrt{(\sigma^a)^2 + (\sigma^b)^2} = \sigma^b$. Notice that firm B's demand of type h has gap δ^b, variability σ^b, and $\lambda(l) = p^b_l$. In comparison, firm D's demand of type lh has gap δ^b, variability $\sqrt{(\sigma^a)^2 + (\sigma^b)^2} = \sigma^b$, and $\lambda(ll) = p^a_l p^b_l < p^b_l = \lambda(l)$. Thus, firm B's type-h demand has the same gap and variability as firm D's type-lh demand but has higher λ. Because of Lemma 1, we know $\pi^b_h > \pi^D_{lh}$; thus, firms A and B have higher combined information rent than does D. By continuity, this is also true when σ^a is sufficiently small.

We now prove the theorem's second part. When σ^a decreases, firm B's information rent will not change, nor will the low-type firm A's information rent (it is constant at zero). Thus, the combined information rents of A and B do not change. However, the information rent of D decreases because when σ^a decreases, the variability of the combined demand $\sqrt{(\sigma^a)^2 + (\sigma^b)^2}$ also decreases and, by Lemma 3, so does D's information rent. Similarly, when p^a_h increases, firms A and B's information rents are not affected, but $\lambda(ll) = p^a_l p^b_l$ decreases and (by Lemma 1) this causes D's information rent to decrease. Q.E.D.

Proof of Theorem 2
We begin with the theorem's first part. When $\delta^a = \delta^b$, types lh and hl are indistinguishable, so bunching occurs between them, and consequently $\pi^D_{hl} = \pi^D_{lh}$. Therefore, the

comparison of π_h^a versus π_{hl}^D is essentially the comparison of π_h^a versus π_{lh}^D. On the other hand, when $\sigma^b = 0$, $\sqrt{(\sigma^a)^2 + (\sigma^b)^2} = \sigma^a$. Notice that firm A with type-h demand has variability σ^a, gap δ^a, and $\lambda(l) = p_l^a$. In comparison, firm D of type-lh demand has the same variability $\sqrt{(\sigma^a)^2 + (\sigma^b)^2} = \sigma^a$, gap $\delta^b = \delta^a$, and lower $\lambda(ll) = p_l^a p_l^b < p_l^a = \lambda(l)$. Because of Lemma 1 we know $\pi_h^a > \pi_{lh}^D$; thus, firms A and B have higher combined information rent than does D. By continuity this is also true when σ^b is sufficiently small and δ^a is sufficiently close to δ^b.

The second part's proof is similar to that of Theorem 1. As σ^b decreases, the information rent of firm A's type-h demand does not change, but firm D's information rent decreases. Also, as p_h^b increases, π^a does not change, but π_{lh}^D decreases because $\lambda(ll) = p_l^a p_l^b$ decreases. Q.E.D.

Proof of Theorem 3

The theorem assumes σ is sufficiently large. However, because of the scalability of the information rents in δ and σ, we can replace this condition by δ that is sufficiently small. To see why, note that if we choose scalar $\kappa > 0$ and set $\pi_{\theta'} - \pi_\theta = r\int_{s-\delta}^{s} F_{(\sigma)}(x)\,dx$ where $\lambda(\theta)(1-q) = F_{(\sigma)}(s) - (1-\lambda(\theta))F_{(\sigma)}(s-\delta)$ and $\pi'_{\theta'} - \pi'_\theta = r\int_{s'-\kappa\delta}^{s'} F_{(\kappa\sigma)}(x)\,dx$ where $\lambda(\theta)(1-q) = F_{(\kappa\sigma)}(s') - (1-\lambda(\theta))F_{(\kappa\sigma)}(s'-\kappa\delta)$, then $s' = \kappa s$, and hence $\pi'_{\theta'} - \pi'_\theta = \kappa(\pi_{\theta'} - \pi_\theta)$. Thus, to study the relative sizes of firm A's, B's, and D's information rents with σ large, we can analyze the related scaled-down setting (where δ is small), which reduces the magnitude of the information rents but leaves their relative sizes unchanged. Consequently, for the remainder of the proof we instead use the condition $\delta \to 0$.

When the demands for a and b are symmetric, bunching occurs between types lh and hl. Thus, to show that firms A and B continue to outperform D when p_l decreases, we need to show that $\pi_h^a + \pi_h^b$'s derivative with respect to p_l is smaller than $(\pi_{hh}^D - \pi_{lh}^D) + \pi_{lh}^D$'s derivative with respect to p_l. We now establish a property that facilitates our comparison of the derivatives.

Because of Proposition 2's proof,

$$\frac{dQ_\theta}{d\lambda(\theta)} = \frac{1-q-F(Q_\theta - \mu_{\theta'})}{f(Q_\theta - \mu_\theta) - (1-\lambda(\theta))f(Q_\theta - \mu_{\theta'})}. \tag{22}$$

Our first step is to show that

$$\frac{dQ_\theta}{d\lambda(\theta)} \to \delta/\lambda^2(\theta) \quad \text{as } \delta \to 0. \tag{23}$$

We begin by noting that a Taylor expansion of F about the point $Q_\theta - \mu_{\theta'}$, together with the fact that $\mu_{\theta'} = \mu_\theta + \delta$, yields

$$F(Q_\theta - \mu_\theta) = F(Q_\theta - \mu_{\theta'}) + \delta f(Q_\theta - \mu_{\theta'}) + o(\delta). \tag{24}$$

Recall FOC (4): $\lambda(\theta)(1-q) = F(Q_\theta - \mu_\theta) - (1-\lambda(\theta)) \cdot F(Q_\theta - \mu_{\theta'})$. With (24), this implies that as $\delta \to 0$ the RHS of (4) approaches $\lambda(\theta)F(Q_\theta - \mu_{\theta'}) + \delta f(Q_\theta - \mu_{\theta'})$; hence,

$$1-q-F(Q_\theta - \mu_{\theta'}) \to \frac{\delta f(Q_\theta - \mu_{\theta'})}{\lambda(\theta)} \quad \text{as } \delta \to 0. \tag{25}$$

Because the RHS of (25) approaches zero as $\delta \to 0$, we also have

$$Q_\theta - \mu_{\theta'} \to F^{-1}(1-q) \quad \text{as } \delta \to 0. \tag{26}$$

By Lipschitz continuity of f, combining (26) and (25) yields

$$1-q-F(Q_\theta - \mu_{\theta'}) \to \frac{\delta f(F^{-1}(1-q))}{\lambda(\theta)} \quad \text{as } \delta \to 0. \tag{27}$$

We now apply these insights to the RHS of (22). Since $\mu_\theta + \delta = \mu_{\theta'}$, (26) implies that the denominator of (22)'s RHS approaches $\lambda(\theta)f(F^{-1}(1-q))$ as $\delta \to 0$. Additionally, (27) implies that the numerator of (22)'s RHS approaches $\delta f(F^{-1}(1-q))/\lambda(\theta)$ as $\delta \to 0$. Upon simplifying, we recover (23).

Furthermore, $d(\pi_{\theta'} - \pi_\theta)/dQ_\theta = r(F(Q_\theta - \mu_\theta) - F(Q_\theta - \mu_{\theta'}))$ (by (5)) $\to r\delta f(F^{-1}(1-q))$ as $\delta \to 0$ (by combining (24) and (26)). Combining this and (23) yields

$$\frac{d(\pi_{\theta'} - \pi_\theta)}{d\lambda(\theta)} = \frac{d(\pi_{\theta'} - \pi_\theta)}{dQ_\theta}\frac{dQ_\theta}{d\lambda(\theta)} \to r\delta^2 f(F^{-1}(1-q))/\lambda^2(\theta) \quad \text{as } \delta \to 0. \tag{28}$$

Let $F_{(\sigma)}$ denote the error distribution of e^i. Then $F_{(\sqrt{2}\sigma)}$ denotes the error distribution of e^D. Generally speaking, for any $\kappa > 0$, since $F_{(\kappa\sigma)}(\kappa s) = F_{(\sigma)}(s)$, we know $(d/dx)F_{(\kappa\sigma)}(x)|_{x=\kappa s} = (1/\kappa)(d/dx)F_{(\sigma)}(x)|_{x=s}$ and $F_{(\kappa\sigma)}^{-1}(1-q) = \kappa F_{(\sigma)}^{-1}(1-q)$. Combining these equalities yields

$$f_{(\kappa\sigma)}(F_{(\kappa\sigma)}^{-1}(1-q)) = \frac{1}{\kappa} f_{(\sigma)}(F_{(\sigma)}^{-1}(1-q)). \tag{29}$$

We now use (28) and (29) to prove the theorem. By (28), $d\pi_h^a/d\lambda(l) = d\pi_h^b/d\lambda(l) \to r\delta^2 f_{(\sigma)}(F_{(\sigma)}^{-1}(1-q))/\lambda^2(l)$ and $d\pi_{lh}^D/d\lambda(ll) \to r\delta^2 f_{(\sqrt{2}\sigma)}(F_{(\sqrt{2}\sigma)}^{-1}(1-q))/\lambda^2(ll) = (1/\sqrt{2})r\delta^2 f_{(\sigma)} \cdot (F_{(\sigma)}^{-1}(1-q))/\lambda^2(ll)$ (by (29)). Similarly, $d(\pi_{hh}^D - \pi_{lh}^D)/d\bar{\lambda} \to (1/\sqrt{2})r\delta^2 f_{(\sigma)}(F_{(\sigma)}^{-1}(1-q))/\bar{\lambda}^2$ ($\bar{\lambda}$ is used because types lh and hl are bunched).

Note that $\lambda(l) = p_l$ for both firms A and B, and $\lambda(ll) = p_l^2$ and $\bar{\lambda} = 2p_l/(1+p_l)$ for firm D. Therefore, by the chain rule, to show that $d(\pi_h^a + \pi_h^b)/dp_l < d[(\pi_{hh}^D - \pi_{hl}^D) + \pi_{hl}^D]/dp_l$, we need only show that $2\sqrt{2}/p_l^2 < (1/p_l^4) \cdot 2p_l + ((1+p_l)^2/(4p_l^2)) \cdot 2/(1+p_l)^2$. The latter can be verified to be true when $p_l < 0.85$. Q.E.D.

Proof of Theorem 4

We first show that bunching between types ll and lh will occur if δ^b is sufficiently small. When $\delta^b \to 0$, (8) converges to $1-q = F^D(Q_{ll}^D - \mu_{ll}^D)$, so its solution $Q^*(8)$ converges to $\mu_{ll}^D + (F^D)^{-1}(1-q)$. Furthermore, since

$$\begin{aligned} p_{lh}^D(1-q) &= (p_{lh}^D + p_{hl}^D + p_{hh}^D)F^D(Q_{lh}^D - \mu_{lh}^D) \\ &\quad - (p_{hl}^D + p_{hh}^D)F^D(Q_{lh}^D - \mu_{hl}^D) \\ &> (p_{lh}^D + p_{hl}^D + p_{hh}^D)F^D(Q_{lh}^D - \mu_{lh}^D) \\ &\quad - (p_{hl}^D + p_{hh}^D)F^D(Q_{lh}^D - \mu_{lh}^D) = p_{lh}^D F^D(Q_{lh}^D - \mu_{lh}^D), \end{aligned}$$

we have $Q^*(9) < \mu_{lh}^D + (F^D)^{-1}(1-q)$. Notice that $Q^*(9) - [\mu_{lh}^D + (F^D)^{-1}(1-q)]$ only depends on μ_{lh}^D and μ_{hl}^D through the gap $\mu_{hl}^D - \mu_{lh}^D = \delta^a - \delta^b$, which converges to δ^a as $\delta^b \to 0$. Thus, as $\delta^b \to 0$, $Q^*(9) - [\mu_{lh}^D + (F^D)^{-1}(1-q)]$ approaches a negative constant. Therefore, for sufficiently small δ^b, $Q^*(8) > Q^*(9)$, violating the MC. This means bunching occurs between types ll and lh as $\delta^b \to 0$.

As in (16) and (17), the information rent of firm i ($i = a$ or b) is described by $\pi_h^i = r\int_{s^i-\delta^i}^{s^i} F^i(x)dx$ and $p_l^i(1-q) =$

$F^i(s^i) - p_h^i F^i(s^i - \delta^i)$, where $s^i \doteq Q_l^i - \mu_l^i$. When bunching occurs between types ll and lh, the information rent of firm D is described by

$$\pi_{lh}^D = r\int_{\underline{s}-\delta^b}^{\underline{s}} F^D(x)\,dx, \quad \pi_{hl}^D = r\int_{\underline{s}-\delta^a}^{\underline{s}} F^D(x)\,dx,$$

$$\pi_{hh}^D = r\int_{\underline{s}-\delta^a}^{\underline{s}} F^D(x)\,dx + r\int_{s_{hl}-\delta^b}^{s_{hl}} F^D(x)\,dx,$$

$$(p_{ll}^D + p_{lh}^D)(1-q) = F^D(\underline{s}) - (p_{hl}^D + p_{hh}^D)F^D(\underline{s} - \delta^a) \tag{30}$$

$$p_{hl}^D(1-q) = (p_{hl}^D + p_{hh}^D)F^D(s_{hl}) - p_{hh}^D F^D(s_{hl} - \delta^b), \tag{31}$$

where $\underline{s} \doteq \underline{Q} - \mu_{ll}^D$ (recall that bunching occurs between types ll and lh) and $s_{hl} \doteq Q_{hl}^D - \mu_{hl}^D$. Since $p_{\theta^a\theta^b}^D = p_{\theta^a}^a p_{\theta^b}^b$, we can further simplify (30) and (31) as $p_l^a(1-q) = F^D(\underline{s}) - p_h^a F^D(\underline{s} - \delta^a)$ and $p_l^b(1-q) = F^D(s_{hl}) - p_h^b F^D(s_{hl} - \delta^b)$. The similar structure of firms A, B, and D's information rent is now evident.

To prove the theorem for firm D of type hl, we must show that $\pi_{hl}^D > \pi_h^a$ (recall that $\pi_l^b = 0$). Observe that the equations determining π_{hl}^D and π_h^a are identical except that the former involves F^D, whereas the latter involves F^a. Because F^D has higher variability, from Lemma 3 we know that $\pi_{hl}^D > \pi_h^a$. Hence, to prove the theorem for firm D of type hh, it is sufficient to show that $\pi_{hh}^D - \pi_{hl}^D > \pi_h^b$. Again, the equations determining $\pi_{hh}^D - \pi_{hl}^D$ and π_h^b are identical except that the former involves F^D while the latter involves F^b. The result follows from Lemma 3. Q.E.D.

Proof of Theorem 5

Proposition 1's proof showed that $FD_{\theta_j}(Q_{\theta_j})$ is first positive and then negative. From the expression for $FD_{\theta_j}(Q_{\theta_j})$, it then follows that Q_{θ_j} is decreasing in $\mu_{\theta_{j+1}}$. Because $\mu_{\theta_j} \le \mu_{\theta_{j+1}} < \infty$, we can obtain upper and lower bounds on Q_{θ_j} by solving (4) with $\mu_{\theta_{j+1}}$ replaced by μ_{θ_j} and ∞, respectively. Doing so establishes that $\mu_{\theta_j} + F^{(-1)}(\lambda(\theta_j)(1-q)) < Q_{\theta_j} \le \mu_{\theta_j} + F^{(-1)}(1-q)$.

The theorem assumes $(r-c)/r > 0.5$ (equivalently, $q < 0.5$), and p_l^a and p_l^b are sufficiently close to 1. The above bounds ensure that $Q_{\theta^i}^i > \mu_{\theta^i}^i + (F^i)^{(-1)}(\lambda(\theta^i)(1-q))$, $i = a, b$, and $Q_{\theta^a\theta^b}^D \le \mu_{\theta^a}^a + \mu_{\theta^b}^b + (F^D)^{(-1)}(1-q)$. Therefore, to show that $Q_{\theta^a}^a + Q_{\theta^b}^b > Q_{\theta^a\theta^b}^D$, it is sufficient to show that $(F^a)^{(-1)}(\lambda(\theta^a)(1-q)) + (F^b)^{(-1)}(\lambda(\theta^b)(1-q)) > (F^D)^{(-1)}(1-q)$. If $\lambda(\theta^a)$ and $\lambda(\theta^b)$ are sufficiently close to 1, $(F^a)^{(-1)}(\lambda(\theta^a) \cdot (1-q)) + (F^b)^{(-1)}(\lambda(\theta^b)(1-q))$ will be sufficiently close to $(F^a)^{(-1)}(1-q) + (F^b)^{(-1)}(1-q)$. Because $F^i(x) = \Phi_{(\sigma^i)}(x)$, we have

$$\begin{aligned}(F^a)^{(-1)}&(1-q) + (F^b)^{(-1)}(1-q)\\ &= \Phi^{-1}_{(\sigma^a+\sigma^b)}(1-q) > \Phi^{-1}_{(\sqrt{(\sigma^a)^2+(\sigma^b)^2})}(1-q) = (F^D)^{(-1)}(1-q).\end{aligned}$$

Therefore, the result follows if $\lambda(\theta^a)$ and $\lambda(\theta^b)$ are sufficiently close to 1.

Recall that when $\theta^i = l$, $\lambda(\theta^i) = p_l^i$, which is close to 1 by assumption. When $\theta^i = h$, by definition $\lambda(\theta^i) \equiv 1$. Therefore, the result follows. Q.E.D.

References

Benjaafar S, Cooper WL, Kim JS (2005) On the benefits of pooling in production-inventory systems. *Management Sci.* 51(4):548–565.

Cachon GP, Zhang F (2006) Procuring fast delivery: Sole sourcing with information asymmetry. *Management Sci.* 52(6):881–896.

Chen F (2007) Auctioning supply contracts. *Management Sci.* 53(10):1562–1576.

Corbett CJ, de Groote X (2000) A supplier's optimal quantity discount policy under asymmetric information. *Management Sci.* 46(3):444–450.

Elmer-DeWitt P (2010) iPhone gross profit margin nears 60%. *CNN Money* (March 2), http://tech.fortune.cnn.com/2010/03/02/what-doth-it-profit-an-iphone/.

Eppen G (1979) Effects of centralization on expected costs in a multi-location newsboy problem. *Management Sci.* 25(5):498–501.

Eppen G, Schrage L (1981) Centralized ordering policies in a multi-warehouse system with lead times and random demand. *TIMS Studies in the Management Sciences. Multi-Level Production/Inventory Control Systems, Theory and Practice*, Vol. 16 (North-Holland, Amsterdam), 51–67.

Gerchak Y, He Q (2003) On the relation between the benefits of risk pooling and the variability of demand. *IIE Trans.* 35(11):1027–1031.

Ha AY (2001) Supplier-buyer contracting: Asymmetric cost information and cutoff level policy for buyer participation. *Naval Res. Logist.* 48:41–64.

Hanany E, Gerchak Y (2008) Nash bargaining over allocations in inventory pooling contracts. *Naval Res. Logist.* 55:541–550.

Hartman BC, Dror M (2003) Optimizing centralized inventory operations in a cooperative game theory setting. *IIE Trans.* 35(3):243–257.

Laffont J-J, Martimort D (2002) *Theory of Incentives: The Principal-Agent Model* (Princeton University Press, Princeton, NJ).

Leonard C (2009) Monsanto stomps down budding seed competitors. *USA Today* (December 14), http://www.usatoday.com/money/industries/food/2009-12-14-monsanto-practices_N.htm.

Lublin JS, Ante SE (2011) Inside Sprint's bet on iPhone. *Wall Street Journal* (October 4), http://online.wsj.com/article/SB10001424052970203405504576603053795839250.html.

Nagarajan M, Sošić G, Zhang H (2009) Stable group purchasing organizations. Working paper, University of British Columbia, Vancouver.

Wan Z, Beil DR (2009) RFQ auctions with supplier qualification screening. *Oper. Res.* 57(4):934–949.

Whitney L (2009) Report details AT&T wait to break even on iPhones. *CNET* (October 9), http://news.cnet.com/8301-13579_3-10371785-37.html.

Yang H, Schrage L (2009) Conditions that cause risk pooling to increase inventory. *Eur. J. Oper. Res.* 192(3):837–851.