DISCUSSION PAPER SERIES

No. 8982

VERTICAL EXCLUSION WITH ENDOGENOUS COMPETITON EXTERNALITIES

Stephen Hansen and Massimo Motta

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8982.asp**

ISSN 0265-8003

VERTICAL EXCLUSION WITH ENDOGENOUS COMPETITON EXTERNALITIES

Stephen Hansen, Universitat Pompeu Fabra and Barcelona GSE
Massimo Motta, ICREA-Universitat Pompeu Fabra, Barcelona GSE and CEPR

Discussion Paper No. 8982
May 2012

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8982

May 2012

ABSTRACT

Vertical Exclusion with Endogenous Competiton Externalities*

In a vertical market in which downstream firms have private information about their productivity and compete for consumers, an upstream firm posts public bilateral contracts. When downstream firms are risk-neutral without wealth constraints, the upstream firm offers the input to all retailers. When they are sufficiently risk averse it sells to one, thereby eliminating externalities among downstream firms that necessitate the payment of risk premia. By similar reasoning exclusion is also optimal with downstream wealth constraints. Thus exclusion arises when contracts are fully observable and downstream firms are ex ante symmetric. The result is robust to a number of extensions.

JEL Classification: D82, L22 and L42
Keywords: adverse selection, exclusive contracts, limited liability and risk

Stephen Hansen
Universitat Pompeu Fabra
C/ Ramon Trias Fargas 25-27
E-08005 Barcelona
SPAIN

Email: stephen.hansen@upf.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=171533

Massimo Motta
Universitat Pompeu Fabra
C/ Ramon Trias Fargas 25-27
E-08005 Barcelona
SPAIN

Email: massimo.motta@upf.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=116312

*We would like to thank Larbi Alaoui for extensive comments, as well as Mark Armstrong, Giacomo Calzolari, Yutian Chen, Simon Cowan, Jan Eeckhout, Péter Eső, Clare Leaver, Patrick Rey, Emanuele Tarantino, John Thanassoulis, participants of the ANR-DFG workshop (Paris), and seminar attendees at Bologna and Oxford for helpful advice. The authors gratefully acknowledge financial aid from the Spanish Ministry of Science and Innovation (project ECO2010-15052).

Submitted 09 May 2012

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] yet their role remains relatively poorly understood.

Perhaps the most common justification for the optimality of exclusive contracts is that given in Hart and Tirole (1990) (see also Segal (1999) and Rey and Tirole (2007)).[2] When downstream firms are homogenous and the upstream firm offers public bilateral contracts, it can always extract the monopoly profit. But when it offers unobservable bilateral contracts, it cannot commit to not renegotiate with downstream firms, and as a result is unable to obtain the monopoly profit. Exclusive contracts serve as a commitment device to restore monopoly profit.

This paper revisits the problem of an upstream firm offering bilateral contracts to downstream firms, and shows that exclusive contracts can be optimal even when such contracts are public information. This arises when downstream firms have private information about their potentially heterogenous productivity levels,[3] and are either sufficiently risk averse or wealth-constrained. The basic message of the paper is that downstream market environments that appear natural in some contexts lead to exclusion in the absence of commitment problems upstream.

Our analysis builds on two basic ideas. First, when downstream firms compete, they impose externalities on one another. If one produces more, the market price goes down, affecting the profits of others. But if productive firms produce more than less productive firms and cost types are private information, downstream firms do not know the size of the externality that competition will impose on them. Second, the size of the externalities is endogenous to the contracts the upstream firm offers. By increasing or decreasing the difference in input levels offered to more and less productive firms, the upstream firm determines the magnitude of the uncertainty that downstream firms face.

When downstream firms are risk-neutral and can absorb losses, the upstream firm offers both of them the input in equilibrium.[4] This helps reduce volatility in aggregate

[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] A separate stream of literature, beginning with Telser (1960), emphasizes the efficiency gains of exclusive contracts.

[3] Firms are ex-ante identical in the sense that their productivities are drawn from the same distribution.

[4] In our baseline model, there are two downstream firms each with two potential cost types, but we

revenue by ensuring that all productive firms are given a share of the market.[5] When firms are risk averse, there is a cost to the upstream firm of contracting two downstream firms: the uncertainty in their realized profit forces it to pay them a risk premium. When risk aversion is sufficiently high, exclusive contracts are optimal because, by offering zero input to one downstream firm, the upstream firm eliminates the competition externality, and with it the uncertainty a downstream firm faces.[6] A similar mechanism operates when downstream firms cannot lose money from competition because they are wealth-constrained or enjoy limited liability. The resulting rents that the upstream firm must pay out when contracting two firms more than offset the gain from output smoothing, which induces it to offer an exclusive contract.

It is worth stressing how our mechanism differs from that in Segal (1999). He shows that when bilateral contracts are public information, inefficiencies only arise if there are externalities on non-traders, which are absent in both the canonical Hart and Tirole (1990) model and ours. In the world of homogenous downstream firms that Segal (1999) considers, each one knows its payoff conditional on observing the full set of posted contracts. In our model firms do not in general know their payoffs conditional on observable contracts, and this feature generates novel insights.[7]

Rey and Tirole (1986) also analyze a model in which the upstream firm suffers from asymmetric information and retailers may be risk averse (either demand realization or distribution costs are uncertain). However, they arrive at the opposite result to ours: when retailers are infinitely risk-averse, the upstream firm wants them to compete, and when they are risk-neutral, it wants to offer them exclusivity. This is because in their paper retailers do not have private information relative to each other. Hence, competition eliminates any uncertainty, because it univocally determines their prices and profits: for instance, when they are undifferentiated, Bertrand competition leads them to set price equal to marginal cost (whatever the realization of distribution cost and of demand turns out to be), leading them to have certain zero profits. In our paper, instead, costs are private information and competition is precisely what creates uncertainty for the retailers.

show in an extension that our results generalize to an arbitrary number of firms and cost types.

[5] One can think of other mechanisms, such as product differentiation, that would also generate the optimality of offering more than one firm the input.

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

[7] Two streams of the mechanism design literature have examined the effect of risk aversion on optimal contracts. They differ in the source of uncertainty for the agent. First, Laffont and Rochet (1998) analyze an environment with one principal and one agent in which the former must design a contract before the realization of the agent's type. Second, Matthews (1983) and Maskin and Riley (1984) consider an auction (i.e. multilateral contracting) environment in which agents know their types, but are uncertain about whether they will win given the uncertainty about others' types. We are not aware of other papers that study the effects of risk aversion with bilateral contracting and multiple agents in which there is uncertainty about realized payoffs given externalities among agents.

Beyond making the theoretical point that non-observability is not necessary for generating exclusivity with bilateral contracting, one might ask the extent to which our model can account for actual exclusion. This in turn depends on whether our assumptions are reasonable. First, key to our results is that contracting is bilateral. If we allowed multilateral contracts, we conjecture (following McAfee and McMillan (1986); Laffont and Tirole (1987); McAfee and McMillan (1987); and Riordan and Sappington (1987)) that the upstream firm would offer the entire market to the most efficient producers in a mechanism that resembled an auction. While we do not doubt that auctions would improve the situation for the upstream firm, the non-linear, price-discriminating bilateral contracts we consider appear substantially more flexible than many real-world vertical contracts. 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.[8] Furthermore, Lafontaine and Shaw (1999) examine a panel data set of franchisors and find that over a 13-year period 75% of them never change their terms.

Second, in our model the upstream firm would like to cancel exclusive contracts offered to non-productive firms ex-post. 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.[9]

Finally, while limited liability is a common component of industrial organization models, firm's risk aversion is not. To the extent that agents are distributors or retailers, or more generally small firms that are unable to diversify risk, it will probably not take much convincing that this assumption is realistic.[10] But there are reasons why even larger firms may be reluctant to take risk. One is that actual decisions are taken by managers who, being individuals, may well be risk averse. Also, Nocke and Thanassoulis (2010) show that firms that face credit constraints subsequent to competing in the downstream market behave as if they were risk averse. Finally, van Eijkel and Moraga-González (2010) present empirical evidence of the relevance of firm risk aversion, as do the references

[8] 27% of those using exclusive contracts cited transaction costs as the main reason for contract uniformity, which are probably also a major factor in explaining why upstream firms do not always allocate market shares through auctions.

[9] 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.

[10] Outside of industrial organization, downstream firms are occasionally modelled as risk averse. For example, Brickley and Dark (1987) assume that franchisors are risk neutral and franchisees are risk averse just as we do, and study the organizational implications of these preferences.

mentioned in Asplund (2002).

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 examines the case where downstream firms are risk neutral but wealth constrained. Section 6 explores three extensions of the baseline model in which exclusion remains optimal. Section 7 concludes the paper. Appendix A contains all 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 $P(Q) + QP'(Q) = \mathrm{MR}(Q)$ is decreasing; and that there exists some finite quantity $\widetilde{Q}$ at which $\mathrm{MR}(\widetilde{Q}) = 0$.

The downstream firms, which may be risk neutral or risk averse, may be heterogeneous in 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 $\mathrm{MR}\,(Q') = 0$, $c < P\,(Q')$.[11]

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

[11] 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 Q', defined by $\mathrm{MR}(Q') = \mathrm{MC} = 0$.

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. We refer to this as a *competition externality.*

Crucial for the model is that firm i does not know the size of the competition externality it faces. 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. M's problem can be expressed as

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

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

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

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

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

$$Q_i(\widehat{c}_i) \geq 0. \tag{NN}$$

Here there are twelve constraints corresponding to participation (denoted PC) and incentive compatibility (denoted IC) constraints for each cost type of each downstream firm, as well as non-negativity constraints on output (denoted NN). $U[0]$ 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 ordered the input. For simplicity we sometimes refer to a contract menu simply as a contract.

There are several relevant definitions for exclusion.

Definition 1 *Let $\{Q_i^*(\widehat{c}_i), T_i^*(\widehat{c}_i)\}_{i=1}^2$ be a solution to (3).*

1. *Firm i is excluded if $Q_i^*(0) = Q_i^*(c) = 0$.*

2. *Firm i is ε-excluded if* $\max\{Q_i^*(0), Q_i^*(c)\} < \varepsilon$.

3. *Firm i is partially excluded if* $Q_i^*(0) < Q_j^*(0)$ *and* $Q_i^*(c) < Q_j^*(c)$ *for* $j \neq i$.

Below we study how different kinds of exclusion depend on U.

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

1. The upstream firm M posts a contract for each downstream firm i.[12]

2. Each downstream firm orders input, which is not observed by the competitor, and commits to pay the corresponding transfer.

3. Downstream firms produce output, the market clears, profits are realized, and they pay the transfer to the upstream firm.

Here downstream firms do not observe each others' actions prior to producing. Were this not so, one could imagine a signalling game being played rather than a simultaneous game of incomplete information. We have also assumed that downstream firms are not wealth-constrained and can pay the agreed transfer even if doing so means negative profits ex-post. We come back to exploring the consequences of wealth constraints in section 5.

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

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

M's constrained optimization problem in (3) can be rewritten as an unconstrained problem in which it chooses quantities to maximize downstream firms' expected revenue net of production costs and information rents. These last two quantities are linear in *aggregate* quantities due to constant returns to scale, so are irrelevant for exclusion.

To understand the effect of the distribution of total contracted output, it is useful to rewrite the contract variables. Define $Q_i(c) \equiv Q_i^H$ and let $Q^H = Q_1^H + Q_2^H$ be the total

[12] We discuss the case in which contracts are the same for both firms in section 6.1.

production of high cost firms; let $\Delta_i = Q_i(0) - Q_i^H$ 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.

Downstream expected revenue can be written as

$$r^2 \left(Q^H + \Delta\right) P \left(Q^H + \Delta\right) + (1-r)^2 Q^H P \left(Q^H\right) +$$
$$r(1-r) \left[\left(Q^H + \Delta_1\right) P \left(Q^H + \Delta_1\right) + \left(Q^H + \Delta - \Delta_1\right) P \left(Q^H + \Delta - \Delta_1\right)\right]. \qquad (5)$$

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

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

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.

Building on this argument, we can show that

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

This immediately implies that both firms sell strictly positive output in equilibrium.

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

To gain further intuition, the linear demand case where $P = 1 - Q$ 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 $Q_i = Q_i^H + \widetilde{x}_i \Delta_i$ 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.

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 input 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 willing 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.[13] 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, so that:

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

In incentive compatible contracts, low cost firms produce more than high cost firms. So, rather than asking downstream firms to pay a transfer equal to expected revenue, M

[13] A common intuition from statistics is that uncertainty should decrease as the number of competitors increases since the distribution of the expected payoff becomes less variable. This does not apply when firms are risk averse because their utility does not only depend on the expected payoff. With infinite risk aversion, what matters is the *worst* possible outcome, which get worse with more competitors.

can now only ask expected payments (net of production costs and information rents) of

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

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 with uncertain costs decreases firm i's utility relative to the risk neutral case, which requires the upstream firm to pay a risk premium.

Because price is decreasing in quantity, (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 $Q_i^H=\Delta_i=0$. The intuition here is that minimizing the risk premium requires the elimination of competition externalities, which exclusive contracts do.[14] Exclusive contracts are *strictly* optimal whenever M wishes to contract a positive amount of high cost production (which it does when r is low); otherwise, $Q^{H*}=0$ and any split of Δ^* across the two downstream firms is optimal.

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 $\Delta_i^*>0$, $Q_i^{H*}>0$, and $\Delta_j^*=Q_j^{H*}=0$.*
2. *Whenever $r \geq r^*$ the optimal contracts are such that $Q_i^{H*}=Q_j^{H*}=0$ and $\Delta_1^*+\Delta_2^*=\Delta^*>0$.*

This result 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.*

[14] Reducing Δ_i would also decrease uncertainty: as $\Delta_i \rightarrow 0$, the upstream firm also pays no risk premium to downstream firms. But this alternative is poor 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 the risk premium, but sacrifices production efficiency. When contracting with one firm, it minimizes the risk premium but also takes advantage of the productivity gains that arise from contracting the low cost firm to produce more than the high cost firm.

This result contains a basic message of the paper. Even if the upstream firm can fully commit to bilateral contracts, it may simply be too costly to include both firms in the downstream market because uncertain competition externalities impose a cost on risk averse firms. 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 $U[L_i(\widehat{c}_i \mid c_i)]$ satisfies the single-crossing condition[15] that greatly reduces the complexity of M's problem. Also, when firms' payoffs are non-linear in profits, characterizing optimal contracts is not straightforward.

This section presents two formulations of intermediate risk preferences that overcome 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

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

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 $U[L_i(\widehat{c}_i \mid c_i)]$ satisfies single-crossing.

The next result shows that 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 $a \geq \overline{a}$.*

Intuitively, 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 with infinite risk aversion.

[15] The condition is that $\frac{\frac{\partial U}{\partial Q_i}}{\frac{\partial U}{\partial T_i}}$ is increasing in c_i.

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. We now turn to an alternative framework to do so.

4.2 Rank-Dependent Utility

The Rank-Dependent Utility (RDU) literature represents preferences over lotteries both in terms of the probability weights attached to wealth outcomes and the utility of wealth function.[16] 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, as suggested by Yaari (1987).[17] 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) $\lim_{r\to 1}\widehat{r}(r) = 1$.[18] Figure 1(a) plots an example of a weighting function.

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

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

or the expected value of the lottery minus $\widehat{r}(r) - r$ times the distance between the two

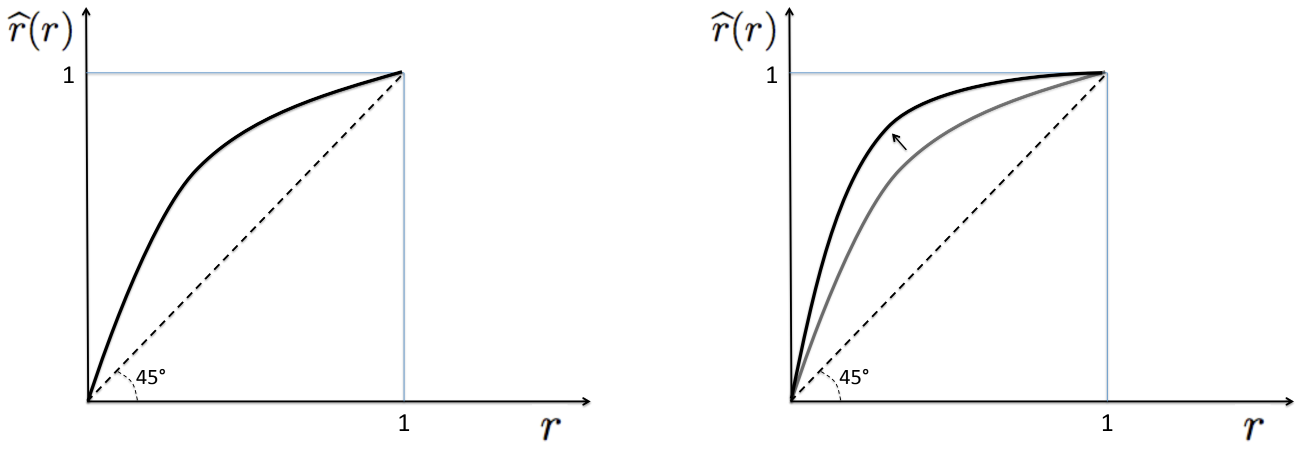
[16] See Quiggin (1982) for a seminal reference.

[17] "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).

[18] 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 $\widehat{r}(r) = 1 - \alpha(1 - r)$.



(a) Weighting Function

(b) Increase in Risk Aversion

Figure 1: Examples of Weighting Functions

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 1(b). We now use this framework to derive optimal contracts.[19]

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. With RDU 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 $\Delta_1 \geq \Delta_2$.

In the appendix we show that the expected payments (net of production costs and

[19] 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.

information rents) M asks with RDU are given by

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

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), and reflects an incentive to minimize the risk premium by eliminating competition externalities by excluding one downstream firm. Moreover, the degree of risk aversion affects the relative importance of each term. When $\widehat{r}$ increases, the incentive to exclude becomes more important, while smoothing becomes less so. Our next result show how the optimal contract resolves this tension.

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

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

Corollary 3

1. *If firms are sufficiently risk averse, exclusive contracts are optimal.*
2. *For any level of risk aversion, there is partial exclusion.*

The point we wish to highlight is that a small amount of risk aversion is sufficient to generate starkly asymmetric contracts in which one firm produces all of the high cost output as well as more of the low cost output.

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.

We next explore the linear demand case $P(Q)=1-Q$ to generate finer predictions.

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 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 also displays another kind of monotonicity. In addition to producing all high cost output, firm 1 produces an increasing share of Δ 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 4 *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. Clearly, $\lim_{c\to 0} \widehat{r}^* = r$, meaning that 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 4 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

As with risk aversion, the existence of wealth constraints or limited liability prevents the upstream firm from fully extracting expected downstream profits when firms compete.

In the baseline case with risk neutrality explored in section 3.1, downstream firms with high cost types are 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 high cost downstream firms (in addition to the information rent left to low cost firms) that again keeps it from fully exploiting the gains of dealing with multiple firms.

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

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

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

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

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

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

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

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

$$Q_i^H \geq 0, Q_i^H + \Delta_i \geq 0. \qquad (NN)$$

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 outweighs the gains from output smoothing. The following provides an answer.[20]

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.[21] Moreover, since wealth constraints are sometimes

[20]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 RDU model, and the analysis proceeds similarly.

[21]The intuition for the condition on r begin sufficiently low is richer with limited liability than infinite

relevant when risk aversion is not, the result also expands the set of situations in which our model predicts exclusion.

Proposition 6 is also important for another theoretical reason. We have shown that the interaction of risk aversion and competition externalities generates the optimality of exclusive contracts. In a more general contracting setup with auctions, competition externalities would not arise because the most efficient firm would produce all output. Nevertheless, it is an open theoretical question whether downstream risk aversion would induce the upstream firm to restrict the number of auction participants to extract more surplus. With limited liability this issue does not arise. 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. So, the upstream firm would not want to restrict the number of limited-liable auction participants. Instead, the *combination* of bilateral contracting (and the associated externalities) with limited liability generates exclusion.

6 Extensions

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

6.1 Standardized contracts

In this section we relax the assumption that M is able to fully discriminate between the downstream firms.[22] When the upstream firm can offer different contracts to different downstream firms, we know that under infinite risk aversion (as long as r is sufficiently small) exclusion of one of them is uniquely optimal. The question is whether M is able to replicate this outcome when it is restricted to offering the same standardized contract to both firms. We show here that the answer is affirmative under certain (possibly very mild) conditions.

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

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.

[22]This is similar to Segal (2003), who examines when the optimal allocation under bilateral contracting with individualized contracts can be replicated with standardized contracts.

1. It offers both firms the same menu of contracts $\{(T(0), Q(0)), (T(c), Q(c)), (0,0)\}$, with $Q(0) = Q^{H*} + \Delta^*$, $Q(c) = Q^{H*}$, $T(0) = (Q^{H*} + \Delta^*)P(Q^* + \Delta^*) - cQ^{H*}$, and $T(c) = Q^{H*}P(Q^{H*}) - cQ^{H*}$, where Q^{H*} and Δ^* are the optimal values found for an exclusive contract.[23]

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

For this to be an equilibrium outcome, each player must find the candidate strategy to be optimal given the choices of the others. For M, optimality follows immediately from the fact that this contract reproduces the optimal outcome obtained under less restrictive assumptions on its strategies (it cannot achieve higher profits when it is obliged to use standardized contracts than when it can discriminate).

In the case of firm 1, the choice is also optimal given that firm 2 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 firm 2 does not sell).

Therefore, we only need to confirm 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 infinite risk aversion, this is true when the following incentive constraints are satisfied:

$$0 \geq Q^{H*}P\left(2Q^{H*} + \Delta^*\right) - cQ^{H*} - T(c) \qquad (IC_{HH})$$
$$0 \geq \left(Q^{H*} + \Delta^*\right)P\left(2Q^{H*} + 2\Delta^*\right) - c\left(Q^{H*} + \Delta^*\right) - T(0) \qquad (IC_{HL})$$

$$0 \geq Q^{H*}P\left(2Q^{H*} + \Delta^*\right) - T(c) \qquad (IC_{LH})$$
$$0 \geq \left(Q^{H*} + \Delta^*\right)P\left(2Q^{H*} + 2\Delta^*\right) - T(0), \qquad (IC_{LL})$$

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

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

Because the demand function is assumed to be decreasing, the first two constraints

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

are always satisfied, so we are left with the last two ICs. We find that

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

Proposition 7 gives a *sufficient* condition for standardized contracts to implement the same exclusionary outcome as individual ones. It is important to note that this condition might be very weak. For instance, under $P = 1 - Q$ with individual contracts, we know that exclusion is optimal for $r < (1-c)/(1+c)$, and that $Q^{H*} = \frac{1-c-r}{2(1-r)}$ and $\Delta^* = \frac{c}{2(1-r)}$. By replacing these values into IC_{LH} and IC_{LL}, one obtains $\overline{c} = 1/2$. Moreover, $c < 1/2$ is exactly what we assumed in the baseline model to ensure that the low cost firms face competition from high cost ones.

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 Production only 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 revenue 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.[24]

In appendix B we extend our model with RDU to allow the upstream firm to choose whether to collect the revenue from the output that each downstream firm produces or whether to let each collect it as in the baseline case. After this decision, it offers optimal contracts given the distribution of revenue income. In all cases, each firm is offered a menu of quantity-transfer pairs as above. 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 the upstream firm's choosing to collect the sales revenue of firm i as "vertically integration" with firm i. We 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 firm 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.[25] Thus, subject to certain conditions, our insights are robust to an environment in which the upstream firm could choose insurance to avoid exclusion if it so wanted. 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 one another's cost types, competition generates uncertainty, leading risk-averse or wealth-constrained agents to require a rent. To save the payment of such rents, the upstream firm sometimes prefers to deal exclusively with one downstream firm, and more generally offers asymmetric contracts in many cases.

[24]One reason why production only contracts might not be efficient is if there were a moral hazard problem whose solution required downstream firms to be exposed to the wealth effects of their actions.

[25]More specifically, this is true with linear demand when $\widehat{r}^* = \frac{r(1+c-r)}{1-c-r+2cr}$, c is sufficiently small, and r is sufficiently large.

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, the introduction of competition will oblige the principal to pay a risk premium whenever agents are risk averse. To save on these, the principal may prefer to contract with a strict subset of the potential agents. This same mechanism should hold in very different settings, such as in a moral hazard model where agents are paid according to relative performance schemes.

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 J.-C. ROCHET (1998): "Regulation of a Risk Averse Firm," *Games and Economic Behavior*, 25(2), 149–173.

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.

MASKIN, E. S., AND J. G. RILEY (1984): "Optimal Auctions with Risk Averse Buyers," *Econometrica*, 52(6), 1473–1518.

MATTHEWS, S. A. (1983): "Selling to risk averse buyers with unobservable tastes," *Journal of Economic Theory*, 30(2), 370–400.

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. TIROLE (1986): "The Logic of Vertical Restraints," *The American Economic Review*, 76(5), 921–939.

——— (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.

Segal, I. (1999): "Contracting With Externalities," *The Quarterly Journal of Economics*, 114(2), 337–388.

——— (2003): "Coordination and discrimination in contracting with externalities: divide and conquer?," *Journal of Economic Theory*, 113(2), 147–181.

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

Telser, L. G. (1960): "Why Should Manufacturers Want Fair Trade?," *Journal of Law and Economics*, 3, 86–105.

van Eijkel, R., and J. L. Moraga-González (2010): "Do Firms Sell Forward for Strategic Reasons? An Application to the Wholesale Market for Natural Gas," Working Paper 864, IESE.

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

A Omitted Proofs

(3) is not a standard adverse selection problem because of externalities. However, if the externalities are separable from types in agents' utility function, one can apply standard tools, as the following result formalizes.

Lemma 1 *Suppose that for $i = 1, 2$ and $j \neq i$*

$$U\left[L_i(\widehat{c}_i \mid c_i)\right] = f\left[Q_i(\widehat{c}_i), Q_j(0), Q_j(c)\right] - c_i Q_i(\widehat{c}_i) - T_i(\widehat{c}_i).$$

Then (3) rewrites as

$$\sum_{i=1}^{2} \left\{ rf\left[Q_i(0), Q_j(0), Q_j(c)\right] + (1-r) f\left[Q_i(c), Q_j(0), Q_j(c)\right] - cQ_i(c) \right\}$$
$$\text{such that } \; Q_i(0) \geq Q_i(c) \geq 0.$$

Proof. IC_i^L and PC_i^H together imply PC_i^L is satisfied. So IC_i^L must bind, since otherwise M could increase $T_i(0)$. Now IC_i^H can be rewritten as[26]

$$c\left[Q_i(0) - Q_i(c)\right] \;\geq\; T_i(c) - T_i(0) \;+\; \left\{ \begin{array}{c} f\left[Q_i(0), Q_j(0), Q_j(c)\right] - \\ f\left[Q_i(c), Q_j(0), Q_j(c)\right] \end{array} \right\} \geq 0. \tag{A.1}$$

So $Q_i(0) \geq Q_i(c)$ is necessary for incentive compatibility, and, under this condition, IC_i^L binding implies IC_i^H is satisfied. This leaves PC_i^H, which must be binding, since otherwise the upstream firm could increase $T_i(c)$. The maximization problem is obtained by substituting in for $T_i(0)$ and $T_i(c)$ and imposing the implementability condition $Q_i(0) \geq Q_i(c)$. ■

A.1 Proof of Proposition 1

Proof. The risk neutrality case satisfies the condition of lemma 1 with

$$f\left[Q_i(c_i), Q_j(0), Q_j(c)\right] = Q_i(\widehat{c}_i) \left\{ rP\left[Q_i(\widehat{c}_i) + Q_j(0)\right] + (1-r)P\left[Q_i(\widehat{c}_i) + Q_j(c)\right] \right\}. \tag{A.2}$$

The resulting objective function is equation (5) 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 $Q_1^H = Q^{H\prime}$ and $\Delta_1 = \Delta_2 = 0$. Now consider the problem in which the upstream maximizes profit within the restricted class of contracts $Q_1^H = Q^H$ and $\Delta_1 = \Delta$. The problem becomes

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

[26] The last inequality follows from IC_i^L.

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

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

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

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 $\mathrm{MR}(0) \leq 0$ which is ruled out by assumption. Suppose not, and that $Q^{H*} > 0$. Then (A.5) gives $\mathrm{MR}\left(Q^{H*}\right) = c > 0$ while (A.4) gives $\mathrm{MR}\left(Q^{H*}\right) < 0$, a contradiction. 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. ■

A.2 Proof of Proposition 2

Proof. The risk neutrality case satisfies the condition of lemma 1 with

$$f\left[Q_i(c_i), Q_j(0), Q_j(c)\right] = Q_i(\widehat{c}_i)P\left[Q_i(\widehat{c}_i) + Q_j(0)\right].$$

The resulting objective function is equation (8) minus cQ^H. By arguments in the text whenever $Q^H > 0$ exclusion is strictly optimal. From (A.4) and (A.5), $Q^{H*} > 0$ if and only if $r < 1 - \frac{c}{\mathrm{MR}(0)} = r^*$. Otherwise $Q^{H*} = 0$, and M's problem becomes $\max_{\Delta \geq 0} \Delta P(\Delta)$. Since $\mathrm{MR}(0) = P(0) > 0$ it is optimal to set $\Delta > 0$. ■

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 $R_i(\widehat{c}_i) = \pi_i(\widehat{c}_i, c, c_i) - \pi_i(\widehat{c}_i, 0, c_i)$. Solving for W gives

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

Utilities expressed as certainty equivalent incomes satisfy lemma 1 by taking

$$f\left[Q_i(c_i), Q_j(0), Q_j(c)\right] = Q_i(\widehat{c}_i)P\left[Q_i(\widehat{c}_i) + Q_j(0)\right] - \frac{\ln\left[(1-r)\exp\left[-aR_i(\widehat{c}_i)\right] + r\right]}{a}.$$

Replacing $\alpha = \frac{1}{a}$, the objective function becomes

$$r\left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) - cQ^H + (1-r)\left[Q_1^H P\left(Q^H + \Delta_2\right) + Q_2^H P\left(Q^H + \Delta_1\right)\right] -$$
$$\sum_i r\alpha \ln\left[(1-r)\exp\left(-\frac{R_i(0)}{\alpha}\right) + r\right] - \sum_i (1-r)\alpha \ln\left[(1-r)\exp\left(-\frac{R_i(c)}{\alpha}\right) + r\right]. \quad \text{(A.6)}$$

Since $\Delta_i \geq 0$, $R_i(0) \geq 0$ and $R_i(c) \geq 0$, meaning that this expression limits to the objective function in the infinite risk aversion case as $\alpha \to 0$.

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 $Q_i^H \geq 0$ 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 bound between $\ln[r]$ and 0 and that

$$r\left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) - cQ^H + (1-r)\left[Q_1^H P\left(Q^H + \Delta_2\right) + Q_2^H P\left(Q^H + \Delta_1\right)\right] \leq$$
$$r\left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) + (1-r)Q^H P\left(Q^H\right) \equiv rQ^1 P\left(Q^1\right) + (1-r)Q^2 P\left(Q^2\right)$$

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

$$\int_0^{\widetilde{Q}+\mu} MR(v)dv + \int_{\widetilde{Q}+\mu}^{Q^k} MR(v)dv \leq \int_0^{\widetilde{Q}+\mu} MR(v)dv + MR(\widetilde{Q}+\mu)(Q^k - \widetilde{Q} - \mu),$$

which is unbounded below as $Q^k \to \infty$ since $MR(\widetilde{Q}+\mu) < 0$. Since M gets 0 by offering 0 input to all firms, the solution correspondence must be bounded. So we can rewrite the constraints as $Q_i^H \in \left[0, \overline{Q}\right]$ 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

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

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. (11) satisfies the conditions of lemma 1 by taking f as in A.2 but substituting $r = \widehat{r}$. The resulting objective function is

$$r\left[\begin{array}{c}\widehat{r}(Q_1^H+\Delta_1)P(Q^H+\Delta)+(1-\widehat{r})(Q_1^H+\Delta_1)P(Q^H+\Delta_1)-cQ_1^H+\\ \widehat{r}(Q_2^H+\Delta_2)P(Q^H+\Delta)+(1-\widehat{r})(Q_2^H+\Delta_2)P(Q^H+\Delta_2)-cQ_2^H\end{array}\right]+$$
$$(1-r)\left[\begin{array}{c}\widehat{r}Q_1^HP(Q^H+\Delta_2)+(1-\widehat{r})Q_1^HP(Q^H+\Delta_2)-cQ_1^H+\\ \widehat{r}Q_2^HP(Q^H+\Delta_1)+(1-\widehat{r})Q_2^HP(Q^H+\Delta_1)-cQ_2^H.\end{array}\right]$$

After removing

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

one is left with

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

which equals

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

and finally

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

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

$$\begin{aligned}\pi(Q^H,\Delta,\Delta_2)=\;&r\widehat{r}R\left(Q^H+\Delta\right)+(1-r)(1-\widehat{r})R\left(Q^H\right)-cQ^H+\\ &r(1-\widehat{r})\left[R\left(Q^H+\Delta-\Delta_2\right)+R\left(Q^H+\Delta_2\right)\right]+\\ &(\widehat{r}-r)Q^HP\left(Q^H+\Delta_2\right).\end{aligned}\tag{A.7}$$

We will solve the relaxed problem

$$\max_{Q^H\geq0,\Delta\geq0,\Delta_2\geq0}\pi(Q^H,\Delta,\Delta_2)\tag{A.8}$$

and show that all solutions satisfy $\Delta_2\leq\frac{\Delta}{2}$.

In general (A.8) is not a concave problem and 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.7) then becomes $Q^HP(Q^H)-cQ^H$, 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 $Q^H = 0$ and $\Delta_2 = 0$. When $Q^H = 0$ (A.7) becomes

$$r\widehat{r}R(\Delta) + (1-r)(1-\widehat{r})R(0) + r(1-\widehat{r})\left[R(\Delta - \Delta_2) + R(\Delta_2)\right] \tag{A.9}$$

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

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

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial Q^H} = \left\{ \begin{array}{c} r\widehat{r}\mathrm{MR}\left(Q^H + \Delta\right) + (1-r)(1-\widehat{r})\mathrm{MR}\left(Q^H\right) + \\ r(1-\widehat{r})\left[\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) + \mathrm{MR}\left(Q^H + \Delta_2\right)\right] + \\ (\widehat{r} - r)\left[P\left(Q^H + \Delta_2\right) + Q^H P'\left(Q^H + \Delta_2\right)\right] \end{array} \right\} - c < 0 \tag{A.10}$$

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial \Delta} = r\widehat{r}\mathrm{MR}\left(Q^H + \Delta\right) + r(1-\widehat{r})\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) = 0 \tag{A.11}$$

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

are all satisfied at $Q^H = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$. (A.12) is clearly satisfied, while (A.10) and (A.11) rewrite as[27]

$$\widehat{r}\mathrm{MR}(\Delta) + (1-\widehat{r})\mathrm{MR}\left(\frac{\Delta}{2}\right) = 0 \tag{A.13}$$

$$(1-r)(1-\widehat{r})\mathrm{MR}(0) + r(1-\widehat{r})\mathrm{MR}\left(\frac{\Delta}{2}\right) + (\widehat{r} - r)P\left(\frac{\Delta}{2}\right) < c. \tag{A.14}$$

Call the solution to A.13 $\Delta^*(\widehat{r})$. Plugging into A.14 gives the condition on r

$$r > \frac{(1-\widehat{r})\mathrm{MR}(0) + \widehat{r}P\left(\frac{\Delta^*(\widehat{r})}{2}\right) - c}{(1-\widehat{r})\left[\mathrm{MR}(0) - \mathrm{MR}\left(\frac{\Delta^*(\widehat{r})}{2}\right)\right] + P\left(\frac{\Delta^*(\widehat{r})}{2}\right)} = \widetilde{r}(\widehat{r}, c). \tag{A.15}$$

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

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

$\widetilde{r}(\widehat{r}, c) > 0$ if[28]

$$c < (1-\widehat{r})P(0) + \widehat{r}P\left(\frac{\Delta^*(\widehat{r})}{2}\right). \tag{A.17}$$

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.13). 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 $Q^H = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$ does not exist when $r < r'$ and $c < c^*$.

[27] Here we have also plugged (A.10) into (A.11).

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

Another potential solution to (A.8) is an exclusive contract in which $\Delta_2 = 0$. By the above arguments an exclusive contract can only be optimal if $Q^H > 0$. Such an exclusive contract satisfies the Kuhn-Tucker first order conditions if

$$\left[\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial Q^H}\right]_{\Delta_2=0} = 0 \tag{A.18}$$

$$\left[\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial \Delta}\right]_{\Delta_2=0} = 0 \tag{A.19}$$

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

which simplifies to

$$r\text{MR}\left(Q^H + \Delta\right) + (1-r)\text{MR}\left(Q^H\right) = c \tag{A.21}$$

$$r\text{MR}\left(Q^H + \Delta\right) = 0 \tag{A.22}$$

$$r(1-\widehat{r})\left[-\text{MR}\left(Q^H + \Delta\right) + \text{MR}\left(Q^H\right)\right] + (\widehat{r} - r)Q^H P'\left(Q^H\right) < 0. \tag{A.23}$$

and further to

$$(1-r)\text{MR}\left(Q^H\right) = c \tag{A.24}$$

$$r\text{MR}\left(Q^H + \Delta\right) = 0 \tag{A.25}$$

$$r(1-\widehat{r})\text{MR}\left(Q^H\right) + (\widehat{r} - r)Q^H P'\left(Q^H\right) < 0. \tag{A.26}$$

Let $Q^{H*}(r)$ be the solution to (A.24). Since $\text{MR}(Q^{H*}) = P(Q^{H*}) + Q^{H*}P'(Q^{H*})$ (A.26) is satisfied whenever

$$\widehat{r} > \frac{rP\left(Q^{H*}(r)\right)}{r\text{MR}\left(Q^{H*}(r)\right) - Q^{H*}(r)P'(Q^{H*}(r))} = f(r) \in (r, 1). \tag{A.27}$$

Let $\widehat{r}' = \max_r f(r)$. Now, because marginal revenue is decreasing, $Q^{H*}(r)$ is decreasing in r. $Q^{H*}(r) > 0$ 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 $Q^{H*}(r'') = 0$. Define $r^* = \min\{r', r''\}$.

The final solution to consider is one in which no boundary solutions to (A.8) exist. The resulting system of equations simplifies to

$$\left\{ \begin{array}{c} (1-r)(1-\widehat{r})\text{MR}\left(Q^H\right) + r(1-\widehat{r})\text{MR}\left(Q^H + \Delta_2\right) + \\ (\widehat{r} - r)\left[P\left(Q^H + \Delta_2\right) + Q^H P'\left(Q^H + \Delta_2\right)\right] \end{array} \right\} = c \tag{A.28}$$

$$\widehat{r}\text{MR}\left(Q^H + \Delta\right) + (1-\widehat{r})\text{MR}\left(Q^H + \Delta - \Delta_2\right) = 0 \tag{A.29}$$

$$r(1-\widehat{r})\left[-\text{MR}\left(Q^H + \Delta - \Delta_2\right) + \text{MR}\left(Q^H + \Delta_2\right)\right] + (\widehat{r} - r)Q^H P'\left(Q^H + \Delta_2\right) = 0. \tag{A.30}$$

Since $P' < 0$, (A.30) implies that $\text{MR}\left(Q^H + \Delta_2\right) > \text{MR}\left(Q^H + \Delta - \Delta_2\right)$ 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.8) 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.30) 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.13) becomes

$$\widehat{r}(1 - 2\Delta) + (1 - \widehat{r})(1 - \Delta) = 0 \tag{A.31}$$

or $\Delta = \frac{1}{1+\widehat{r}}$. Plugging this expression in (A.15) gives

$$\begin{aligned}\widetilde{r}(\widehat{r}, c) = \frac{(1-\widehat{r}) + \widehat{r}\left(1 - \frac{1}{2(1+\widehat{r})}\right) - c}{(1-\widehat{r})\left[1 - \left(1 - \frac{1}{1+\widehat{r}}\right)\right] + \left(1 - \frac{1}{2(1+\widehat{r})}\right)} = \frac{1 - \frac{\widehat{r}}{2(1+\widehat{r})} - c}{(1-\widehat{r})\left(\frac{1}{1+\widehat{r}}\right) + 1 - \frac{1}{2(1+\widehat{r})}} = \\ \frac{2(1+\widehat{r}) - \widehat{r} - 2(1+\widehat{r})c}{1 + 2(1+\widehat{r}) - 2\widehat{r}} = \frac{2 + \widehat{r} - 2(1+\widehat{r})c}{3}.\end{aligned} \tag{A.32}$$

Next consider the exclusive contract solution. (A.24) 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.27) gives

$$f(r) = \frac{r\left[\frac{2-2r-(1-r-c)}{2(1-r)}\right]}{r\left[\frac{1-r-(1-r-c)}{(1-r)}\right] + \frac{1-r-c}{2(1-r)}} = \frac{r(1-r+c)}{1-r-c+2cr}. \tag{A.33}$$

Now finally consider the solution with partial exclusion. Expressions (A.28)-(A.30) solve as

$$Q^* = \frac{r(1-\widehat{r})(2 - 3r + \widehat{r} - 2c(1+\widehat{r}))}{4r - 2r\widehat{r} - \widehat{r}^2 + r^2(-5 + 4\widehat{r})} > 0 \tag{A.34}$$

$$\Delta^* = \frac{-r^2 + r(2 + 6c - 6c\widehat{r}) + \widehat{r}(-2 + 2c(1-\widehat{r}) + \widehat{r})}{8r - 4r\widehat{r} - 2\widehat{r}^2 + 2r^2(-5 + 4\widehat{r})} > 0 \tag{A.35}$$

$$\Delta_2^* = \frac{-r^2 - (1-c)\widehat{r} + r(1 + c + \widehat{r} - 2c\widehat{r})}{4r - 2r\widehat{r} - \widehat{r}^2 + r^2(-5 + 4\widehat{r})} > 0. \tag{A.36}$$

Whenever $r < \frac{2+\widehat{r}-2(1+\widehat{r})c}{3}$ the numerator of (A.34) is positive. The condition for the denominator to be positive is that

$$r \leq \widehat{r} \leq -r(1-2r) + 2(1-r)\sqrt{r(1+r)} < 1. \tag{A.37}$$

(A.35) is positive when

$$r \leq \widehat{r} \leq \frac{1-c+3cr-\sqrt{(1-r)^2(1-2c)+c^2(1+3r)^2}}{1-2c}, \tag{A.38}$$

which is more stringent than the condition in (A.37). Finally, the condition for (A.36) positive is

$$r \leq \widehat{r} \leq \frac{r(1+c-r)}{1-c-r+2cr} \equiv \widehat{r}^*. \tag{A.39}$$

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

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

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 $T(\widehat{r}^*) = 0$ 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.

$$\max_{\{Q_i(\widehat{c}_i),T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^{2} rT_i(0)+(1-r)T_i(c) \text{ such that} \tag{A.40}$$

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

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

$$Q_i^H\geq 0, Q_i^H+\Delta_i\geq 0 \qquad (NN)$$

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

One can express transfers as

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

and

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

Plugging back into the objective function gives

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

which rewrites as

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

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

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

Substituting $\Delta_1 = \Delta - \Delta_2$ gives the expression

$$\begin{aligned}\pi(Q^H, \Delta, \Delta_2) = {} & r^2 R\left(Q^H + \Delta\right) - r(1-r) R\left(Q^H\right) - cQ^H + \\ & r(1-r)\left[R\left(Q^H + \Delta - \Delta_2\right) + R\left(Q^H + \Delta_2\right)\right] + \\ & (1-r) Q^H P\left(Q^H + \Delta_2\right)\end{aligned} \tag{A.41}$$

We will solve the relaxed problem

$$\max_{Q^H \geq 0, \Delta \geq 0, \Delta_2 \geq 0} \pi(Q^H, \Delta, \Delta_2) \tag{A.42}$$

and ignore the constraint $\Delta_2 \leq \frac{\Delta}{2}$, and show that all solutions 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 $Q^H = 0$ and $\Delta_2 = \frac{\Delta}{2}$ satisfies the Kuhn-Tucker first order conditions if

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial Q^H} = \left\{ \begin{array}{c} r^2 \mathrm{MR}\left(Q^H + \Delta\right) - r(1-r)\mathrm{MR}\left(Q^H\right) + \\ r(1-r)\left[\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) + \mathrm{MR}\left(Q^H + \Delta_2\right)\right] + \\ (1-r)\left[P\left(Q^H + \Delta_2\right) + Q^H P'\left(Q^H + \Delta_2\right)\right] \end{array} \right\} - c < 0 \tag{A.43}$$

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial \Delta} = r^2 \mathrm{MR}\left(Q^H + \Delta\right) + r(1-r)\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) = 0 \tag{A.44}$$

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

are all satisfied at $Q^H = 0$ and $\Delta_2 = \frac{\Delta}{2} > 0$. (A.45) is clearly satisfied, while (A.43) and (A.44) rewrite as[29]

$$r\mathrm{MR}\left(\Delta\right) + (1-r)\mathrm{MR}\left(\frac{\Delta}{2}\right) = 0 \tag{A.46}$$

$$-r(1-r)\mathrm{MR}\left(0\right) + r(1-r)\mathrm{MR}\left(\frac{\Delta}{2}\right) + (1-r)P\left(\frac{\Delta}{2}\right) < c. \tag{A.47}$$

This solution cannot exist for r sufficiently small 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.42) is an exclusive contract in which $\Delta_2 = 0$. By the above arguments an exclusive contract can only be optimal if $Q^H > 0$. Such an exclusive contract satisfies the Kuhn-Tucker first order conditions if

$$\left[\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial Q^H}\right]_{\Delta_2 = 0} = 0 \tag{A.48}$$

$$\left[\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial \Delta}\right]_{\Delta_2 = 0} = 0 \tag{A.49}$$

[29]Here we have also plugged (A.43) into (A.44).

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

which simplifies to

$$r\text{MR}\left(Q^H+\Delta\right)+(1-r)\text{MR}\left(Q^H\right)=c \tag{A.51}$$

$$r\text{MR}\left(Q^H+\Delta\right)=0 \tag{A.52}$$

$$r(1-r)\left[-\text{MR}\left(Q^H+\Delta\right)+\text{MR}\left(Q^H\right)\right]+(1-r)Q^HP'\left(Q^H\right)<0. \tag{A.53}$$

and further to

$$(1-r)\text{MR}\left(Q^H\right)=c \tag{A.54}$$

$$\text{MR}\left(Q^H+\Delta\right)=0 \tag{A.55}$$

$$r\text{MR}\left(Q^H\right)+Q^HP'\left(Q^H\right)<0. \tag{A.56}$$

This solution clearly exists when r is small.

The final solution to consider is one in which no boundary solutions to (A.42) exist. The resulting system of equations simplifies to

$$\left\{\begin{array}{c}-r(1-r)\text{MR}\left(Q^H\right)+r(1-r)\text{MR}\left(Q^H+\Delta_2\right)+\\(1-r)\left[P\left(Q^H+\Delta_2\right)+Q^HP'\left(Q^H+\Delta_2\right)\right]\end{array}\right\}=c \tag{A.57}$$

$$r\text{MR}\left(Q^H+\Delta\right)+(1-r)\text{MR}\left(Q^H+\Delta-\Delta_2\right)=0 \tag{A.58}$$

$$r(1-r)\left[-\text{MR}\left(Q^H+\Delta-\Delta_2\right)+\text{MR}\left(Q^H+\Delta_2\right)\right]+(1-r)Q^HP'\left(Q^H+\Delta_2\right)=0. \tag{A.59}$$

Since $P'<0$, (A.59) implies that $\text{MR}\left(Q^H+\Delta_2\right)>\text{MR}\left(Q^H+\Delta-\Delta_2\right)$ 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.59) 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 $\bar{c}_1=\inf\left(P(Q^{H*})-P\left(2Q^{H*}+\Delta^*\right)\right)$. For any $c\leq\bar{c}_1$, 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 $\bar{c}_2=\inf\left[P(Q^{H*}+\Delta^*)-P\left(2Q^{H*}+2\Delta^*\right)\right]$. For any $c\leq\bar{c}_2$, the IC_{LL}

is satisfied.

Define $\bar{c} = \min\{\bar{c}_1, \bar{c}_2\}$. If $c \leq \bar{c}$, 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 $Q(c_j)$ be total output of cost type c_j; let $\Delta_i^j = Q_i(c_j) - Q_i(c_{j-1})$ (defined for $j > 1$); 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 let $\Delta_{-i}^j = \sum_{k \neq i} \Delta_k^j$ 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

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

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

$$Q_i(c_j)\left\{P\left[Q_i(c_j) + \overline{Q}_{-i}\right] - c_j\right\} - T_i(c_j) \geq Q_i(c_l)\left\{P\left[Q_i(c_l) + \overline{Q}_{-i}\right] - c_j\right\} - T_i(c_l) \qquad (IC_i^{jl})$$

$$Q_i(c_j) \geq 0. \qquad (NN)$$

where $\overline{Q}_{-i}$ is the maximum amount produced by all firms other than i.

Lemma 2 *The optimal transfers are given by*

$$T_i(c_1) = Q_i(c_1)\left\{P\left[Q_i(c_1) + \sum_{k \neq i} Q_k(c_N)\right] - c_1\right\}$$
$$T_i(c_j) = Q_i(c_j)\left\{P\left[Q_i(c_j) + \sum_{k \neq i} Q_k(c_N) - c_j\right]\right\} - \sum_{l=1}^{j-1} Q_i(c_l)(c_l - c_{l+1}) \ \forall j > 1$$

Proof. Following steps similar to those in the proof of lemma 1, one can show that a necessary condition for incentive compatibility is $Q_i(c_{j+1}) \geq Q_i(c_j)$. So $\overline{Q}_{-i} = \sum_{k \neq i} Q_k(c_N)$. Moreover, the optimal transfers make the participation constraint of type c_1 bind (from which the expression for $T_i(c_1)$ follows directly), as well as the upward local incentive compatibility constraint $(IC_i^{j,j+1})$ for all other types. The transfer for cost type c_2 is obtained from

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

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

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

So the proof follows by induction. ■

Clearly aggregate production costs and information rents are independent of output distribution. Expected revenue is given by

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

The distribution of output only depends on the first two terms. Observe that

$$\sum_{i=1}^{F} Q_i(c_1) P \left[Q(c_1) + \sum_{j=2}^{N} \Delta_{-i}^{j} \right] \leq Q(c_1) P \left[Q(c_1) \right] \tag{A.60}$$

and

$$\begin{aligned}
& \sum_{j=2}^{N-1} r_j \sum_{i=1}^{F} \left(Q_i(c_1) + \sum_{l=2}^{j} \Delta_i^l \right) P \left[Q(c_1) + \sum_{l=2}^{j} \Delta^l + \sum_{l=j+1}^{N} \Delta_{-i}^l \right] \leq \\
& \sum_{j=2}^{N-1} r_j \left(Q(c_1) + \sum_{l=2}^{j} \Delta^l \right) P \left[Q(c_1) + \sum_{l=2}^{j} \Delta^l \right].
\end{aligned} \tag{A.61}$$

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 both upper bounds, so exclusive contracts are optimal. Moreover the inequality in (A.61) is strict whenever $\Delta^{N-1} > 0$, so the only circumstance under which exclusive contracts are not *strictly* optimal is when M only wants to contract positive aggregate output from cost type c_N.

Let $Q^j = Q(c_j)$. When contracting with one firm, M chooses Q^j to maximize

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

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

$$\begin{aligned} \mathrm{MR}(Q^N) - c_N &= 0 \qquad \text{(A.62)} \\ r_j\left[\mathrm{MR}(Q^j) - c_j\right] - (c_j - c_{j+1}) \sum_{l=0}^{N-j-1} r_{N-l} &= \\ r_j \mathrm{MR}(Q^j) - c_j \sum_{l=0}^{N-j} r_{N-j} + c_{j+1} \sum_{l=0}^{N-j-1} r_l &< 0, \ \forall\, j < N. \qquad \text{(A.63)} \end{aligned}$$

Now, plugging in $Q^j = 0$ for $j < N$ and adding together equations (A.63) gives

$$\sum_{j=1}^{N-1} r_j \mathrm{MR}(0) - c_1 + c_N r_N < 0. \qquad \text{(A.64)}$$

As $r_N \to 0$ we must have $\sum_{j=1}^{N-1} r_j \to 1$ and (A.64) 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 Production Only 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

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

and the lottery it faces is

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

Given an organization, M's problem is

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

$$\begin{aligned} \text{such that} \qquad & U\left[L_i(0 \mid 0, x_i)\right] \geq U[0] && (PC_i^L) \\ & U\left[L_i(c \mid c, x_i)\right] \geq U[0] && (PC_i^H) \\ & U\left[L_i(0 \mid 0, x_i)\right] \geq U\left[L_i(c \mid 0, x_i)\right] && (IC_i^L) \\ & U\left[L_i(c \mid c, x_i)\right] \geq U\left[L_i(0 \mid c, x_i)\right]. && (IC_i^H) \end{aligned}$$

$Q_i(\widehat{c}_i) P\left[Q_i(\widehat{c}_i) + Q_j(\widehat{c}_j)\right]$ 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) solve (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 3 *Suppose that* $r < \frac{1-c}{1+c}$.

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

Proof. Let $(x_1, x_2) = (\text{IND}, \text{IND})$. The optimal exclusive contract is given by equations (A.54) and (A.55) 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

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

which when plugged into M's objective function gives

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

which rewrites as

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

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

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

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

$$r^2(Q^H+\Delta)(1-Q^H-\Delta)+(1-r)^2Q^H(1-Q^H)+2r(1-r)\left(Q^H+\frac{\Delta}{2}\right)\left(1-Q^H-\frac{\Delta}{2}\right)$$
$$-cQ^H-\frac{r(\widehat{r}-r)\Delta^2}{4}. \tag{B.4}$$

The first order conditions for Q^H and Δ are

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

Solving these two together gives

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

Let $(x_1,x_2)=(\text{INT},\text{INT})$. The derivation of the optimal contracts is identical to the previous taking $\widehat{r}=r$. ■

Now that we have established optimal contracts in each organization, we can provide conditions under which $(x_1,x_2)=(\text{INT},\text{INT})$ maximizes joint utility.

Proposition 9 *There exists a $\bar{c}>0$ and $\underline{r}<1$ such that $(x_1^*,x_2^*)=(IND,IND)$ when $c<\bar{c}$ and $r>\underline{r}$.*

Proof. We will first compare organizations (IND, IND) and (INT, IND). M's profit in (IND, IND) is equivalent to its profit from contracting with one risk-neutral firm since it always offers an exclusive contract. As explained on page 7, M's profit in this case is

$$\pi_M(\text{IND},\text{IND})=\mathbb{E}[\,Q(\text{IND},\text{IND})\,]-\mathbb{E}[\,Q(\text{IND},\text{IND})\,]^2-r(1-r)\Delta(\text{IND},\text{IND})-cQ^H(\text{IND},\text{IND}).$$

Expression (B.4) shows that M's profit in the (INT, IND) organization is equal to its profit from contracting two risk-neutral firms minus a risk premium term, so that M's profit can be written as

$$\pi_M(\text{IND}, \text{IND}) = \mathbb{E}[\,Q(\text{INT}, \text{IND})\,] - \mathbb{E}[\,Q(\text{INT}, \text{IND})\,]^2 - r(1-r)\sum_i \Delta_i^2(\text{INT}, \text{IND}) - \frac{r(\widehat{r}-r)\Delta(\text{INT}, \text{IND})^2}{4} - cQ^H(\text{INT}, \text{IND}).$$

Since $\mathbb{E}[\,Q(\text{IND}, \text{IND})\,] = \mathbb{E}[\,Q(\text{INT}, \text{IND})\,]$ the difference in profits only depends on the Δ and Q^H terms.

Consider now downstream profits. In (IND, IND), each downstream firm earns zero profit in all events except the one in which it is low cost and receives the exclusive contract, in which case it earns the information rent $cQ^H(\text{IND}, \text{IND})$. This event occurs with probability $0.5r$. So, following the general RDU formulation presented in footnote 18, joint downstream profits are

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

In (INT, IND) the integrated firm earns an information rent with probability r and the independent firm always earns 0. So joint downstream profits are

$$\pi_D(\text{INT}, \text{IND}) = \left[1 - \widehat{r}(1-r)\right] c \left(\frac{1-c}{2} - \frac{rc}{1-r+\frac{\widehat{r}-r}{2}}\right).$$

Note that $1 - \widehat{r}(1-r) = \frac{r(r-c)}{c+r-2cr}$, $2\left[1-\widehat{r}(1-0.5r)\right] = \frac{r(0.5r-c)}{c+0.5r-cr}$, and

$$\frac{r(r-c)}{c+r-2cr} - \frac{r(0.5r-c)}{c+0.5r-cr} = \frac{r^2c(1-c)}{(c+r-2cr)(c+0.5r-cr)} > 0.$$

So

$$\pi_D(\text{IND}, \text{IND}) - \pi_D(\text{INT}, \text{IND}) > \frac{rc^2(0.5r-c)}{c+0.5r-cr}\left(\frac{r}{1-r+\frac{\widehat{r}-r}{2}} - \frac{r}{1-r}\frac{1}{2}\right) - \frac{1}{2}\frac{r^2c^2(1-c)^2}{(c+r-2cr)(c+0.5r-cr)}.$$

Now observe that

$$\lim_{c\to 0}\left[\begin{array}{c}\frac{\pi_M(\text{IND},\text{IND})-\pi_M(\text{INT},\text{IND})}{c^2}+ \\ \frac{r(0.5r-c)}{c+0.5r-cr}\left(\frac{r}{1-r+\frac{\widehat{r}-r}{2}} - \frac{r}{1-r}\frac{1}{2}\right) - \frac{1}{2}\frac{r^2(1-c)^2}{(c+r-2cr)(c+0.5r-cr)}\end{array}\right] =$$

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

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

This last expression is clearly positive for high enough r. So there exists some $\overline{c}^1$ and $\underline{r}^1$ such that (IND, IND) produces higher joint utility than (INT, IND) whenever $c < \overline{c}^1$ and $r > \underline{r}^1$.

$\pi_M(\text{INT}, \text{INT})$ and $\pi_D(\text{INT}, \text{INT})$ are obtained by evaluating $\pi_M(\text{INT}, \text{IND})$ and $\pi_D(\text{INT}, \text{IND})$ at $\widehat{r} = r$. One can follow exactly the same steps to show that there exists some $\overline{c}^2$ and $\underline{r}^2$ such that (IND, IND) produces higher joint utility than (INT, INT) whenever $c < \overline{c}^2$ and $r > \underline{r}^2$. The proof is completed by taking $\underline{r} = \max\{\underline{r}^1, \underline{r}^2\}$ and $\overline{c} = \min\{\overline{c}^1, \overline{c}^2\}$. ■