Single sourcing from a supplier with unknown efficiency and capacity

Christophe Bernard — *Toulouse Business School*

Sébastien Mitraille — *Toulouse Business School*

PRELIMINARY AND INCOMPLETE

Abstract

When a supplier cost to scale up production is unknown except to this firm, we show that a retailer's optimal purchase policy depends on market demand: when demand is large, the retailer over-purchases from types who face decreasing returns, under-purchases from types who face constant returns, and purchases an inflexible quantity independent of demand from a set of intermediate types. Such policy prevents small scale suppliers to under-state the degree of decreasing returns, and large scale suppliers to over-state their capacity constraint. When demand is low enough, the retailer under-purchases to all types except to the most efficient one, and an inflexible rule consisting in not purchasing from large scale producers can be used. The second best purchase policy is such that the retail price can be above or below that of an integrated monopoly.

JEL Classification: D82; D86; L23

Keywords: Adverse selection; soft capacity constraints; countervailing incentives; Spence-Mirrlees condition.

Acknowledgements: We are indebted to Bruno Jullien and David Martimort for many helpful discussions. We thank Serge Moresi for his comments, as well as the participants to the IIOC 2021, ASSET 2021 and INFORMS 2021 conferences.

1 Introduction

As pointed out by many empirical analysis, firms display an enormous amount of heterogeneity even within the same industry, be it on their production technology or on the cost of their inputs. This heterogeneity also prevails geographically, no matter whether U.S. firms or firms from developing countries are considered[1]. Productivity or cost functions estimations[2] also show that firms operate neither at the same scale nor with the same level of efficiency, and often face increasing marginal costs of production, that is decreasing returns-to-scale[3].

Although it is not their focus, all these estimations also allow to emphasize an important issue. Once a firm has chosen its short-run capacity of production and once its operations are planned, generally based on demand forecasts and given longer run investments whose costs are sunk[4], a firm's short run cost structure is fixed. When the production and sales phases start, given the orders already confirmed, each new order a firm receives exhausts its planned capacity. Once it is fully used, additional orders force this firm to increase its capacity by acquiring extra inputs (e.g. temporary labor or raw materials/components from spot markets), and/or by increasing the number of workers shifts (if the installed production tools or the labour legislation allow to), and/or by subcontracting part of the additional orders received. These solutions can be more costly at the margin than what the firm had been planned ex-ante. How low and how steep is the marginal cost curve of a firm depend therefore on the technological choice made (whose cost is sunk), as well as on the orders received. Differences in efficiencies as well as in residual capacities may coexist and appear in the changes of the marginal cost of production when the level of production changes. As the most efficient operators can attract customers more easily than the least efficient ones, efficiencies and residual capacities

[1]See for example Baily et al [4], Bartelsman and Dhrymes [5], and Roberts and Tybout [22].

[2]See amongst others Beard et al [6], Röller [23], Van Biesebroeck [25] or Kim and Knittel [16].

[3]For example Van Biesebroeck [25] points out that these differences, and in particular decreasing returns, may come from the use of a lean manufacturing system instead of a mass production system which generates more economies of scale.

[4]In capital, such as in a plant or in the machines needed. Planning production is of concern to management scientists but also to economists, at least as early as in Holt, Modigliani, Muth and Simon [12]. The rationale for sharing production plans with suppliers in the automobile industry is studied in Doyle and Snyder [10].

can evolve in opposite directions: a firm's residual capacity, embodied by how steep its marginal production cost curve is, can be larger (i.e. the marginal cost can be flatter) for the least efficient firms than for the most efficient ones at any point in time.

Whether a supplier's residual capacity is large or not, and whether a firm's marginal cost is low or not, is generally not observable to a buyer[5]. This is notably the case for perishable products or services. For example in consulting services, the composition of the labour force of a firm by level of seniority can generate all sorts of patterns for the marginal cost of production of achieving a particular mission. In the fast fashion industry, the total cost to produce a collection results from a combination of human labor and machines which is not observable to the buyer[6]. When deciding how much to order, a buyer may face an efficient but capacity constrained seller, or on the contrary an inefficient but capacity unconstrained seller, or any combination between these two extremes. These differences in marginal costs of production result in differences in total cost of production of these different types of suppliers, which may give each of them the opportunity to raise some profits. For example a supplier efficient at the margin but experiencing large decreasing returns-to-scale could pretend it is less efficient but less capacity constrained. A supplier less efficient at the margin but less capacity constrained could pretend it is more capacity constrained. Depending on the demand it faces, a buyer (be it a retailer, a downstream manufacturer, or a client) therefore faces a tension between procuring the quantity of product it needs as efficiently as possible (i.e. which limits the informational rents of its supplier), and ordering an optimal quantity which allows to serve the downstream market optimally.

In this paper, we characterize a retailer's optimal purchase strategy to a supplier whose marginal cost of production is unknown to all but itself. The supplier either faces decreasing returns-to-scale and is more efficient at the margin than other types of suppliers to produce at a small scale, or faces constant returns-to-scale and is more efficient at the margin than other types at producing at a large scale, or any combination of these two situations such that the steeper the marginal cost of production the smaller its intercept and reciprocally. A unique type represents these

[5]For example see again Van Biesebroeck [25].

[6]See Cachon and Zhang [9].

differences across suppliers: low types have low marginal costs for low output levels (and high marginal costs for high output levels), while high types have low marginal costs for high output levels (and high marginal costs for low output levels). Hence the marginal costs of production do not have the same ranking across types when the production increases.

Two fundamental assumptions satisfied by the basic textbook principal-agent model do not hold any more in our setting. First, the marginal costs of production of the different types of supplier, which rotate around a single output level when the type changes, do not allow the supplier's payoff to satisfy the Spence-Mirrlees condition[7]. Second, as total costs of production obtain from the summation of all the marginal costs a type faces to produce an output level, low types can produce more cheaply small output levels than high types, while high types can produce more cheaply large output levels. Therefore total costs of production are equal to each other for a given output, but do not rank identically across types depending on the production level considered. To say it differently, the supplier's payoff is not monotonic in its type, and hence countervailing incentives are present[8].

We characterize the distortions the retailer acting as a principal chooses on the quantity it purchases, compared to what a fully informed monopoly would purchase and resell. We demonstrate that these distortions depend on how the market demand the retailer faces compares to the determinants of a supplier's cost of production.

When demand is large, the retailer's purchase policy must prevent small scale suppliers to over-state their capacity by over-stating their type, as well as large scale suppliers to under-state their type (and hence their capacity). When paid as the opposite type of supplier, such untruthful reports would allow small scale producers to gain on the production of all the initial units (on which they are more efficient at the margin than large scale producers), even if they loose on the last units produced. Reciprocally, large scale producers would gain on the production of the last units produced, even if they loose on the initial units where small scale producers are more efficient at the margin. To prevent this, the retailer must over-purchase from small scale producers whose marginal cost is steep, under-purchase

[7]See Araujo and Moreira [1], [2] and Schottmüller [24].

[8]See Lewis and Sappington [19], Biglaiser and Mezzeti [7], Maggi and Rodriguez-Clare [21], Jullien [15], and Boone and Schottmüller [8].

from large scale producers whose marginal cost is flat, and purchase a fixed quantity which is independent of demand from a set of intermediate types. This inflexible rule is set at the output level at which total costs of production are equal to each other across the various types. Under this second best purchase strategy, only the extreme types as well as an interior one are required to produce the first best level, and the inflexible purchase policy leaves no rent to the intermediate types. The retailer's second best purchase policy is such that the retail price can be above or below that of an integrated monopoly in this case.

When demand is small, small scale producers are more attractive to the retailer, and to induce truthful type reporting the retailer must under-purchase to all types except to the one whose marginal cost is the steepest. Doing so, the retailer prevents small scale producers to benefit from over-stating their capacity of production and be paid as a large scale producer required to product a small batch. By introducing an inflexible rule which consists in not purchasing from the largest scale producers, the retailer is able to reduce the informational rents left to all the small scale producers which are required to produce a positive quantity. In this case the retail price of the product is definitely above what an integrated perfectly informed monopoly would choose, and the product is not always marketed.

How a firm shall source the input it uses has been the subject of many studies, which either belong to the literature studying the regulation of firms under asymmetric information[9], or which belong to the literature studying spit-award auctions[10]. Our paper crosses two streams of research in the principal-agent literature which, to our knowledge, have been examined separately so far. In Lewis and Sappington [19], affine total costs of production change with the agent's type and countervailing incentives follow from the tension between misreporting one's fixed and one's unit cost of production. This tension occurs more generally when the agent's outside option depends on its type, which has been analyzed comprehensively in Jullien [15]. We show that countervailing incentives may also follow from differences between the determinants of the variable cost of production of a firm, namely the intercept and the slope of the marginal cost, in a model where fixed costs are sunk and hence outside

[9]See Laffont and Tirole [18]

[10]Our primary focus is not to discuss whether single or multiple sourcing should occur - we assume single sourcing. However single sourcing is a theoretical equilibrium prediction of the split-award auctions literature (see Inderst [13]).

options can be normalized to zero for all types. Over-production, under-production, and an inflexible purchase rule can occur in equilibrium. This mirrors in our setting the seminal result obtained by Lewis and Sappington [19] and [20].

As we argued, countervailing incentives result from the fact that the marginal cost of production is not monotonic in a supplier's type. The Spence-Mirrlees condition fails to be satisfied, and our paper therefore relates to Araujo and Moreira [1] and [2], and Schottmüller [24]), which demonstrate that a global (non local) incentive compatibility constraint must be taken into account. Under the assumption that marginal costs functions rotate around each other as the supplier's type changes, we demonstrate that the global (non local) incentive constraints never bind when the purchase order is monotonic with respect to the supplier's type.

The monotonicity of the quantity purchased in the supplier's type, as well as the concavity of the virtual surplus, interact with the participation constraints which can be binding for any type due to the presence of countervailing incentives (as in Jullien [15]). We characterize sufficient conditions under which monotonic second best purchase orders are chosen at equilibrium by the retailer[11]. This is in contrast with Schottmüller [24], who analyzes the case where quadratic cost functions depend on the agent's type through the variable total cost of production and also through type-dependent fixed costs, in such a manner that the agent's payoff is monotonic in its type. In such a setting, he characterizes the effect of non local incentive constraints; in particular he shows that distortions may occur including for the type which realizes the first best.

No matter whether the market demand is large or small, asymmetric information results in additional social losses compared to a perfectly informed monopoly, even if the retailer is able to contract with its supplier and double marginalization is absent. Our results have several striking testable implications: first, to stop high types from pretending they are more capacity constrained than what they truly are, the retailer must purchase a smaller quantity than what would occur if information was symmetric. This strategy is chosen by the retailer when the market demand is large enough, case in which the downstream retail price will be above what an integrated or informed monopoly would choose. This effect worsens the natural

[11]With the further specificity that the quantity purchased at the second best is not continuous in the agent's type, and jumps upward to reach the inflexible rule.

price increase which occurs during an economic "boom". On the contrary to stop low types to pretend they are less capacity constrained than what they truly are, the retailer must purchase a larger quantity than what would occur if information was symmetric. This strategy is chosen by the retailer when the market demand is small enough, case in which the downstream retail price will be below what an integrated or informed monopoly would choose. This effect worsens the natural fall in prices which occurs in an economic downturn. Last but not least, the retailer can be better off with an inflexible purchase rule, which consists in ordering a quantity which can be produced at the same total cost by the different types. Doing so, informational rents are reduced to 0 but the quantity ordered does not depend on the market demand anymore. We show that such a policy is optimal in the absence of exogenous administrative costs, but results rather from asymmetric information.

The rest of the paper is organized as follows: section 2 presents our model, and section 3 presents some preliminary results. Then section 4 characterizes the second best purchase policy when demand is large, while section 5 when it is small.

2 The Model

A downstream retailer D sells to its customers a (non-negative) quantity of product q, which it procures from a single upstream supplier U for a payment T. The product is perishable and cannot be stored[12]. The consumers' inverse demand is denoted $P(q)$, which is linear and strictly decreasing in q,

$$P(q) = \max\{a - bq, 0\} \quad \text{with } a > 0,\ b > 0. \tag{1}$$

We let $P'(q)$ denote the first order derivative of the inverse demand[13]. The retailer's profit is therefore given by

$$\pi_D(q; T) = P(q)q - T. \tag{2}$$

[12]Therefore the product traded can be a service the downstream firm purchases from the upstream one, as e.g. a number of hours of subcontracting. Strictly speaking to match the consulting interpretation we discussed briefly in the introduction, we should assume that the service generates a strictly concave revenue $R(q)$ to firm D, with $R'' < 0$ and constant. The analysis of the non-perishable good case is left for another study.

[13]We use a $'$ to indicate the total derivative of a function with respect to a variable.

The upstream supplier U cannot directly access the market, and produces the quantity q with a technology of production whose investment costs are sunk and normalized to 0, and whose variable cost of production is continuous and convex in q. This variable cost also depends on a parameter θ, which is private information to firm U. We denote $C(q;\theta)$ the total cost of production of supplier U, given by[14]

$$C(q;\theta) = \theta q + \frac{1}{2}d(\theta)q^2 \quad \text{for all } q \geq 0, \text{ with } \theta \geq 0, \ d(\theta) \geq 0, \tag{3}$$

so that its profit writes

$$\pi_U(q;T;\theta) = T - C(q;\theta). \tag{4}$$

The parameter θ is the realization of a random variable Θ which is drawn according to a cumulative distribution function $F(\theta)$ on a bounded support $[0,\bar{c}]$, and is revealed only to U. This parameter θ determines the value of the function $d(\theta)$, which is continuous and strictly decreasing in θ. Whereas θ is unknown to D, the function $d(\cdot)$ is common knowledge to U and D. We assume

$$d(\theta) = \bar{d}\left(1 - \frac{\theta}{\bar{c}}\right), \tag{5}$$

which belongs to $[0,\bar{d}]$ as θ varies in $[0,\bar{c}]$. Therefore once θ is drawn, U learns θ and $d(\theta)$ but D does not.

We denote $C_q(q;\theta)$ the first order partial derivative with respect to q, i.e. the marginal cost of supplier U to produce q when it is of type θ

$$C_q(q;\theta) = \theta + \bar{d}\left(1 - \frac{\theta}{\bar{c}}\right)q. \tag{6}$$

The first order and the cross-partial derivatives with respect to θ (and respectively θ and q) are equal to

$$C_\theta(q;\theta) = q - \frac{\bar{d}}{2\bar{c}}q^2 \text{ and } C_{q\theta}(q;\theta) = 1 - \frac{\bar{d}}{\bar{c}}q. \tag{7}$$

Finally the second and third order derivatives of the cost function are

$$C_{qq}(q;\theta) = \bar{d}\left(1 - \frac{\theta}{\bar{c}}\right) \geq 0, \ C_{q\theta\theta}(q;\theta) = 0, \ \text{ and } C_{qq\theta}(q;\theta) = -\frac{\bar{d}}{\bar{c}} < 0. \tag{8}$$

[14]Assuming that fixed costs are sunk allows us to focus on countervailing incentives which come from increasing marginal costs of production, and not from type dependent participation constraints as in Jullien [15]. This cost function also differs from the example in Schottmüller [24], in which non sunk type-dependent fixed costs are introduced to ensure that informational rents are monotonic in the agent's type and hence that countervailing incentives are absent.

The first order and the cross-partial derivatives defined in (7) do not have a constant sign as q varies: $C_\theta(q;\theta) > 0$ if $q < \frac{2\bar{c}}{\bar{d}} \equiv 2q^0$ and strictly negative if $q > 2q^0$, while $C_{q\theta}(q;\theta) > 0$ if $q < \frac{\bar{c}}{\bar{d}} \equiv q^0$ and is strictly negative if $q > q^0$. Therefore given a payment T offered for a purchase of q, $\pi_U(T,q) = T - C(q;\theta)$ is such that $\frac{\partial^2 \pi_U(q;\theta)}{\partial q \partial \theta} = -C_{q\theta}(q;\theta) > 0$ if $q > q^0$ and negative else. Therefore when $q > q^0$ we are working under the $CS+$ assumption, while when $q < q^0$ we are working under $CS-$[15]. Moreover $\frac{\partial \pi_U(q;\theta)}{\partial \theta} = -C_\theta(q;\theta) > 0$ if $q > 2q^0$, and negative else, so that to carry on with the terminology Guesnerie and Laffont [11], rotations of marginal costs imply that there are 3 regions of concern in the production space: $q \geq 2q^0$, in which the supplier's profit verifies $CS++$, $q^0 \leq q \leq 2q^0$ in which it verifies $CS+-$, and $q \leq q^0$ in which it verifies $CS--$. The fact that the derivative of the supplier profit function with respect to θ changes sign in $2q^0$ generates countervailing incentives in our model[16].

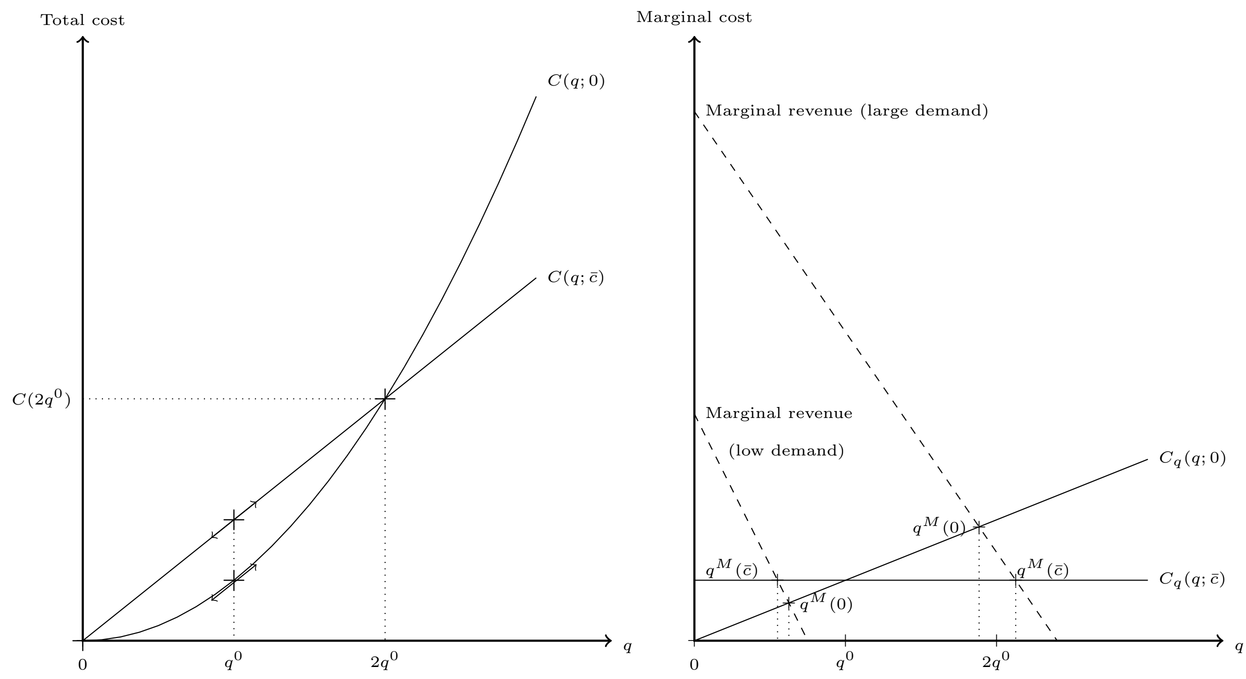
Figure 1 (left panel) illustrates the dependance of the total cost function to the supplier's type: the total costs intersect twice, at $q = 0$ where there are all nil, and at $2q^0$, where they are all equal to $C(2q^0)$. When θ tends to 0, the slope of the marginal cost tends to $\bar{d}$ while, when θ tends to $\bar{c}$ the slope of the marginal cost tends to 0. All marginal costs rotate around the same value q^0. Under these assumptions, a very efficient supplier ($\theta = 0$) is strongly capacity constrained ($d(0) = \bar{d}$) and operates therefore very efficiently at a small scale, while the least efficient supplier faces constant returns-to-scale ($\theta = \bar{c}$ and $d(\bar{c}) = 0$) and therefore operates more efficiently than other types of suppliers at a large scale[17]. Moreover, three regions of interest matter to the determination of the incentive contract offered by D as we demonstrate below: around $q = 2q^0$, where total costs are large but are all close to each other, and where marginal costs differ a lot; around $q = 0$, where total costs are small but are again all close to each other, and where again marginal costs differ

[15]See Guesnerie and Laffont [11], and Laffont and Martimort [17].

[16]See Lewis and Sappington [19], Maggi and Rodriguez-Clare [21] and Jullien [15].

[17]Johnson and Myatt [14] study how monopoly pricing reacts to changes in the shape of demand, when shifts in demand occur through rotations and in the absence of private information. Araujo and Moreira [2] study monopoly pricing when the inverse demands of different types of consumers rotate around a single value, which causes the single crossing condition to not hold. They assume that countervailing incentives are absent. We analyze the incentives a monopolistic retailer must deal with when it orders to a supplier whose scale and efficiency are inversely related, in the simplest model where these features appear.

a lot; last around $q = q^0$, where total costs are intermediate but differ a lot from each other, while marginal costs are all close to each other.



(a) Total costs for $\theta = \{0, \bar{c}\}$ (b) Monopolistic purchases for $\theta \in \{0, \bar{c}\}$

Figure 1: Supplier's costs and monopoly production as types and demand change

The cumulative distribution function and the density function of θ are continuous on $[0, \bar{c}]$ and given respectively by

$$F(\theta) \in [0,1] \text{ and } f(\theta) \geq 0 \quad \text{for } \theta \in [0, \bar{c}], \tag{9}$$

where $F(\theta)$ verifies[18]

$$\frac{\partial}{\partial \theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0 \geq \frac{\partial}{\partial \theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right) \quad \text{for } \theta \in [0, \bar{c}]. \tag{10}$$

To procure the quantity q it markets, retailer D is able to offer a menu of contracts to its upstream supplier U, from which U can choose. From the Revelation Principle, any contract can be mimicked by a direct revelation mechanism in which

[18] As explained below, the cost structure we consider is such that the objective function of the principal (here firm D) is not necessary concave in the purchase order q. Under the assumption we make on $F(\theta)$, there could exist some types strictly greater than a threshold type $\bar{\theta}$ such that the objective function is convex for $\theta > \bar{\theta}$, and concave else. This assumption on $F(\theta)$ is identical to that in Jullien [15], and satisfied if $F(\theta)$ is log-concave (see Bagnoli and Bergstrom [3]).

the agent truthfully report its type. A direct revelation mechanism consists in a pair of functions of the type $\tilde{\theta}$ reported by U, $\left(T(\tilde{\theta}), q(\tilde{\theta})\right)$, which correspond to the payment and the purchase order D sends to supplier U, and to which D commits. Supplier U announces a type $\tilde{\theta} \in [0, \bar{c}]$, and truth-telling occurs when the contract offered by the retailer satisfies the following set of incentive-compatibility constraints:

$$\pi_U(q(\theta); \theta) \geq \pi_U(q(\tilde{\theta}); \theta) \text{ for } \tilde{\theta} \neq \theta, \tag{11}$$

where the ex-post profit of supplier U of type θ reporting $\tilde{\theta}$ is given by:

$$\pi_U(q(\tilde{\theta}); \theta) = T(\tilde{\theta}) - C(q(\tilde{\theta}); \theta) \text{ for } \theta \in [0, \bar{c}]. \tag{12}$$

If it does not accept the contract offered by the downstream retailer D, the upstream supplier U earns no profit.

The retailer D earns an expected profit[19] equal to

$$\pi_D^e(\tilde{\theta}) = \mathrm{E}\left(P(q(\tilde{\theta}))q(\tilde{\theta}) - T(\tilde{\theta})\right) \tag{13}$$

where the expectation is computed on the distribution of types $F(\theta)$. The timing of the game is the following:

1. Nature draws the type θ of supplier U and informs this firm;
2. The retailer D offers a menu of binding contracts to U, $\left(T(\tilde{\theta}); q(\tilde{\theta})\right)_{\tilde{\theta} \in [0, \bar{c}]}$;
3. U reports its type;
4. U produces $q(\tilde{\theta})$ which is then sold by D, and payoffs are realized.

The profit $\pi_U(q(\theta); \theta)$ and the quantity ordered to each supplier's $q(\theta)$ must be such that reporting $\tilde{\theta} = \theta$ is optimal for each type of supplier U. Replacing $T(\theta) = \pi_U(\theta; \theta) + C(q(\theta); \theta)$ into the expression of D's expected profit, the retailer optimization problem consists in choosing the quantity ordered $q(\theta)$ and a profit level for its supplier $\pi_U(\theta; \theta)$ which maximizes its expected profit

$$\pi_D^e(\theta) = E\left(P(q(\theta))q(\theta) - \pi_U(q(\theta); \theta) - C(q(\theta); \theta)\right) \tag{14}$$

[19]Since the product is perishable and leftover inventories could be infinitely costly to dispose (economically or for reputation reasons, as the recent scandals on Amazon leftovers inventories showed), D cannot sell less than the quantity purchased. That is, D's sales are exactly equal to its purchases q. The case of a storable good is left for another paper.

subject to the supplier's individual rationality (IR) and the incentive compatibility (IC) constraints :

$$(IR)\ \pi_U(q(\theta);\theta) \geq 0 \text{ for all } \theta \in [0,\bar{c}] \tag{15}$$

$$(IC)\ \pi_U(q(\theta);\theta) \geq \pi_U(q(\tilde{\theta});\theta) \ \text{ for all } (\tilde{\theta},\theta) \in [0,\bar{c}] \times [0,\bar{c}]. \tag{16}$$

Moreover the retailer never obtains a profit below her reservation level normalized to zero:

$$q(\theta) \leq \frac{2(a-\theta)}{2b+\bar{d}(1-\frac{\theta}{\bar{c}})} \equiv q_{max}(\theta). \tag{17}$$

The value $q_{max}(\theta)$ is the maximum quantity ordered by D for each possible value of θ.

We can define the symmetric information benchmark. Let the industry profit, and its derivative with respect to q, when supplier U's total cost is common knowledge, be given by

$$\Pi(q;\theta) = P(q)q - C(q;\theta) \text{ and } \Pi_q(q;\theta) = P(q) + qP'(q) - C_q(q;\theta). \tag{18}$$

Under our assumptions, $\Pi(q;\theta)$ is continuous, differentiable and concave in q. If D cannot discriminate consumers, we have:

Definition 1 (Monopolistic purchases) *The integrated monopoly production $q^M(\theta)$ is the unique solution of $\Pi_q(q^M(\theta);\theta) = 0$, equal to*

$$q^M(\theta) = \frac{a-\theta}{2b+\frac{\bar{d}}{\bar{c}}(\bar{c}-\theta)}, \quad \textit{with derivative} \quad q^{M'}(\theta) = \frac{-2b-\bar{d}+\frac{\bar{d}}{\bar{c}}a}{\left(2b+\frac{\bar{d}}{\bar{c}}(\bar{c}-\theta)\right)^2}.$$

We refer to this threshold as being the *monopoly* one hereafter[20]. This quantity $q^M(\theta)$ is equal to half of the maximum quantity $q_{max}(\theta)$ for each θ. Figure 1 (right panel) illustrates for $\theta = 0$ and $\theta = \bar{c}$ the cost function of the supplier and the quantity an integrated monopoly would produce and sell, which varies with the size of market demand: when demand is large, monopolistic purchases increase with the intercept of the supplier's marginal cost of production θ, while when demand is low, monopolistic purchases decrease with θ.

[20]The first best level of production for the entire economy obtains when the market price is equal to the marginal cost θ, $P(q^{FB}(\theta)) - C_q(q^{FB}(\theta);\theta) = 0$, which gives here $q^{FB}(\theta) = \frac{a-\theta}{b+\frac{d}{\bar{c}}(\bar{c}-\theta)}$.

3 Preliminary results

In this section, we establish some preliminary results which allow us to characterize some properties of the contract offered by the retailer at equilibrium. The first subsection presents the optimality conditions which are derived from the supplier's maximization problem. The second subsection studies the properties of the virtual surplus which follow from our assumptions on the marginal cost of production of the supplier.

3.1 Incentive compatibility conditions for the supplier

The first two Lemmas below follow directly from the changes in the sign of $\frac{\partial^2 \pi_U}{\partial q \partial \theta}$ and $\frac{\partial \pi_U}{\partial \theta}$ when q varies, and adapt to our setting the standard results of adverse selection models with a continuum of types. The third Lemma provides a condition on the quantity scheme $q(\theta)$ offered by the retailer at equilibrium, under which the producer has no opportunity to deviate globally and hence under which only the local incentive constraints matter.

When $C_{q\theta}$ changes sign as the quantity varies, satisfying locally the incentive compatibility constraints implies that the purchase order $q(\theta)$ must be increasing or decreasing in the supplier's type θ depending on how the order compares to q^0. We have:

Lemma 1 *In any local optimum of the retailer which satisfies the supplier's (IC) constraints, $q(\theta)$ must be (weakly) decreasing with θ ($q'(\theta) \leq 0$) when $q(\theta) \leq q^0$, and (weakly) increasing with θ ($q'(\theta) \geq 0$) when $q(\theta) \geq q^0$.*

Proof. See Appendix A.1.‖

A second result characterizes the non-monotonicity of the net profit of supplier U when θ changes: incentive compatible contracts must leave a minimal rent to the type to which a contract purchasing $2q^0$ is offered (should the menu of incentive compatible contracts include such a quantity). We have:

Lemma 2 *In any local optimum of the retailer which satisfies the supplier's (IC) constraints, the derivative of $\pi_U(q(\theta);\theta)$ with respect to θ is $\pi_U'(q(\theta);\theta) = -C_\theta(q(\theta);\theta)$,* *which is strictly positive when $q(\theta) > 2q^0$, strictly negative when $q(\theta) < 2q^0$, and nil when $q(\theta) = 2q^0$ or $q(\theta) = 0$.*

Proof. See Appendix A.2.‖

Therefore to forbid the supplier to lie locally, the retailer must choose a quantity scheme $q(\theta)$ and a rent $\pi_U(q(\theta); \theta)$ which are both increasing with θ when the quantity ordered exceed $2q^0$. When the quantity ordered belongs to $[q^0, 2q^0]$, the quantity scheme must increase with θ but the rent must decrease with θ. Finally when the quantity ordered is lower than q^0, both the rent and the quantity scheme must decrease with θ.

Since the Spence-Mirrlees condition is not satisfied, the contract offered by the retailer to its supplier must not only satisfy its local incentive constraints, but it must also satisfy the non-local ones[21]: supplier U of type θ must not find profitable to announce to be a type $\hat{\theta}$ "far" from θ in $[0, \bar{c}]$. As the proof of Lemma 3 below shows, the difference of profit supplier U earns from announcing $\hat{\theta}$ instead of θ when its true type is θ rewrites as a function of the cross-partial derivative of the total cost of production $C_{q\theta}(q; \theta)$:

$$\pi_U(q(\theta); \theta) - \pi_U(q(\hat{\theta}); \theta) = -\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q}; \tilde{\theta}) d\tilde{q} d\tilde{\theta} = \int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} \left(\frac{q}{q^0} - 1 \right) d\tilde{q} d\tilde{\theta}. \tag{19}$$

Then the non-local incentive constraints states that this profit difference must be positive for any announcement $\hat{\theta}$ different from supplier U's true type θ,

$$\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} \left(\frac{q}{q^0} - 1 \right) d\tilde{q} d\tilde{\theta} \geq 0 \quad \forall (\theta, \hat{\theta}) \in [0, \bar{c}] \times [0, \bar{c}], \hat{\theta} \neq \theta. \tag{20}$$

As established when presenting our model, when marginal costs rotate around a unique value q^0, the cross-partial derivative of the total cost of production $C_{q\theta}(q; \theta)$ becomes strictly negative when q is strictly larger $q^0 = \frac{\bar{c}}{d}$, which is independent of θ. Therefore:

Lemma 3 *If the contract offered by the retailer D to its supplier U is such that the quantity scheme $q(\theta)$ verifies either (i), $\min_\theta q(\theta) > q^0$ and $q'(\theta) \geq 0$, or (ii), $\max_\theta q(\theta) < q^0$ and $q'(\theta) \leq 0$, then the supplier's non-local incentive compatibility constraint is always satisfied with strict inequality.*

[21]See Araujo and Moreira [1] and Schotmüller [24].

Proof. See Appendix A.3.‖

Lemma 3 implies that if the contract offered by the retailer is such that $\min_\theta q^*(\theta) > q^0$ or $\max_\theta q^*(\theta) < q^0$, then the supplier's non-local incentive compatibility constraint can be neglected. Consequently we can determine the solution of a relaxed maximization program of the retailer D, in which the non-local incentive compatibility constraint is absent, and then check that this relaxed solution verifies the conditions in Lemma 3.

3.2 The retailer's relaxed problem

The retailer's relaxed optimization problem writes

$$\max \pi_D^e = \int_0^{\bar{c}} \Pi(q(\theta);\theta) - \pi_U(q(\theta);\theta) dF(\theta) \tag{21}$$

with respect to $(q(\theta), \pi_U(q(\theta);\theta))$ for all $\theta \in [0, \bar{c}]$, subject to

$$\pi_U(q(\theta);\theta) \geq 0 \quad \forall \theta \in [0, \bar{c}] \tag{IR}$$

$$\pi_U'(q(\theta);\theta) = -C_\theta(q(\theta);\theta) \quad \forall \theta \in [0, \bar{c}] \tag{LIC}$$

$$q'(\theta) \leq 0 \text{ if } q \leq q^0, \text{ and } q'(\theta) \geq 0 \text{ if } q \geq q^0 \tag{MON}$$

$$q(\theta) \leq q_{max}(\theta) \quad \forall \theta \in [0, \bar{c}].$$

The solution to this problem is denoted $(q^*(\theta), \pi_U^*(\theta))$ for $\theta \in [0, \bar{c}]$. The expected virtual surplus can be determined and simplified as in Jullien [15], starting from the maximization of the expected profit of the retailer π_D^e in (21) with respect to the (IR) and (LIC) constraints. The presence of countervailing incentives implies that the (IR) constraints of a subset of types interior to $[0, \bar{c}]$ can be binding, and hence we focus for the moment on the determination of the expected virtual surplus without introducing the monotonicity constraint into its expression. As we point out below, the fact that marginal costs of production are not ranked across types interacts with the fact that (IR) constraints may bind anywhere in $[0, \bar{c}]$. It implies two difficulties: first, the virtual surplus is not concave in q; second, its cross partial derivative with respect to q and θ has not a constant sign.

Let $\mu(\theta)$ be the non negative multiplier of the (IR) constraint of a type θ, which we assume to be an integrable function of θ[22]. We denote $M(\theta)$ the integral of $\mu(\theta)$

$$M(\theta) = \int_0^\theta \mu(t)dt, \tag{22}$$

and we let $1 - F(\theta)$ be a primitive of $-f(\theta)$ and $1 - M(\theta)$ be a primitive of $-\mu(\theta)$. Integrating by parts the expected virtual surplus gives

$$\begin{aligned} V_D^e &= \int_0^{\bar{c}} \Pi(q(\theta);\theta) - \pi_U(q(\theta);\theta)dF(\theta) + \int_0^{\bar{c}} \mu(\theta)\pi_U(q(\theta);\theta)d\theta \\ &= \int_0^{\bar{c}} \Pi(q(\theta);\theta)dF(\theta) + [(1-F(\theta))\pi_U(q(\theta);\theta)]_0^{\bar{c}} - \int_0^{\bar{c}} \frac{1-F(\theta)}{f(\theta)}\pi'_U(q(\theta);\theta)dF(\theta) \\ &\quad - [(1-M(\theta))\pi_U(q(\theta);\theta)]_0^{\bar{c}} + \int_0^{\bar{c}} \frac{1-M(\theta)}{f(\theta)}\pi'_U(q(\theta);\theta)dF(\theta). \end{aligned} \tag{23}$$

The multiplier of a type θ (IR) constraint $\mu(\theta)$ interprets as the opportunity gain for the retailer to reduce $\pi_U(q(\theta);\theta)$ from an infinitesimal (positive) amount to 0, holding the quantity $q(\theta)$ unchanged. As $\mu(\theta)$ is positive or nil, $M(\theta)$ cannot decrease, and interprets as the opportunity gain D obtains by reducing uniformly the profits left to all types between 0 and θ, from an infinitesimal (positive) amount to 0, holding all quantities unchanged. Then, keeping quantities unchanged, a uniform reduction of profits across all types continuously distributed over $[0, \bar{c}]$ has a cumulated opportunity gain given by $M(\bar{c}) = 1$, and consequently $M(\theta)$ has the property of a cumulated distribution function[23].

Using the local incentive constraint (LIC) to substitute $\pi'_U(q(\theta);\theta)$ into V_D^e above, and assuming for the moment that $M(0) = 0$[24], the expected virtual surplus the retailer maximizes with respect to $q(\theta)$ simplifies into:

$$V_D^e = \int_0^{\bar{c}} \Pi(q(\theta);\theta) - \frac{F(\theta) - M(\theta)}{f(\theta)} C_\theta(q(\theta);\theta)dF(\theta) \tag{24}$$

It is the surplus when the informational rents induced by incentive compatibility are taking into account. The point-wise optimization with respect to $q(\theta)$ for each

[22]For the moment, we neglect the cases where only the (IR) constraint of $\theta = 0$ or of $\theta = \bar{c}$ bind. We analyze these two situations as "corner cases" of Propositions 2 and 3 below.

[23]Moreover $M(\theta)$ remains constant on every interval of types for which $\mu(\theta) = 0$, and hence when $\mu(\theta) > 0$, the support of types for which $M(\theta) \in (0,1)$ is included in $[0, \bar{c}]$. When only the IR constraint of $\theta = 0$ binds, then $M(0) = 1$ and 0 else, while when the IR constraint of $\theta = \bar{c}$ binds, then $M(\bar{c}) = 1$ and 0 else.

[24]I.e. we exclude for the moment the case in which $M(0) = 1$ i.e. in which only the IR constraint of the lowest type, $\theta = 0$, binds. We discuss this case separately at the end of the section.

θ gives the first and second order conditions

$$V_q(q(\theta);\theta) = \Pi_q(q(\theta);\theta) - \frac{F(\theta) - M(\theta)}{f(\theta)} C_{\theta q}(q(\theta);\theta) = 0 \tag{25}$$

and

$$V_{qq}(q(\theta);\theta) = \Pi_{qq}(q(\theta);\theta) - \frac{F(\theta) - M(\theta)}{f(\theta)} C_{\theta qq}(q(\theta);\theta) \leq 0. \tag{26}$$

We need to address a first difficulty: the expected virtual surplus is not always concave in the purchased quantity $q(\theta)$[25], even if $\Pi_{qq}(q(\theta);\theta) = 2P'(q(\theta) - C_{qq}(q(\theta);\theta) = -2b - d(\theta)$ is negative under our assumptions. Indeed, the second term of (26) above is not always negative. On the one hand our assumptions on the marginal costs curves impose that $C_{\theta qq}(q(\theta);\theta) = d'(\theta)$ is strictly negative. On the other hand, since the support of $M(\theta)$ belongs to the support of $F(\theta)$ then when $M(\theta) = 1$, $F(\theta) - 1 \leq 0$ and the second term of (26) is negative, but when $M(\theta) = 0$, the second term of (26) is positive. Consequently the second order condition (26) is not always negative, as $F(\theta) - M(\theta)$ may be positive and may exceed the second order derivative of the industry profit $\Pi_{qq}(q(\theta);\theta)$ (in absolute value).

A second difficulty that must be discussed is the possibility that the scheme $q^*(\theta)$ which solves $V_q(q(\theta);\theta) = 0$ does not necessarily satisfy the "piecewise" monotonicity constraint, i.e. does not satisfy Lemma 1. There may be a contradiction between the sign of $C_{\theta q}(q(\theta);\theta)$ which depends only on how $q(\theta)$ compares to q^0, implying the sign of $q'(\theta)$, and the solution to the retailer's optimization problem, which could require $q(\theta)$ to evolve in the opposite direction as θ varies. Indeed, the retailer is better off choosing the quantity scheme $q^*(\theta)$ whose derivative is equal to $\frac{dq^*(\theta)}{d\theta} = -\frac{V_{q\theta}(q(\theta);\theta)}{V_{qq}(q(\theta);\theta)}$.

Let us further differentiate (25) with respect to θ. Under our assumptions on the cost $C(q;\theta)$, $C_{\theta\theta q} = 0$, it comes:

$$\begin{aligned}
V_{q\theta}(q(\theta);\theta) =& \Pi_{q\theta}(q(\theta);\theta) - \frac{d}{d\theta}\left(\frac{F(\theta) - M(\theta)}{f(\theta)}\right) C_{\theta q}(q(\theta);\theta) \\
=& - C_{q\theta}(q(\theta);\theta) - \frac{d\frac{F(\theta)}{f(\theta)}}{d\theta} C_{\theta q}(q(\theta);\theta) + \frac{\mu(\theta) f(\theta) - M(\theta) f'(\theta)}{(f(\theta))^2} C_{q\theta}(q(\theta);\theta) \\
=& - C_{q\theta}(q(\theta);\theta) \left(1 + \frac{d\left(\frac{F(\theta)}{f(\theta)}\right)}{d\theta} - \frac{\mu(\theta)}{f(\theta)} + \frac{M(\theta) f'(\theta)}{(f(\theta))^2}\right)
\end{aligned} \tag{27}$$

[25] And therefore our model does not satisfy assumption 2 in Jullien [15], so that the property called *potential separation* cannot be verified here.

Consider the value of the expected virtual marginal surplus at the output level q^0 where the Spence-Mirrlees condition changes sign, $V_q(q^0;\theta) = \Pi_q(q^0;\theta)$ as $C_{\theta q}(q^0;\theta) = 0$. As $C_q(q^0;\theta) = \bar{c}$ for all types $\theta \in [0,\bar{c}]$ and $P'(q^0)q^0 + P(q^0)$ is independent of θ, $V_q(q^0;\theta)$ does not depend on θ. Moreover, it can be positive or negative depending on the size of the market demand. That is, the expected virtual marginal surplus has the same value at q^0 for all θ, which follows directly from the fact that all marginal costs of production are equal to each other at q^0. Let us denote $V_q(q^0;\theta) \equiv V_q(q^0)$. Suppose first that $V_q(q;\theta)$ is declining in q implying that $\frac{dq^*(\theta)}{d\theta}$ is of the same sign as $V_{q\theta}(q(\theta);\theta)$. When $V_q(q^0) > 0$, and as V_q is linear (and decreasing) in q, then the quantity scheme $q^*(\theta)$ solution to the retailer's relaxed optimization problem is such that $q^*(\theta) > q^0$ for all types[26]. This requires from Lemma 1 that $q'(\theta) \geq 0$. In that case $C_{q\theta}(q(\theta);\theta) < 0$ for all types and the sign of $V_{q\theta}(q(\theta);\theta)$ is given by the sign of

$$1 + \frac{d\left(\frac{F(\theta)}{f(\theta)}\right)}{d\theta} - \frac{\mu(\theta)}{f(\theta)} + \frac{M(\theta)f'(\theta)}{(f(\theta))^2}, \tag{28}$$

which depends on the values of $\mu(\theta)$ and $M(\theta)$. If this expression is negative for some types, then $V_{q\theta}(q(\theta);\theta) < 0$, and $\frac{dq^*(\theta)}{d\theta}$ is negative. This contradicts Lemma 1 and implies that ironing the quantity scheme can be required. This reasoning can be mirrored to the case where $V_q(q^0) < 0$, case in which the quantity scheme $q^*(\theta)$ solution to the retailer's relaxed optimization problem is such that $q^*(\theta) < q^0$ for all types[27]. Symmetrically, if we suppose that $V_q(q;\theta)$ is increasing in q, implying that $\frac{dq^*(\theta)}{d\theta}$ is of the opposite sign as $V_{q\theta}(q(\theta);\theta)$, we obtain that when $V_q(q^0) > 0$, $q^*(\theta) = q_{max} \geq q^0$ for each θ, and by Lemma 1 $q(\theta)$ must be increasing. As $q^*(\theta) \geq q^0$, $C_{q\theta}(q(\theta);\theta) < 0$ for all types, and the sign of $-V_{q\theta}(q(\theta);\theta)$ is given by the opposite of the sign of 28 which can be positive or negative. In this case again, the reasoning can be mirrored to the case where $V_q(q^0) < 0$.

In the two sections which follow, we analyze separately the two cases in which either $V_q(q^0) > 0$ or in which $V_q(q^0) \leq 0$. Using the expressions of the demand, $V_q(q^0) > 0$ is equivalent to $a > \left(1 + \frac{2b}{d}\right)\bar{c}$. This case, addressed in section 4, corresponds to a large demand situation. The mirror case addressed in section 5, $V_q(q^0) \leq 0$ (or $a \leq \left(1 + \frac{2b}{d}\right)\bar{c}$), corresponds to a low demand situation. In both cases, we start by determining the solution $q^*(\theta)$ first neglecting the monotonicity

[26]The optimal value $q^*(\theta)$ must be lower than $q_{max}(\theta)$.

[27]$q^*(\theta)$ must be greater than 0.

constraint and the non-local incentive compatibility constraint. Then we determine sufficient conditions under which these solutions are the actual global optima of the retailer. Then when these sufficient conditions are not met, we determine the optimal quantity scheme in which bunching occurs.

4 Optimal sourcing when demand is large

In this section, we consider the case of a large demand relative to the support of the distribution of types/cost intercept: $q^M(0) \geq q^0$, which is equivalent to $V_q(q^0) > 0$.

We demonstrate that the minimal quantity the retailer D orders to its supplier when information is asymmetric is such that $C_{q\theta}(q, \theta)$ is strictly negative. Under this result, Lemma 3 applies, the quantity scheme $q^*(\theta)$ does not trigger a non local deviation of any type. We address at the end of the section the possibility to bunch types in order for the scheme offered to satisfy the monotonicity constraint.

Let us denote $M^*(\theta)$ the cumulated multiplier of the (IR) constraints. Two caricatural quantity schemes are useful to determine $q^*(\theta)$ and $M^*(\theta)$. The first one assumes away the (IR) constraints by supposing that they do not bind for any type ($M(\theta) = 0$), while the second one assumes that $M(\theta) = 1$ for all types. The equilibrium $q^*(\theta)$ we are searching for is in between or can match one of these two solutions for some value of θ. Let us denote the quantities $\tilde{q}(\theta, 1)$ and $\tilde{q}(\theta, 0)$ which maximize V_D^e for respectively $M(\theta) = 1$ and $M(\theta) = 0$, for all $\theta \in [0, \bar{c}]$. To determine $\tilde{q}(\theta, 1)$ and $\tilde{q}(\theta, 0)$, we need to impose the following assumption:

Assumption 1 *In our model, the demand is such that* $P(0) - G(\bar{c}) \geq 0$, *where* $G(\theta) = \theta + \frac{F(\theta)}{f(\theta)}$.

This assumption ensures that all types participate: we demonstrate below that it implies that the quantity ordered by the retailer to any type is never nil. As $F(\theta)/f(\theta)$ is increasing in θ, $G(\theta)$ is increasing in θ ~~and the second part of~~ this assumption can be satisfied by a demand function $P(q)$ if (necessary condition) $f(\bar{c}) > 0 \cancel{\neq 0}$ so that $G(\bar{c})$ is bounded. To say it differently, Assumption 1 ensures that the quantity ordered by the retailer exists for every θ even if the retailer's profit is convex in q.

Under Assumption 1, we can prove:

Lemma 4 *The unique values $\tilde{q}(\theta,1)$ and $\tilde{q}(\theta,0)$ which maximize V_D^e respectively for $M=1$ and $M=0$ are such that:*

(i) $\tilde{q}(\theta,1)$ *is continuous and increasing in θ.*

(ii) $\tilde{q}(\theta,0)=\min\{\tilde{\tilde{q}}(\theta,0),q_{max}(\theta)\}$ *is continuous and increasing in θ, and kinked at $\bar{\bar{\theta}}$ such that $\tilde{\tilde{q}}(\bar{\bar{\theta}},0)=q_{max}(\bar{\bar{\theta}})$.*

(iii) $\tilde{q}(\theta,0)\geq q^M(\theta)\geq\tilde{q}(\theta,1)\ \forall\ \theta\in[0,\bar{c}]$, *with $\tilde{q}(0,0)=q^M(0)$ and $\tilde{q}(\bar{c},1)=q^M(\bar{c})$.*

Proof. See Appendix A.4‖.

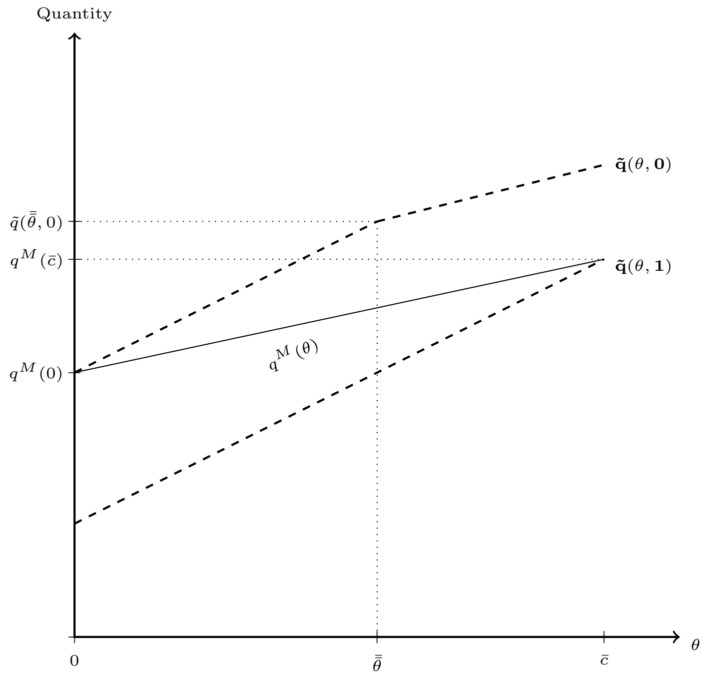


Figure 2: Monopolistic purchases (solid line) and virtual surplus optima (dashed lines) when $M=1$ or $M=0$ for all $\theta\in[0,\bar{c}]$

Lemma 4 is illustrated in Figure 2. The upper and lower bounds on the equilibrium quantity scheme $q^*(\theta)$ the retailer offers to its supplier depend on the types whose (IR) constraints bind: $q^*(\theta)$ belongs to $[\tilde{q}(\theta,1),\tilde{q}(\theta,0)]$, and it will equate $\tilde{q}(\theta,0)$ or $\tilde{q}(\theta,1)$ for some types.

Note that $2q^0$, the quantity around which the derivative of the agent profit with respect to θ changes sign, can be part of a contract which leaves no rent to the supplier. Indeed, the contract $(q(\theta),\pi_U(q(\theta);\theta))=(2q^0,0)$ is implementable and satisfies the individual rationality constraint of all types $\theta \in [0, \bar{c}]$, i.e. is feasible[28]. This contract is such that the IR constraint of any type offered this contract is binding. Moreover as the industry profit $\Pi(2q^0; \theta) \geq 0$ for all θ, the retailer can purchase the product from any type of supplier by offering such a contract. From Definition 1, it exists a unique value θ^0 such that $2q^0 P'(2q^0) + P(2q^0) = C_q(2q^0; \theta^0)$. Depending on how large the market demand is, the type θ^0 such that the corresponding industry marginal profit is nil when the production is exactly equal to $2q^0$, belongs to $[0, \bar{c}]$ or not. As the industry marginal profit is strictly decreasing in q, $\theta^0 \in [0, \bar{c}]$ if $2q^0 P'(2q^0) + P(2q^0) - C_q(2q^0; 0) \leq 0$ and $2q^0 P'(2q^0) + P(2q^0) - C_q(2q^0; \bar{c}) \geq 0$, that is if the marginal revenue at $2q^0$ is in between the marginal cost of the lowest type $\theta = 0$ computed at $2q^0$, and the marginal cost of the highest type $\theta = \bar{c}$. We start to consider this case and address the two other cases ($\theta^0 < 0$ or $\theta^0 > \bar{c}$) at the end of this section.

In what follows we maintain first our approach which consists in assuming that $V_{\theta q} \geq 0$, ensuring that ironing the quantity scheme is not required, as a companion assumption to $V_q(q^0) > 0$. Proposition 4 below establishes a sufficient condition for this assumption to hold for every θ.

The comparison between θ^0 and the bounds of the interval of types $[0, \bar{c}]$ is equivalent to compare $2q^0$ to the maximal and minimal quantities ordered to the supplier. As Lemma 4 demonstrates, the minimal quantity ordered is equal to $\tilde{q}(0, 1)$, and the largest order is $q_{max}(\bar{c}, 0)$ when $\bar{\bar{\theta}} \leq \bar{c}$ and $\tilde{q}(\bar{c}, 0)$ otherwise.

We consider first the case where the demand is such that $\tilde{q}(0, 1) \leq 2q^0 \leq \tilde{q}(\bar{c}, 1)$. In this configuration, the (IR) constraints of a subset of types interior to $[0, \bar{c}]$ bind.

Proposition 1 *When $C_q(2q^0; \bar{c}) \leq P(2q^0) + 2q^0 P'(2q^0) \leq C_q(2q^0; 0)$, i.e. $q^0 \leq q^M(0) \leq 2q^0 \leq 2q^M(0)$ and $2q^0 \leq \tilde{q}(\bar{c}, 1)$, the quantity ordered $q^*(\theta)$ is such that the IR of all types $\theta \in [\min\{\theta_1, \bar{\theta}\}, \theta_2]$ bind, where θ_1 (respectively θ_2) is the unique type*

[28] Here $q(\theta) = 2q^0$ is constant and the supplier's reservation utility is nil for all θ. Moreover the rate of growth of this reservation utility is nil, and the derivative of the supplier's profit with respect to θ, equal to $-C_\theta(q; \theta)$, evaluated at $q = 2q^0$, is nil too. Hence Jullien [15]'s *homogeneity* property is (weakly) verified in our model.

which solves $\tilde{q}(\theta_1, 0) = 2q^0$ (resp. $\tilde{q}(\theta_2, 1) = 2q^0$). We have:

$$(q^*(\theta), M^*(\theta)) = \begin{cases} (\min\{\tilde{q}(\theta,0), q_{max}(\theta)\}, 0) & \text{if } \theta < \theta_1 \\ (2q^0, F(\theta) - f(\theta)\frac{\Pi_q(2q^0;\theta)}{C_{\theta q}(2q^0;\theta)}) & \text{if } \theta \in [\theta_1, \theta_2] \\ (\tilde{q}(\theta,1), 1) & \text{if } \theta > \theta_2 \end{cases}$$

Proof. See Appendix A.5.||

From the previous analysis, the following corollary comes immediately.

Corollary 1 (to Proposition 1) *The order $q^*(\theta)$ is larger (respectively lower) than $q^M(\theta)$ for θ lower (resp. larger) than θ^0, where θ^0 solves $2q^0P'(2q^0) + P(2q^0) = C_q(2q^0;\theta^0)$.*

Figure **??** illustrates the equilibrium described in proposition 1 above, by representing the scheme $q(\theta)$ with a thick black line and the monopoly purchases $q^M(\theta)$ by a thin black line going from $q^M(0)$ to $q^M(\bar{c})$. We can provide the intuition for this result. When the market demand is such that the marginal revenue intersects the supplier marginal cost at $q < 2q^0$ for low types, and at $q > 2q^0$ for high types, the retailer must forbid low types to over-state their type and high types to under-state their type. Over-stating one's type allows a low type to gain on the production of $[0, q^0]$ where it is paid as a higher type while it produces at a lower marginal cost, even if it looses on the production of $q > q^0$ where its marginal cost of production is higher than the payment received by a higher type. The reverse logic explains why a high type can be tempted to under-state its type: a high type gains on $q > q^0$ where its marginal cost is lower than the payment it receives, and looses on $[0, q^0]$ where it is higher. Then the retailer forbids low types to over-state their type by enlarging its purchases compared what would happen with symmetric information, at every type level i.e. at every level of the marginal cost θ which corresponds to a small scale producer. On the contrary, it forbids the high types to understate their types by reducing its purchases at every level of the marginal cost of production corresponding to a large scale producer.

The distortions in the purchases the retailer chooses reduce the rents it leaves to the different types of supplier. To satisfy the participation constraint of the intermediate types who earn less and less as the distortion increases compared to the monopoly, the retailer may have to offer these types an inflexible rule[29]. This rule consists in purchasing a fixed quantity $2q^0$ against the payment of $C(2q^0)$ which is identical for all types. In this situation, asymmetric information forces the uninformed party to renounce to procure a quantity which varies with the type of the supplier or with the market demand.

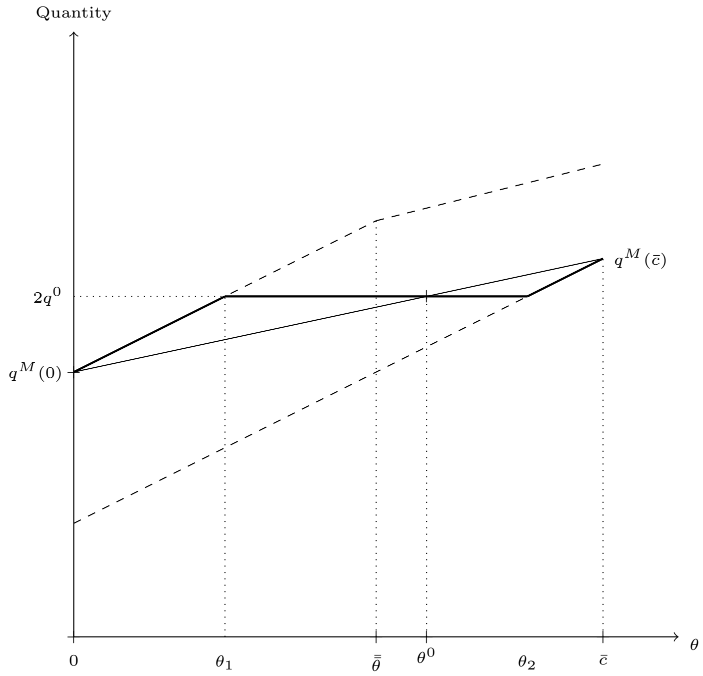


Figure 3: Equilibrium when $C_q(2q^0;\bar{c}) \leq P(2q^0) + 2q^0 P'(2q^0) \leq C_q(2q^0;0)$

The next two propositions characterize the retailer's procurement strategies in the cases where the market demand is such that only one type produces the first best, while all other types receive distorted purchase orders from the retailer.

The type who realizes the first best can be either the highest $\bar{c}$ when demand is very large, or on the contrary the lowest 0 when demand is not so large (but still satisfies the assumption $V_q(q^0) \geq 0$. In both cases the distortions on the quantities procured can force the retailer to use an inflexible purchase rule. First, consider the case in which demand is large enough, so that the retailer is better off employing optimally a supplier with a large capacity of production. We have,

Proposition 2 *When* $P(2q^0)+2q^0P'(2q^0) > C_q(2q^0;0)$ *and* $2q^0 \geq \tilde{q}(0,1)$, $(q^*(\theta), M^*(\theta))$ *is such that the IR constraint of all types* $\theta \in [0, \max\{\theta_2, 0\}]$ *bind, where* θ_2 *solves*

[29] This terminology first appeared in Lewis and Sappington [20].

$\tilde{q}(\theta_2, 1) = 2q^0$. *It is given by:*

$$(q^*(\theta), M^*(\theta)) = \begin{cases} (2q^0, F(\theta) + f(\theta)\frac{\Pi_q(2q^0;\theta)}{C_{\theta q}(2q^0;\theta)}) & \text{if } \theta \in [0, \max\{\theta_2, 0\}) \\ (\tilde{q}(\theta, 1), 1) & \text{if } \theta \in [\theta_2, \bar{c}]. \end{cases}$$

When $P(2q^0) + 2q^0 P'(2q^0) > C_q(2q^0; 0)$ *and* $2q^0 < \tilde{q}(0, 1)$, $(q^*(\theta), M^*(\theta))$ *is such that only the IR constraint of the type* 0 *binds, and it is given by* $((\tilde{q}(\theta, 1), 1)$, *for any* $\theta \in [0, \bar{c}]$.

In the case of a large demand, the most attractive supplier for retailer D is the one with the largest capacity, $\theta = \bar{c}$, and when the demand is very large, $\tilde{q}(0, 1) > 2q^0$ the IR constraint of a type $\theta = 0$ is the only one binding: $M^*(0) = 1$[30].

Corollary 2 (to Proposition 2) *The order* $q^*(\theta)$ *is strictly lower than (respectively equal to)* $q^M(\theta)$ *when* θ *is strictly lower than (respectively equal to)* $\bar{c}$.

Proof. See Appendix A.6.‖

Oppositely, consider the case demand is small enough so that the retailer is better off employing optimally the supplier whose marginal cost of production has the lowest intercept. We have,

Proposition 3 *When* $P(2q^0) + 2q^0 P'(2q^0) < C_q(2q^0; \bar{c})$, $(q^*(\theta), M^*(\theta))$ *is such that only the IR constraint of the type* $\theta = \bar{c}$ *binds It is given by:*

$$(q^*(\theta), M^*(\theta)) = \begin{cases} (\tilde{q}(\theta, 0), 0) & \text{if } \theta \in [0, \theta_1] \\ (2q^0, F(\theta) + f(\theta)\frac{\Pi_q(2q^0;\theta)}{C_{\theta q}(2q^0;\theta)}) & \text{if } \theta \in [\theta_1, \bar{c}]. \end{cases}$$

In this case of a large but small enough demand $2q^0 > \tilde{q}(\bar{c}, 1)$, the most attractive supplier for retailer D is the most efficient one but the most capacity constrained, $\theta = 0$, and the IR constraint of a type $\theta = \bar{c}$ is the only one binding.

Corollary 3 (to Proposition 3) *The order* $q^*(\theta)$ *is strictly larger than (respectively equal to)* $q^M(\theta)$ *when* θ *is strictly larger than (respectively equal to)* 0.

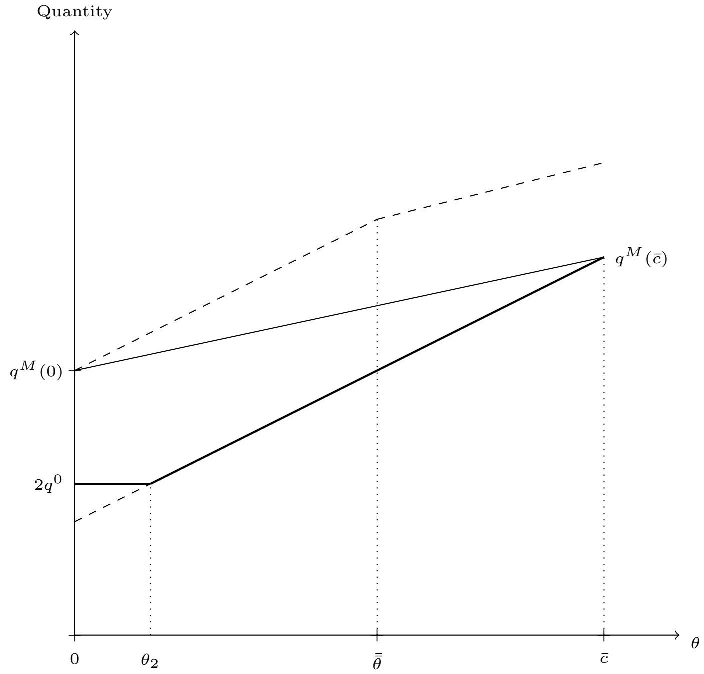


Figure 4: $P(2q^0) + 2q^0 P'(2q^0) > C_q(2q^0; 0)$

Figure 5: Equilibria when the market size points to one extreme type

Proof. See Appendix A.7.‖

We must now examine if the monotonicity condition of the solutions $q^*(\theta)$ is satisfied for each of the three cases considered above. Whether the expected virtual surplus is concave or convex, $\Pi_q(q^0; \theta) > 0$ implies that the solution $q^*(\theta)$ to $V_q(q(\theta); \theta) = 0$ is strictly larger than q^0 for all θ, and hence must be increasing in θ to respect the envelope conditions of Lemma 1 and 2.

From $\frac{dq^*(\theta)}{d\theta} = -\frac{V_{q\theta}(q(\theta);\theta)}{V_{qq}(q(\theta);\theta)}$ and (27) above we know that for the monotonicity to hold it must be the case that

$$1 + \frac{d\left(\frac{F(\theta)}{f(\theta)}\right)}{d\theta} - \frac{\mu(\theta)}{f(\theta)} + \frac{M(\theta)f'(\theta)}{(f(\theta))^2} \begin{cases} \geq 0 & \text{when V is concave} \\ \leq 0 & \text{when V is convex.} \end{cases} \tag{29}$$

When $V(q(\theta), \theta)$ is concave, the condition (29) is granted for all types for which $\mu(\theta) = 0$ and $M(\theta) = 0$, i.e. the types $\theta \leq \theta_1$ such that $q^*(\theta) = q(\tilde{\theta}, 0)$. Indeed

[30] This corresponds to the standard case in the literature, where $M(\theta)$ is a Dirac at $\theta = 0$ inducing no distortion at the top and only the lowest type participation constraint which is binding.

we assumed $\frac{d\left(\frac{F(\theta)}{f(\theta)}\right)}{d\theta} \geq 0$ and therefore all the terms on the left hand side of the inequality are positive or nil. However, for types such that (i) $\mu(\theta) > 0$ for which $M(\theta) > 0$, or such that (ii) $M(\theta) > 0$ and $\mu(\theta) = 0$, ~~it may not be the case~~, the monotonicity constraint may not be satisfied: it depends on the properties and the shapes of $F(\theta)$ and $f(\theta)$.

Consider case (i), i.e. $\theta \in [\theta_1, \theta_2]$: the informational rent of all the types are binding, thus for any θ, $\mu(\theta) > 0$ and the same quantity $q^*(\theta) = 2q^0$ is offered: the solution is constant, i.e. does not depend on $F(\theta)$ and $f(\theta)$, and therefore monotonic.

Consider case (ii): In this case, (29) could be negative, and the retailer could be interested in reducing $q(\theta)$ as θ increases, which would contradict the monotonicity requirement. Notice that if the distribution of types is uniform or increasing, $f'(\theta) \geq 0$, the monotonicity constraint is satisfied by the quantity scheme.

Therefore, when $V(q(\theta), \theta)$ is concave, there is no need to "iron" the quantity as long as $q^*(\theta)$ is smaller or equal to $2q^0$ in Propositions 1, 2 and 3. Moreover, when $P(2q^0) + 2q^0 P'(2q^0) < C_q(2q^0; \bar{c})$, the quantity asked to all types $\theta \in [\theta_1, \bar{c}]$ is constant and equal to $2q^0$, i.e. such that $\mu(\theta) > 0$, while the quantity asked to $\theta \in [0, \theta_1]$ is strictly lower than $2q^0$ and therefore such that $\mu(\theta) = M(\theta) = 0$. So $q^*(\theta)$ defined in Proposition 3 satisfies the monotonicity constraint of Lemma 1 for any $\theta \in [0, \bar{c}]$.

The case $V(q(\theta), \theta)$ convex only appears for values of θ strictly greater than $\bar{\theta}$. Thus, we must verify if the monotonicity constraint is satisfied by $q^*(\theta)$ only for the values of θ such that $\theta \geq \bar{\theta}$. As $\theta_1 < \bar{\theta}$, the optimal quantity $q^*(\theta)$ is greater or equal to $2q^0$. Two cases are possible: either $\theta_1 < \theta_2 < \bar{\theta}$, or $\theta_1 < \bar{\theta} < \theta_2$. The monotonicity constraint may not be verified for $\theta \geq max\{\theta_2; \bar{\theta}\}$, i.e. types such that $M(\theta) > 0$ and $\mu(\theta) = 0$. In this case, it may be possible to have (29) positive depending on the sign of $f'(\theta)$ implying that the monotonicity condition is not satisfied. A decreasing distribution of types, $f'(\theta) \leq 0$, is necessary but not sufficient for the monotonicity constraint to be satisfied. Therefore,

Proposition 4 *(i) The contract defined in Proposition 3 satisfies the monotonicity constraint of Lemma 1 no matter the distribution of types, and whatever the convexity of $V(q(\theta), \theta)$;*

(ii) The contracts defined in Propositions 1 and 2 satisfy the monotonicity con-

straint of Lemma 1 for any $\theta \in [0, \bar{c}]$ if the distribution of types is such that $f'(\theta) \geq 0$ for any $\theta \in [0, \bar{c}]$ (resp. ≤ 0, for any $\theta \in [\theta_2, \bar{c}]$), when $V(q(\theta), \theta)$ is concave (resp. convex).

A uniform distribution $f(\theta) = \frac{1}{\bar{c}}$ on $[0, \bar{c}]$ satisfies Proposition 4 (ii) when $V(q(\theta), \theta)$ is concave. More generally, when the distribution $F(\theta)$ has a density function $f'(\theta) < 0$ such that the monotonicity constraint is not satisfied by the optimal scheme proposed to types $\theta > \theta_2$, then the retailer must "iron" the contract it offers. We leave the investigation of this possibility for another paper.

5 Optimal sourcing when demand is low

In this section we consider the case where the demand is low relative to the support of the distribution of types/cost intercept:

$$V_q(q^0) \leq 0 \Leftrightarrow a \leq \left(1 + \frac{2b}{\bar{d}}\right) \bar{c}. \tag{30}$$

To proceed, we distinguish in particular two situations: first the caricatural situation in which $V_q(q^0) = 0$ for all θ, and second the situation in which $V_q(q^0) < 0$, case in which the marginal profit at $q = q^0$ is strictly negative for all types.

When $V_q(q^0) = 0$ for all θ, the indifference curves of all types of suppliers are tangent at q^0: indeed when $q = q^0$, $C_q(q^0, \theta) = C_q(q^0, \theta')$ for all $\theta \neq \theta'$. Then the retailer offers a single pooling contract such that all types participate,

$$(q^*(\theta), T^*(\theta)) = (q^0, C(q^0, \bar{c})) \text{ for all } \theta \in [0, \bar{c}]. \tag{31}$$

When such a contract is offered to the supplier, there are no distortions of the quantity ordered compared to the first best no matter its type, and the lower θ the higher the rent the supplier earns[31].

When $V_q(q^0) < 0$, the first best orders are all smaller than q^0, and the supplier of type $\theta = 0$ is the most efficient producer at every level of production $q < q^0$. The derivative of the total cost of production with respect to θ, $C_\theta(q, \theta)$, is strictly positive for all $q \in (0, q^0]$, and is nil for $q = 0$. The cross-partial derivative $C_{q\theta}(q, \theta)$ is strictly positive as long as $q < q^0$. From Lemma 1 and 2, the supplier's rent decreases with

[31] This case is a particular case of Araujo and Moreira [1]: the U-shaped condition (Theorem 1 p. 1120) holds at q^0 for all types, and hence the pooling contract is implementable.

θ, and the quantity ordered $q(\theta)$ is also strictly decreasing (i.e. monotonic and hence implementable). Consequently the scale at which high types ($\theta = \bar{c}$) are working is smaller than the scale at which low types ($\theta = 0$) are working, and to limit the rent left to low types, the retailer D must distort downwards the order it sends to higher types. Depending on how small the market demand is, the possibility to limit the participation of the higher types must be examined: indeed, the exclusion of these types limit the informational rents left to the lower ones. Therefore this case can be again analyzed as in the previous section and in Jullien [15], but the downward distortion which results from $C_{q\theta}(q;\theta) > 0$ may jeopardize the full participation of all types[32].

The first step of this analysis consists in adapting Lemma 4. We have:

Lemma 5 *When $V_q(q^0) < 0$, the unique values $\hat{q}(\theta, 1)$ and $\hat{q}(\theta, 0)$ which maximize V_D^e respectively for $M = 1$ and $M = 0$ are such that:*

(i) $\hat{q}(\theta, 1)$ is continuous and decreasing in θ.

(ii) $\hat{q}(\theta, 0)$ is continuous and strictly decreasing in θ, for $\theta \in [0, \underline{\theta}]$, is nil for $\theta \geq \underline{\theta}$, where $\underline{\theta}$ solves $\frac{F(\theta)}{f(\underline{\theta})} = \Pi_q(0; \underline{\theta})$. The threshold value $\underline{\theta}$ could be lower or greater than $\bar{c}$.

(iii) $\hat{q}(\theta, 0) \leq q^M(\theta) \leq \hat{q}(\theta, 1)\ \forall\, \theta \in [0, \bar{c}]$, with $\hat{q}(0, 0) = q^M(0)$ and $\hat{q}(\bar{c}, 1) = q^M(\bar{c})$.

Proof. See Appendix A.8$\|$.

Using (25), $\hat{q}^0(\theta, 0)$ solves

$$\Pi_q(\hat{q}(\theta, 0); \theta) - \frac{F(\theta)}{f(\theta)} C_{q\theta}(\hat{q}(\theta, 0); \theta) = 0. \tag{32}$$

When the quantity ordered to supplier U is nil, the payment it receives is obviously nil too, and the supplier's (IR) constraint is binding. Moreover when $\theta > \bar{\theta}$, the multiplier $M^*(\theta)$ depends also on the value of the multiplier of the constraint $q'(\theta) = 0$ which is strictly positive[33]. We have:

[32] Assuming that $C_{q\theta}(q;\theta) > 0$ implies that the cross-partial derivative of the supplier's profit with respect to the order and the type is negative, which contradicts Jullien [15]'s assumption 1.

[33] We do not determine it as it is not central to our analysis, and we focus only on $q^*(\theta)$ in the first part of the proposition which follows. The second part corresponds to the standard case with a cross partial derivative of the profit of the supplier with respect to type and quantity strictly negative (a "CS^-" configuration in the terminology of Guesnerie and Laffont [11]).

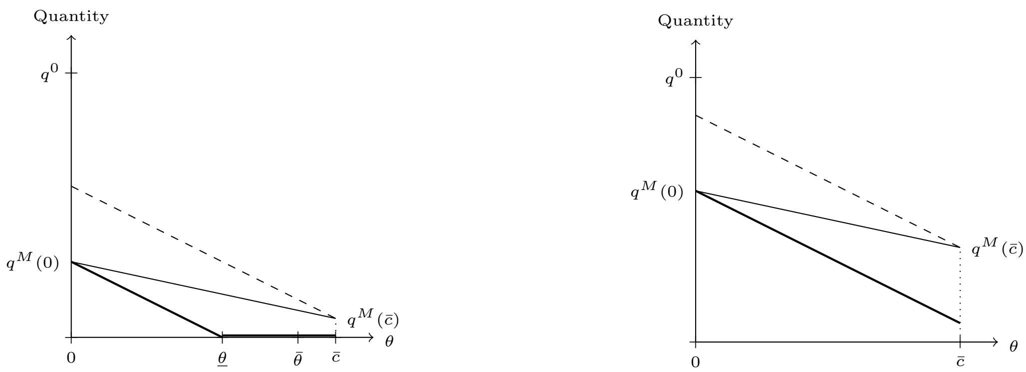
Proposition 5 *When $V_q(q^0) < 0$ and $\underline{\theta} < \bar{c}$, $q^*(\theta)$ is such that the (IR) constraints of all types $\theta \in [\underline{\theta}, \bar{c}]$ bind. It is given by*

$$q^*(\theta) = \begin{cases} \hat{q}(\theta, 0) & \text{if } \theta \in [0, \underline{\theta}] \\ 0 & \text{if } \theta \in (\underline{\theta}, \bar{c}] \end{cases}$$

Else when $\Pi_q(q^0, 0) < 0$ and $\underline{\theta} \geq \bar{c}$ only the (IR) constraint of $\theta = \bar{c}$ binds, $M(\bar{c}) = 1$ and 0 for all $\theta \neq \bar{c}$, and all types produce according to $\hat{q}(\theta, 0)$.

Proof. See Appendix A.9‖.

Therefore when $C_{\theta q}(q;\theta) > 0$, the order is distorted downwards for all types $\theta > 0$, and the retailer orders less than what a monopoly fully informed would do. Doing so, it avoids the low types who are very efficient on a small scale to lie and pretend they are larger scale producers. We can illustrate graphically this result. The optimal order implemented by the retailer corresponds to the black thick line, while the fully informed monopoly corresponds to the black thin line. The left panel illustrates the case of a very low demand, corresponding to $\underline{\theta} < \bar{c}$. In that case the order sent to an external supplier can be nil, and to preserve the possibility to sort types, the retailer must further impose $q^*(\theta) = 0$ for $\theta > \bar{\theta}$, i.e. "iron" the quantity ordered in the plan (θ, q). In the right panel, the demand is larger and all types produce a strictly positive quantity. The only (IR) constraint which binds is the one of the type $\bar{c}$.



(a) $\underline{\theta} < \bar{c}$ (b) $\underline{\theta} \geq \bar{c}$

Figure 6: Equilibria when the market demand is low, $\Pi_q(q^0, 0) < 0$

When $V_q(q^0) < 0$, the solution $q(\theta)$ to $V_q(q(\theta);\theta) = 0$ is strictly lower than q^0 for all θ, and hence $q(\theta)$ must be decreasing in θ to respect the envelope conditions of Lemma 1 and 2. In this situation, $q(\theta) < q^0$ for all types and $-C_{q\theta}(q(\theta);\theta) < 0$ for all types. Again it must also be the case that $V_{\theta q}(q(\theta);\theta) < 0$ which requires

$$1 + \frac{d\left(\frac{F(\theta)}{f(\theta)}\right)}{d\theta} - \frac{\mu(\theta)}{f(\theta)} + \frac{M(\theta)f'(\theta)}{(f(\theta))^2} > 0 \tag{33}$$

which is again granted for all types for which $\mu(\theta) = 0$ and $M(\theta) = 0$. However for types such that $\mu(\theta) > 0$ and $M(\theta) > 0$ (sole situation which occurs as the local analysis showed) it may not be the case: it depends again on the properties and the shapes of $F(\theta)$ and $f(\theta)$. However all the types for which $\mu(\theta) > 0$ are offered the same quantity $q(\theta) = 0$, and hence the solution is monotonic. So there is no need to iron the quantity $q(\theta)$ and the contract described in Proposition 5 is the global optimum of the retailer.

6 Discussion and concluding remarks

We have demonstrated that when demand is large, a retailer chooses to buy from its supplier a quantity lower than what an informed monopoly would order. This purchase policy causes the price to exceed that of an informed monopoly. In this case, asymmetric information results in additional social losses compared to a perfectly informed monopoly, in a context where purchases are done through non linear contracts where double marginalization is absent. On the contrary when demand is small, the retailer chooses to buy from its supplier a quantity larger than what an informed monopoly would order, which causes the price to be smaller than that of an informed monopoly. Then the social losses decrease compared to the monopoly. In this case, the presence of asymmetric information and of capacity constraints constrains the exercise of downstream monopoly power (when the product is a service or a consumption good is not storable). To establish these findings, we demonstrated that the contract offered by the retailer at equilibrium requires the use of an inflexible purchase rule for which a constant payment is made. Moreover this contract is potentially not continuous with the quantity purchased. The optimal contract is therefore more sophisticated than an affine contract, which is the classical solution to the double marginalization issue in vertical relationships models. Part of the social losses which follow from the presence of a chain of monopolies are partly resolved.

Our results have several striking testable implications to situations in which the market demand varies: first, to stop high types from pretending they are more capacity constrained than what they truly are, the retailer must purchase a smaller quantity than what would occur if information was symmetric. This strategy is chosen by the retailer when the market demand is large enough, case in which the downstream retail price will be above what an integrated or informed monopoly would choose. This effect worsens the natural price increase which occurs during an economic “boom”. On the contrary to stop low types to pretend they are less capacity constrained than what they truly are, the retailer must purchase a larger quantity than what would occur if information was symmetric. This strategy is chosen by the retailer when the market demand is small enough, case in which the downstream retail price will be below what an integrated or informed monopoly would choose. This effect worsens the natural fall in prices which occurs in an economic downturn. Last but not least, the retailer can be better off with an inflexible purchase rule, which consists in ordering a quantity which can be produced at the same total cost by the different types. Doing so, informational rents are reduced to 0 but the quantity ordered does not depend on the market demand anymore. We show that such a policy is optimal in the absence of exogenous administrative costs, but results rather from asymmetric information. Last, these findings have been characterized under a non disposability assumption. Studying the purchasing behaviour of a retailer when the product is storable is the object of another paper.

References

[1] Aloisio Araujo and Humberto Moreira. Adverse selection problems without the spence–mirrlees condition. *Journal of Economic Theory*, 145(3):1113–1141, 2010.

[2] Aloisio Araujo, Humberto Moreira, and S Vieira. The marginal tariff approach without single-crossing. *Journal of Mathematical Economics*, 61:166–184, 2015.

[3] Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. *Economic theory*, 26(2):445–469, 2005.

[4] Martin Neil Baily, Charles Hulten, David Campbell, Timothy Bresnahan, and Richard E Caves. Productivity dynamics in manufacturing plants. *Brookings papers on economic activity. Microeconomics*, 1992:187–267, 1992.

[5] Eric J Bartelsman and Phoebus J Dhrymes. Productivity dynamics: Us manufacturing plants, 1972–1986. *Journal of productivity analysis*, 9(1):5–34, 1998.

[6] T Randolph Beard, Steven B Caudill, and Daniel M Gropper. Finite mixture estimation of multiproduct cost functions. *The review of economics and statistics*, pages 654–664, 1991.

[7] Gary Biglaiser and Claudio Mezzetti. Principals competing for an agent in the presence of adverse selection and moral hazard. *Journal of Economic Theory*, 61(2):302–330, 1993.

[8] Jan Boone and Christoph Schottmüller. Procurement with specialized firms. *The RAND Journal of Economics*, 47(3):661–687, 2016.

[9] Gérard P Cachon and Fuqiang Zhang. Procuring fast delivery: Sole sourcing with information asymmetry. *Management Science*, 52(6):881–896, 2006.

[10] Maura P Doyle and Christopher M Snyder. Information sharing and competition in the motor vehicle industry. *Journal of Political Economy*, 107(6):1326–1364, 1999.

[11] Roger Guesnerie and Jean-Jacques Laffont. A complete solution to a class of principal-agent probems with an application to the control of a self-managed firm. *Journal of public Economics*, 25(3):329–369, 1984.

[12] Charles C Holt, Franco Modigliani, John F Muth, and Herbert A Simon. Production planning, inventories, and workforce. *Englewood Cliffs, N]: Prentice-Hall*, 1960.

[13] Roman Inderst. Single sourcing versus multiple sourcing. *The Rand journal of economics*, 39(1):199–213, 2008.

[14] Justin P Johnson and David P Myatt. On the simple economics of advertising, marketing, and product design. *American Economic Review*, 96(3):756–784, 2006.

[15] Bruno Jullien. Participation constraints in adverse selection models. *Journal of Economic Theory*, 93(1):1–47, 2000.

[16] Dae-Wook Kim and Christopher R Knittel. Biases in static oligopoly models? evidence from the california electricity market. *The Journal of Industrial Economics*, 54(4):451–470, 2006.

[17] Jean-Jacques Laffont and David Martimort. *The theory of incentives: the principal-agent model*. Princeton University Press, 2009.

[18] Jean-Jacques Laffont and Jean Tirole. *A theory of incentives in procurement and regulation*. MIT press, 1993.

[19] Tracy R Lewis and David EM Sappington. Countervailing incentives in agency problems. *Journal of economic theory*, 49(2):294–313, 1989.

[20] Tracy R Lewis and David EM Sappington. Inflexible rules in incentive problems. *The American Economic Review*, pages 69–84, 1989.

[21] Giovanni Maggi and Andres Rodriguez-Clare. On countervailing incentives. *Journal of Economic Theory*, 66(1):238–263, 1995.

[22] Mark J Roberts and James R Tybout. Industrial evolution in developing countries: a preview. *Industrial Evolution in Developing Countries*, pages 1–14, 1996.

[23] Lars-Hendrik Röller. Proper quadratic cost functions with an application to the bell system. *The review of economics and statistics*, pages 202–210, 1990.

[24] Christoph Schottmüller. Adverse selection without single crossing: Monotone solutions. *Journal of Economic Theory*, 158:127–164, 2015.

[25] Johannes Van Biesebroeck. Productivity dynamics with technology choice: An application to automobile assembly. *The Review of Economic Studies*, 70(1):167–198, 2003.

A Proofs

A.1 Proof of Lemma 1

When maximizing its profit with respect to the type $\tilde{\theta}$ it reports, given its type θ, supplier U faces the following necessary and sufficient conditions. The derivative of its profit $\pi_U(q(\tilde{\theta});\theta)$ with respect to $\tilde{\theta}$ must be nil at $\tilde{\theta}=\theta$,

$$T'(\tilde{\theta}) - q'(\tilde{\theta})C_q(q(\tilde{\theta});\theta)|_{\tilde{\theta}=\theta} = 0, \tag{34}$$

and this solution must be maximizing its profit,

$$T''(\tilde{\theta}) - q''(\tilde{\theta})C_q(q(\tilde{\theta});\theta) - (q'(\tilde{\theta}))^2 C_{qq}(q(\tilde{\theta});\theta) \leq 0. \tag{35}$$

The necessary condition at $\tilde{\theta}=\theta$ rewrites as $T'(\theta) - q'(\theta)C_q(q(\theta);\theta) = 0$. Differentiating this expression with respect to θ gives

$$T''(\theta) - q''(\theta)C_q(q(\theta);\theta) - (q'(\theta))^2 C_{qq}(q(\theta);\theta) - q'(\theta)C_{q\theta}(q(\theta);\theta) = 0. \tag{36}$$

Using the sufficient condition (35) above rewritten for $\tilde{\theta}=\theta$, and substituting (36), we obtain

$$q'(\theta)C_{q\theta}(q(\theta);\theta) \leq 0. \tag{37}$$

The cross partial derivative of total cost with respect to q and θ evaluated at $q(\theta)$ is equal to $C_{q\theta}(q(\theta);\theta) = 1 - \frac{\bar{d}}{\bar{c}}q(\theta)$, whose sign is given by

$$C_{q\theta}(q(\theta);\theta) \geq 0 \Leftrightarrow 1 - \frac{\bar{d}}{\bar{c}}q(\theta) \geq 0 \Leftrightarrow q(\theta) \leq q^0 \tag{38}$$

while $C_{q\theta}(q(\theta);\theta) \leq 0$ when $q(\theta) \geq q^0$. Therefore $q'(\theta) \leq 0$ for θ such that $q(\theta) \leq q^0$ while $q'(\theta) \geq 0$ for θ such that $q(\theta) \geq q^0$.

A.2 Proof of Lemma 2

The derivative of the supplier's profit with respect its type θ when U reports it truthfully is equal to

$$\pi_U'(\theta;\theta) = T'(\theta) - q'(\theta)C_q(q(\theta);\theta) - C_\theta(q(\theta);\theta) = -C_\theta(q(\theta);\theta) \tag{39}$$

once the first order condition of the reporting game has been cancelled out. Using the expression of the total cost, it comes

$$C_\theta(q(\theta);\theta) = q(\theta) + \frac{1}{2}\bar{d}\left(-\frac{1}{\bar{c}}\right)(q(\theta))^2 = q(\theta)\left(1 - \frac{\bar{d}}{2\bar{c}}q(\theta)\right). \tag{40}$$

As $q(\theta) \geq 0$, the sign of this derivative is given by the sign of $1-\frac{\bar{d}}{2\bar{c}}q(\theta)$. Consequently $\pi'_U(c;c) \geq 0$ if $1-\frac{\bar{d}}{2\bar{c}}q(\theta) \leq 0$, i.e. if $q(\theta) \geq \frac{2\bar{c}}{\bar{d}} = 2q^0$, and $\pi'_U(\theta;\theta) \leq 0$ if $1-\frac{\bar{d}}{2\bar{c}}q(\theta) \geq 0$, i.e. if $q(\theta) \leq \frac{2\bar{c}}{\bar{d}} = 2q^0$.

A.3 Proof of Lemma 3

The profit of supplier U of type θ, $\pi_U(q(\theta);\theta)$, rewrites as a function of the partial derivative of its cost function: using the first order condition of the revelation game,

$$\pi'_U(q(\theta);\theta) = T'(\theta) - q'(\theta)C_q(q(\theta);\theta) - C_\theta(q(\theta);\theta) = -C_\theta(q(\theta);\theta), \tag{41}$$

so that

$$\begin{aligned}\pi_U(q(\theta);\theta) =& T(\theta) - C(q(\theta);\theta) = \pi_U(q(\bar{c});\bar{c}) - \int_\theta^{\bar{c}} \pi'_U(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} \\ =& \pi_U(q(\bar{c});\bar{c}) + \int_\theta^{\bar{c}} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta}.\end{aligned} \tag{42}$$

There are no non-local (or global) deviations if a type θ is better off announcing its true type θ than any other type $\hat{\theta}$ in $[0,\bar{c}]$[34]. That is,

$$\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = \pi_U(q(\theta);\theta) - T(\hat{\theta}) + C(q(\hat{\theta});\theta) \geq 0 \quad \forall \hat{\theta} \neq \theta. \tag{43}$$

We can use (42) above, and then rewrite this deviation as a function of $C_{q\theta}(q;\theta)$. We have:

$$\begin{aligned}&\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = \pi_U(q(\theta);\theta) - T(\hat{\theta}) + C(q(\hat{\theta});\theta) \\ =& \pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\hat{\theta}) - C(q(\hat{\theta});\hat{\theta}) + C(q(\hat{\theta});\theta) \\ =& \int_\theta^{\bar{c}} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} - \int_{\hat{\theta}}^{\bar{c}} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} - C(q(\hat{\theta});\hat{\theta}) + C(q(\hat{\theta});\theta).\end{aligned} \tag{44}$$

Then suppose that $\hat{\theta} < \theta$ and $q'(\theta) > 0$, so that $q(\hat{\theta}) < q(\theta)$. (44) simplifies into

$$\begin{aligned}&\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = -\int_{\hat{\theta}}^{\theta} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} + \int_{\hat{\theta}}^{\theta} C_\theta(q(\hat{\theta});\tilde{\theta})d\tilde{\theta} \\ =& -\int_{\hat{\theta}}^{\theta} \left(C_\theta(q(\tilde{\theta});\tilde{\theta}) - C_\theta(q(\hat{\theta});\tilde{\theta})\right) d\tilde{\theta} = -\int_{\hat{\theta}}^{\theta}\int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta}.\end{aligned} \tag{45}$$

[34] See Araujo and Moreira [1] and Schottmüller [24].

The same simplification can be done for $\hat{\theta} > \theta$ and $q'(\theta) > 0$, so that $q(\hat{\theta}) > q(\theta)$: (44) simplifies into

$$\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = \int_\theta^{\hat{\theta}} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} - \int_\theta^{\hat{\theta}} C_\theta(q(\hat{\theta});\tilde{\theta})d\tilde{\theta}$$

$$= -\int_\theta^{\hat{\theta}} \left(C_\theta(q(\hat{\theta});\tilde{\theta}) - C_\theta(q(\tilde{\theta});\tilde{\theta})\right) d\tilde{\theta} = -\int_\theta^{\hat{\theta}} \int_{q(\tilde{\theta})}^{q(\hat{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta}$$

$$= -\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta} \tag{46}$$

For $\hat{\theta} < \theta$ and $q'(\theta) < 0$, so that $q(\hat{\theta}) > q(\theta)$, (44) simplifies into

$$\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = -\int_{\hat{\theta}}^{\theta} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} + \int_{\hat{\theta}}^{\theta} C_\theta(q(\hat{\theta});\tilde{\theta})d\tilde{\theta}$$

$$= \int_{\hat{\theta}}^{\theta} \left(C_\theta(q(\hat{\theta});\tilde{\theta}) - C_\theta(q(\tilde{\theta});\tilde{\theta})\right) d\tilde{\theta} = \int_{\hat{\theta}}^{\theta} \int_{q(\tilde{\theta})}^{q(\hat{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta}$$

$$= -\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta} \tag{47}$$

Last for $\hat{\theta} > \theta$ and $q'(\theta) < 0$, so that $q(\hat{\theta}) < q(\theta)$, (44) simplifies into

$$\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = \int_\theta^{\hat{\theta}} C_\theta(q(\tilde{\theta});\tilde{\theta})d\tilde{\theta} - \int_\theta^{\hat{\theta}} C_\theta(q(\hat{\theta});\tilde{\theta})d\tilde{\theta}$$

$$= \int_\theta^{\hat{\theta}} \left(C_\theta(q(\tilde{\theta});\tilde{\theta}) - C_\theta(q(\hat{\theta});\tilde{\theta})\right) d\tilde{\theta} = \int_\theta^{\hat{\theta}} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta}$$

$$= -\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta} \tag{48}$$

Therefore no matter the deviation considered and the sign of $q'(\theta)$, the non local incentive constraint requires

$$\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta) = -\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} C_{q\theta}(\tilde{q};\tilde{\theta})d\tilde{q}d\tilde{\theta} = \int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(\tilde{\theta})} \left(\frac{\tilde{q}}{q^0} - 1\right) d\tilde{q}d\tilde{\theta} \geq 0 \tag{49}$$

for any θ and $\hat{\theta}$ in $[0, \bar{c}]$. Then suppose that in the contract offered by D to U, the quantity scheme $q(\theta)$ is always strictly larger than q^0 for all $\theta \in [0, \bar{c}]$ and increasing. All the pairs quantities above which the cross-partial derivative $C_{q\theta}(q;\theta)$ is integrated belong to the half-plane $q > q^0$ in which $C_{q\theta}(q;\theta) < 0$, and hence $\frac{\tilde{q}}{q^0} - 1 > 0$. Consequently the non-local incentive constraint, which is equal to an integral below a strictly positive function, is strictly positive too. The same holds true when in the contract offered by D to U, the quantity scheme $q(\theta)$ is always strictly lower than q^0 for all $\theta \in [0, \bar{c}]$ and decreasing. All the pairs of quantities above which the cross-partial derivative $C_{q\theta}(q;\theta)$ is integrated belong to the half-plane $q < q^0$ in which $C_{q\theta}(q;\theta) > 0$, and hence $\frac{\tilde{q}}{q^0} - 1 < 0$. As $q(\hat{\theta}) > q(\tilde{\theta})$ when $\hat{\theta} < \theta$, the integration bounds of the second integral with respect to q must be permuted, which reverts the sign of $\pi_U(q(\theta);\theta) - \pi_U(q(\hat{\theta});\theta)$ to positive.

A.4 Proof of Lemma 4

As the analysis of (26) demonstrates, the virtual surplus is concave in q for $M = 1$ and hence the solution $\tilde{q}(\theta, 1)$ solves the first order condition which is necessary and sufficient:

$$\Pi_q(\tilde{q}(\theta,1);\theta) - \frac{F(\theta)-1}{f(\theta)} C_{\theta q}(\tilde{q}(\theta,1);\theta) = 0. \tag{50}$$

It comes

$$\tilde{q}(\theta,1) = \frac{a - \theta - \frac{F(\theta)-1}{f(\theta)}}{2b + \overline{d} - \frac{\overline{d}}{\overline{c}}(\theta + \frac{F(\theta)-1}{f(\theta)})}. \tag{51}$$

which is continuous in θ, and

$$\frac{d\tilde{q}(\theta,1)}{d\theta} = \frac{[-1 + \frac{d(\frac{1-F(\theta)}{f(\theta)})}{d\theta}][2b + \overline{d} + \frac{\overline{d}}{\overline{c}}(-\theta + \frac{1-F(\theta)}{f(\theta)})] - [a - \theta + \frac{1-F(\theta)}{f(\theta)}][\frac{\overline{d}}{\overline{c}}(\frac{d(\frac{1-F(\theta)}{f(\theta)})}{d\theta} - 1)]}{[2b + \overline{d} + \frac{\overline{d}}{\overline{c}}(-\theta + \frac{1-F(\theta)}{f(\theta)})]^2}. \tag{52}$$

The sign of (52) is the sign of its numerator:

$$N = \left(\frac{2b\overline{c} + \overline{d}\overline{c} - a\overline{d}}{\overline{c}}\right)\left(\frac{d(\frac{1-F(\theta)}{f(\theta)})}{d\theta} - 1\right), \tag{53}$$

The first factor in N is negative when $V_q(q^0) \geq 0$. The second factor is also negative by the assumption made on the distribution $F(\theta)$. Therefore $\tilde{q}(\theta, 1)$ is increasing in θ when the demand is large. This proves (i).

Consider now $M = 0$: in this case the virtual surplus is not concave in q for every θ. When $M(\theta) \geq 0$ the second order condition for V to be concave writes:

$$G(\theta) - \frac{2b + \bar{d}}{\bar{d}}\bar{c} \leq \frac{M(\theta)}{f(\theta)}, \tag{54}$$

with $G(\theta) = \theta + \frac{F(\theta)}{f(\theta)}$ defined in Assumption 1. When $M = 0$, (54) rewrites

$$G(\theta) \leq \frac{2b + \bar{d}}{\bar{d}}\bar{c}. \tag{55}$$

Let $G^{-1}(x)$ the reciprocal of $G(x)$. Then (55) is equivalent to

$$\theta \leq G^{-1}\left((2b+\bar{d})\frac{\bar{c}}{\bar{d}}\right) \equiv \bar{\theta}. \tag{56}$$

As the virtual surplus is concave for $\theta = 0$ where $F(0) = 0$, the threshold value $\bar{\theta}$ is strictly positive. As $\bar{\theta}$ can be larger or smaller than $\bar{c}$, the virtual surplus may be convex in some case. Indeed, when $\bar{\theta} \geq \bar{c}$, the virtual surplus is concave for each θ in $[0, \bar{c}]$. However, when $\bar{\theta} < \bar{c}$, the second order derivative of V_D^e evaluated at $\theta = \bar{c}$ is equal to $-2b + \frac{\bar{d}}{\bar{c}f(\bar{c})}$ and can be positive when $\bar{d}$ is large enough (i.e. when cost functions are very convex in q), or negative when $\bar{d}$ is not sufficiently large.

Let $\tilde{q}(\theta, 0)$ be the solution of the first order condition

$$\Pi_q(\tilde{q}(\theta,0);\theta) - \frac{F(\theta)}{f(\theta)}C_{\theta q}(\tilde{q}(\theta,0);\theta) = 0. \tag{57}$$

It comes

$$\tilde{q}(\theta,0) = \frac{a - \theta - \frac{F(\theta)}{f(\theta)}}{2b + \bar{d} - \frac{\bar{d}}{\bar{c}}(\theta + \frac{F(\theta)}{f(\theta)})}, \tag{58}$$

with $\lim_{\theta \to \bar{\theta}} \tilde{q}(\theta, 0) = +\infty$.

Calculating the derivative of $\tilde{q}(\theta, 0)$ gives

$$\frac{d\tilde{q}(\theta,0)}{d\theta} = \frac{[-1 - \frac{d(\frac{F(\theta)}{f(\theta)})}{d\theta}][2b + \bar{d} - \frac{\bar{d}}{\bar{c}}(\theta + \frac{F(\theta)}{f(\theta)})] - [a - \theta - \frac{F(\theta)}{f(\theta)}][-\frac{\bar{d}}{\bar{c}}(\frac{d(\frac{F(\theta)}{f(\theta)})}{d\theta} + 1)]}{[2b + \bar{d} - \frac{\bar{d}}{\bar{c}}(\theta + \frac{F(\theta)}{f(\theta)})]^2}, \tag{59}$$

which is positive if

$$\left[2b + \bar{d} - a\frac{\bar{d}}{\bar{c}}\right]\left[-1 - \frac{d(\frac{F(\theta)}{f(\theta)})}{d\theta}\right] \geq 0. \tag{60}$$

Under our assumption on the distribution F, (60) is positive if and only if $[2b + \bar{d} - a\frac{\bar{d}}{\bar{c}}]$ is negative, which is always true when the demand is large, i.e. $V_q(q^0) \geq 0$. Therefore, $\tilde{q}(\theta, 0)$ is continuous and increasing in θ from $\tilde{q}(\theta, 0)$ to $+\infty$, for $\theta \in [0, \bar{\theta}]$. It follows that there is a unique value $\bar{\bar{\theta}} < \bar{\theta}$ such that $\tilde{q}(\bar{\bar{\theta}}, 0) = q_{max}(\bar{\bar{\theta}})$. We must distinguish the cases $\bar{\bar{\theta}}$ lower or greater than $\bar{c}$. In the first case, we have $\tilde{q}(\theta, 0) \geq q_{max}(\theta)$ for any $\theta \geq \bar{\bar{\theta}}$. In the second case, $\tilde{q}(\theta, 0)$ is always lower than $q_{max}(\theta)$. This proves (ii).

The comparison of $\tilde{q}(\theta, 0)$ with $\tilde{q}(\theta, 1)$ is straightforward from the comparison of the first order conditions which determine these quantities. The virtual surplus is concave in q for $M = 1$ and $M = 0$ for any θ when $\bar{\bar{\theta}} \leq \min\{\bar{\theta}, \bar{c}\}$: both marginal virtual surpluses defining these two solutions are strictly decreasing in q. Therefore, under the assumption $V_q(q^0) \geq 0$, the quantity produced is greater than q^0, i.e. such that $C_{\theta q} < 0$ for all θ. The left-hand-side of the first order condition which determines $\tilde{q}(\theta, 1)$ defines a function of θ which is below the function of θ defined by the first order condition which determines $\tilde{q}(\theta, 0)$. When $M = 0$ and $\bar{\bar{\theta}} \geq \bar{c}$, both marginal surpluses are again concave, and again $\tilde{q}(\theta, 1)$ is below $\tilde{q}(\theta, 0)$. When $\theta \geq \bar{\bar{\theta}}$, the quantity ordered for $M = 0$ is $q_{max}(\theta)$ and is *a fortiori* above $\tilde{q}(\theta, 1)$. Therefore $\min\{\tilde{q}(\theta, 0), \bar{q}\} > \tilde{q}(\theta, 1)$.

When the demand is linear, the calculation of the difference $\tilde{q}(\theta, 0) - \tilde{q}(\theta, 1)$ gives

$$\frac{\frac{1}{f(\theta)}(a\frac{\bar{d}}{\bar{c}} - 2b + \bar{d})}{[2b + \bar{d} - \frac{\bar{d}}{\bar{c}}G(\theta)]^2 + \frac{\bar{d}}{\bar{c}f(\theta)}[2b + \bar{d} - \frac{\bar{d}}{\bar{c}}G(\theta)]}. \tag{61}$$

When $\theta < \bar{\theta}$ and $V_q(q^0) \geq 0$, the numerator and the denominator are positive. Indeed, $V_q(q^0) \geq 0$ is equivalent to $a\frac{\bar{d}}{\bar{c}} - 2b + \bar{d} \geq 0$, and $2b + \bar{d} - \frac{\bar{d}}{\bar{c}}G(\theta) \geq 0$ is equivalent to $G(\theta) \leq G(\bar{\theta})$ which is always true when $\theta < \bar{\theta}$ as $G(\theta)$ is increasing. When $\theta \geq \bar{\theta}$, the quantity ordered for $M = 0$ is $q_max(\theta)$. Last the monopoly solution $q^M(\theta)$ is strictly increasing with θ, and such that $\Pi_q(q^M(\theta); \theta) = 0$. Inspecting the first order conditions determining $\tilde{q}(\bar{c}, 1)$ and $\tilde{q}(0, 0)$ above, it is immediate to verify that these solutions coincide with $q^M(\bar{c})$ and $q^M(0)$ respectively. Therefore $q^M(\theta)$, which increases with θ, lies in between the solutions for $M = 1$ and $M = 0$. This proves (iii). Lemma 2 established that the supplier payoff decreases with θ when the quantity is lower than $2q^0$, while it increases with θ when q exceeds $2q^0$. Moreover for every $\theta \in [0, \theta_1)$, the equilibrium order $q^*(\theta)$ which belongs to $[\tilde{q}(\theta, 1), \tilde{q}(\theta, 0)]$ is strictly lower than $2q^0$.

A.5 Proof of Proposition 1 and its corollary

When $P(2q^0) + 2q^0 P'(2q^0) \in [C_q(2q^0; \bar{c}), C_q(2q^0; 0)]$, two cases arise depending on wether $\tilde{q}(\bar{\bar{\theta}}, 0)$ is greater or lower than $\tilde{q}(\bar{c}, 1)$. In the first case, as here we consider a demand such that $\tilde{q}(0, 1) \leq 2q^0 \leq \tilde{q}(\bar{c}, 1)$, it follows that $2q^0$ is always lower than $\tilde{q}(\bar{\bar{\theta}}, 0)$, and therefore for $M = 0$, $q(\theta) = \tilde{q}(\theta, 0)$ for any $\theta \in [0, \bar{\theta}]$. In the second case, it is possible to have $2q^0 \geq \tilde{q}(\bar{\bar{\theta}}, 0)$ and therefore for $M = 0$, $q(\theta) = \tilde{q}(\theta, 0)$ for $\theta \leq \bar{\bar{\theta}}$, and $q(\theta) = q_{max}(\theta)$ for $\theta \geq \bar{\bar{\theta}}$.

The proof is similar for each case, we make it in the case $\tilde{q}(\bar{\bar{\theta}}, 0) \geq \tilde{q}(\bar{c}, 1)$, illustrated in Figure ??. Lemma 4 established that $\tilde{q}(\theta, 0)$ and $\tilde{q}(\theta, 1)$ are ~~strictly~~ increasing in θ. Therefore it exists θ_1 and θ_2 such that $\theta_1 \leq \theta_2$ which solve respectively $\tilde{q}(\theta_1, 0) = 2q^0$ and $\tilde{q}(\theta_2, 1) = 2q^0$. Therefore as $\Pi_U(2q^0, \theta) \geq 0$ and decreases in θ when $q < 2q^0$, the profit of supplier U, $\Pi_U(q^*(\theta), \theta)$, is strictly larger than $\Pi_U(2q^0, \theta)$ and hence strictly positive for $\theta < \theta_1$. Therefore $\mu^*(\theta) = 0$ for every $\theta \in [0, \theta_1)$ and $M^*(\theta) = 0$ for every $\theta \in [0, \theta_1)$. The optimization of the virtual surplus, which is concave in q for $\theta \leq \bar{\bar{\theta}} < \bar{\theta}$, directly leads the retailer to offer $q^*(\theta) = \tilde{q}(\theta, 0) < 2q^0$.

A symmetric argument applies to every $\theta \in (\theta_2, \bar{c}]$: the equilibrium order $q^*(\theta)$ which belongs to $[\tilde{q}(\theta, 1), \tilde{q}(\theta, 0)]$ is strictly larger than $2q^0$. Therefore as $\Pi_U(2q^0, \theta) \geq 0$ and increases for every θ such that $q > 2q^0$, then $\Pi(q^*(\theta), \theta) > 0$ for $\theta \in (\theta_2, \bar{c}]$. Consequently $\mu^*(\theta) = 0$ for every $\theta \in (\theta_2, \bar{c}]$, and $M^*(\theta) = 1$ for every $\theta \in (\theta_2, \bar{c}]$. The optimization of the virtual surplus, which is concave in q for every θ, directly leads the retailer to offer $q^*(\theta) = \tilde{q}(\theta, 1) > 2q^0$.

It remains to examine the case where $\theta \in [\theta_1, \theta_2]$. First, let θ^0 be the value of θ such that $\Pi_q(2q^0, \theta^0) = 0$, i.e. $q^M(\theta^0) = 2q^0$. By (25), the retailer can offer the contract $(q(\theta), \pi_U(q; \theta)) = (2q^0, 0)$ to its supplier θ^0 if $M(\theta^0) = F(\theta^0) > 0$. This contract does not leave a rent to this supplier, and requires this type to supply the first best. Then (50) and (25) are both satisfied (with strict inequality for the latter) if $\theta = \theta^0$. Therefore $(q^*(\theta^0), M^*(\theta^0)) = (2q^0, F(\theta^0))$ maximizes the virtual surplus the retailer obtains from a type θ^0. Then from Lemma 1 the quantity scheme offered to the supplier must be weakly increasing in θ to respect the local incentive compatibility constraint. Therefore for $\theta \in [\theta_1, \theta_2]$ the quantity offered must be constant and equal to $2q^0$: if it was different from $2q^0$, the local incentive compatibility constraint would be violated, by forcing the quantity $q(\theta)$ to decrease either compared to its level for $\theta \in [0, \theta_1)$, or within the interval $[\theta_1, \theta_2]$. The contract $(q(\theta), \pi_U(q; \theta)) = (2q^0, 0)$ can therefore be offered to all types $\theta \in [\theta_1, \theta_2]$. Then $M^*(\theta)$ must be equal to

$$M^*(\theta) = F(\theta) - f(\theta)\frac{\Pi_q(2q^0; \theta)}{C_{\theta q}(2q^0; \theta)} \tag{62}$$

which is such that equates the first order condition of the optimization of the virtual surplus is equal to 0 for $q(\theta) = 2q^0$.

It remains to establish the continuity of $q^*(\theta)$, by verifying the continuity of $M^*(\theta)$ at θ_1 and θ_2. The functions $\Pi_q(q;\theta)$ and $C_{q\theta}(q;\theta)$ are continuous in θ. Therefore from the determination of $\tilde{q}(\theta,0)$ in Lemma 4, we have

$$\frac{\Pi_q(2q^0,\theta_1)}{C_{q\theta}(2q^0,\theta_1)} = \frac{F(\theta_1)}{f(\theta_1)}. \tag{63}$$

We can plug this expression into $M^*(\theta_1)$ which gives immediately $M^*(\theta_1) = 0$. The same analysis can be performed at θ_2, where from Lemma 4 again we have

$$\frac{\Pi_q(2q^0,\theta_2)}{C_{q\theta}(2q^0,\theta_2)} = \frac{F(\theta_2)-1}{f(\theta_2)}, \tag{64}$$

and therefore $M^*(\theta_2) = 1$. Since the industry marginal profit $\Pi_q(q;\theta)$ is strictly decreasing in q, and since $\Pi_q(2q^0;\theta^0) = 0$, $\Pi_q(2q^0;\theta) > 0$ for $\theta > \theta^0$ (as the first best order $q^M(\theta)$ exceeds $2q^0$) and $\Pi_q(2q^0;\theta) < 0$ for $\theta < \theta^0$ (as the first best order $q^M(\theta)$ is strictly lower than $2q^0$). Therefore $M^*(\theta) \leq F(\theta)$ for $\theta \leq \theta^0$ and $M^*(\theta) \geq F(\theta)$ for $\theta \geq \theta^0$.

The corollary can be proved immediately: the comparison between $q^*(\theta)$ and $q^M(\theta)$ is a straightforward consequence of Lemma 4: since $q_{max}(\theta) \geq \tilde{q}(\theta,0) > q^M(\theta)$, then $q^*(\theta) > q^M(\theta)$ for $\theta < \theta^0$. Conversely, since $\tilde{q}(\theta,1) < q^M(\theta)$, we have $q^*(\theta) < q^M(\theta)$ for $\theta > \theta^0$.

A.6 Proof of Proposition 2 and its corollary

Start with Proposition 2. When $2q^0P'(2q^0) + P(2q^0) - C_q(2q^0;0) > 0$, and $\tilde{q}(0,1) \leq 2q^0$, for any $\theta \in [\theta_2; \min\{\bar{\theta};\bar{c}\}]$, $q^*(\theta)$ is strictly larger than $2q^0$ and by the same reasoning as in the proof of proposition 1, the optimization of the virtual surplus leads the retailer to offer $q^*(\theta) = q(\tilde{\theta},1) \geq 2q^0$.

For any $\theta \in [0,\theta_2]$, we must distinguish when $V(q(\theta);\theta)$ is concave or convex. When $V(q(\theta);\theta)$ is concave, V_q is decreasing in q. As $V_q(q^0) \geq 0$, $q^*(\theta) \geq q^0$ for any $\theta \in [0,\theta_2]$, and by Lemma 1, $q^*(\theta)$ is increasing in θ. As $q^*(\theta_2) = 2q^0$, $q^*(\theta)$ must be such that $\tilde{q}(0,1) \leq q^*(\theta) \leq 2q^0$. On one hand, $q^*(\theta) \leq 2q^0$ implies from Lemma 2 that the rent of the supplier is decreasing in θ: $\Pi_U(q^*(\theta),\theta) \geq \Pi_U(2q^0,\theta)$. On the other hand, $q^*(\theta)$ is here always lower than $q^M(\theta)$ implying that $\Pi_q(q^*(\theta))$ is positive for any $\theta \in [0,\theta_2]$ ($R_m(q^*(\theta)) \geq C_q(q^*(\theta))$). It follows that the expected virtual surplus of the retailer is increasing in $q^*(\theta)$. Therefore the retailer is better off by setting for any $\theta \in [0,\theta_2]$ an order $q^*(\theta) = 2q^0$ which eliminates informational rents and increases its profit .

When $V(q(\theta);\theta)$ is convex, which is possible depending on the sign of (28), $V(q(\theta);\theta)$ is positive and increasing in q with $V_q(q^0) \geq 0$. Therefore the retailer would like to order the maximal possible quantity: $\tilde{q}(\theta,0)$ when $\theta \leq \bar{\bar{\theta}}$, and $q_{max}(\theta) > q^0$ otherwise. However, $\tilde{q}(\theta,0)$ and $q_{max}(\theta)$ are greater than $2q^0$ the value of $q^*(\theta)$ when $\theta = \theta_2$. And as $q'(\theta) \geq 0$ for $q(\theta) \geq q^0$ it is not possible for the retailer to order a quantity greater than $2q^0$ over $[0,\theta_2]$: the monotonicity condition would be violated. The retailer is better off ordering $2q^0$.

When $2q^0P'(2q^0) + P(2q^0) - C_q(2q^0;0) > 0$, and $\tilde{q}(0,1) > 2q^0$, $q^*(\theta)$ which belongs to $[\tilde{q}(\theta,0),\tilde{q}(\theta,1)]$ is greater than $2q^0$. From Lemma 2, $\Pi_U(q^*(\theta),\theta)$ and $q^*(\theta)$ are strictly increasing too. The retailer must distort downward the quantity asked to the supplier from the efficient one to decrease the informational rents. For any θ, the minimal possible quantity is the closest to $2q^0$, i.e. $\tilde{q}(\theta,1)$. Moreover, as $q^*(\theta)$ is strictly increasing, only the participation constraint of the type 0 supplier can be binding with $\Pi_U(q^*(0)) = 0$ for the quantity $q^*(0) = \tilde{q}(0,1)$, all other types benefiting from a positive rent. As $M(\theta) = 1$ for any θ, $V(q(\theta);\theta)$ is always concave. Finally, in both cases $\tilde{q}(0,1) \leq 2q^0$ and $\tilde{q}(0,1) > 2q^0$, only the type $\bar{c}$ realizes the first best: $\tilde{q}(\bar{c},1) = q^M(\bar{c})$. The proof of the corollary is immediate: $2q^0$ and $\tilde{q}$ are lower than $q^M(\theta)$ for any $\theta \in [0,\bar{c}]$.

A.7 Proof of Proposition 3 and its corollary

Consider now Proposition 3: as $2q^0P'(2q^0)+P(2q^0)-C_q(2q^0;\bar{c}) < 0$, i.e. $R_m(2q^0) \leq C_q(q(\theta),\bar{c})$, we have $2q^0 \geq q^M(\theta)$ for any $\theta \in [0,\bar{c}]$. Moreover, the large demand implies $q^0 \leq q^M(0) = \tilde{q}(0,0)$, therefore $q^0 \leq q^M(\theta) \leq 2q^0$. For any $\theta \in [0,\theta_1]$, with θ_1 such that $\tilde{q}(\theta_1,0) = 2q^0$, $q^*(\theta) \leq 2q^0$ and the rent of the supplier is decreasing in q with a minimum at $\Pi_U(2q^0,\theta) \geq 0$. It follows that $\mu^*(\theta) = 0$ and $M^*(\theta) = 0$ for every $\theta \in [0,\theta_1]$, with θ_1 always lower than $\bar{\theta}$. The optimal quantity ordered to the types lower than θ_1 is $\tilde{q}(\theta,0)$.

As $\tilde{q}(\theta_1,0) = 2q^0$ and $q^*(\theta) \geq 0$ is increasing over $[0,\bar{c}]$, the optimal quantity over $[\theta_1,\bar{c}]$ must be larger than $2q^0$. This implies that $\Pi_U(q^*(\theta),\theta)$ is increasing in q for every $\theta \in [\theta_1,\bar{c}]$. To decrease the rent of the supplier, the retailer set the orders of the different types of supplier as close as possible to $2q^0 \geq q^M(\bar{c})$. Moreover, decreasing the quantity ordered to the supplier allows him to increase its profit. Indeed, $R_m(q)$ is decreasing in q therefore $q^*(\theta) \geq 2q^0$ implies that $R_m(q^*(\theta)) \leq R_m(2q^0)$. And here we have $R_m(2q^0) \leq C_q(q(\theta), \bar{c}) \leq C_q(q(\theta), \theta)$ for any $\theta \in [0, \bar{c}]$, therefore $R_m(q^*(\theta)) \leq C_q(q^*(\theta), \theta)$, i.e. $\Pi_q(q^*(\theta), \theta) \leq 0$ for any $\theta \in [\theta_1, \bar{c}]$. By decreasing $q^*(\theta)$, the retailer increases its profit and minimizes the rent of the supplier. It follows that the retailer sets $q^*(\bar{c}) = 2q^0$ and as $\tilde{q}(\theta_1, 0) = 2q^0$ and $q^*(\theta)$ must be increasing in θ (monotonicity constraint), $q^*(\theta) = 2q^0$ for any $\theta \in [\theta_1, \bar{c}]$.

The proof of the corollary is immediate: $\tilde{q}(\theta, 0)$ and $2q^0$ are larger than $q^M(\theta)$. And as only the IR constraint of the type $\bar{c}$ is binding, $M^*(\bar{c}) = 1$ and $M^*(\theta) = 0$ for all $\theta \neq \bar{c}$.

A.8 Proof of Lemma 5

As the proof of Lemma 4 established, the virtual surplus is concave in q for $M = 1$ and hence the solution $\hat{q}(\theta, 1)$ solves the first order condition which is necessary and sufficient:

$$\Pi_q(\hat{q}(\theta, 1); \theta) - \frac{F(\theta) - 1}{f(\theta)} C_{\theta q}(\hat{q}(\theta, 1); \theta) = 0, \tag{65}$$

where $\hat{q}(\theta, 1)$ is continuous in θ. Similarly to the proof of Lemma 4, we obtain

$$\hat{q}(\theta, 1) = \frac{a - \theta - \frac{F(\theta)-1}{f(\theta)}}{2b + \bar{d} - \frac{\bar{d}}{\bar{c}}(\theta + \frac{F(\theta)-1}{f(\theta)})}. \tag{66}$$

which is continuous in θ. And the sign of $\frac{d\tilde{q}(\theta,1)}{d\theta}$ is given by the sign of:

$$N = \left(\frac{2b\bar{c} + \bar{d}\bar{c} - a\bar{d}}{\bar{c}}\right)\left(\frac{d(\frac{1-F(\theta)}{f(\theta)})}{d\theta} - 1\right), \tag{67}$$

The first factor in N is positive when $V_q(q^0) \leq 0$. The second factor is negative by the assumption made on the distribution $F(\theta)$. Therefore $\hat{q}(\theta, 1)$ is decreasing in θ when the demand is large. This proves (i).

We have also demonstrated that the virtual surplus is not concave in q for $M = 0$ and every θ. When $M = 0$, using again $G(\theta) = \theta + \frac{F(\theta)}{f(\theta)}$ which is an increasing and continuous function, with G^{-1} its reciprocal, the second order condition is negative if and only if

$$-2b - \bar{d} + \frac{\bar{d}}{\bar{c}}\left(\theta + \frac{F(\theta)}{f(\theta)}\right) \leq 0 \Leftrightarrow \theta \leq G^{-1}\left((2b + \bar{d})\frac{\bar{c}}{\bar{d}}\right) \equiv \bar{\theta}. \tag{68}$$

As the virtual surplus is concave for $\theta = 0$, the threshold value $\bar{\theta}$ is strictly positive. However it can be larger or smaller than $\bar{c}$: when $\theta = \bar{c}$, the second order derivative of V_D^e is equal to $-2b + \frac{\bar{d}}{\bar{c}f(\bar{c})}$ which can be positive when $\bar{d}$ is large (i.e. when cost functions are very convex in q), or negative when $\bar{d}$ is not sufficiently large. The virtual surplus V_D^e ~~is therefore~~ may not be concave for $M = 0$ when $\theta \in [\min\{\bar{c};\bar{\theta}\}, \bar{c}]$.

When $\theta \in [0, \min\{\bar{c};\bar{\theta}\}]$, the first order condition is sufficient to determine $\hat{q}(\theta, 0)$ which solves

$$\Pi_q(\hat{q}(\theta,0);\theta) - \frac{F(\theta)}{f(\theta)} C_{\theta q}(\hat{q}(\theta,0);\theta) = 0, \tag{69}$$

where $\hat{q}(\theta, 0)$ is continuous in θ for $\theta \in [0, \min\{\bar{c};\bar{\theta}\}]$. Since $C_{\theta q} > 0$, $\hat{q}(\theta,0) < q^M(\theta)$ except for $\theta = 0$ where they coincide. It comes

$$\hat{q}(\theta,0) = \frac{a - \theta - \frac{F(\theta)}{f(\theta)}}{2b + \bar{d} - \frac{\bar{d}}{\bar{c}}(\theta + \frac{F(\theta)}{f(\theta)})}, \tag{70}$$

with $\lim_{\theta \to \bar{\theta}} \tilde{q}(\theta, 0) = -\infty$, as $a - G(\bar{\theta})$ is negative when $V_q(q^0) \leq 0$. Moreover, the derivative in θ of $\hat{q}(\theta, 0)$ is negative if

$$\left[2b + \bar{d} - a\frac{\bar{d}}{\bar{c}}\right] \left[-1 - \frac{d(\frac{F(\theta)}{f(\theta)})}{d\theta}\right] \leq 0. \tag{71}$$

Under our assumption on the distribution F, (71) is negative if and only if $[2b + \bar{d} - a\frac{\bar{d}}{\bar{c}}]$ is positive, which is always true when the demand is small, i.e. $V_q(q^0) \leq 0$. Therefore, $\hat{q}(\theta, 0)$ is continuous and decreasing in θ from $\hat{q}(0,0)$ to 0, for $\theta \in [0, \bar{\theta}]$, with $\hat{q}(\theta,0) = 0$ for $\theta \in [\underline{\theta}, \bar{\theta}]$, and $\underline{\theta}$ such that $a = G(\underline{\theta})$, i.e. $\underline{\theta} = G^-1(a)$, which is always positive since $G(0) = 0$. Consequently $\hat{q}(\theta, 0) > 0$ for $\theta < \underline{\theta}$, and $\hat{q}(\theta,0) = 0$ else, and $\hat{q}(\theta, 0)$ is continuous at $\underline{\theta}$. If $\bar{\theta} > \bar{c}$, $\hat{q}(\theta, 0)$ is strictly positive. This proves (ii).

Moreover, when $\theta = \underline{\theta}$, the marginal virtual surplus is nil,

$$\Pi_q(0;\underline{\theta}) - \frac{F(\underline{\theta})}{f(\underline{\theta})} C_{\theta q}(0;\underline{\theta}) = 0 \Leftrightarrow P(0) - C_q(0,\underline{\theta}) - \frac{F(\underline{\theta})}{f(\underline{\theta})} C_{\theta q}(0;\underline{\theta}) = 0,$$
$$\Leftrightarrow a - \underline{\theta} - \frac{F(\underline{\theta})}{f(\underline{\theta})} = 0. \tag{72}$$

. and therefore as the marginal virtual surplus is decreasing here, it is strictly negative for all $\theta \in [\underline{\theta}, \bar{\theta}]$.

When $\theta \in [\min\{\bar{c};\bar{\theta}\}, \bar{c}]$, the virtual surplus is not concave, its derivative increases with q. To respect the incentive compatibility constraints, the quantity ordered must be weakly decreasing in types. Since $\hat{q}(\theta, 0) = 0$ for $\theta \geq \underline{\theta}$, and $\underline{\theta} < \bar{\theta}$ the quantity scheme must be "ironed" as in Guesnerie and Laffont [11], with $\hat{q}(\theta, 0) = 0$ for $\theta \in [\min\{\bar{c};\bar{\theta}\}, \bar{c}]$.

Last the first order condition (25) decreases with M when $C_{q\theta} > 0$, then the solution $\hat{q}(\theta, 0)$ must be lower than $\hat{q}(\theta, 1)$, and it is immediate to verify that the solutions $\hat{q}(\bar{c}, 1)$ and $\hat{q}(0, 0)$ coincide with $q^M(\bar{c})$ and $q^M(0)$ respectively. Therefore $q^M(\theta)$, which decreases with θ when $a < (2b + \bar{d})\frac{\bar{c}}{d}$, lies in between the solutions for $M = 1$ and $M = 0$. This proves (iii).

A.9 Proof of Proposition 5

Let us start with the case where $\underline{\theta} < \bar{c}$. From Lemma 2, the profit of the supplier is strictly decreasing in θ when $q < 2q^0$, and moreover it is nil for $\theta = \underline{\theta}$ which is producing nothing. Therefore $\pi_U(0, \underline{\theta}) = 0$, and the profit of all types $\theta < \underline{\theta}$ must be strictly positive. Therefore $M(\theta) = 0$ for all types smaller than $\underline{\theta}$. Consequently the retailer is better off by ordering $\hat{q}(\theta, 0)$ to all types $\theta < \underline{\theta}$, which from Lemma 5 is the unique value which maximizes the virtual surplus. When $\theta > \underline{\theta}$, the supplier's profit must still be decreasing in θ, and to respect the implementability constraint $q^*(\theta)$ must also be weakly decreasing. Therefore the optimum for the retailer consists in ordering nothing to types $\theta > \underline{\theta}$. In that case the multiplier $M^*(\theta)$ must be equal to $M^*(\theta) = F(\theta) - f(\theta)\frac{\Pi_q(0;\theta)}{C_{\theta q}(0;\theta)} = F(\theta) - f(\theta)(a - \theta)$ for every $\theta < \bar{\theta}$ for which ironing does not occur. When ironing occurs (for $\theta > \bar{\theta}$, the multiplier $M^*(\theta)$ depends also on the value of the multiplier of the constraint $q'(\theta) = 0$ which is strictly positive. We do not determine it as it is not central to our paper.

In the case where $\underline{\theta} \geq \bar{c}$, all productions must be positive, and for the rent to be strictly decreasing in θ the retailer must set the (IR) constraint of a type $\bar{c}$ to 0. No ironing occurs Therefore $M(\bar{c}) = 1$ and all other types earn a rent, $M(\theta) = 0$ for all $\theta < \bar{c}$. Then the optimum of the retailer consists in ordering $\hat{q}(\theta; 0)$ for all types.