Vertical Exclusion with Endogenous Competition Externalities*

Stephen Hansen
Universitat Pompeu Fabra
Barcelona GSE
stephen.hansen@upf.edu

Massimo Motta
ICREA-UPF
Barcelona GSE
massimo.motta@upf.edu

March 1, 2012

PRELIMINARY DRAFT

Abstract

In a vertical market in which downstream firms have private information about their productivity and compete for consumers, an upstream firm posts a public contract for each firm. When downstream firms are risk-neutral without wealth constraints, the upstream firm offers all the input, but when they are sufficiently risk averse or wealth-constrained it sells to one. This eliminates externalities arising from competition and keeps the upstream firm from having to pay downstream firms a rent. Thus exclusion arises when contracts are fully observable and downstream firms are ex ante symmetric. The result is robust to a number of extensions.

Keywords: Exclusive Contracts, Adverse Selection, Risk, Limited Liability

JEL Codes: D82, L22, L42

*We would like to thank Larbi Alaoui for extensive comments, as well as Jan Eeckhout, Patrick Rey, participants of the ANR-DFG workshop (Paris), and seminar attendees at Bologna and Oxford for helpful advice.

1 Introduction

Very often, a manufacturer has to decide whether to sell its products through one or several retailers, a franchisor whether to have one or multiple franchisees, the owner of a patent whether to license its technology to one or more licensees. As a result, exclusive clauses may be signed whereby in a given geographical area only one agent will deal with the principal's product, brand, or technology. Such clauses are quite common in vertical markets.[1] Exclusivity clauses are a form of vertical restraint which has received the attention of antitrust agencies and courts,[2] as well as raised the interest of economic and legal scholars. The literature has so far identified pro-competitive and anti-competitive reasons for such exclusive clauses, which effectively foreclose all but one downstream agent from the market.

On the pro-competitive side, exclusive territorial protection clauses may have the effect of promoting investments by retailers. For instance, when local advertising and promotion activities matter, there may be under-provision of such activities because of free-riding among retailers of the same brand. By allocating exclusive selling rights in a given region to only one retailer, incentives to invest are restored.

Probably, the main reason why exclusive territories may have anticompetitive effects is due to Hart and Tirole (1990) (see also Rey and Tirole (2007)).[3] When contract offers are unobservable a monopolistic manufacturer may not be able to earn monopoly profits: anticipating its incentives to improve contract terms to any retailer at the expenses of the others, retailers' willingness to pay for the product lowers. The manufacturer can solve its commitment problem—conceptually similar to that of a durable good monopolist—by assigning exclusive rights to one dealer. As a result, it will be able to implement the monopolist prices and profits, but to the detriment of consumers.

We revisit the problem of an upstream firm to trade with one or several retailers without relying on any of the above-mentioned motives for exclusive clauses. In particular,

[1] Azoulay and Shane (2001) collect an original dataset of newly founded business-franchise systems. They find that 142 out of 170 systems adopt exclusive contracts. Blair and Lafontaine (2011) analyze a large dataset of franchise contracts, and show that in 17 out of 18 sectors, more than 50% of franchisors adopt exclusive territories. In the context of licensing deals, Anand and Khanna (2000) show that over 30% of the agreements in their dataset are exclusive.

[2] In the US, the Supreme Court has moved from a per se prohibition rule of exclusive territorial protection (*US v. Arnold Schwinn & Co.*, 1967) to a rule of reason (*Continental TV v. GTE Sylvania*, 1977), although it is probably fair to say that it is unlikely that such practices would be found today in violation of the Sherman Act. In the EU, such restraints would be allowed but subject to certain provisions; for instance, a manufacturer could not prohibit its retailers in one country from selling to unsolicited customers from another country.

[3] Exclusive territories may harm welfare also through strategic interbrand effects, as identified by Rey and Stiglitz (1995). By giving exclusivity to one retailer, a manufacturer will suppress intra-brand competition and induce a less aggressive price of its brand, in order to relax inter-brand competition. While industry profits will increase, consumers will end up paying higher prices, and total welfare will decrease.

while Hart and Tirole (1990) assume that contracts are private information, we show that exclusivity can arise when contracts are public information, but downstream firms have private information and are risk-averse or wealth constrained.

More precisely, we analyze an adverse selection model where two downstream firms (the "agents") competing for final consumers may differ in (privately-known) production costs. The upstream firm (the "principal") posts a public menu of contracts designed for each cost type. When downstream firms are risk-neutral and can absorb losses, both of them will be offered the input at equilibrium, because resorting to multiple agents improves the manufacturer's profits.[4]

We first show that when agents are sufficiently risk-averse, the upstream firm will optimally sell to only one firm, and exclude the other. This is because each firm's payoff is affected by the uncertainty about the cost of the rival, and it will want a rent which compensates for the risk of having to compete with a low-cost rival. By supplying only one downstream firm, the source of uncertainty is removed and the principal is better off, despite losing the benefit of having several agents.[5] A similar mechanism works when agents are wealth constrained. Again, the upstream firm must leave a rent to downstream firms that compete in the market, and it does better by eliminating the competition and trading with one firm.

There are several assumptions in the model that are somewhat restrictive in theory but less so in reality. First, since our model has many agents with hidden types, the most general contract would allow transfers and allocations for each agent to depend on type announcements of all agents. The equilibrium of such a mechanism (studied in McAfee and McMillan (1986); Laffont and Tirole (1987); McAfee and McMillan (1987); and Riordan and Sappington (1987)) resembles an auction in which the most efficient agent produces for the entire market. In our model instead the transfer and allocation of each agent only depend on its own type announcement. Still, even this approach appears far more flexible than many real-world vertical contracts in which uniformity is the norm. For example, Lafontaine (1992) surveys 130 business-format chains, and finds that 42% offered a single contract on a take-it-or-leave-it basis, with 38% more only allowing negotiations for non-monetary terms.[6] Furthermore, Lafontaine and Shaw (1999) examine a panel data set of franchisors and find that over a 13-year period 75%

[4]In our model with cost uncertainty, the principal benefits from having multiple agents because this helps smoothing output (intuitively, the principal reduces uncertainty by contracting with both firms). But one can think of other mechanisms, such as product differentiation, whereby contracting with both downstream firms raises profits.

[5]We shall also show that for intermediate risk preferences partial exclusion arises, with one downstream firm being offered no trade if it is high cost, and lower trade than the other firm if it is low cost.

[6]27% of those using exclusive contracts cited transaction costs as the main reason for contract uniformity, which are probably also major factor in explaining why upstream firms do not always allocation market shares through auctions.

of them never change their terms.

Second, in our model there is a positive probability that the upstream firm ends up offering an exclusive contract to an inefficient downstream firm. After doing so, the upstream firm would like to cancel the contract and deal with a new firm. To the extent that exclusive contracts are long-term and binding, our static contracting approach is appropriate; otherwise a more complex dynamic model would be needed. Blair and Lafontaine (2011) report that the average duration of a franchise contract in their dataset is 10.7 years, with over 90% of contracts renewed. Also, they describe rather stringent legal requirements for breaking a franchise deal or not renewing it.[7] Hence our framework appears realistic.

Finally, while limited liability is a common component of industrial organization models, firm risk aversion is not.[8] To the extent that agents are distributors or retailers, or more generally small firms which are unable to diversify risk, it will probably not take much convincing that this assumption is realistic. But there are several reasons why even larger firms may be reluctant to take risk, starting from the fact that actual decisions are taken by managers who, being individuals, may well be risk averse. Nocke and Thanassoulis (2010) also endogenously explain risk aversion when firms are credit constrained. Asplund (2002) also mentions empirical research pointing to the risk aversion of firms.

The paper is organized thus. Section 2 describes the baseline model, which is then solved first for the cases of risk-neutral and infinitely risk-averse agents in section 3. Section 4 then proposes two ways of modelling intermediate risk aversion and provides conditions for exclusive contracts to be optimal. Section 5 then examines the case where downstream firms are risk neutral but wealth constrained. Section 6 explores three extensions of the baseline model under which exclusion remains optimal under certain conditions. Section 7 concludes the paper. Appendix A contains all omitted proofs.

2 Model

We consider an industry in which a risk-neutral upstream firm M supplies an input that is transformed in a one-to-one relationship by two downstream firms $i = 1, 2$ whose product is homogenous. Aggregate demand for the product is $P(Q)$, where $Q \geq 0$ is aggregate quantity. We assume that $P'(Q) < 0$; that marginal revenue $P(Q) + QP'(Q) = \text{MR}(Q)$ is decreasing; and that there exists some finite quantity $\widetilde{Q}$ at which $\text{MR}(\widetilde{Q}) = 0$.

The downstream firms, which may be risk neutral or risk averse, are heterogenous in

[7]For example, in most US states an automobile manufacturer cannot refuse to renew a dealership contract. More generally, common law requires good cause both for termination and non-renewal, which is defined as a substantial breach of the material terms of the contract.

[8]An exception is Brickley and Dark (1987), who assume that franchisees are risk averse just as we do.

their productivity. Each has a constant marginal cost of production $c_i \in \{0, c\}$ where $c > 0$, $\Pr[c_i = 0] = r$, and c_1 and c_2 are independent. We also assume that c_i is private information for firm i. The constant returns to scale embedded in the constant marginal cost assumption keeps aggregate production costs independent of the number of firms in the market, allowing one to focus on revenue volatility as the main driver of the exclusion results. We assume that at the quantity Q' at which $\mathrm{MR}(Q') = 0$, $c < P(Q')$ also holds.[9]

M offers the downstream firms contract menus $T_i(Q_i)$, where $Q_i \geq 0$ is the amount of input that firm i uses (and which it converts to output Q_i) and $T_i(Q_i) \in \mathbb{R}$ is the transfer that firm i pays to M for using Q_i. $(0, 0)$ is included as an element of $T_i(Q_i)$. The fact that in general $T_1 \neq T_2$ for the same input level means that M can price discriminate between the firms. Unlike in Hart and Tirole (1990), we assume that these contracts are publicly observable. Given the posted contract menus, firms play a simultaneous game of incomplete information. When a pure strategy equilibrium exists, its outcome is $\{Q_i(c_i), T_i(c_i)\}_{i=1}^2$, or a quantity and transfer choice made by each cost type of each downstream firm. Using the standard revelation principle argument, one can without loss of generality focus on M offering each downstream firm a two-point, incentive-compatible contract menu. To formalize this idea, one can denote the menu offered to firm i as $[Q_i(\widehat{c}_i), T_i(\widehat{c}_i)]$ for $\widehat{c}_i \in \{0, c\}$. Here $\widehat{c}_i$ corresponds to the cost type that firm i reports at the stage it must choose an element from the menu. Let

$$\pi_i(\widehat{c}_i, \widehat{c}_j, c_i) = Q_i(\widehat{c}_i)\left\{P\left[Q_i(\widehat{c}_i) + Q_j(\widehat{c}_j)\right] - c_i\right\} - T_i(\widehat{c}_i) \tag{1}$$

be firm i's profit from reporting cost type $\widehat{c}_i$ when firm $j \neq i$ reports cost type $\widehat{c}_j$ and firm i has marginal cost c_i. Here one can see the externality in the model: firm j's choice of $\widehat{c}_j$ affects π_i through the market price. Suppose it is common knowledge that firm j reports $\widehat{c}_j = 0$ with probability r and $\widehat{c}_j = c$ with probability $1 - r$. This induces the lottery

$$L_i(\widehat{c}_i \mid c_i) = \left\{\left[\pi_i(\widehat{c}_i, 0, c_i), \pi_i(\widehat{c}_i, c, c_i)\right]; (r, 1 - r)\right\} \tag{2}$$

for firm i. Let U be the common utility function over such lotteries. This function embeds downstream firms' risk preferences, for which we will provide specific functional forms.

[9]This is a standard condition. It implies that low cost firms face competition from high cost firms in the sense that the high cost firm can still profitably produce when the low cost firm chooses its monopoly quantity.

M's problem can be expressed as

$$\max_{\{Q_i(\widehat{c}_i),T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^{2} rT_i(0) + (1-r)T_i(c) \text{ such that} \tag{3}$$

$$U\left[L_i(0 \mid 0)\right] \geq U[0] \tag{PC_i^L}$$

$$U\left[L_i(c \mid c)\right] \geq U[0] \tag{PC_i^H}$$

$$U\left[L_i(0 \mid 0)\right] \geq U\left[L_i(c \mid 0)\right] \tag{IC_i^L}$$

$$U\left[L_i(c \mid c)\right] \geq U\left[L_i(0 \mid c)\right]. \tag{IC_i^H}$$

Here there are eight constraints corresponding to participation (denoted PC) and incentive compatibility (denoted IC) constraints for each cost type of each downstream firm. $U[0]$ corresponds to the utility from receiving the certain wealth level 0. We can focus without loss of generality on contracts that induce participation since $(0,0)$ is an element of each contract. Since optimal contracts are incentive compatible, one can interpret M as posting contract menus but not actually observing which element is chosen until after each downstream firm has undertaken production. For simplicity we sometimes refer to a contract menu simply as a contract. We want to establish the conditions under which M chooses to deal with one or both of the downstream firms. Let $\{Q_i^*(\widehat{c}_i), T_i^*(\widehat{c}_i)\}_{i=1}^2$ be a solution to (3). We say that firm i is *excluded* if $Q_i^*(0) = Q_i^*(c) = 0$.

To summarize, the timing of the game is the following:

1. The upstream firm M posts a contract for each downstream firm i.

2. Downstream firms order input, commit to pay the corresponding transfer, and produce output.

3. The output market clears, downstream firms' profits are realized, and they pay the transfer to the upstream firm.

The assumption on the timing of the transfer payments is not essential, but we make it to ensure that neither downstream firm observes the action of the competitor. Observing the competitor's action would in fact not violate incentive compatibility in our model, but would violate the participation constraints. Here we have also assumed that downstream firms are not credit-constrained and can pay the agreed transfer even if doing so means negative profits ex-post. We come back to exploring the consequences of limited liability in section 5. With the basic structure of the model in place, we now solve it for various specifications of the utility function U.

3 Exclusion and Attitudes Towards Risk

To begin our analysis of the relationship between risk aversion and exclusion, we consider the extreme cases of risk neutrality and infinite risk aversion. Here we show that under risk neutrality, the upstream firm gains from dealing with both firms, while under infinite risk aversion it prefers to exclude one of the two firms.

3.1 Risk Neutrality

When downstream firms are risk neutral, their utility from facing the lottery induced by the competitor selling in the same market is simply the expected profit, so that

$$U\left[L_i(\widehat{c}_i \mid c_i)\right] = r\pi_i(\widehat{c}_i, 0, c_i) + (1-r)\pi_i(\widehat{c}_i, c, c_i) \tag{4}$$

Using standard arguments, one can show that M's constrained optimization problem in (3) collapses to an unconstrained problem in which it chooses quantities to maximize expected revenue net of production costs and information rents. These last two quantities are linear in aggregate quantities due to the constant returns to scale assumption, so are irrelevant for exclusion.

To understand the effect of the distribution of total contracted output between the two firms, it is useful to rewrite the contract variables. Let $Q^H = Q_1(c) + Q_2(c)$ be the total production of high cost firms; let $\Delta_i = Q_i(0) - Q_i(c)$ be the difference between the quantity produced by the low and high cost types of downstream firm i; and let $\Delta = \Delta_1 + \Delta_2$. One can easily establish that $\Delta_i \geq 0$ is a necessary condition for incentive compatibility, meaning that meeting a low cost competitor is worse for profits that meeting a high cost one. When these variables carry asterisk superscripts, they should be understood to represent optimal values.

The value of two firms is established by the result that

Proposition 1 *In the optimal contracts,* $\Delta^* > 0$ *and* $\Delta_1^* = \Delta_2^* = \frac{\Delta^*}{2}$.

which immediately implies that

Corollary 1 *Under the optimal contracts, neither firm is excluded.*

We refer to M's *revenue* from a contract as the total amount of money it can ask from downstream firms net of expected production costs and information rent. In the case of risk neutrality, M's revenue is exactly equal to the expected revenue that downstream firms will earn in the market, or

$$\begin{aligned} & r^2\left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) + (1-r)^2 Q^H P\left(Q^H\right) + \\ & r(1-r)\left[\left(Q^H + \Delta_1\right) P\left(Q^H + \Delta_1\right) + \left(Q^H + \Delta - \Delta_1\right) P\left(Q^H + \Delta - \Delta_1\right)\right] .. \end{aligned} \tag{5}$$

This only depends on the distribution of output between the two firms via the term in square brackets, which is the total revenue conditional on there being one high cost and one low cost firm downstream. So, for a fixed level of Δ, the optimal value of Δ_1 is defined by

$$\mathrm{MR}\left(Q^H + \Delta_1^*\right) = \mathrm{MR}\left(Q^H + \Delta - \Delta_1^*\right). \tag{6}$$

In other words, the optimal allocation of Δ across firms 1 and 2 equates the marginal revenue from increasing Δ_i for each firm. To see why, suppose we begin with a situation in which $\Delta_1 = 0$. Increasing Δ_1 has two offsetting effects. First, it increases total revenue when firm 1 is low cost and firm 2 is high cost. At the same time, it decreases total revenue when the opposite is true. But, because marginal revenue is decreasing in aggregate quantity, the first effect dominates the second. One can use similar arguments for any allocation of the aggregate low cost output gap for which $\Delta_1 \neq \Delta_2$: there is always a possibility to increase expected revenue by distributing Δ more evenly.

To see the connection between two firms and revenue uncertainty more clearly, the linear demand case is helpful. In this situation, aggregate revenue is given by $\mathbb{E}[\,Q(1-Q)\,] = \mathbb{E}[\,Q\,] - \mathbb{E}[\,Q\,]^2 - V[\,Q\,]$. So distributing output between the two firms should be done to decrease the variance in aggregate output. This is not because M is risk averse (in fact, it is risk-neutral), but because aggregate revenue is concave in aggregate output. Now, firm i's output in an incentive compatible contract is the random variable $Q_i = Q_i(c) + \widetilde{x}_i\Delta_i$ where $\widetilde{x}_i$ is a Bernoulli random variable with mean r and variance $r(1-r)$. So $V[\,Q\,] = r(1-r)\sum_i \Delta_i^2$, which is clearly minimized by equating Δ_i across firms. In other words, equal low cost output gaps maximally reduce aggregate output volatility, which makes M better off.

Proposition 1 pins down the distribution of Δ across downstream firms, but not that of Q^H. As long as $\Delta_1 = \Delta_2$ holds, any split of Q^H between the firms is optimal. In particular, asymmetric contracts are optimal. The important point is that both firms have a positive probability of producing under the optimal contract menus in order to smooth output. The model breaks the indeterminacy of the optimal number of firms in complete information models of vertical markets in which the upstream firm can fully commit to contracts. In such models there is one known monopoly quantity that maximizes the upstream firm's profits, and it can be distributed arbitrarily among any number of downstream firms. Here the upstream firm is uncertain about the optimal quantity: it can either be high or low depending on the realizations of the downstream firms' cost types. Having two firms in the market helps it "hedge its bets" by making sure that when one of the two firms is the low cost type it gets a piece of the market.

3.2 Infinite Risk Aversion

If serving two firms is useful for the upstream firm because of a reduction in the uncertainty about the aggregate output level, the opposite is true for the downstream firms. If a downstream firm knows that it alone produces, it knows for certain what will be its profits conditional on producing. On the other hand, when it knows that the other firm is offered a menu with different output levels for high and low cost type realizations, its profit conditional on producing is uncertain. In the case of risk neutrality, this has no effect since downstream firms are happy to pay a transfer equal to expected profit in order to enter the market. In reality, however, one might imagine that downstream firms have some aversion to the uncertainty that competition creates. To begin the analysis of how this affects the optimal menus, we make the extreme assumption—which we later relax—that downstream firms are *infinitely* risk averse in the sense that their utility from a lottery is its worst realization:

$$U\left[L_i(\widehat{c}_i \mid c_i)\right] = \min\{\pi_i(\widehat{c}_i, 0, c_i), \pi_i(\widehat{c}_i, c, c_i)\} \tag{7}$$

The implication for the upstream firm's profits of infinite risk aversion is that downstream firms' certainty equivalent income of participating in the market no longer equals their expected income, so M must pay a risk premia. This has important implications for exclusion.

Proposition 2 *There exists an $r^* \in (0,1)$ such that for $i, j = 1, 2$, $i \neq j$:*

1. *Whenever $r < r^*$ the optimal contracts are such that $Q_i^*(0) > 0$, $Q_i^*(c) > 0$, and $Q_j^*(0) = Q_j^*(c) = 0$.*
2. *Whenever $r \geq r^*$ the optimal contracts are such that $Q_i^*(c) = Q_j^*(c) = 0$ and $\Delta_1^* + \Delta_2^* = \Delta^* > 0$.*

Which immediately implies that

Corollary 2 *Under infinite downstream risk aversion, exclusion of one downstream firm always solves M's profit maximization problem, and whenever $r < r^*$, exclusion is the only solution to its problem.*

This result contains a basic message of the paper. Even if the upstream firm can fully commit to contracts, it may simply be too costly to include both firms in the downstream market if they are sufficiently risk averse. Rather than pay each one a risk premium, the upstream firm avoids paying any premium by choosing an exclusive contract.

To see this more formally, M's revenue is now

$$\begin{aligned}
& r\left[\left(Q^H+\Delta_1\right) P\left(Q^H+\Delta\right)+\left(Q^H+\Delta_2\right) P\left(Q^H+\Delta\right)\right]+ \\
& (1-r)\left[Q_1(c) P\left(Q^H+\Delta_2\right)+Q_2(c) P\left(Q^H+\Delta_1\right)\right] \\
= \; & r\left(Q^H+\Delta\right) P\left(Q^H+\Delta\right)+(1-r)\left[Q_1(c) P\left(Q^H+\Delta_2\right)+Q_2(c) P\left(Q^H+\Delta_1\right)\right]. \qquad (8)
\end{aligned}$$

The difference between (8) and (5) comes from two sources. First, since downstream firms' certainty equivalent income is equal to the worst profit realization, two firms can no longer smooth output. Second, the presence of a competitor increases uncertainty, which must be compensated. To see why exclusion is optimal, note that (8) has an upper bound of

$$r\left(Q^H+\Delta\right) P\left(Q^H+\Delta\right)+(1-r) Q^H P\left(Q^H\right), \qquad (9)$$

which is exactly the profit level achieved through an exclusive contract in which some firm i is offered the contract $Q_i(c)=\Delta_i=0$. Such contracts are *strictly* optimal whenever M wishes to contract a positive amount of high cost production, which is the case if and only if r is sufficiently low; otherwise, $Q^{H*}=0$ and any split of Δ^* across the two downstream firms is optimal.[10]

While the case of infinite risk aversion provides a clear illustration of the main idea of the paper, it is admittedly an extreme case. The next section explores the optimal contracts with intermediate risk aversion to generate a more subtle understanding of risk and exclusion.

4 Intermediate Risk Preferences

The standard way of modelling risk preferences is to represent utility over lotteries in terms of the expected value of some utility of wealth function u. Such a formulation presents major challenges in our framework. Perhaps most seriously, it is unclear whether the resulting expected utility $U\left[L_i(\widehat{c}_i \mid c_i)\right]$ satisfies the single-crossing condition[11] that greatly reduces the complexity of M's problem. Also, when firms' payoffs are non-linear in profits, obtaining closed-form solutions for optimal contracts is not straightforward.

This section presents two formulations of intermediate risk preferences that overcome

[10] At this point, one might ask why the upstream firm chooses to contract just one firm to produce output when reducing Δ_i would also decrease uncertainty. In fact, as $\Delta_i \to 0$, the upstream firm also pays no risk premia to downstream firms. This is a poor alternative because when Δ_i is small, M does not take advantage of the greater productivity of low cost firms. It keeps both downstream firms in the market and minimizes risk premia, but sacrifices production efficiency. By contrast, when contracting with just one firm, it minimizes risk premia but also takes advantage of the productivity gains that arise from contracting the low cost firm to produce more than the high cost firm.

[11] The condition is that $\left(\frac{\partial^2 U}{\partial T_i \partial Q_i}\right)$ is monotonic in c_i.

these problems. First, we examine the case in which downstream firms have CARA preferences; next, we examine the case in which they have rank-dependent utility.

4.1 CARA preferences

We now assume firms' preferences are given by

$$U\left[L_i(\widehat{c}_i \mid c_i)\right] = -r\exp\left[-a\pi_i(\widehat{c}_i, 0, c_i)\right] - (1-r)\exp\left[-a\pi_i(\widehat{c}_i, c, c_i)\right] \tag{10}$$

where a represents the constant level of absolute risk aversion. CARA utility has the property that the certainty equivalent income from a lottery is independent of initial wealth, which implies that $U\left[L_i(\widehat{c}_i \mid c_i)\right]$ satisfies single-crossing.

We say that an optimal contract satisfies ε-exclusion if there exists some firm i for which $\max\{Q_i^*(0), Q_i^*(c)\} < \varepsilon$. The next result shows that such "nearly" exclusive contracts can be optimal with sufficient risk aversion.

Proposition 3 *Suppose that $Q^{H*} > 0$ under infinite risk aversion. Then for every $\varepsilon > 0$, there exists an $\overline{a}$ such that ε-exclusion is optimal for all $a \geq \overline{a}$.*

As a grows large, downstream firms become close to infinitely risk averse. Since M's objective function is continuous in a, the solution to its problem when a is large is close to the solution to its problem when downstream firms are indeed infinitely risk averse.

The importance of this result is to show that the optimality of exclusion does not depend on the knife-edge assumption of infinite risk aversion. One can observe contracts arbitrarily close to exclusive contracts with less extreme preferences. At the same time, proposition 3 is a limit result, and as such does not characterize optimal contracts for a wide range of intermediate cases. Moreover, the curvature introduced by CARA utility makes such a characterization challenging, so we now turn to an alternative framework.

4.2 Rank-dependent utility

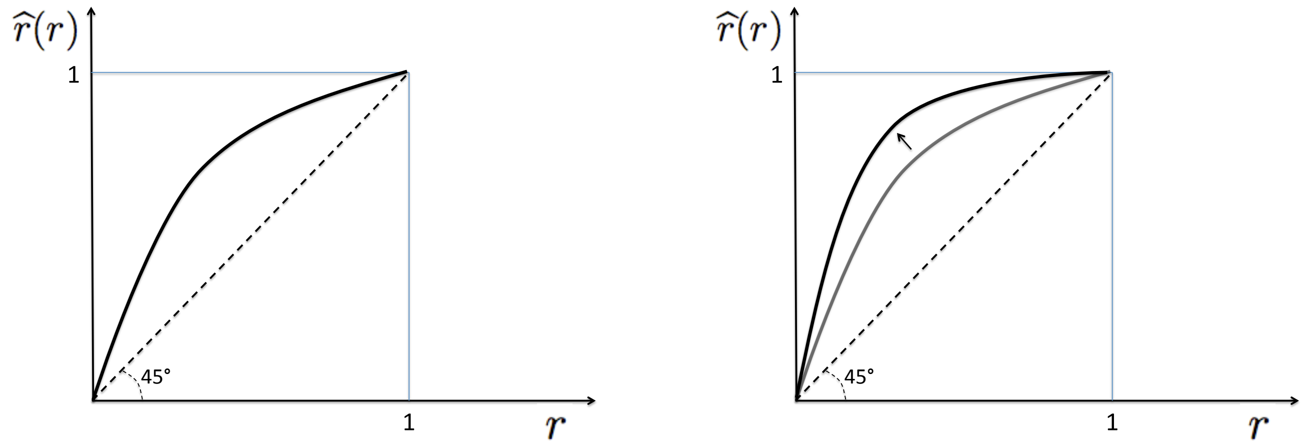
The rank-dependent utility literature represents preferences over lotteries both in terms of the probability weights attached to wealth outcomes and the utility of wealth function.[12] When utility is linear in wealth, the probability weights alone embody an individual's attitudes towards risk. We use this framework to model downstream firms' utility functions,

[12] See Quiggin (1982) for a seminal reference.

as suggested by Yaari (1987).[13] More specifically we assume that

$$U\left[L_i(\widehat{c}_i \mid c_i)\right] = \widehat{r}(r)\pi_i(\widehat{c}_i, 0, c_i) + \left[1 - \widehat{r}(r)\right]\pi_i(\widehat{c}_i, c, c_i) \tag{11}$$

where $\widehat{r}(r)$ is a *weighting function* defined on $r \in (0,1)$ with the following properties: (1) $\widehat{r}(r) > r$; (2) $\widehat{r}'(r) > 0$; (3) $\lim_{r\to 0}\widehat{r}(r) = 0$; and (4) $\lim_{r\to 1}\widehat{r}(r) = 1$.[14] Figure 1a plots an example of a weighting function.



(a) Weighting Function

(b) Increase in Risk Aversion

Figure 1: Examples of Weighting Functions

This formulation of risk aversion nests the two extreme cases from section 3. Risk neutrality is obtained as $\widehat{r}(r)$ approaches r and infinite risk aversion is obtained as $\widehat{r}(r)$ approaches 1. Essentially the weighting function captures pessimism since the weight attached to the worse profit realization—the one corresponding to meeting a low cost competitor—is higher than the probability of meeting such a firm. To see that this

[13] "In studying the behavior of firms, linearity in payments may in fact be an appealing feature. Under the dual theory, maximization of a linear function of profits can be entertained simultaneously with risk aversion. How often has the desire to retain profit maximization led to contrived arguments about firms' risk neutrality?" (Yaari (1987), page 96).

[14] The general formulation of rank-dependent utility is the following. Let n wealth outcomes be ordered so that $w_1 \geq w_2 \geq \ldots \geq w_n$ where p_i is the probability of w_i being realized. Then utility is

$$\sum_{i=2}^{n}\left[\alpha\left(\sum_{j=1}^{i} p_j\right) - \alpha\left(\sum_{j=1}^{i-1} p_j\right)\right] u(w_i) + \alpha(p_1)u(w_1)$$

where α is a weighting function such that $\alpha(0) = 0$ and $\alpha(1) = 1$. In our formulation $u(w_i) = w_i$, $w_1 = \pi_i(\widehat{c}_i, c, c_i)$, $w_2 = \pi_i(\widehat{c}_i, 0, c_i)$, and $p_1 = 1 - r$. So $\alpha(p_1) = \alpha(1-r)$ describes preferences, which we have represented by the mirror function $\widehat{r}(r) = 1 - \alpha(1-r)$.

notion is closely related to risk aversion, one can rewrite $U\left[L_i(\widehat{c}_i \mid c_i)\right]$ in (11) as

$$r\pi_i(\widehat{c}_i, 0, c_i) + [1-r]\,\pi_i(\widehat{c}_i, c, c_i) - [\widehat{r}(r) - r]\,[\pi_i(\widehat{c}_i, c, c_i) - \pi_i(\widehat{c}_i, 0, c_i)]\,, \qquad (12)$$

or the expected value of the lottery minus $\widehat{r}(r) - r$ times the distance between the two wealth outcomes. This is similar to the standard approach of endowing firms with mean-variance preferences, but replaces variance with an alternative measure of outcome dispersion. This alternative representation also makes clear that $\widehat{r}(r) - r$ measures risk aversion. Suppose some firm x has a weighting function $\widehat{r}_x(r)$ such that $\widehat{r}_x(r) - r = \chi_x(r)$ and some firm y has a weighting function $\widehat{r}_y(r)$ such that $\widehat{r}_y(r) - r = \chi_y(r) < \chi_x(r)\ \forall r \in (0,1)$. Then whenever firm x accepts an uncertain bet, firm y does too. Thus one can model an increase in risk aversion as an increase in $\widehat{r}(r)$ for all values of r, as shown in figure 1b. We now use this framework to derive optimal contracts.[15]

4.2.1 Optimal contracts

With CARA preferences, we showed that sufficient risk aversion led to optimal contracts that were at least very close to exclusive. In the rank-dependent framework, we go further by showing that sufficient risk aversion leads to full exclusion. At the same time, optimal contracts are asymmetric on the full range of intermediate cases. In this section, we normalize firm 1 to be the one for which $\Delta_1 \geq \Delta_2$.

Proposition 4 *There exist values c, r^*, and $\widehat{r}^*$ such that, whenever $c < c^*$ and $r < r^*$:*

1. $(Q_1^{H*}, \Delta_1^*) = (Q^{H*}, \Delta^*)$ *uniquely optimal whenever* $\widehat{r} > \widehat{r}^*$
2. $Q_1^{H*} = Q^{H*}$ *and* $\Delta_2^* < \frac{\Delta^*}{2}$ *is uniquely optimal whenever* $\widehat{r} \leq \widehat{r}^*$.

In the appendix we show that M's revenue with rank-dependent utility is given by

$$\begin{aligned}
&r\widehat{r}(Q^H + \Delta)P(Q^H + \Delta) + (1-r)(1-\widehat{r})Q^H P(Q^H) + \\
&r(1-\widehat{r})\left[(Q^H + \Delta_1)P(Q^H + \Delta_1) + (Q^H + \Delta_2)P(Q^H + \Delta_2)\right] + \\
&(\widehat{r} - r)\left[Q_1(c)P(Q^H + \Delta_2) + Q_2(c)P(Q^H + \Delta_1)\right] \qquad (13)
\end{aligned}$$

The two relevant quantities for distributing output across the two firms with intermediate risk aversion emerged in the previous section with extreme preferences. The term in the second line of (13) is also found in (5), and reflects an incentive to smooth output by splitting Δ even across downstream firms. The term in the third line is found in (8),

[15] Another interpretation of the RDU framework with two cost types is simply that the upstream firm and downstream firms have different subjective beliefs about the probability that each firm will meet a low-cost competitor. We rely specifically on the RDU structure in appendix B when we must compute downstream firms' utility over more than two wealth outcomes.

and reflects an incentive to minimize the payment of risk premia by concentrating all output on one downstream firm. Moreover, the degree of risk aversion affects the relative importance of each term. When $\widehat{r}$ increases, the incentive to concentrate output becomes more important, while that to smooth output less so. This tension resolves itself by fully excluding one firm when $\widehat{r}$ is sufficiently high. But, even for low values of $\widehat{r}$, the term in the third line induces an symmetry. First, all high cost output is undertaken by firm 1. Second, more of the additional output produced by low cost firms is given to firm 1 than to firm 2. One can interpret this as "partial exclusion" of firm 2. The point we wish to highlight is that a small amount of risk aversion is sufficient to generate starkly asymmetric contracts even though firms are ex ante identical.

The conditions that r and c take low values are imposed in order to ensure that the upstream firm wishes to contract high cost firms to produce positive output, which will be the case when the probability of meeting high cost firms is high (so that r is low) and when high cost firms are relatively efficient. When no high cost output is contracted, the optimal contract splits Δ evenly across firms.

In order to get finer predictions from the model, one can use the linear demand case of $P(Q) = 1 - Q$.

Proposition 5 *With linear demand, whenever* $r < \min\left\{1 - c, \frac{2+\widehat{r}-2(1+\widehat{r})c}{3}\right\}$, $Q_1^{H*} = Q^{H*} > 0$ *and:*

1. $\Delta_2^* = 0$ *whenever* $\widehat{r} > \widehat{r}^* = \frac{r(1+c-r)}{1-c-r+2cr}$
2. $0 < \Delta_2^* < \frac{\Delta^*}{2}$ *whenever* $\widehat{r} \leq \widehat{r}^*$
3. $\frac{\Delta_2^*}{\Delta^*}$ *is declining in* $\widehat{r}$ *on* $\widehat{r} \in (r, \widehat{r}^*)$

The condition on r contained in the result is again made to ensure that positive high cost output is contracted. Conditional on Q^H being positive, exclusive contracts are optimal *if and only if* firms' risk aversion surpasses a certain threshold, whereas proposition 5 shows the weaker result that exclusive contracts arise when firms are sufficiently risk averse. This gives the model some empirical content—at least in the linear demand case. It predicts that in markets in which firms are more risk averse (e.g. because they are smaller, have less access to capital markets, etc.) exclusion is more likely to be observed.

The solution to the linear case also displays another kind of monotonicity. We already know from the general demand case that firm 1 produces all of the high cost output, but this result shows that, in addition, the production share of the aggregate low cost output gap of firm 2 is declining when risk aversion increases. Hence a more refined empirical prediction of the model is that the difference in average output between two risk averse downstream firms is increasing in their risk aversion.

Proposition 5 also provides insights as to how exclusion depends on the heterogeneity of downstream firms as measured by c. One can easily show that $\frac{\partial \widehat{r}^*}{\partial c} > 0$. This means that when heterogeneity increases, the threshold for exclusion becomes higher.

Corollary 3 *In the linear demand case, the probability of exclusion is decreasing in the heterogeneity of downstream firms.*

The model also delivers another surprising result when c takes on low values: $\lim_{c \to 0} \widehat{r}^* = r$; in words, when heterogeneity between downstream firms declines sufficiently far, almost *any* level of risk aversion is sufficient to guarantee the optimality of exclusive contracts. This further reinforces the point that the optimality of exclusive contracts arises in a potentially wide range of circumstances. At least a partial intuition for corollary 3 arises from the trade-offs identified in the general demand case. When c is lower and downstream firms are more homogenous, the gains from contracting a larger Δ are reduced. At the same time, when Δ is lower, the gains from spreading it across firms is lower, so exclusion becomes more likely.

5 Exclusion and Limited Liability

Risk aversion keep the upstream firm from being able to fully extract expected downstream profits when firms compete. This section explores another related friction that keeps the upstream firm from being able to fully extract downstream profits when contracting with two firms: wealth constraints.

In the baseline case with risk neutrality explored in section 3.1, downstream firms with high cost types were held to their participation constraints. This implies that high-cost downstream firms lose money when meeting a low cost competitor. In reality downstream firms may not be able to suffer losses, either because they have insufficient funds and have limited access to credit markets, or because of limited liability laws. In this case, contracting two firms forces the upstream firm to leave a rent to downstream firms which again keeps it from fully exploiting the gains of multiple firms.

We modify the baseline model and suppose that the upstream firm solves the problem

$$\max_{\{Q_i(\widehat{c}_i),T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^{2} rT_i(0) + (1-r)T_i(c) \text{ such that} \tag{14}$$

$$(Q_i(c) + \Delta_i)\, P\left[Q^H + \Delta\right] - T_i(0) \geq 0 \qquad (LL_i^{LL})$$

$$(Q_i(c) + \Delta_i)\, P\left[Q^H + \Delta_i\right] - T_i(0) \geq 0 \qquad (LL_i^{LH})$$

$$Q_i(c)\left(P\left[Q^H + \Delta_j\right] - c\right) - T_i(c) \geq 0 \qquad (LL_i^{HL})$$

$$Q_i(c)\left(P\left[Q^H\right] - c\right) - T_i(c) \geq 0 \qquad (LL_i^{HH})$$

$$(Q_i(c) + \Delta_i)\left\{rP\left[Q^H + \Delta\right] + (1-r)P\left[Q^H + \Delta_i\right]\right\} - T_i(0) \geq$$
$$Q_i(c)\left\{rP\left[Q^H + \Delta_j\right] + (1-r)P\left[Q^H\right]\right\} - T_i(c) \qquad (IC_i^L)$$

$$Q_i(c)\left\{rP\left[Q^H + \Delta_j\right] + (1-r)P\left[Q^H\right] - c\right\} - T_i(c) \geq$$
$$(Q_i(c) + \Delta_i)\left\{rP\left[Q^H + \Delta\right] + (1-r)P\left[Q^H + \Delta_j\right] - c\right\} - T_i(0). \qquad (IC_i^H)$$

Here the participation constraints in (3) have been replaced by limited liability constraints that ensure that downstream firms never lose money in the market. We are interested in conditions under which the cost of contracting two downstream firms wipes out the gains from output smoothing. The following provides an answer.[16]

Proposition 6 *There exists an r^* such that for all $r < r^*$ only exclusive contracts are optimal with limited liability.*

We find a remarkable similarity between the case of infinite risk aversion and limited liability. In both cases, if the probability of finding an efficient firm is sufficiently low, exclusive contracts alone are optimal. Hence the competition externalities that lie at the core of our theoretical setup map into optimal contracts in a similar way across two different downstream environments.[17] Moreover, since wealth constraints are sometimes relevant when risk aversion is not, the result also expands the set of situations in which our model would predict the emergence of the optimality of exclusion.

Proposition 6 is also important for another theoretical reason. As mentioned in the introduction, a more general contracting setup would allow the upstream firm to auction off the downstream market to the most efficient downstream firm. In previous sections we

[16]The strategy of the proof is to ignore all constraints except IC_i^L and LL_i^{HL} and show that the only global maxima of the relaxed problem are exclusive contracts when r is sufficiently low. The firm's objective function takes essentially the same form as in the rank-dependent utility model, and the analysis proceeds in a similar way.

[17]The intuition for the condition on r begin sufficiently low is richer with limited liability than infinite risk aversion. When r is low the probability of meeting a low cost firm is low. But, since the profit of meeting a low cost firm must be zero, the upstream firm pays out a rent with high probability when r is low. When r is low enough these rents payments wipe out the gain from contracting two firms.

have both restricted this possibility *and* introduced risk aversion, and the contribution of each feature to generating exclusion is not entirely clear. With limited liability this is not the case. Since firms have independent private values in the model, we suspect a wide range of auction formats (for example, a second price auction) could be used to allocate the downstream market optimally without imposing losses on any auction participant. It is the combination of limited liability and restricted contracts that generates exclusion.

6 Extensions

We use this section to explore three extensions of the model. We consider situations in which the upstream firm cannot price discriminate between the downstream firms; there are more than two firms and cost types; and the upstream firm can choose whether to insure downstream firms (or, alternatively in the limited liability case, provide liquidity) prior to offering optimal contracts. In each extension, we continue to find conditions in which exclusive contracts are optimal.

6.1 Uniform contracts

In this section we relax the assumption that the principal is able to discriminate between the two agents.[18] When the upstream firm can make discriminatory offers, we know that under infinite risk aversion (and as long as r is sufficiently small) exclusion of one of the two downstream firms is uniquely optimal. The question is whether it is able to replicate this outcome when it is restricted to offering uniform contracts. We show here that the answer is affirmative under certain (possibly very mild) conditions.

More precisely, the principal would like to implement the following equilibrium outcome:

1. She offers both agents the same menu of contracts $\{(T(0), Q(0)), (T(c), Q(c)), (0, 0)\}$, with $Q(0) = Q^{H*} + \Delta^*$, $Q(c) = Q^{H*}$, $T(0) = (Q^{H*} + \Delta^*)P(Q^* + \Delta^*) - cQ^{H*}$, $T(c) = Q^{H*}P(Q^{H*}) - cQ^{H*}$, where Q^{H*} and Δ^* are the optimal values found for the case where the principal offers and exclusive contracts.;[19]

2. One agent (say agent 1) chooses $(T(0), Q(0))$ if low cost and $(T(c), Q(c))$ if high cost; and the other agent (say agent 2) chooses $(0, 0)$ for both cost types.

[18] This is an important question, since we know that in some jurisdictions there may be legal constraints to price discrimination among retailers (think of the Robinson-Patman Act in the US), and that in any case the vast majority of franchisors offer uniform contract conditions to their franchisees (see e.g. Blair and Lafontaine, 2011 pp. 54-56).

[19] These are the values which solve: $\text{MR}\left(Q^{H*} + \Delta^*\right) = 0$ and $(1 - r)\text{MR}\left(Q^{H*}\right) = c$.

For this to be the equilibrium outcome, it is necessary for each player to find the candidate strategy to be optimal given the choices of the other. In the case of the principal, optimality follows immediately from the fact that this contract reproduces the optimal outcome obtained under less restrictive assumptions on her strategies (she cannot achieve higher profits when she is obliged to set uniform contract than when she can discriminate).

In the case of agent 1, the choice is also optimal given that the other agent chooses not to participate in the market: since $(0,0)$ effectively amounts to exclusion, we know that the participation and incentive constraints are all satisfied when agent 2 doe snot sell anything).

Therefore, the only thing we should check is whether firm 2 prefers not to participate (i.e., chooses $(0,0)$) given that firm 1 chooses $(T(0),Q(0))$ if low cost and $(T(c),Q(c))$ if high cost. Under the assumption of infinite risk aversion, requiring that the choice of agent 2 is optimal amounts to verifying that the following incentive constraints are satisfied:

$$
\begin{aligned}
&0 \geq Q^{H*}P\left(2Q^{H*}+\Delta^*\right) - cQ^{H*} - T(c) && (IC_{HH})\\
&0 \geq \left(Q^{H*}+\Delta^*\right)P\left(2Q^{H*}+2\Delta^*\right) - c\left(Q^{H*}+\Delta^*\right) - T(0) && (IC_{HL})\\
&0 \geq Q^{H*}P\left(2Q^{H*}+\Delta^*\right) - T(c) && (IC_{LH})\\
&0 \geq \left(Q^{H*}+\Delta^*\right)P\left(2Q^{H*}+2\Delta^*\right) - T(0), && (IC_{LL})
\end{aligned}
$$

where IC_{HH} and IC_{HL} refer to the possible deviation of a high cost agent 2 (respectively, he might want to pick the contract $(T(c),Q(c))$ designed for a high cost firm or the contract $(T(0),Q(0))$ designed for a low cost firm); and similarly IC_{LH} and IC_{LL} refer to the possible deviation of a low cost agent 2. After substitution, the four ICs become:

$$
\begin{aligned}
&0 \geq P\left(2Q^{H*}+\Delta^*\right) - P(Q^{H*}) && (IC_{HH})\\
&0 \geq P\left(2Q^{H*}+2\Delta^*\right) - P\left(Q^*+\Delta^*\right) && (IC_{HL})\\
&0 \geq P\left(2Q^{H*}+\Delta^*\right) - P(Q^{H*}) + c && (IC_{LH})\\
&0 \geq \left(Q^{H*}+\Delta^*\right)\left(P\left(2Q^{H*}+2\Delta^*\right) - P\left(Q^{H*}+\Delta^*\right)\right) + cQ^{H*}. && (IC_{LL})
\end{aligned}
$$

It is immediate to notice that the first two constraints are always satisfied, because the demand function is assumed to be decreasing. We are left with the last two ICs. We can state the following proposition.

Proposition 7 *There exists a $\bar{c} > 0$ such that for $c \leq \bar{c}$ the principal is able to implement the (optimal) exclusionary outcome by making use of non-discriminatory contracts.*

This Proposition gives a *sufficient* condition for the uniform offers to implement the same exclusionary outcome as discriminatory offers. It is important to note that this condition might be very weak. For instance, if we assumed a linear demand function $P = 1 - Q$, we know that the exclusionary contract under discrimination is optimal for $r < (1-c)/(1+c)$ and that $Q^{H*} = \frac{1-c-r}{2(1-r)}; \Delta^* = \frac{c}{2(1-r)}$. By replacing these values into IC_{LH} and IC_{LL} one can check that $\widehat{c} = 1/2$: the sufficient condition for these constraints to be satisfied is $c < 1/2$, which is exactly the competition condition we assumed in the baseline model.

6.2 Multiple firms and types

Throughout the paper we have maintained a simple two by two setup for expositional clarity. We now adopt a more general environment in which there are $F \in \mathbb{N}$ firms each with $N \in \mathbb{N}$ possible cost types. We maintain symmetry by assuming that the distribution of cost types is iid, and let r_j be the probability that firm i has cost type c_j. We adopt the convention that $c_j > c_{j+1}$, assume that $c_N \geq 0$, and that $P(0) > c_1$.

Proposition 8 *There exists an r_N^* such that, when all downstream firms are infinitely risk averse and $r_N < r_N^*$, only exclusive contracts are optimal.*

This result says that as long as the probability of firms' being the most efficient type is not too large, only exclusive contracts solve M's problem. The condition placed on r_N ensures that the upstream firm wants at least the second most efficient type to produce output. Thus the optimality of exclusive contracts does not rely on the least efficient type producing, merely *any* type that is not the most efficient producing. We view this as a weak restriction. For example, if one assumed that the cost distribution were uniform and the values of c_1 and c_N fixed, the condition would be satisfied for high enough N. Moreover, by employing the same logic as in the proof of proposition 3, one can show that ε-exclusion is optimal when downstream firms have CARA utility, $r_N < r_N^*$, and the level of absolute risk aversion is sufficiently high.

6.3 Fixed wage contracts

We have argued that competition is a source of uncertainty for downstream firms, and that this leads the upstream firm to offer an exclusive contract and sacrifice the benefits of two firms if risk aversion is sufficiently high. A seemingly straightforward solution is for the upstream firm to pay downstream firms to produce output, but then itself collect the revenue from selling the output. This arrangement insures downstream firms against profit volatility, and allows the upstream firm to replicate the payoff it obtains from contracting with two risk neutral firms. Since fairly simple contracts appear to address the tension in our model, one might question the relevance of our story.

In appendix B we extend our baseline model to allow the upstream firm to choose whether or not to offer each downstream a contract in which the upstream firm collects the sales revenue or whether to let the downstream firms collect the revenue as in the baseline case. Since some definitions of ownership are formulated in terms of who collects the residual income from an asset (in this case the output to be sold), we interpret the former as the upstream firm's "vertically integrating" with the downstream firm. We then solve for the organizational arrangement that maximizes the joint ex-ante expected utility of all three firms.

Two key results emerge. First, unsurprisingly, the upstream firms is always better off when integrated with both firms since it no longer has to pay any risk premium. More surprising is that joint downstream revenue is lower when both firms are integrated and contracted to produce compared to when both are independent and only one offered an exclusive contract. The reason is that aggregate information rents decrease with an additional firm. This allows us to construct an example in which the organization that maximizes joint ex-ante utility leaves both firms independent with one excluded.[20] Thus our insights are robust to an environment in which the upstream firm could choose insurance to avoid exclusion if it so wants. If one instead interpreted vertical integration as a means of granting wealth-constrained downstream firms access to internal capital markets, this example would also show that maintaining such constraints and excluding can be preferable to providing downstream access to funding.

7 Summary and conclusions

This paper identifies a new rationale for using exclusivity provisions: when agents compete downstream, and do not observe each other's cost type, competition generates uncertainty, leading risk-averse or wealth-constrained agents to require a rent as a compensation for facing a more efficient rival. To save the payment of such rents, the principal may prefer to deal exclusively with one agent.

Beyond the specific model used here, we believe that this mechanism offers a general reason why a principal may endogenously restrict the number of agents with whom it wants to deal. Whenever the payoff of one agent depends on the actions or the type of other agents, and there is imperfect information, if agents are risk averse the principal will be obliged to pay risk premia. To save on such rents, the principal may prefer to contract with a strict subset of the potential agents. We plan to show that the same mechanism still holds good in very different settings, such as for instance models of moral hazard where agents are paid according to relative performance schemes.

[20] More specifically, this is true in the rank-dependent utility framework with linear demand when $\widehat{r}^* = \frac{r(1+c-r)}{1-c-r+2cr}$; is sufficiently small; and r is sufficiently large.

References

ANAND, B. N., AND T. KHANNA (2000): "The Structure of Licensing Contracts," *Journal of Industrial Economics*, 48(1), 103–35.

ASPLUND, M. (2002): "Risk-averse firms in oligopoly," *International Journal of Industrial Organization*, 20(7), 995–1012.

AZOULAY, P., AND S. SHANE (2001): "Entrepreneurs, Contracts, and the Failure of Young Firms," *Management Science*, 47(3), 337–358.

BLAIR, R. D., AND F. LAFONTAINE (2011): *Economics of Franchising*. Cambridge University Press, first paperback edn.

BRICKLEY, J. A., AND F. H. DARK (1987): "The choice of organizational form: the case of franchising," *Journal of Financial Economics*, 18(2), 401–420.

HART, O., AND J. TIROLE (1990): "Vertical Integration and Market Foreclosure," *Brookings Papers on Economic Activity (Microeconomics)*, 1990, 205–286.

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

LAFFONT, J.-J., AND J. TIROLE (1987): "Auctioning Incentive Contracts," *Journal of Political Economy*, 95(5), 921–37.

LAFONTAINE, F. (1992): "How and Why do Franchisors do What They do: A Survey Report," in *Sixth Annual Proceedings of the Society of Franchising*, ed. by P. J. Kaufmann.

LAFONTAINE, F., AND K. L. SHAW (1999): "The Dynamics of Franchise Contracting: Evidence from Panel Data," *Journal of Political Economy*, 107(5), 1041–1080.

MCAFEE, R. P., AND J. MCMILLAN (1986): "Bidding for Contracts: A Principal-Agent Analysis," *RAND Journal of Economics*, 17(3), 326–338.

——— (1987): "Competition for Agency Contracts," *RAND Journal of Economics*, 18(2), 296–307.

NOCKE, V., AND J. THANASSOULIS (2010): "Vertical Relations under Credit Constraints," Discussion Paper 7636, Centre for Economic Policy Research.

QUIGGIN, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3(4), 323–343.

REY, P., AND J. STIGLITZ (1995): "The Role of Exclusive Territories in Producers' Competition," *RAND Journal of Economics*, 26(3), 431–451.

REY, P., AND J. TIROLE (2007): "A Primer on Foreclosure," in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, pp. 2145–2220. North-Holland.

RIORDAN, M. H., AND D. E. M. SAPPINGTON (1987): "Awarding Monopoly Franchises," *American Economic Review*, 77(3), 375–87.

Sundaram, R. K. (1996): *A First Course in Optimization Theory*. Cambridge University Press.

Yaari, M. E. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95–115.

A Omitted Proofs

A.1 Proof of Proposition 1

Proof. Given the representation of utility in (5), one can compute $\frac{\partial^2 U}{\partial c \partial Q_i} = -1$ so that utility satisfies the single-crossing condition. So by standard arguments, (for example, those in Laffont and Martimort (2002)) (1) whenever IC_i^L is binding, IC_i^H is satisfied; (2) $\Delta_i \geq 0$ is necessary and sufficient for incentive compatibility; and (3) whenever PC_i^H and IC_i^L are satisfied, PC_i^L is satisfied. So optimal transfers are set to make IC_i^H and PC_i^L bind, which implies that

$$T_i^*(0) = (Q_i(c) + \Delta_i) \left\{ rP\left[Q^H + \Delta\right] + (1-r)P\left[Q^H + \Delta_i\right] \right\} - cQ_i(c) \tag{A.1}$$

$$T_i^*(c) = Q_i(c) \left\{ rP\left[Q^H + \Delta_j\right] + (1-r)P\left[Q^H\right] - c \right\}. \tag{A.2}$$

Plugging back these into the objective function yields a term equal to equation (5) in the text minus cQ^H.

It remains to be shown that the optimal value of Δ is positive. Suppose not, and let the optimal value of Q^H be $Q^{H\prime}$. The total profit of the upstream firm from this solution is $Q^{H\prime}\left[P\left(Q^{H\prime}\right) - c\right]$. Without loss of generality, this can be achieved through contracts in which $Q_1(c) = Q^{H\prime}$ and $\Delta_1 = \Delta_2 = 0$. Now consider the problem in which the upstream maximizes profit within the restricted class of contracts $Q_1(c) = Q^H$ and $\Delta_1 = \Delta$. The problem becomes

$$\max_{Q^H \geq 0, \Delta \geq 0} r\left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) + (1-r) Q^H P\left(Q^H\right) - cQ^H. \tag{A.3}$$

The optimal values (Q^{H*}, Δ^*) for this problem solve the first order conditions

$$\text{MR}\left(Q^{H*} + \Delta^*\right) \leq 0 \tag{A.4}$$

$$r\text{MR}\left(Q^{H*} + \Delta^*\right) + (1-r)\text{MR}\left(Q^{H*}\right) - c \leq 0 \tag{A.5}$$

where (A.4) holds with equality if $\Delta^* > 0$ and (A.5) holds with equality if $Q^{H*} > 0$. These conditions together imply that $\Delta^* > 0$. Suppose not, and that $Q^{H*} = 0$. Then, from (A.4), it must be the case that MR < 0 which is ruled out by assumption. Suppose not, and that $Q^{H*} > 0$. Then (A.5) gives $\text{MR}\left(Q^{H*}\right) = c > 0$ while (A.4) gives $\text{MR}\left(Q^{H*}\right) < 0$, a contradiction. Now since the contracts $Q^H = Q^{H\prime}$ and $\Delta_1 = \Delta = 0$ are within the set of feasible contracts for (A.3) and are not chosen, we have arrived at a contradiction of their optimality.

Finally, since $\Delta^* > 0$, $\Delta_1^* = \Delta_2^* = \frac{\Delta^*}{2} > 0$, so that the solutions to the unconstrained problem are incentive compatible. ■

A.2 Proof of Proposition 2

Proof. (7) again satisfies the single crossing condition, so that optimal transfers are set to make IC_i^H and PC_i^L bind. Since incentive compatibility implies $\Delta_i \geq 0$, these transfers become

$$\begin{aligned} T_i^*(0) &= Q_i^*(0)P\left[Q_i^*(0)+Q_j^*(0)\right] - cQ_i^*(c) \\ T_i^*(c) &= Q_i^*(c)\left\{P\left[Q_i^*(c)+Q_j^*(0)\right] - c\right\}. \end{aligned}$$

In the text we argued that whenever it is optimal to contract $Q^H > 0$, it is optimal to contract with only one firm. To establish whether $Q^{H*} > 0$, observe again expressions (A.4) and (A.5). We have established that (A.4) always holds with equality. Under this condition (A.5) becomes $(1-r)\text{MR}\left(Q^{H*}\right) - c \leq 0$. In order for the upstream firm to contract high cost output, it must be the case that $r < 1 - \frac{c}{\text{MR}(0)} = r^*$. $r^* > 0$ since by assumption $c < P(0) = \text{MR}(0)$. When $r > r^*$, $Q^{H*} = 0$ and M's problem becomes $\max_{\Delta \geq 0} \Delta P(\Delta)$. Since $\text{MR}(0) = P(0) > 0$ it is optimal to set a positive aggregate low cost output gap in this case. ■

A.3 Proof of Proposition 3

Proof.

Let W be firm i's certainty equivalent income from the lottery $L_i(\widehat{c}_i \mid c_i)$. This is defined by

$$-r\exp\left[-a\pi_i(\widehat{c}_i, 0, c_i)\right] - (1-r)\exp\left[-a\pi_i(\widehat{c}_i, 0, c_i) - aR_i(\widehat{c}_i)\right] = -\exp\left[-aW\right].$$

where

$$R_i(\widehat{c}_i) = Q_i(\widehat{c}_i)\left\{\begin{array}{c} P\left[Q_i(\widehat{c}_i)+Q_j(c)\right] - \\ P\left[Q_i(\widehat{c}_i)+Q_j(0)\right] \end{array}\right\} = \pi_i(\widehat{c}_i, c, c_i) - \pi_i(\widehat{c}_i, 0, c_i).$$

Solving for W gives

$$W = \pi_i(\widehat{c}_i, 0, c_i) - \frac{\ln\left[(1-r)\exp\left[-aR_i(\widehat{c}_i)\right] + r\right]}{a}$$

So, by expressing utilities in terms of certainty equivalent income, the principal's problem in (3) writes as

$$\max_{\{Q_i(\widehat{c}_i), T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^2 rT_i(0) + (1-r)T_i(c) \text{ such that}$$

$$\begin{aligned} &\pi_i(0,0,0) - \frac{\ln\left[(1-r)\exp\left[-aR_i(0)\right]+r\right]}{a} \geq 0 \\ &\pi_i(c,0,c) - \frac{\ln\left[(1-r)\exp\left[-aR_i(c)\right]+r\right]}{a} \geq 0 \\ &\pi_i(0,0,0) - \frac{\ln\left[(1-r)\exp\left[-aR_i(0)\right]+r\right]}{a} \geq \pi_i(c,0,0) - \frac{\ln\left[(1-r)\exp\left[-aR_i(c)\right]+r\right]}{a} \\ &\pi_i(c,0,c) - \frac{\ln\left[(1-r)\exp\left[-aR_i(c)\right]+r\right]}{a} \geq \pi_i(0,0,c) - \frac{\ln\left[(1-r)\exp\left[-aR_i(0)\right]+r\right]}{a}. \end{aligned}$$

Since the R_i terms are independent of cost type, this is again a supermodular problem, so IC_i^H and PC_i^L bind. Solving for $T_i(0)$ and $T_i(c)$, and replacing $\alpha = \frac{1}{a}$ allows one to write the principal's unconstrained objective function as

$$\begin{cases} \begin{array}{c} r\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)-cQ^H+ \\ (1-r)\left[Q_1(c)P\left(Q^H+\Delta_2\right)+Q_2(c)P\left(Q^H+\Delta_1\right)\right]- \\ \sum_i r\alpha\ln\left[(1-r)\exp\left[-\frac{R_i(0)}{\alpha}\right]+r\right]- \\ \sum_i(1-r)\alpha\ln\left[(1-r)\exp\left[-\frac{R_i(c)}{\alpha}\right]+r\right] \end{array} & \text{if } \alpha>0 \\ \\ \begin{array}{c} r\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)-cQ^H+ \\ (1-r)\left[Q_1(c)P\left(Q^H+\Delta_2\right)+Q_2(c)P\left(Q^H+\Delta_1\right)\right] \end{array} & \text{if } \alpha=0. \end{cases} \tag{A.6}$$

Since incentive compatibility requires that $Q_i(0) \geq Q_i(c)$, one can conclude that $R_i(0) \geq 0$ and $R_i(c) \geq 0$ must also hold. So (A.6) is continuous at $\alpha = 0$.

Denote by $S(\alpha) = \left(Q_i^{H*}(\alpha), \Delta_i^*(\alpha)\right)_{i=1}^2 \in \mathbb{R}_+^4$ the solution correspondence to the problem of maximizing (A.6) such that $Q_i(c) \geq 0$ and $\Delta_i \geq 0$ for $i = 1, 2$. We now argue that this correspondence must be bounded for all $\alpha \in [0, \infty)$. Note that $\ln\left[(1-r)\exp\left[-\frac{R_i(c)}{\alpha}\right]+r\right]$ and $\ln\left[(1-r)\exp\left[-\frac{R_i(0)}{\alpha}\right]+r\right]$ are bounded between $\ln[r]$ and 0 and that

$$\begin{aligned} &r\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)-cQ^H+(1-r)\left[Q_1(c)P\left(Q^H+\Delta_2\right)+Q_2(c)P\left(Q^H+\Delta_1\right)\right] \leq \\ &r\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)+(1-r)Q^HP\left(Q^H\right)=rQ^1P\left(Q^1\right)+(1-r)Q^2P\left(Q^2\right) \end{aligned}$$

We claim that $\lim_{Q^k\to\infty} Q^kP\left(Q^k\right) = -\infty$. One can write

$$\begin{aligned} Q^kP\left(Q^k\right) &= \int_0^{\widetilde{Q}+\mu} MR(v)dv + \int_{\widetilde{Q}+\mu}^{Q^k} MR(v)dv \\ &\leq \int_0^{\widetilde{Q}+\mu} MR(v)dv + MR(\widetilde{Q}+\mu)(Q^k-\widetilde{Q}-\mu), \end{aligned}$$

which is unbounded below as $Q^k \to \infty$ since $MR(\widetilde{Q}+\mu) < 0$. Since the upstream firm can guarantee itself a payoff of 0 by offering 0 output to all firms, the solution correspondence must be bounded. So we can rewrite the constraints as $Q_i(c) \in \left[0, \overline{Q}\right]$ and $\Delta_i \in \left[0, \overline{\Delta}\right]$ for some $\overline{Q} \leq \infty$ and $\overline{\Delta} \leq \infty$.

Since (A.6) is continuous and the constraint set is compact-valued, the Maximum Theorem applies (see Sundaram (1996) theorem 9.14 for details), and so $S(\alpha)$ is upper-semicontinuous. Now, by proposition 2, whenever $r < r^*$ $S(0) = \left\{\left(Q^{H*}, \Delta^*, 0, 0\right), \left(0, 0, Q^{H*}, \Delta^*\right)\right\}$. Now define the open set $V \in \mathbb{R}^4$ as

$$V = \begin{pmatrix} (Q^{H*} - \varepsilon, Q^{H*} + \varepsilon) \\ (\Delta^* - \varepsilon, \Delta^* + \varepsilon) \\ (-\varepsilon, \varepsilon) \\ (-\varepsilon, \varepsilon) \end{pmatrix} \cup \begin{pmatrix} (-\varepsilon, \varepsilon) \\ (-\varepsilon, \varepsilon) \\ (Q^{H*} - \varepsilon, Q^{H*} + \varepsilon) \\ (\Delta^* - \varepsilon, \Delta^* + \varepsilon) \end{pmatrix}$$

Clearly $S(0) \subset V$. So by usc there exists some $\alpha' > 0$ such that $S(\alpha) \subset V$ whenever $\alpha \in (0, \alpha')$.
■

A.4 Proof of Proposition 4

Proof. Clearly (11) satisfies supermodularity, so that we can again take PC_i^H and IC_i^L as the binding constraints and the transfers write as

$$T_i^*(0) = Q_i^*(0)\left\{\widehat{r}P\left[Q_i^*(0) + Q_j^*(0)\right] + (1-\widehat{r})P\left[Q_i^*(0) + Q_j^*(c)\right]\right\} - cQ_i^*(c) \quad \text{(A.7)}$$

$$T_i^*(c) = Q_i^*(c)\left\{\widehat{r}P\left[Q_i^*(c) + Q_j^*(0)\right] + (1-\widehat{r})P\left[Q_i^*(c) + Q_j^*(c)\right] - c\right\}. \quad \text{(A.8)}$$

One can then write

$$\begin{aligned} & r[T_1(c) + T_2(0)] + (1-r)[T_1(c) + T_2(c)] = \\ & r\begin{bmatrix} \widehat{r}[Q_1(c) + \Delta_1]P(Q^H + \Delta) + (1-\widehat{r})[Q_1(c) + \Delta_1]P(Q^H + \Delta_1) - cQ_1(c) + \\ \widehat{r}[Q_2(c) + \Delta_2]P(Q^H + \Delta) + (1-\widehat{r})[Q_2(c) + \Delta_2]P(Q^H + \Delta_2) - cQ_2(c) \end{bmatrix} + \\ (1-r) & \begin{bmatrix} \widehat{r}Q_1(c)P(Q^H + \Delta_2) + (1-\widehat{r})Q_1(c)P(Q^H + \Delta_2) - cQ_1(c) + \\ \widehat{r}Q_2(c)P(Q^H + \Delta_1) + (1-\widehat{r})Q_2(c)P(Q^H + \Delta_1) - cQ_2(c). \end{bmatrix} \end{aligned}$$

After removing the term

$$r\widehat{r}(Q^H + \Delta)P(Q^H + \Delta) + (1-r)(1-\widehat{r})Q^H P(Q^H) - cQ^H$$

from this expression one is left with

$$\begin{gathered} r(1-\widehat{r})\left[(Q_1(c) + \Delta_1)P(Q^H + \Delta_1) + (Q_2(c) + \Delta_2)P(Q^H + \Delta_2)\right] + \\ (1-r)\widehat{r}\left[Q_1(c)P(Q^H + \Delta_2) + Q_2(c)P(Q^H + \Delta_1)\right] \end{gathered}$$

which equals

$$\begin{gathered} r(1-\widehat{r})\left[(Q_1(c) + \Delta_1)P(Q^H + \Delta_1) + (Q_2(c) + \Delta_2)P(Q^H + \Delta_2)\right] + \\ r(1-\widehat{r})\left[Q_1(c)P(Q^H + \Delta_2) + Q_2(c)P(Q^H + \Delta_1)\right] + \\ (1-r)\widehat{r}\left[Q_1(c)P(Q^H + \Delta_2) + Q_2(c)P(Q^H + \Delta_1)\right] - \\ r(1-\widehat{r})\left[Q_1(c)P(Q^H + \Delta_2) + Q_2(c)P(Q^H + \Delta_1)\right] \end{gathered}$$

and finally

$$r(1-\widehat{r})\left[(Q^H+\Delta_1)P(Q^H+\Delta_1)+(Q^H+\Delta_2)P(Q^H+\Delta_2)\right]+$$
$$(\widehat{r}-r)\left[Q_1(c)P(Q^H+\Delta_2)+Q_2(c)P(Q^H+\Delta_1)\right].$$

Whenever $\Delta_1 \geq \Delta_2$, $Q_1^H = Q$ and $Q_2^H = 0$ is (weakly) optimal. Substituting these conditions in profits along with the condition $\Delta_1 = \Delta - \Delta_2$ gives

$$\begin{aligned}\pi(Q,\Delta,\Delta_2) = {} & r\widehat{r}R(Q+\Delta)+(1-r)(1-\widehat{r})R(Q)-cQ+ \\ & r(1-\widehat{r})\left[R(Q+\Delta-\Delta_2)+R(Q+\Delta_2)\right]+ \\ & (\widehat{r}-r)QP(Q+\Delta_2).\end{aligned} \tag{A.9}$$

We will solve the relaxed problem

$$\max_{Q\geq 0,\Delta\geq 0,\Delta\geq 0} \pi(Q,\Delta,\Delta_2) \tag{A.10}$$

and ignore the constraint $\Delta_2 \leq \frac{\Delta}{2}$, and show that all solution to the relaxed problem satisfy this ignored constraint.

In general (A.44) is not a concave problem and in general there can be multiple solutions to the Kuhn-Tucker first order conditions, some of which are not global maximum. We can rule out the case where $\Delta = 0$ (and therefore $\Delta_1 = \Delta_2 = 0$) since (A.43) then becomes $QP(Q) - cQ$, which, by the arguments from the proof of proposition 1, can be improved by some $\Delta_1 > 0$. One can also rule out solutions in which $Q = 0$ and $\Delta_2 = 0$. When $Q = 0$ (A.43) becomes

$$r\widehat{r}R(\Delta)+(1-r)(1-\widehat{r})R(0)+r(1-\widehat{r})\left[R(\Delta-\Delta_2)+R(\Delta_2)\right] \tag{A.11}$$

which is maximized at $\Delta_2 = \frac{\Delta}{2}$.

This solution (with $Q = 0$ and $\Delta_2 = \frac{\Delta}{2}$) satisfies the Kuhn-Tucker first order conditions if

$$\frac{\partial\pi(Q,\Delta,\Delta_2)}{\partial Q} = \left(\begin{array}{c} r\widehat{r}\mathrm{MR}(Q+\Delta)+(1-r)(1-\widehat{r})\mathrm{MR}(Q)+ \\ r(1-\widehat{r})\left[\mathrm{MR}(Q+\Delta-\Delta_2)+\mathrm{MR}(Q+\Delta_2)\right]+ \\ (\widehat{r}-r)\left[P(Q+\Delta_2)+QP'(Q+\Delta_2)\right] \end{array}\right) - c < 0 \tag{A.12}$$

$$\frac{\partial\pi(Q,\Delta,\Delta_2)}{\partial \Delta} = r\widehat{r}\mathrm{MR}(Q+\Delta)+r(1-\widehat{r})\mathrm{MR}(Q+\Delta-\Delta_2) = 0 \tag{A.13}$$

$$\frac{\partial\pi(Q,\Delta,\Delta_2)}{\partial \Delta_2} = \left(\begin{array}{c} r(1-\widehat{r})\left[-\mathrm{MR}(Q+\Delta-\Delta_2)+\mathrm{MR}(Q+\Delta_2)\right]+ \\ (\widehat{r}-r)QP'(Q+\Delta_2) \end{array}\right) = 0 \tag{A.14}$$

are all satisfied at $Q = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$. (A.47) is clearly satisfied, while (A.45) and (A.46) rewrite as[21]

$$\widehat{r}\mathrm{MR}\,(\Delta) + (1-\widehat{r})\mathrm{MR}\left(\frac{\Delta}{2}\right) = 0 \tag{A.15}$$

$$(1-r)(1-\widehat{r})\mathrm{MR}\,(0) + r(1-\widehat{r})\mathrm{MR}\left(\frac{\Delta}{2}\right) + (\widehat{r}-r)P\left(\frac{\Delta}{2}\right) < c. \tag{A.16}$$

Call the solution to A.48 $\Delta^*(\widehat{r})$. Plugging into A.49 gives the condition on r

$$r > \frac{(1-\widehat{r})\mathrm{MR}(0) + \widehat{r}P\left(\frac{\Delta^*(\widehat{r})}{2}\right) - c}{(1-\widehat{r})\left[\mathrm{MR}(0) - \mathrm{MR}\left(\frac{\Delta^*(\widehat{r})}{2}\right)\right] + P\left(\frac{\Delta^*(\widehat{r})}{2}\right)} = \widetilde{r}(\widehat{r}, c). \tag{A.17}$$

$\widetilde{r}(\widehat{r}, c) < 1$ since

$$(1-\widehat{r})\mathrm{MR}\left(\frac{\Delta^*(\widehat{r})}{2}\right) + (\widehat{r}-1)P\left(\frac{\Delta^*(\widehat{r})}{2}\right) = (1-\widehat{r})\left(\mathrm{MR}\left(\frac{\Delta^*(\widehat{r})}{2}\right) - P\left(\frac{\Delta^*(\widehat{r})}{2}\right)\right) < 0 < c. \tag{A.18}$$

$\widetilde{r}(\widehat{r}, c) > 0$ if[22]

$$c < (1-\widehat{r})P(0) + \widehat{r}P\left(\frac{\Delta^*(\widehat{r})}{2}\right). \tag{A.19}$$

The right-hand side of this expression is positive since $P\left(\frac{\Delta^*(\widehat{r})}{2}\right) > R\left(\frac{\Delta^*(\widehat{r})}{2}\right) > 0$, where the last inequality follows from (A.48). Now define $c^* \min_{\widehat{r}}(1-\widehat{r})P(0) + \widehat{r}P\left(\frac{\Delta^*(\widehat{r})}{2}\right) \in (0, P'(0))$ and $r' = \min_{\widehat{r}} \widetilde{r}(\widehat{r}, c^*)$. So one can conclude that the solution with $Q = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$ does not exist when $r < r'$ and $c < c^*$.

Another potential solution to (A.44) is an exclusive contract in which $\Delta_2 = 0$. By the above arguments an exclusive contract can only be optimal if $Q > 0$. Such an exclusive contract satisfies the Kuhn-Tucker first order conditions if

$$\left[\frac{\partial\pi(Q,\Delta,\Delta_2)}{\partial Q}\right]_{\Delta_2=0} = 0 \tag{A.20}$$

$$\left[\frac{\partial\pi(Q,\Delta,\Delta_2)}{\partial \Delta}\right]_{\Delta_2=0} = 0 \tag{A.21}$$

$$\left[\frac{\partial\pi(Q,\Delta,\Delta_2)}{\partial \Delta_2}\right]_{\Delta_2=0} = 0. \tag{A.22}$$

which simplifies to

$$r\mathrm{MR}\,(Q+\Delta) + (1-r)\mathrm{MR}\,(Q) = c \tag{A.23}$$

$$r\mathrm{MR}\,(Q+\Delta) = 0 \tag{A.24}$$

$$r(1-\widehat{r})\left[-\mathrm{MR}\,(Q+\Delta) + \mathrm{MR}\,(Q)\right] + (\widehat{r}-r)QP'\,(Q) < 0. \tag{A.25}$$

[21] Here we have also plugged (A.45) into (A.46).

[22] Recall that $\mathrm{MR}(0) = P(0)$.

and further to

$$(1-r)\text{MR}\,(Q) = c \tag{A.26}$$

$$r\text{MR}\,(Q+\Delta) = 0 \tag{A.27}$$

$$r(1-\widehat{r})\text{MR}\,(Q) + (\widehat{r}-r)QP'\,(Q) < 0. \tag{A.28}$$

Let $Q^*(r)$ be the solution to (A.56). Since $\text{MR}(Q^*) = P(Q^*) + Q^*P'(Q^*)$ (A.58) is satisfied whenever

$$\widehat{r} > \frac{rP\,(Q^*(r))}{r\text{MR}\,(Q^*(r)) - Q^*(r)P'(Q^*(r))} = f(r) \in (r,1). \tag{A.29}$$

Let $\widehat{r}' = \max_r f(r)$. Now, because marginal revenue is decreasing, $Q^*(r)$ is decreasing in r. $Q^*(r) > 0$ for r near 0 since by assumption $\text{MR}(0) = P'(0) > c$. But, assuming that $\text{MR}(0) < \infty$, there will exist some point $r'' > 0$ at which $Q^*(r'') = 0$. Define $r^* = \min\{r', r''\}$.

The final solution to consider is one in which no boundary solutions to (A.44) exist. The resulting system of equations simplifies to

$$\left(\begin{array}{c} (1-r)(1-\widehat{r})\text{MR}\,(Q) + r(1-\widehat{r})\text{MR}\,(Q+\Delta_2) + \\ (\widehat{r}-r)\left[P\,(Q+\Delta_2) + QP'\,(Q+\Delta_2)\right] \end{array}\right) = c \tag{A.30}$$

$$\widehat{r}\text{MR}\,(Q+\Delta) + (1-\widehat{r})\text{MR}\,(Q+\Delta-\Delta_2) = 0 \tag{A.31}$$

$$r(1-\widehat{r})\left[-\text{MR}\,(Q+\Delta-\Delta_2) + \text{MR}\,(Q+\Delta_2)\right] + (\widehat{r}-r)QP'\,(Q+\Delta_2) = 0. \tag{A.32}$$

Since $P' < 0$, (A.61) implies that $\text{MR}\,(Q+\Delta_2) > \text{MR}\,(Q+\Delta-\Delta_2)$ which in turn implies $\Delta_2 < \Delta - \Delta_2$ and $\Delta_2 < \frac{\Delta}{2}$. So the original claim that the solution to the relaxed problem in (A.44) is also the solution to the problem with the constraint $\Delta_1 \geq \Delta_2$ is validated. Also notice that as $\widehat{r}$ approaches 1, the left hand side of (A.61) must be strictly negative. So there exists some value $\widehat{r}''$ such that this solution does not exist for $\widehat{r} > \widehat{r}''$. Let $\widehat{r}^* = \max\{\widehat{r}', \widehat{r}''\}$.

We have show that for $r < r^*$ and $c < c^*$ only the exclusive contract and above solution exist. Moreover, within this parameter space, when $\widehat{r} > \widehat{r}^*$, only the exclusive contract solution exists. When $\widehat{r} \leq \widehat{r}^*$, either the exclusive contract or above solution exist. For both solutions we find that $\Delta_2^* < \frac{\Delta^*}{2}$. ■

A.5 Proof of Proposition 5

Proof. The strategy for the first part of the proof is to utilize the expressions for the existence of the three solutions derived in the proof of proposition 4. First consider the solution in which $Q = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$. Equation (A.48) becomes

$$\widehat{r}(1-2\Delta) + (1-\widehat{r})(1-\Delta) = 0 \tag{A.33}$$

or $\Delta = \frac{1}{1+\widehat{r}}$. Plugging this expression in (A.17) gives

$$\widetilde{r}(\widehat{r},c) = \frac{(1-\widehat{r}) + \widehat{r}\left(1-\frac{1}{2(1+\widehat{r})}\right) - c}{(1-\widehat{r})\left[1-\left(1-\frac{1}{1+\widehat{r}}\right)\right] + \left(1-\frac{1}{2(1+\widehat{r})}\right)} = \frac{1-\frac{\widehat{r}}{2(1+\widehat{r})} - c}{(1-\widehat{r})\left(\frac{1}{1+\widehat{r}}\right) + 1 - \frac{1}{2(1+\widehat{r})}} = \frac{2(1+\widehat{r}) - \widehat{r} - 2(1+\widehat{r})c}{1+2(1+\widehat{r}) - 2\widehat{r}} = \frac{2+\widehat{r}-2(1+\widehat{r})c}{3}. \tag{A.34}$$

Next consider the exclusive contract solution. (A.56) gives $Q^*(r) = \frac{1-r-c}{2(1-r)}$ which is positive as long as $r < 1-c$. So whenever $r < \min\left\{1-c, \frac{2+\widehat{r}-2(1+\widehat{r})c}{3}\right\}$ the exclusive contract solution exists and the solution above does not. Plugging $Q^*(r)$ into (A.29) gives

$$f(r) = \frac{r\left[\frac{2-2r-(1-r-c)}{2(1-r)}\right]}{r\left[\frac{1-r-(1-r-c)}{(1-r)}\right] + \frac{1-r-c}{2(1-r)}} = \frac{r(1-r+c)}{1-r-c+2cr}. \tag{A.35}$$

Now finally consider the solution with partial exclusion. Expressions (A.59)-(A.61) solve as

$$Q^* = \frac{r(1-\widehat{r})(2-3r+\widehat{r}-2c(1+\widehat{r}))}{4r-2r\widehat{r}-\widehat{r}^2+r^2(-5+4\widehat{r})} > 0 \tag{A.36}$$

$$\Delta^* = \frac{-r^2+r(2+6c-6c\widehat{r})+\widehat{r}(-2+2c(1-\widehat{r})+\widehat{r})}{8r-4r\widehat{r}-2\widehat{r}^2+2r^2(-5+4\widehat{r})} > 0 \tag{A.37}$$

$$\Delta_2^* = \frac{-r^2-(1-c)\widehat{r}+r(1+c+\widehat{r}-2c\widehat{r})}{4r-2r\widehat{r}-\widehat{r}^2+r^2(-5+4\widehat{r})} > 0. \tag{A.38}$$

Whenever $r < \frac{2+\widehat{r}-2(1+\widehat{r})c}{3}$ the numerator of (A.36) is positive. The condition for the denominator to be positive is that

$$r \leq \widehat{r} \leq -r(1-2r) + 2(1-r)\sqrt{r(1+r)} < 1. \tag{A.39}$$

(A.37) is positive when

$$r \leq \widehat{r} \leq \frac{1-c+3cr-\sqrt{(1-r)^2(1-2c)+c^2(1+3r)^2}}{1-2c}, \tag{A.40}$$

which is more stringent than the condition in (A.39). Finally, the condition for (A.38) positive is

$$r \leq \widehat{r} \leq \frac{r(1+c-r)}{1-c-r+2cr} \equiv \widehat{r}^*. \tag{A.41}$$

Since $\widehat{r}^* < \frac{1-c+3cr-\sqrt{(1-r)^2(1-2c)+c^2(1+3r)^2}}{1-2c}$, $r \leq \widehat{r} \leq \frac{r(1+c-r)}{1-c-r+2cr} \equiv \widehat{r}^*$ is the condition for the partial exclusion solution to hold.

To prove the final statement, let

$$T = \frac{\Delta_2^*}{\Delta^*} = \frac{2\left[-r^2-(1-c)\widehat{r}+r(1+c+\widehat{r}-2c\widehat{r})\right]}{-r^2+r(2+6c-6c\widehat{r})+\widehat{r}(-2+2c(1-\widehat{r})+\widehat{r})}.$$

Differentiating this expression yields

$$\frac{\partial T}{\partial \widehat{r}} = \frac{-2\left[A\left(c,r\right) + B(c,r)\widehat{r} - C(c,r)\widehat{r}^2\right]}{\left[-r^2 + r(2+6c-6c\widehat{r}) + \widehat{r}(-2+2c(1-\widehat{r})+\widehat{r})\right]^2},$$

where

$$A(c,r) = 4cr - 4c^2r - r^2 - 9cr + 6c^2r^2 + r^3 + 4cr^3,$$

$$B(c,r) = 2r\left(1 - c - 2c^2 - r + 2cr\right) > 0,$$

and

$$C(c,r) = (1-2c)\left[(1-c) - (1-2c)r\right] > 0.$$

To show that $\frac{\partial T}{\partial \widehat{r}} \leq 0$ we have to show that $C(c,r)\widehat{r}^2 - B(c,r)\widehat{r} - A(c,r) \leq 0$. This inequality is solved by: $\widehat{r}_1(c,r) \leq \widehat{r} \leq \widehat{r}_2(c,r)$, where: $\widehat{r}_1(c,r) = \frac{B-\sqrt{B^2-4AC}}{2C}$ and $\widehat{r}_2(c,r) = \frac{B+\sqrt{B^2-4AC}}{2C}$. Now

$$\left[\frac{\partial T}{\partial \widehat{r}}\right]_{\widehat{r}=r} = -\frac{1-c-r(1+c)}{8cr(1-r)} < 0.$$

We also know that $T\left(\widehat{r}^*\right) = 0$ and that, for $\widehat{r} > \widehat{r}^*$, $T(\widehat{r}) \leq 0$. If $\widehat{r}_2(c,r)$ were lower than $\widehat{r}^*$, then it would also be true that $\left[\frac{\partial T}{\partial \widehat{r}}\right]_{\widehat{r}=\widehat{r}^*} > 0$, which is a contradiction since it would imply that there exists an ε such that $T(\widehat{r}) > 0$ for $\widehat{r} = \widehat{r}^* + \varepsilon$. Hence, it must be that $\left[\frac{\partial T}{\partial \widehat{r}}\right]_{\widehat{r}=\widehat{r}^*} \leq 0$, and that $T(\widehat{r})$ is decreasing on $\widehat{r} \in (r, \widehat{r}^*)$. ■

A.6 Proof of Proposition 6

Proof. We consider the relaxed problem.

$$\max_{\{Q_i(\widehat{c}_i),T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^{2} rT_i(0) + (1-r)T_i(c) \text{ such that} \tag{A.42}$$

$$Q_i(c)\left(P\left[Q^H + \Delta_j\right] - c\right) - T_i(c) \geq 0 \qquad (LL_i^{HL})$$

$$\begin{aligned}(Q_i(c) + \Delta_i)\left\{rP\left[Q^H + \Delta\right] + (1-r)P\left[Q^H + \Delta_i\right]\right\} - T_i(0) \geq \\ Q_i(c)\left\{rP\left[Q^H + \Delta_j\right] + (1-r)P\left[Q^H\right]\right\} - T_i(c)\end{aligned} \qquad (IC_i^L)$$

If an exclusive contract is the only solution this relaxed problem, an exclusive contract must be the only solution to (14). This is because an exclusive contract satisfies the constraints we have ignored.

Clearly

$$\begin{aligned} T_i(c) &= Q_i(c)\left(P\left[Q^H + \Delta_j\right] - c\right) \\ &= Q_i(c)\left(rP[Q^H + \Delta_j] + (1-r)P[Q^H] - c\right) - Q_i(c)(1-r)\left(P[Q^H] - P[Q^H + \Delta_j]\right) \end{aligned}$$

The incentive compatibility constraint can be written as

$$\begin{aligned}
T_i(0) &= (Q_i(c)+\Delta_i)\left\{rP\left[Q^H+\Delta\right]+(1-r)P\left[Q^H+\Delta_i\right]\right\}- \\
&\quad Q_i(c)\left\{rP\left[Q^H+\Delta_j\right]+(1-r)P\left[Q^H\right]\right\}+T_i(c) \\
&= (Q_i(c)+\Delta_i)\left\{rP\left[Q^H+\Delta\right]+(1-r)P\left[Q^H+\Delta_i\right]\right\}- \\
&\quad Q_i(c)\left\{rP\left[Q^H+\Delta_j\right]+(1-r)P\left[Q^H\right]\right\}+Q_i(c)\left(P\left[Q^H+\Delta_j\right]-c\right) \\
&= (Q_i(c)+\Delta_i)\left\{rP\left[Q^H+\Delta\right]+(1-r)P\left[Q^H+\Delta_i\right]\right\}- \\
&\quad Q_i(c)(1-r)\left(P\left[Q^H\right]-P\left[Q^H+\Delta_j\right]\right)-cQ_i(c)
\end{aligned}$$

Plugging back into the objective function gives

$$\begin{aligned}
&r^2\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)+(1-r)^2Q^HP\left(Q^H\right)-cQ^H+ \\
&r(1-r)\left[\left(Q^H+\Delta_1\right)P\left(Q^H+\Delta_1\right)+\left(Q^H+\Delta_2\right)P\left(Q^H+\Delta_2\right)\right]- \\
&(1-r)Q_1(c)\left[P\left(Q\right)-P\left(Q+\Delta_2\right)\right]-(1-r)Q_2(c)\left[P\left(Q\right)-P\left(Q+\Delta_1\right)\right].
\end{aligned}$$

which rewrites as

$$\begin{aligned}
&r^2\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)-r(1-r)Q^HP\left(Q^H\right)-cQ^H+ \\
&r(1-r)\left[\left(Q^H+\Delta_1\right)P\left(Q^H+\Delta_1\right)+\left(Q^H+\Delta_2\right)P\left(Q^H+\Delta_2\right)\right]+ \\
&(1-r)Q_1(c)P\left(Q^H+\Delta_2\right)+(1-r)Q_2(c)P\left(Q^H+\Delta_1\right).
\end{aligned}$$

We normalize $\Delta_1 \geq \Delta_2$, so that $Q_2(c)=0$ is optimal. With this restriction the above becomes

$$\begin{aligned}
&r^2\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)-r(1-r)Q^HP\left(Q^H\right)-cQ^H+ \\
&r(1-r)\left[\left(Q^H+\Delta_1\right)P\left(Q^H+\Delta_1\right)+\left(Q^H+\Delta_2\right)P\left(Q^H+\Delta_2\right)\right]+ \\
&(1-r)QP\left(Q+\Delta_2\right).
\end{aligned}$$

Substituting $Q^H=Q$ and $\Delta_1=\Delta-\Delta_2$ gives the expression

$$\begin{aligned}
\pi(Q,\Delta,\Delta_2) = {} & r^2R\left(Q+\Delta\right)-r(1-r)R\left(Q\right)-cQ+ \\
& r(1-r)\left[R\left(Q+\Delta-\Delta_2\right)+R\left(Q+\Delta_2\right)\right]+ \\
& (1-r)QP\left(Q+\Delta_2\right) \qquad \text{(A.43)}
\end{aligned}$$

We will solve the relaxed problem

$$\max_{Q\geq 0,\Delta\geq 0,\Delta\geq 0}\pi(Q,\Delta,\Delta_2) \qquad \text{(A.44)}$$

and ignore the constraint $\Delta_2 \leq \frac{\Delta}{2}$, and show that all solution to the relaxed problem satisfy this ignored constraint. As explained in the proof of proposition 4, one need only consider three solutions to this problem.

The solution with $Q = 0$ and $\Delta_2 = \frac{\Delta}{2}$ satisfies the Kuhn-Tucker first order conditions if

$$\frac{\partial \pi(Q,\Delta,\Delta_2)}{\partial Q} = \begin{pmatrix} r^2 \mathrm{MR}\,(Q+\Delta) - r(1-r)\mathrm{MR}\,(Q) + \\ r(1-r)\left[\mathrm{MR}\,(Q+\Delta-\Delta_2) + \mathrm{MR}\,(Q+\Delta_2)\right] + \\ (1-r)\left[P\,(Q+\Delta_2) + QP'\,(Q+\Delta_2)\right] \end{pmatrix} - c < 0 \tag{A.45}$$

$$\frac{\partial \pi(Q,\Delta,\Delta_2)}{\partial \Delta} = r^2 \mathrm{MR}\,(Q+\Delta) + r(1-r)\mathrm{MR}\,(Q+\Delta-\Delta_2) = 0 \tag{A.46}$$

$$\frac{\partial \pi(Q,\Delta,\Delta_2)}{\partial \Delta_2} = \begin{pmatrix} r(1-r)\left[-\mathrm{MR}\,(Q+\Delta-\Delta_2) + \mathrm{MR}\,(Q+\Delta_2)\right] + \\ (1-r)QP'\,(Q+\Delta_2) \end{pmatrix} = 0 \tag{A.47}$$

are all satisfied at $Q = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$. (A.47) is clearly satisfied, while (A.45) and (A.46) rewrite as[23]

$$r\mathrm{MR}\,(\Delta) + (1-r)\mathrm{MR}\left(\frac{\Delta}{2}\right) = 0 \tag{A.48}$$

$$-r(1-r)\mathrm{MR}\,(0) + r(1-r)\mathrm{MR}\left(\frac{\Delta}{2}\right) + (1-r)P\left(\frac{\Delta}{2}\right) < c. \tag{A.49}$$

This solution cannot exist for r since $\mathrm{MR}\left(\frac{\Delta}{2}\right) = 0$ and $P\left(\frac{\Delta}{2}\right) < c$ cannot hold simultaneously by assumption.

Another potential solution to (A.44) is an exclusive contract in which $\Delta_2 = 0$. By the above arguments an exclusive contract can only be optimal if $Q > 0$. Such an exclusive contract satisfies the Kuhn-Tucker first order conditions if

$$\left[\frac{\partial \pi(Q,\Delta,\Delta_2)}{\partial Q}\right]_{\Delta_2=0} = 0 \tag{A.50}$$

$$\left[\frac{\partial \pi(Q,\Delta,\Delta_2)}{\partial \Delta}\right]_{\Delta_2=0} = 0 \tag{A.51}$$

$$\left[\frac{\partial \pi(Q,\Delta,\Delta_2)}{\partial \Delta_2}\right]_{\Delta_2=0} < 0. \tag{A.52}$$

which simplifies to

$$r\mathrm{MR}\,(Q+\Delta) + (1-r)\mathrm{MR}\,(Q) = c \tag{A.53}$$

$$r\mathrm{MR}\,(Q+\Delta) = 0 \tag{A.54}$$

$$r(1-r)\left[-\mathrm{MR}\,(Q+\Delta) + \mathrm{MR}\,(Q)\right] + (1-r)QP'\,(Q) < 0. \tag{A.55}$$

and further to

$$(1-r)\mathrm{MR}\,(Q) = c \tag{A.56}$$

$$\mathrm{MR}\,(Q+\Delta) = 0 \tag{A.57}$$

$$r\mathrm{MR}\,(Q) + QP'\,(Q) < 0. \tag{A.58}$$

[23] Here we have also plugged (A.45) into (A.46).

The final solution to consider is one in which no boundary solutions to (A.44) exist. The resulting system of equations simplifies to

$$\begin{pmatrix} -r(1-r)\mathrm{MR}\,(Q) + r(1-r)\mathrm{MR}\,(Q+\Delta_2) + \\ (1-r)\left[P\,(Q+\Delta_2) + QP'\,(Q+\Delta_2)\right] \end{pmatrix} = c \tag{A.59}$$

$$r\mathrm{MR}\,(Q+\Delta) + (1-r)\mathrm{MR}\,(Q+\Delta-\Delta_2) = 0 \tag{A.60}$$

$$r(1-r)\left[-\mathrm{MR}\,(Q+\Delta-\Delta_2) + \mathrm{MR}\,(Q+\Delta_2)\right] + (1-r)QP'\,(Q+\Delta_2) = 0. \tag{A.61}$$

Since $P' < 0$, (A.61) implies that $\mathrm{MR}\,(Q+\Delta_2) > \mathrm{MR}\,(Q+\Delta-\Delta_2)$ which in turn implies $\Delta_2 < \Delta - \Delta_2$ and $\Delta_2 < \frac{\Delta}{2}$. As r approaches 0, the left hand side of (A.61) must be strictly negative. So this solution cannot exist for r sufficiently small. ■

A.7 Proof of Proposition 7

Proof. We know that the optimal exclusionary solution can be implemented if IC_{LH} and IC_{LL} are satisfied. First, consider IC_{LH}. It can be rewritten as $P(Q^{H*}) - P\left(2Q^{H*} + \Delta^*\right) \geq c$. The LHS is always strictly positive because of decreasing demand. (Note that Q^{H*} and Δ^* in general depend on c, but as $c \to 0$, Q^{H*} will always be strictly positive whereas $\lim_{c\to 0}\Delta^* = 0$. This follows immediately by inspection of the FOCs implicitly defining Q^{H*} and Δ^*.) Define $\inf\left(P(Q^{H*}) - P\left(2Q^{H*} + \Delta^*\right)\right)$ as the lowest value that the LHS can attain, and set $\overline{c}_1 = \inf\left(P(Q^{H*}) - P\left(2Q^{H*} + \Delta^*\right)\right)$. For any $c \leq \overline{c}_1$, the IC_{LH} is satisfied.

Next, consider IC_{LL}. It can be rewritten as $\left[P(Q^{H*} + \Delta^*) - P\left(2Q^{H*} + 2\Delta^*\right)\right](Q^{H*} + \Delta^*)/Q^{H*} \geq c$. Since $(Q^{H*} + \Delta^*)/Q^{H*} \geq 1$, if $\left[P(Q^{H*} + \Delta^*) - P\left(2Q^{H*} + 2\Delta^*\right)\right] \geq c$, the IC_{LL} will be satisfied. The LHS of the last inequality is always strictly positive because of decreasing demand. Define $\inf\left[P(Q^{H*} + \Delta^*) - P\left(2Q^{H*} + 2\Delta^*\right)\right]$ as the lowest value that the LHS can attain, and set $\widehat{c}_2 = \inf\left[P(Q^{H*} + \Delta^*) - P\left(2Q^{H*} + 2\Delta^*\right)\right]$. For any $c \leq \widehat{c}_2$, the IC_{LL} is satisfied.

Define $\widehat{c} = \min\{\widehat{c}_1, \widehat{c}_2\}$. If $c \leq \widehat{c}$, both ICs are satisfied: the principal is able to implement the exclusionary solution, with only one firm selling in equilibrium, by making use of uniform contracts. ■

A.8 Proof of Proposition 8

Proof. Let $Q_i(c_j)$ be the quantity offered to the cost type c_j of firm i; let $\Delta_i^j = Q_i(c_j) - Q_i(c_{j-1})$ (which is defined for $j > 1$); and let $\Delta^j = \sum_{i=1}^{F}\Delta_i^j$ be the total output difference between cost types c_j and c_{j-1} and $\Delta_{-i}^j = \sum_{k\neq i}\Delta_k^j$ be the total output difference between cost types c_j and c_{j-1} that is not produced by firm i.

The maximization problem for M is

$$\max_{\{Q_i(\widehat{c}_j),T_i(\widehat{c}_j)\}_{i=1,j=1}^{i=F,j=N}} \sum_{i=1}^{F}\sum_{j=1}^{N} r_j T_i(\widehat{c}_j) \text{ such that}$$

$$Q_i(c_j)\left\{P\left[Q_i(c_j)+\overline{Q}_{-i}\right]-c_j\right\}-T_i(c_j)\geq 0 \qquad (PC_i^j)$$

$$Q_i(c_j)\left\{P\left[Q_i(c_j)+\overline{Q}_{-i}\right]-c_j\right\}-T_i(c_j)$$
$$\geq Q_i(c_l)\left\{P\left[Q_i(c_l)+\overline{Q}_{-i}\right]-c_j\right\}-T_i(c_l) \quad \forall j,l\neq j \qquad (IC_i^j)$$

where $\overline{Q}_{-i}$ is the maximum possible output produced by all firms but i. Since agent's utilities satisfy single crossing, $\overline{Q}_{-i}=\sum_{j\neq i}Q(c_N)$ and the problem becomes

$$\max_{\{Q_i(\widehat{c}_j),T_i(\widehat{c}_j)\}_{i=1,j=1}^{i=F,j=N}} \sum_{i=1}^{F}\sum_{j=1}^{N} r_j T_i(\widehat{c}_j) \text{ such that}$$

$$Q_i(c_1)\left\{P\left[Q_i(c_1)+\overline{Q}_{-i}\right]-c_1\right\}-T_i(c_1)=0 \qquad (PC_i^1)$$

$$Q_i(c_j)\left\{P\left[Q_i(c_j)+\overline{Q}_{-i}\right]-c_j\right\}-T_i(c_j)=$$
$$Q_i(c_{j-1})\left\{P\left[Q_i(c_{j-1})+\overline{Q}_{-i}\right]-c_j\right\}-T_i(c_{j-1}). \quad \forall j>1 \qquad (IC_i^j)$$

Lemma 1

$$T_i^*(c_1)=Q_i(c_1)\left\{P\left[Q_i(c_1)+\overline{Q}_{-i}\right]-c_1\right\}$$
$$T_i^*(c_j)=Q_i(c_j)\left\{P\left[Q_i(c_j)+\overline{Q}_{-i}-c_j\right]\right\}-\sum_{l=1}^{j-1}Q_i(c_l)(c_l-c_{l+1}) \ \forall j>1$$

Proof. The expression for $T_i^*(c_1)$ comes directly from the participation constraint. The expression for the other cost types can be obtained via induction. The transfer for cost type c_2 is obtained from

$$Q_i(c_2)\left\{P\left[Q_i(c_2)+\overline{Q}_{-i}\right]-c_2\right\}-T_i(c_2)$$
$$=Q_i(c_1)\left\{P\left[Q_i(c_1)+\overline{Q}_{-i}\right]-c_2\right\}-T_i(c_1)$$
$$=(c_1-c_2)Q_i(c_1).$$

Now suppose the stated form is true for some $j \in \{3, \ldots, F-1\}$. Then we obtain

$$\begin{aligned}
& Q_i(c_{j+1}) \left\{ P\left[Q_i(c_j) + \overline{Q}_{-i} \right] - c_{j+1} \right\} - T_i(c_{j+1}) \\
= & \; Q_i(c_j) \left\{ P\left[Q_i(c_j) + \overline{Q}_{-i} \right] - c_{j+1} \right\} - T_i(c_j) \\
= & \sum_{l=1}^{j-1} Q_i(c_l)(c_l - c_{l+1}) + Q_i(c_j)(c_j - c_{j+1}) = \sum_{l=1}^{j} Q_i(c_l)(c_l - c_{l+1}).
\end{aligned}$$

■

M's profit depends on the distribution of output between firms by

$$\begin{aligned}
& r_1 \sum_{i=1}^{F} Q_i(c_1) P\left[Q(c_1) + \sum_{j=2}^{N} \Delta_{-i}^{j} \right] + \\
& \sum_{j=2}^{N-1} r_j \sum_{i=1}^{F} \left(Q_i(c_1) + \sum_{l=2}^{j} \Delta_i^l \right) P\left[Q(c_1) + \sum_{l=2}^{j} \Delta^l + \sum_{l=j+1}^{N} \Delta_{-i}^{l} \right] + \\
& r_N \left(Q(c_1) + \sum_{l=2}^{N} \Delta^l \right) P\left[Q(c_1) + \sum_{l=2}^{N} \Delta^l \right].
\end{aligned}$$

In fact the distribution of output only depends on the first two terms of this expression. Now clearly

$$\sum_{i=1}^{F} Q_i(c_1) P\left[Q(c_1) + \sum_{j=2}^{N} \Delta_{-i}^{j} \right] \leq Q(c_1) P\left[Q(c_1) \right] \tag{A.62}$$

and

$$\begin{aligned}
& \sum_{j=2}^{N-1} r_j \sum_{i=1}^{F} \left(Q_i(c_1) + \sum_{l=2}^{j} \Delta_i^l \right) P\left[Q(c_1) + \sum_{l=2}^{j} \Delta^l + \sum_{l=j+1}^{N} \Delta_{-i}^{l} \right] \leq \\
& \sum_{j=2}^{N-1} r_j \left(Q(c_1) + \sum_{l=2}^{j} \Delta^l \right) P\left[Q(c_1) + \sum_{l=2}^{j} \Delta^l \right].
\end{aligned} \tag{A.63}$$

An exclusive contract in which some firm i is offered the contract $Q_i(c_1) = Q(c_1)$ and $\Delta_i^j = \Delta^j$ achieves this upper bound. Moreover, such an exclusive contract will be strictly optimal whenever $Q(c_1) > 0$ or there is some j for which $\Delta^j > 0$.

Let $Q^j = Q(c_j)$. M's optimal exclusive contract chooses Q^j to maximize

$$r_1 \left[Q^1 P[Q^1] - c_1 Q^1 \right] + \sum_{j=2}^{N} \left(r_j \left[Q^j P(Q^j) - c_j \right] - \sum_{l=1}^{j-1} Q^l (c_l - c_{l+1}) \right).$$

The first order conditions for a solution in which all quantities except Q^N are 0 is

$$\begin{aligned} \mathrm{MR}(Q^N) - c_N &= 0 \quad (A.64) \\ r_j\left[\mathrm{MR}(Q^j) - c_j\right] - (c_j - c_{j+1}) \sum_{l=0}^{N-j-1} r_{N-l} &= \\ r_j\mathrm{MR}(Q^j) - c_j \sum_{l=0}^{N-j} r_{N-j} + c_{j+1} \sum_{l=0}^{N-j-1} r_l &< 0, \ \forall\, j < N \quad (A.65) \end{aligned}$$

Now, plugging in $Q^j = 0$ for $j < N$ and adding together equations (A.65) gives

$$\sum_{j=1}^{N-1} r_j\mathrm{MR}(0) - c_1 + c_N r_N < 0. \quad (A.66)$$

As $r_N \to 0$ we must have $\sum_{j=1}^{N-1} r_j \to 1$ and (A.66) becomes $\mathrm{MR}(0) - c_1 < 0$ which is ruled out by assumption. So there exists some r_N^* such that, whenever $r_N < r_N^*$ the firm contracts some cost type c_j, $j > N$, to produce positive output, in which case only exclusive contracts are optimal. ■

B Fixed Wage Contracts

We introduce a period 0 in which the upstream firm chooses $x_i \in \{\mathrm{INT}, \mathrm{IND}\}$ for $i = 1, 2$. $x_i = \mathrm{INT}$ means firm i is integrated, while $x_i = \mathrm{IND}$ means firm i is independent. We refer to (x_1, x_2) as an *organization*. Firm i's profit is

$$\pi_i(\widehat{c}_i, \widehat{c}_j, c_i, x_i) = \mathbb{1}(x_i = \mathrm{IND})Q_i(\widehat{c}_i)P\left[Q_i(\widehat{c}_i) + Q_j(\widehat{c}_j)\right] - cQ_j(\widehat{c}_j) - T_i(\widehat{c}_i) \quad (B.1)$$

and the lottery it faces is

$$L_i(\widehat{c}_i \mid c_i, x_i) = \left\{\left[\pi_i(\widehat{c}_i, 0, c_i, x_i), \pi_i(\widehat{c}_i, c, c_i, x_i)\right]; (r, 1-r)\right\}. \quad (B.2)$$

Given an organization, M's problem is

$$\max_{\{Q_i(\widehat{c}_i), T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^{2} \left\{rT_i(0) + (1-r)T_i(c) + \mathbb{1}(x_i = \mathrm{INT})Q_i(\widehat{c}_i)P\left[Q_i(\widehat{c}_i) + Q_j(\widehat{c}_j)\right]\right\} \quad (B.3)$$

such that

$$U\left[L_i(0 \mid 0, x_i)\right] \geq U[0] \quad (PC_i^L)$$

$$U\left[L_i(c \mid c, x_i)\right] \geq U[0] \quad (PC_i^H)$$

$$U\left[L_i(0 \mid 0, x_i)\right] \geq U\left[L_i(c \mid 0, x_i)\right] \quad (IC_i^L)$$

$$U\left[L_i(c \mid c, x_i)\right] \geq U\left[L_i(0 \mid c, x_i)\right]. \quad (IC_i^H)$$

$Q_i(\widehat{c}_i)P\left[Q_i(\widehat{c}_i)+Q_j(\widehat{c}_j)\right]$ represents the revenue that accrues from selling $Q_i(\widehat{c}_i)$ units when firm $j \neq i$ sells $Q_j(\widehat{c}_j)$ units. When a downstream firm i is integrated, this revenue accrues to the upstream firm, and $L_i(\widehat{c}_i \mid c, x_i)$ is a degenerate lottery. When firm i is independent, this revenue accrues directly to it.

We denote by (x_1^*, x_2^*) the organization that maximizes the sum of M's and the two downstream firms' joint ex-ante expected utility (that is, before downstream firms' types are realized), assuming that the contracts chosen under organization (x_1, x_2) solves (B.3). We imagine M committing to an organization in exchange for a fixed payment from each downstream firm that holds it to its ex-ante participation constraint. The outcome of this Coasean bargain is (x_1^*, x_2^*). Since (INT, IND) and (IND, INT) produce the same ex-ante expected payoffs by symmetry, we do not consider the latter. To simplify the problem, we assume that demand is linear; that downstream firms have rank-dependent utility; and that $\widehat{r}(r)=\frac{r(1+c-r)}{1-c-r+2cr}$ so that exclusive contracts solve (B.3) when $(x_1, x_2) = (\text{IND}, \text{IND})$. Under (IND, IND) each downstream firm gets the exclusive contract with probability 0.5.

Lemma 2 *Suppose that* $r < \frac{1-c}{1+c}$.

1. *Under (IND, IND),* $Q^{H*}(A) = \frac{1-c}{2} - \frac{c}{2}\frac{r}{1-r}$, *and* $\Delta^*(A) = \frac{c}{2(1-r)}$.
2. *Under (INT, IND),* $Q_1^*(c) = Q^{H*}(B) = \frac{1-c}{2} - \frac{cr}{1-r+\frac{\widehat{r}-r}{2}}$, *and* $\Delta_1^*(B) = \Delta_2^*(B) = \frac{c}{2}\frac{1}{1-r+\frac{\widehat{r}-r}{2}}$.
3. *Under (INT, INT),* $Q^{H*}(C) = \frac{1-c}{2} - \frac{cr}{1-r}$, *and* $\Delta_1^*(C) = \Delta_2^*(C) = \frac{c}{2(1-r)}$.

Proof. Let $(x_1, x_2) = (\text{IND}, \text{IND})$. The optimal exclusive contract is given by equations (A.56) and (A.57) which rewrite as

$$\begin{aligned}(1-r)\left(1-2Q^{H*}\right) &= c \\ 1-2\left(Q^{H*}+\Delta^*\right) &= 0\end{aligned}$$

and from which 1. emerges immediately.

Let $(x_1, x_2) = (\text{INT}, \text{IND})$. The optimal transfers are given by

$$\begin{aligned}T_1(0) &= T_1(c) = cQ_1(c) \\ T_2(0) &= (Q_2(c)+\Delta_2)\left\{\widehat{r}P\left[Q^H+\Delta\right]+(1-\widehat{r})P\left[Q^H+\Delta_1\right]\right\} - cQ_2(c) \\ T_2(c) &= Q_2(c)\left\{\widehat{r}P\left[Q^H+\Delta_2\right]+(1-\widehat{r})P\left[Q^H\right]-c\right\}\end{aligned}$$

which when plugged into M's objective function gives

$$\begin{aligned}&r[T_1(c)+T_2(0)]+(1-r)[T_1(c)+T_2(c)] = \\ &r\left[\begin{array}{c} r[Q_1(c)+\Delta_1]P(Q^H+\Delta)+(1-r)[Q_1(c)+\Delta_1]P(Q^H+\Delta_1)-cQ_1(c)+ \\ \widehat{r}[Q_2(c)+\Delta_2]P(Q^H+\Delta)+(1-\widehat{r})[Q_2(c)+\Delta_2]P(Q^H+\Delta_2)-cQ_2(c)\end{array}\right]+ \\ &(1-r)\left[\begin{array}{c} rQ_1(c)P(Q^H+\Delta_2)+(1-r)Q_1(c)P(Q^H+\Delta_2)-cQ_1(c)+ \\ \widehat{r}Q_2(c)P(Q^H+\Delta_1)+(1-\widehat{r})Q_2(c)P(Q^H+\Delta_1)-cQ_2(c)\end{array}\right]\end{aligned}$$

which rewrites as

$$
\begin{aligned}
&r^2 \left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) + (1-r)^2 Q^H P\left(Q^H\right) - cQ^H + \\
&r(1-r)\left[\left(Q^H + \Delta_1\right) P\left(Q^H + \Delta_1\right) + \left(Q^H + \Delta - \Delta_1\right) P\left(Q^H + \Delta - \Delta_1\right)\right] + \\
&r(\widehat{r} - r)\left(Q_2(c) + \Delta_2\right)\left(P(Q^H + \Delta) - P(Q^H + \Delta_2)\right) + \\
&(1-r)(\widehat{r} - r) Q_2(c)\left(P(Q^H + \Delta_1) - P(Q^H)\right)
\end{aligned}
$$

The last two lines of this expression are clearly negative, so $Q_1(c) = Q^H$ and $Q_2(c) = 0$ is optimal. When demand is linear, the part of this expression that depends on the distribution of output between firms becomes

$$
-(1-r)\Delta_1^2 - (1-r)\Delta_2^2 - (\widehat{r} - r)\Delta_1\Delta_2 = -(1-r)\Delta^2 + \Delta_1\Delta_2(2 - \widehat{r} - r)
$$

which is clearly maximized at $\Delta_1 = \Delta_2 = \frac{\Delta}{2}$. M's objective function then becomes

$$
\begin{gathered}
r^2(Q+\Delta)(1-Q-\Delta) + (1-r)^2 Q(1-Q) + 2r(1-r)\left(Q + \frac{\Delta}{2}\right)\left(1 - Q - \frac{\Delta}{2}\right) \\
-cQ - \frac{r(\widehat{r} - r)\Delta^2}{4}.
\end{gathered}
$$

The first order conditions for Q and Δ are

$$
\begin{aligned}
1 - 2Q^* - c - 2r\Delta^* &= 0 \\
r^2(1 - 2Q^* - 2\Delta^*) + r(1-r)(1 - 2Q^* - \Delta^*) - \frac{r(\widehat{r} - r)\Delta^*}{2} &= 0.
\end{aligned}
$$

Solving these two together gives

$$
\begin{aligned}
\Delta^* &= \frac{c}{1 - r + \frac{\widehat{r} - r}{2}} \\
Q^* &= \frac{1-c}{2} - \frac{rc}{1 - r + \frac{\widehat{r} - r}{2}}
\end{aligned}
$$

Let $(x_1, x_2) = (\mathrm{INT}, \mathrm{INT})$. The derivation of the optimal contracts is identical to the case in which $(x_1, x_2) = (\mathrm{INT}, \mathrm{IND})$ taking $\widehat{r} = r$. ■

Now that we have established optimal contracts in each organization, we can provide conditions under which exclusive contract within the organization $(x_1, x_2) = (\mathrm{INT}, \mathrm{INT})$ maximize joint utility.

Proposition 9 *There exists a $\overline{c}$ and $\overline{r}$ such that $(x_1^*, x_2^*) = (IND, IND)$ when $c < \overline{c}$ and $r > \overline{r}$.*

Proof. We begin the analysis by comparing M's profit under the organizations $(\mathrm{INT}, \mathrm{INT})$ and $(\mathrm{IND}, \mathrm{IND})$ which is

$$
\mathbb{E}[\,Q(1-Q)\,] - cQ^H = \mathbb{E}[\,Q\,] - \mathbb{E}[\,Q\,]^2 - V[\,Q\,] = -r(1-r)\sum_i \Delta_i^2 - cQ^H.
$$

Since $\mathbb{E}[Q]$ is the same in both organizations.

Moving from (IND, IND) to (INT, INT) increases M's payoff by

$$r(1-r)\left[\frac{c}{2(1-r)}\right]^2 - 2r(1-r)\left[\frac{c}{2(1-r)}\right]^2 + c\left(\frac{1-c}{2} - \frac{r}{1-r}\frac{c}{2}\right) - c\left(\frac{1-c}{2} - \frac{r}{1-r}c\right) =$$
$$\frac{r}{1-r}\frac{c^2}{4}.$$

Moving from (IND, IND) to (INT, INT) changes joint downstream utility by an amount

$$[1-\widehat{r}(1-r)]\,c\left(\frac{1-c}{2} - \frac{r}{1-r}c\right) - 2\,[1-\widehat{r}(1-0.5r)]\,c\left(\frac{1-c}{2} - \frac{r}{1-r}\frac{c}{2}\right).$$

Simple calculations reveal that $1-\widehat{r}(1-r) = \frac{r(r-c)}{c+r-2cr}$, $2\,[1-\widehat{r}(1-0.5r)] = \frac{r(0.5r-c)}{c+0.5r-cr}$ and that $\frac{r(r-c)}{c+r-2cr} > \frac{r(0.5r-c)}{c+0.5r-cr}$. So (IND, IND) produces higher joint expected utility than (INT, INT) whenever

$$\frac{r(0.5r-c)}{c+0.5r-cr}\left(\frac{1-c}{2} - \frac{r}{1-r}\frac{c}{2}\right) - \frac{r(r-c)}{c+r-2cr}\left(\frac{1-c}{2} - \frac{cr}{1-r}\right) - \frac{r}{1-r}\frac{c}{4} > 0.$$

A stronger condition given that $\frac{r(r-c)}{c+r-2cr} > \frac{r(0.5r-c)}{c+0.5r-cr}$ is that

$$\frac{r(0.5r-c)}{c+0.5r-cr}\frac{r}{1-r}\frac{c}{2} - \frac{r}{1-r}\frac{c}{4} > \frac{1-c}{2}\left(\frac{r(r-c)}{c+r-2cr} - \frac{r(0.5r-c)}{c+0.5r-cr}\right).$$

By simple algebra

$$\frac{r(r-c)}{c+r-2cr} - \frac{r(0.5r-c)}{c+0.5r-cr} = \frac{r^2c(1-c)}{(c+r-2cr)(c+0.5r-cr)}.$$

So the above condition writes as

$$\frac{r(0.5r-c)}{c+0.5r-cr} - \frac{1}{2} > \frac{r(1-r)(1-c)^2}{(c+r-2cr)(c+0.5r-cr)}.$$

As $c \to 0$ this condition becomes

$$r - \frac{1}{2} > \frac{2(1-r)}{r}$$

which is clearly satisfied for r sufficiently high. So there exists some $\overline{c}^1$ and $\overline{r}^1$ such that (IND, IND) produces higher joint utility than (INT, INT) whenever $c < \overline{c}^1$ and $r > \overline{r}^1$.

Now we will compare the organizations (IND, IND) and (INT, IND). Following similar steps as above, the joint payoff under (IND, IND) is higher whenever

$$2\left[1-\widehat{r}(1-0.5r)\right]c\left(\frac{1-c}{2}-\frac{r}{1-r}\frac{c}{2}\right)-\left[1-\widehat{r}(1-r)\right]c\left(\frac{1-c}{2}-\frac{rc}{1-r+\frac{\widehat{r}-r}{2}}\right)+$$
$$2r(1-r)\left[\frac{c}{1-r+\frac{\widehat{r}-r}{2}}\frac{1}{2}\right]^2-\frac{rc^2}{1-r+\frac{\widehat{r}-r}{2}}+\frac{r(\widehat{r}-r)}{4}\left[\frac{c}{1-r+\frac{\widehat{r}-r}{2}}\right]^2-$$
$$-r(1-r)\left(\frac{c}{2(1-r)}\right)^2+\frac{r}{1-r}\frac{c^2}{2}>0.$$

As explained above, a stronger condition is that

$$\frac{r(0.5r-c)}{c+0.5r-cr}\left(\frac{rc}{1-r+\frac{\widehat{r}-r}{2}}-\frac{r}{1-r}\frac{c}{2}\right)+$$
$$2r(1-r)c\left[\frac{1}{1-r+\frac{\widehat{r}-r}{2}}\frac{1}{2}\right]^2-\frac{rc}{1-r+\frac{\widehat{r}-r}{2}}+\frac{cr(\widehat{r}-r)}{4}\left[\frac{1}{1-r+\frac{\widehat{r}-r}{2}}\right]^2-$$
$$cr(1-r)\left(\frac{1}{2(1-r)}\right)^2+\frac{r}{1-r}\frac{c}{2}>\frac{1}{2}\frac{r^2c(1-c)^2}{(c+r-2cr)(c+0.5r-cr)}.$$

First dividing through by c, and then taking the limit as $c\to 0$ gives

$$r\left[\frac{r}{1-r}-\frac{r}{2(1-r)}\right]+2r(1-r)\left[\frac{1}{1-r}\frac{1}{2}\right]^2-\frac{r}{1-r}-r(1-r)\left(\frac{1}{2(1-r)}\right)^2+\frac{r}{1-r}\frac{1}{2}>1$$

which again produces the condition

$$r-\frac{1}{2}>\frac{2(1-r)}{r}.$$

So there exists some $\overline{c}^2$ and $\overline{r}^2$ such that (IND, IND) produces higher joint utility than (INT, IND) whenever $c<\overline{c}^2$ and $r>\overline{r}^2$. The proof is completed by taking $\overline{r}=\max\{\overline{r}^1,\overline{r}^2\}$ and $\overline{c}=\min\{\overline{c}^1,\overline{c}^2\}$. ∎