Vertical Exclusion with Endogenous Competition Externalities*

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

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

February 17, 2016

Abstract

An upstream firm with full commitment bilaterally contracts with two ex ante identical downstream firms. Each observes its own cost shock, and faces uncertainty from its competitor's shock. When they are risk neutral and can absorb losses, the upstream firm contracts symmetric outputs for production efficiency. However, when they are risk averse, competition requires the payment of a risk premium due to revenue uncertainty. Moreover, when they enjoy limited liability, competition requires the upstream firm to share additional surplus. To resolve these trade-offs, the upstream firm offers exclusive contracts in many cases.

Keywords: Exclusive Contracts, Risk, Limited Liability

JEL Codes: L22, L42, D86

*The authors gratefully acknowledge financial aid from the Spanish Ministry of Science and Innovation (projects ECO2010-15052 and ECO2013-43011-P). We would like to thank Larbi Alaoui for extensive comments on previous versions of this paper, as well as Mark Armstrong, Giacomo Calzolari, Yutian Chen, Simon Cowan, Jan Eeckhout, Péter Eső, Ricard Gil, Liliane Karlinger, Clare Leaver, Marco Ottaviani, Patrick Rey, Monika Schnitzer, Emanuele Tarantino, John Thanassoulis, participants of the ANR-DFG workshop (Paris), and seminar attendees at Bologna, Carlos III, Mannheim, and Oxford for helpful advice.

†Massimo Motta is currently Chief Competition Economist at the European Commission. Obviously, the views expressed here are personal, and cannot be interpreted as representing those of the European Commission.

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. In many cases, a situation arises in which only one agent will deal with the principal's product, brand, or technology.[1] This paper provides a novel rationale for the optimality of such exclusive relationships and, more generally, asymmetric market shares at the retail level. We argue that competition with imperfectly correlated and privately observed cost or demand shocks necessarily creates uncertainty for downstream retailers, and that exclusive contracts can therefore benefit upstream firms when these retailers are subject to risk aversion or limited liability.

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 production levels depend on the realization of individual shocks that are not fully observable to competitors, 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 firms, the upstream firm determines the magnitude of the uncertainty that downstream firms face. In short, much of the literature focuses on the effect of competition on average downstream profits, or the first moment of the payoff distribution. We instead study the effect of competition on the second moment, and the resulting implications for market structure.

When downstream firms are risk neutral and can absorb losses, the upstream firm has an inherent incentive to offer all of them the input because doing so ensures that whenever a more productive firm exists it serves some of the market.[2] When downstream firms are risk averse, there is a cost to the upstream firm of contracting more than one: 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 all but one downstream firm, the upstream firm eliminates the competition externality, and with it the uncertainty from competition that downstream firms face.[3] A similar mechanism operates when downstream firms enjoy limited liability. The resulting

[1] According to Lafontaine and Slade (2008), one third of retail sales through independent outlets occur in exclusive relationships. Further evidence on this point comes from Blair and Lafontaine (2011), who 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] One can think of other mechanisms, such as product differentiation, that would also generate the optimality of offering more than one firm the input.

[3] We also show that, for intermediate risk preferences and two firms, partial exclusion arises, with one downstream firm producing more for all shock realizations.

surplus that the upstream firm must pay out when contracting with two firms more than offsets the gain from production efficiency, which induces it to offer an exclusive contract under certain conditions.

This paper builds on a long literature on bilateral contracting[4] in vertical markets (Hart and Tirole 1990, McAfee and Schwartz 1994, Segal 1999, Rey and Tirole 2007) that largely focuses on exclusion as a response to commitment problems. When the upstream firm can commit to public bilateral contracts, it can always extract the monopoly profit; but when it offers unobservable bilateral contracts, it cannot commit not to renegotiate with downstream firms and does not obtain the monopoly profit. Exclusive contracts serve as a commitment device to restore monopoly profit. Our mechanism is wholly different from this one. In the model the upstream firm has full commitment, but chooses exclusive outcomes to reduce uncertainty in the downstream market.

The most closely related paper to ours in the literature is Rey and Tirole (1986), who also study a vertical market with bilateral contracting in which downstream firms are subject to shocks. In their model, when downstream firms are infinitely risk averse the upstream firm allows them to compete, while with risk neutrality it offers exclusive territories. The key difference with our paper is that in Rey and Tirole (1986) downstream shocks are perfectly correlated so that market structure does not affect downstream uncertainty. We instead show that competition creates uncertainty when shocks are imperfectly correlated, and as a result arrive at the exact opposite conclusion to that of Rey and Tirole (1986). In our model, when downstream firms are risk neutral the upstream firm allows competition, but when they are infinitely risk averse it provides a fully exclusive contract. Hence we argue that competition uncertainty creates fundamentally new effects in vertical markets as yet unrecognized in the literature.

A nascent literature has begun to study mechanism design with bilateral contracting. For example, Dequiedt and Martimort (2015) study an environment in which risk-neutral downstream firms with privately observed, heterogeneous costs can make type reports to an upstream firm, which then chooses an allocation for each without any ex ante commitment.[5] When such communication is allowed, the upstream firm allocates an exclusive contract to the most efficient firm in what resembles an all-pay auction.[6] In

[4]A bilateral contract between an upstream and downstream firm cannot directly depend on the outputs or messages of other firms. Motivations for this restriction include the transaction costs associated with writing and enforcing multilateral contracts, and the possibility that multilateral contracts might facilitate collusion. Although vertical contracts are typically regarded as having less anticompetitive potential than horizontal contracts, antitrust authorities are often concerned with *contracts that reference rivals*, that is, vertical contracts between a buyer and a seller whose terms may depend on information or contract terms pertaining to the buyer's rivals (as in the situation depicted in our paper) or the seller's rivals (Scott Morton 2012).

[5]The typical mechanism design literature studies multilateral contracting problems in which the upstream firm can commit to a mapping from the vector of type reports to a vector of allocations.

[6]This outcome is similar to that with multilateral contracting. See McAfee and McMillan (1986),

our paper we rule out such communication and adopt what Dequiedt and Martimort (2015) call "simple bilateral contracts" in which each downstream firm's allocation only depends on its reported type. This allows us to cleanly study the impact of competition on uncertainty, but we return to this restriction in the discussion at the end of the paper.

Finally, while limited liability is a rather standard assumption in the industrial-organization theory literature, risk aversion is not. However, the empirical literature has long recognized its importance. For example, the majority of exclusive retailing occurs in franchise networks (Lafontaine and Slade 2008), within which the owners of retail outlets are typically small and undiversified; some franchisors even explicitly seek out retailers whose incomes are highly correlated with their outlets' performance (Kaufmann and Lafontaine 1994). Asplund (2002) and Banal-Estañol and Ottaviani (2006) also present numerous references to empirical studies documenting the relevance of firm risk aversion.[7] More generally, Nocke and Thanassoulis (2014) provide theoretical foundations for introducing curvature into downstream firms' payoff functions. They show that when downstream firms face credit constraints subsequent to competing in the downstream market, they behave as if they were risk averse even if they are risk neutral.

The paper is organized thus. Section 2 describes the model and presents a solution of the baseline case without risk aversion or limited liability. Section 3 then considers the impact of risk aversion on exclusive arrangements, while section 4 considers the impact of limited liability. Section 5 provides a discussion and concludes. *Appendix A contains all proofs.*

2 Model

Consider a vertical market in which an upstream firm supplies an input that is transformed into output in a one-to-one relationship by two downstream firms $i = 1, 2$. Aggregate demand for the product is $P(Q)$ where $Q \geq 0$ is aggregate quantity. We assume that $P'(Q) < 0$; that marginal revenue $MR(Q) \equiv P(Q) + QP'(Q)$ is decreasing; and that there exists some finite quantity $\overline{Q}$ at which $MR(\overline{Q}) = 0$. This last assumption implies that $QP(Q) \to -\infty$ as $Q \to \infty$.

Each downstream firm has a constant marginal cost of production $c_i \in \{0, c\}$ where $c > 0$. Each firm observes the realization of its own cost shock, but not that of its competitor. The upstream firm observes neither shock. Shocks are iid with $\Pr[\, c_i = 0 \,] = r$. Constant returns to scale keeps aggregate production costs independent of the distribution of output

Laffont and Tirole (1987), McAfee and McMillan (1987), Riordan and Sappington (1987), and Dasgupta and Spulber (1989).

[7]They also make a related point to ours in horizontal markets by observing that mergers have a role to play in reducing the uncertainty that firms face.

across downstream firms, so that we can isolate the impact of revenue uncertainty. The interpretation of c_i as a cost shock is simply for concreteness, as it can equally represent a firm-specific demand shock.[8] Finally, we assume that $c < P(\overline{Q})$.[9]

The upstream firm offers a contract to downstream firm i that is equivalent to a nonlinear tariff of the quantity Q_i firm i purchases. Given these contracts, firms engage in Cournot competition and simultaneously choose outputs. Profits are then realized and payments are made to the upstream firm. Downstream firms can guarantee zero profits by exiting the market without producing.

By the revelation principle, we can focus on the upstream firm's offering downstream firm i an incentive-compatible two-point contract $[Q_i(\hat{c}_i), T_i(\hat{c}_i)]$ for $\hat{c}_i \in \{0, c\}$ in which i prefers to truthfully report its realized marginal cost. Let

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

be firm i's profit from reporting cost type $\widehat{c}_i$ when the competitor reports cost type $\widehat{c}_j$ and firm i has marginal cost c_i. We term the impact of firm j's choice of $\widehat{c}_j$ on π_i a *competition externality.* A key feature of the model is the uncertainty regarding the competitor's cost shock which in turn generates revenue uncertainty due to the competition externality.

In incentive-compatible contracts, firm i faces the lottery

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

when reporting cost type $\widehat{c}_i$. We assume firms use the constant absolute risk aversion (CARA) utility function $u(\pi_i) = -\exp(-a\pi_i)$ to evaluate the expected utility of the lottery, which we denote $U\left[L_i(\widehat{c}_i \mid c_i)\right]$. The a parameter is the coefficient of absolute risk aversion, and higher values indicate more risk aversion. a is common knowledge and shared by both downstream firms.

The upstream firm's problem can be written as

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

$$U\left[L_i(c_i \mid c_i)\right] \geq 0 \tag{PC}$$

$$U\left[L_i(c_i \mid c_i)\right] \geq U\left[L_i(\hat{c}_i \mid c_i)\right] \ \text{ for } \hat{c}_i \neq c_i \tag{IC}$$

$$Q_i(c_i) \geq 0. \tag{NN}$$

[8]Consider a differentiated products model in which firm i's demand is $p_i = v_i - Q_i - \gamma \sum_{j \neq i} Q_j$ and its profit is $\pi = (p_i - c_i)Q_i$. For $\gamma \to 1$, whether the shock is on v_i or c_i is formally equivalent.

[9]This is a standard condition. It implies that low cost firms face competition from high cost firms in the sense that the high cost firm can still profitably produce when the low cost firm chooses its monopoly quantity, which is precisely $\overline{Q}$.

Since firms can earn zero profit from exiting the market, their contracts must provide them at least 0 to ensure participation. The IC constraints ensure firms report their realized cost shock truthfully, while the NN constraints express non-negativity constraints on output.

In an alternative formulation of the model, we assume that firms are risk neutral ($a = 0$) but cannot be forced to absorb losses. Absent limited liability, the participation constraints allow firm i profits to be negative when facing a low-cost competitor and positive when facing a high-cost competitor so long as the expected utility of both events provides a payoff equivalent to leaving the market. In practice limited-liability constraints might make this infeasible. To incorporate these into the model, we follow the approach of Demougin and Garvie (1991) and introduce non-negativity constraints on profits, which we analyze in section 4. More specifically, we replace the PC constraints in (3) with

$$\min\{\pi_i(c_i, 0, c_i), \pi_i(c_i, c, c_i)\} \geq 0 \tag{LL}$$

for each c_i. One can interpret these as *ex post* participation constraints that allow downstream firms to exit the market and receive a zero payoff after observing the realization of profits. In contrast, the PC constraints in (3) are interim participation constraints that allow exit after a negative cost shock, but not after a negative revenue shock induced by an efficient competitor.[10]

Both downstream firms are symmetric. They have the same distribution over cost shocks and the same utility function over lotteries. We are interested in situations in which the upstream firm nevertheless induces asymmetric outcomes in the downstream market due to revenue uncertainty. There are two relevant definitions of 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 fully excluded if* $Q_i^*(0) = Q_i^*(c) = 0$.
2. *Firm i is partially excluded if* $Q_i^*(c_i) < Q_j^*(c_j)$ *for* $j \neq i$ *whenever* $c_i = c_j$.

We emphasize that both types of exclusion are implicit in the sense that the firm chooses contracts that induce asymmetric outcomes in the Cournot game played between downstream firms. Moreover, these outcomes need not be implemented by the direct mechanism we study in the paper.[11]

[10] This discussion presumes that more efficient firms produce more in equilibrium, which we will indeed show is true in the formal results below.

[11] We provide one example of indirect implementation of fully exclusive contracts in Hansen and Motta (2012). The upstream firm can design uniform contracts that do not price discriminate between downstream firms in such a way that one chooses to enter the market and produce while the other best responds by exiting.

2.1 Simplified program

In lottery (2), the uncertainty comes from the revenue side, whereas the production cost $Q_i(\widehat{c}_i)c_i$ and transfer $T_i(\widehat{c}_i)$ are deterministic. CARA utility allows one to linearly separate these and represent expected utility as

$$U\left[L_i(\widehat{c}_i \mid c_i)\right] = \mathrm{CertRev}_i(\widehat{c}_i) - Q_i(\widehat{c}_i)c_i - T_i(\widehat{c}_i).$$

Here $\mathrm{CertRev}_i(\widehat{c}_i)$ is certainty-equivalent revenue, or the fixed payment net of production and transfer costs that gives the firm the same expected payoff as (2).

This representation of expected utility allows one to write program (3) in a simpler way using standard arguments from the mechanism design literature. The basic idea is to solve the principal's program considering only the participation constraints of the high-cost firms and incentive-compatibility constraints of the low-cost firms.

Lemma 1 *Program* (3) *is equivalent to maximizing*

$$\sum_i \left[r\,\mathrm{CertRev}_i(0) + (1-r)\,\mathrm{CertRev}_i(c)\right] - \sum_i cQ_i(c) \tag{4}$$

such that $Q_i(0) \geq Q_i(c) \geq 0$.

The representation of the upstream firm's objective function in lemma 1 has an intuitive form. The first summation is expected certainty-equivalent revenue. The second summation represents the part of revenue that cannot be appropriated with the transfer payments. This consists of expected production costs, which only high-cost downstream firms incur, and so equal $\sum_i (1-r)cQ_i(c)$. Also, high-cost firms make zero profit (their participation constraints are binding), but low-cost downstream firm i makes a profit, or information rent, of $cQ_i(c)$. So expected downstream profit is $\sum_i rcQ_i(c)$. As for the constraints, $Q_i(0) \geq Q_i(c)$ is a necessary condition for incentive compatibility, and means that efficient firms indeed produce more than inefficient ones in equilibrium.

2.2 Baseline solution

In section 3 we study how exclusion depends on risk aversion, and in section 4 how it depends on limited liability. First, though, we explore a baseline in which neither friction operates. In other words, certainty equivalent revenue is simply expected revenue (which corresponds to the case where $a \to 0$)

$$\mathrm{CertRev}_i(\widehat{c}_i) = rQ_i(\widehat{c}_i)P[Q_i(\widehat{c}_i) + Q_j(0)] + (1-r)Q_i(\widehat{c}_i)P[Q_i(\widehat{c}_i) + Q_j(c)]$$

and inefficient firms can make losses ex post.

In deriving optimal contracts, it is useful to define $Q_i^H \equiv Q_i(c)$, $\Delta_i \equiv Q_i(0) - Q_i^H$, $Q^H \equiv Q_1^H + Q_2^H$, and $\Delta \equiv \Delta_1 + \Delta_2$. Q^H and Δ are aggregate production variables, while Q_i^H and Δ_i are distribution variables. When these variables carry asterisk superscripts, they should be understood to represent optimal values.

With this notation, one can write the upstream firm's objective function in (4) as

$$\begin{gathered} r^2\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)+(1-r)^2Q^HP\left(Q^H\right)+ \\ r(1-r)\left(Q^H+\Delta_1\right)P\left(Q^H+\Delta_1\right)+r(1-r)\left(Q^H+\Delta-\Delta_1\right)P\left(Q^H+\Delta-\Delta_1\right)- \\ cQ^H \end{gathered} \tag{5}$$

Upstream profits depend on the distribution of output only through Δ_1. Notice that its impact arises in the events in which downstream firms differ in their costs. When firm 1 is low cost and firm 2 is high cost, the marginal impact of raising Δ_1 is $\mathrm{MR}\left(Q^H+\Delta_1\right)$, while in the opposite case the marginal impact is $-\mathrm{MR}\left(Q^H+\Delta-\Delta_1\right)$. For a fixed value of Δ, the optimal Δ_1 is defined by

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

Given the assumption of decreasing marginal revenue, we can represent the two marginal revenue curves as in figure 1. There is a unique solution to the optimal distribution problem of $\Delta_1^* = \Delta/2$. In other words, the upstream firm fully smooths the additional output of low cost firms—i.e. Δ—across downstream firms. Otherwise, marginal revenues across states in which firms are ex post heterogeneous diverge, which is inefficient. Building on this argument, we can show that

Proposition 1 *When firms are risk neutral and do not face limited liability,* $\Delta_1^* = \Delta_2^* = \frac{\Delta^*}{2} > 0$*, while profits are independent of the distribution of* Q^H *across firms.*

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 the upstream firm is risk averse (in fact, it is risk neutral), but because aggregate revenue is concave in aggregate output. Now, firm i's output 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. Essentially, having two firms in the market helps the upstream firm "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.

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 is optimal. However, there is

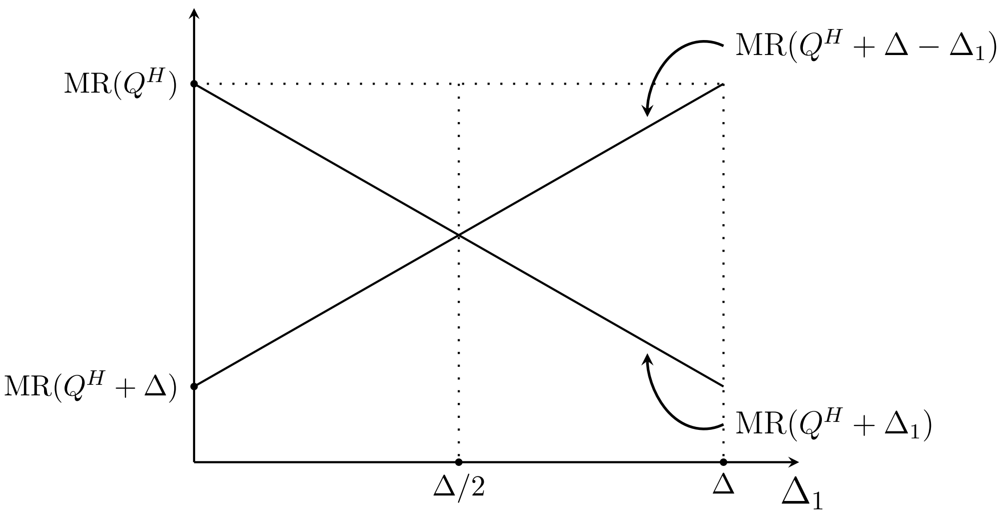


Figure 1: Effects of varying Δ_1 on upstream profits.

This figure illustrates the impact on upstream profits given by (5) of changes in Δ_1. An optimum occurs where the marginal revenue curve when firm 1 is high cost and firm 2 is low cost intersects with the marginal revenue curve in the opposite case. Given we assume decreasing marginal revenue, there is a unique intersection at $\Delta_1 = \Delta/2$. The linearity of the curves is for concreteness and is not assumed in the baseline model.

no fundamental force driving asymmetric outcomes, and we view the natural contracts as those that offer both cost types of both firms the same output. In any case, the important insight of this baseline exercise is that the upstream firm supplies through both downstream firms when they are symmetric and risk neutral.

3 Exclusion and Risk Aversion

We now depart from the baseline case and assume that downstream firms are risk averse. 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 downstream firms. If a downstream firm knows that it alone produces, it knows for certain what its profits will be. On the other hand, when the other firm's production level varies with its cost shock, profits are uncertain. In the case of risk neutrality, this has no effect on downstream firms' utility. In reality, however, one might imagine that downstream firms have some aversion to the uncertainty that competition creates.

3.1 High risk aversion

To begin the analysis of how risk affects the optimal contracts, we consider the starkest case in which firms are infinitely risk averse, and evaluate the outcome of a lottery according to its worst realization. This corresponds to the limit of CARA utility as $a \to \infty$.

Effectively, downstream firms behave as if they will meet an efficient firm for certain rather than with probability r. The objective function in (4) becomes

$$\begin{aligned} r\left[\left(Q_1^H+\Delta_1\right)P\left(Q^H+\Delta\right)+\left(Q^H-Q_1^H+\Delta-\Delta_1\right)P\left(Q^H+\Delta\right)\right]+ \\ (1-r)\left[Q_1^H P\left(Q^H+\Delta-\Delta_1\right)+(Q^H-Q_1^H)P\left(Q^H+\Delta_1\right)\right]-cQ^H \\ = \quad r\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)+ \\ (1-r)\left[Q_1^H P\left(Q^H+\Delta-\Delta_1\right)+(Q^H-Q_1^H)P\left(Q^H+\Delta_1\right)\right]-cQ^H. \end{aligned} \tag{7}$$

Figure 2 shows that risk aversion has dramatic consequences for the optimal distribution of output. Figure 2a shows the impact on upstream profits of changes in Q_1^H and Δ_1 with risk neutrality holding Q^H and Δ constant. As discussed in the baseline solution, Q_1^H has no effect while moving Δ_1 closer to the even split $\Delta_1 = \frac{\Delta}{2}$ always increases them. In contrast, figure 2b shows the equivalent impact for infinite risk aversion. First, when $\Delta_1 > \frac{\Delta}{2}$, the upstream firm benefits from increasing Q_1^H. Intuitively, in this case firm 2 is facing more risk than firm 1 because the loss from meeting an efficient competitor is higher since it will produce more. This makes serving firm 2 relatively more expensive due to the associated risk premium that must be paid out, so the upstream firm gains from shifting more production to firm 1. Similarly, we can show that if Q_1^H is sufficiently large (but smaller than Q^H), increasing Δ_1 increases upstream profits.[12] Again, the intuition is that the downstream firm that produces more should face less risk.

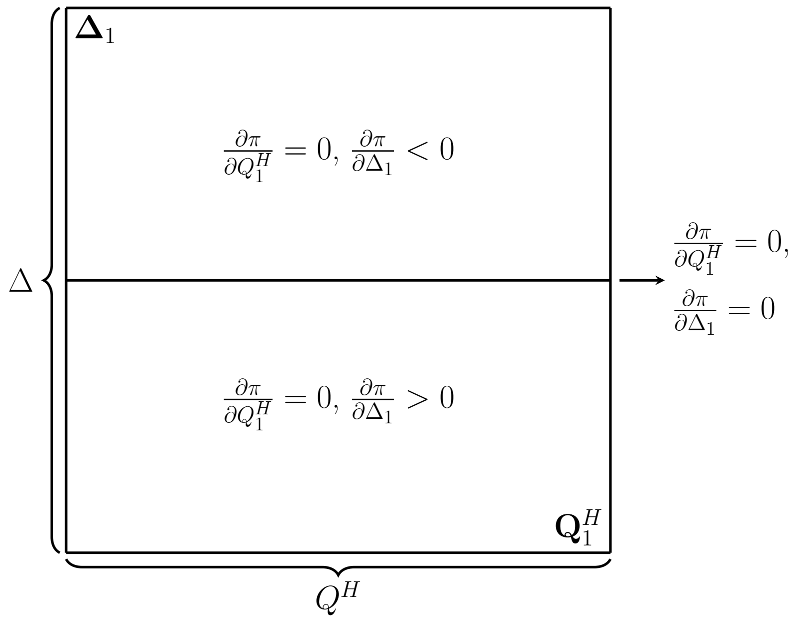
An immediate consequence of this reasoning is that whenever $Q^H > 0$ and $\Delta > 0$, the optimal distribution is a fully exclusive contract.[13] In contrast, the worst distribution is for the firm that produces all the high-cost output to face all the risk, i.e. $Q_i^H = Q^H$ and $\Delta_i = 0$. The only remaining issue is to determine when $Q^{H*} > 0$, since when $Q^{H*} = 0$ any split of Δ^* across the downstream firms is optimal. In the appendix we show that $Q^{H*} > 0$ if and only if $r < 1 - \frac{c}{P(0)}$. When high-cost firms are sufficiently productive (c low) and common in the population (r low), the upstream firm wishes them to participate in the market. While the logic so far rests on infinite risk aversion, we show using continuity arguments that full exclusion arises away from the limiting case.

Proposition 2 *Whenever $r < 1 - \frac{c}{P(0)}$, there exists a finite $\overline{a}$ such that fully exclusive contracts are strictly optimal for all $a > \overline{a}$.*

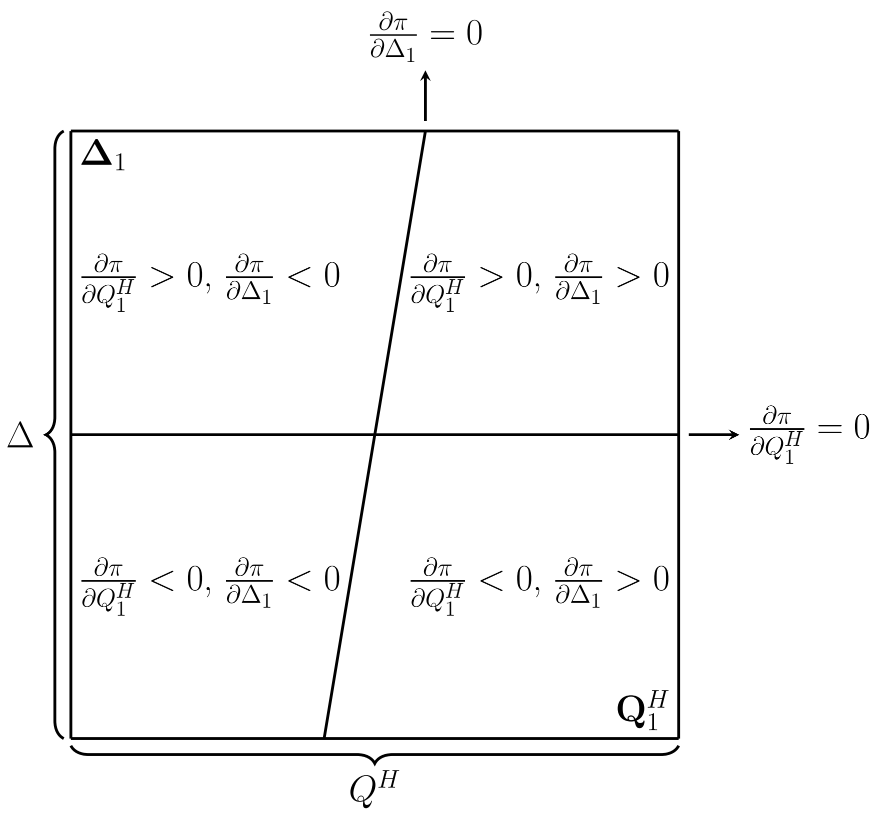
This result contains a basic message of the paper. Even if the upstream firm can fully

[12] We formally derive the properties of the derivatives plotted in figure 2 in the appendix. The point in figure 2b at which $\frac{\partial\pi}{\partial Q_1^H} = \frac{\partial\pi}{\partial\Delta_1^H} = 0$ is a saddle point, and profits are minimized when one firm produces no high cost output but all the low cost output (the northwest and southeast corners of the figure).

[13] One can also establish this result without any derivatives. Because price is decreasing in quantity, (7) has an upper bound of $r\left(Q^H+\Delta\right)P\left(Q^H+\Delta\right)+(1-r)Q^H P\left(Q^H\right)-cQ^H$, which is exactly the profit level achieved through a fully exclusive contract in which some firm i is offered the contract $Q_i^H = \Delta_i = 0$.



(a) Risk neutrality



(b) Infinite risk aversion

Figure 2: Derivatives of upstream profits π in distribution variables

This figure illustrates the signs of the derivative of the upstream firm's objective function in Q_1^H and Δ_1 for fixed $Q^H > 0$ and $\Delta > 0$ in the cases of risk neutrality and infinite risk aversion, respectively. In each subfigure, Q_1^H is plotted on the horizontal axis, whose length is Q^H and Δ_1 on the vertical, whose length is Δ. The locus of points at which $\frac{\partial \pi}{\partial \Delta_1} = 0$ in the infinite risk aversion case is for illustration only; in general it is not linear.

commit to bilateral contracts, it may simply be too costly to include both firms in the downstream market because uncertain competition externalities hurt risk-averse firms.

3.2 Intermediate risk aversion

We now explore the model for all values of a. In the risk neutral objective function (5) and infinite risk aversion objective function (7), upstream profits are linear in various revenue terms. With CARA utility and $a \in (0, \infty)$, there is instead curvature in revenue in the upstream objective function. To avoid the additional complication of curvature in the demand function, we analyze the linear demand case $P(Q) = 1 - Q$.

In the analysis, we focus on parameter values for which high-cost firms produce a positive quantity in aggregate (i.e. $Q^H > 0$). In other words, negative productivity shocks are not large enough to force both downstream firms to cease production altogether.

Lemma 2 *For all parameter values $\Delta^* > 0$. There exist values of r^* and c^* such that $Q^{H*} > 0$ whenever $r < r^*$ and $c < c^*$.*

We assume that $r < r^*$ and $c < c^*$ since the main question of interest is not whether production occurs at all, but how production is distributed across downstream firms given risk aversion.[14]

The baseline case of section 2.2 showed that the upstream firm chooses symmetric downstream outputs with risk neutrality to enhance production efficiency, while with high risk aversion it chooses starkly asymmetric outcomes to minimize risk. In the intermediate case, the upstream firm responds to both production efficiency and risk premia, but the implications of this are not immediately obvious. For example, for low levels of risk aversion, risk premia are also small, so one might plausibly think that symmetric outcomes remain optimal. Our next result shows that asymmetric outcomes arise for *any* level of risk aversion.

Proposition 3 *When $P = 1 - Q$, partially exclusive contracts are optimal for all $a > 0$.*

To gain an intuition for the result, recall that with risk neutrality the marginal impact on upstream profits of varying Δ_1 around its optimal value $\Delta/2$ is essentially zero since profits are strictly concave in Δ_1. On the other hand, for $a > 0$, adjusting Δ_1 away from $\Delta/2$ to expose the firm that produces relatively more to less risk (i.e. increasing Δ_1 when $Q_1^H > Q^H/2$, and decreasing it otherwise) has a positive impact on profits through reducing aggregate risk premia. The optimal Δ_1 with intermediate risk aversion is thus bound away from the symmetric outcome. In terms of the optimal Q_1^H, with risk neutrality it has no impact on profits, but with a small amount of risk aversion the

[14]When $Q^{H*} = 0$, the optimal distribution is $\Delta_1^* = \Delta_2^* = \Delta^*/2$; proof available on request.

upstream firm again uses it to reduce risk premia.[15] Formally speaking, we show that the signs of the partial derivatives of upstream profits in a neighborhood around symmetric contracts in figure 2b hold for any level of risk aversion.

Rather than being a knife-edge case, the linear demand condition should be understood as implying that partial exclusion arises subject to a bound on the curvature of the demand function. It is a technical condition in the sense that linearity is useful for completing the proof of proposition 3; we leave open the question of how substantial curvature in demand affects the optimal distribution for small values of a.

Combining the results so far together also gives a global prediction on how the degree of risk aversion affects the extent to which optimal distribution contracts are asymmetric. We know from proposition 1 that full symmetry is optimal for $a = 0$. Proposition 3 then shows that as we increase a, both firms continue to produce, but one firm produces more than another. Finally, proposition 2 shows that as a passes a critical threshold, one firm stops producing altogether. So with linear demand, a prediction of the model is that higher levels of downstream risk aversion should be associated with more asymmetry. Intuitively, this is because the risk premium begins to dominate the efficiency gains of output smoothing.

Proposition 3 is also relevant from an empirical perspective. Suppose one has a dataset in which downstream firms produce different levels of output on average. This is consistent with our model of partial exclusion, but also any model in which there is heterogeneity in some unobserved firm characteristic. But in our model, even if one could in theory condition on this characteristic (the productivity shock), output differences remain as a result of the upstream firm's choice of market structure. That is, *ex post* identical firms produce different output levels. Moreover, this is true whenever downstream firms have even a small amount of risk aversion.

4 Exclusion and Limited Liability

Instead of assuming that downstream firms are risk averse, we now assume they enjoy limited liability in the manner described in section 2. To see the negative effect of competition, recall that the optimal contract with risk neutrality and unlimited liability featured (1) $\Delta^* > 0$, (2) $\Delta_1^* = \Delta_2^*$, and (3) binding participation constraints for high-cost firms. The fact that participation constraints bind and $\Delta_i > 0$ means that high-cost firms earn a profit when facing a high-cost competitor and lose money when facing a low-cost competitor, such that on average profits are zero. Clearly this contract violates

[15]Notice that when $Q_i^H > Q_j^H$, then $Q_i(c) > Q_j(c)$ by definition. Also, $\Delta_i > \Delta_j$ implies $Q_i(0) - Q_i^H > Q_j(0) - Q_j^H$ by definition, or $Q_i(0) - Q_j(0) > Q_i^H - Q_j^H$. So $Q_i^H > Q_j^H$ and $\Delta_i > \Delta_j$ together imply partial exclusion.

the limited liability constraints (LL). Moreover, if the upstream firm wishes to maintain $\Delta_1 = \Delta_2$, ensuring that high-cost firms earn zero profit when facing a low-cost competitor implies they earn a profit when facing a high-cost competitor. This profit represents surplus the upstream firm loses from downstream competition. Instead, a fully exclusive contract avoids this lost surplus albeit at the cost of lost production efficiency. This section explores how the upstream firm trades off these two forces.[16]

Replacing (PC) with (LL) in the program (3) clearly has no additional implications for incentive compatibility, so one can again ignore the IC constraint for high-cost firms and focus on contracts with $\Delta_i \geq 0$. This in turn implies that $\pi_i(c_i, 0, c_i) \leq \pi_i(c_i, c, c_i)$. In other words, if the upstream firm ensures that downstream firm i earns 0 when the competitor has a cost $c_j = 0$, it typically pays out a positive amount when $c_j = c$. We show in the appendix that $\pi_i(c, 0, c) = 0$, i.e. that high-cost firms make zero profit when meeting a low-cost competitor. Suppose furthermore for the sake of argument that the IC constraint of the low-cost firm binds.[17] These constraints imply the equilibrium transfers

$$\begin{aligned} T_i(c) &= Q_i^H[P(Q^H + \Delta_j) - c] && (8) \\ T_i(0) &= (Q_i^H + \Delta_i)[rP(Q^H + \Delta) + (1-r)P(Q^H + \Delta_i)] - \\ &\quad Q_i^H[rP(Q^H + \Delta_j) + (1-r)P(Q^H)] + T_i(c). && (9) \end{aligned}$$

Plugging into the upstream objective then yields

$$\begin{aligned} & r^2 \left(Q^H + \Delta\right) P\left(Q^H + \Delta\right) - r(1-r) Q^H P\left(Q^H\right) - cQ^H + \\ & r(1-r)\left[\left(Q_1^H + \Delta_1\right) P\left(Q^H + \Delta_1\right) + \left(Q_2^H + \Delta_2\right) P\left(Q^H + \Delta_2\right)\right] + \\ & (1-r^2)[Q_1^H P\left(Q^H + \Delta_2\right) + Q_2^H P\left(Q^H + \Delta_1\right)]. \qquad (10) \end{aligned}$$

Whenever $\Delta_i \geq \Delta_j$, the upstream firm maximizes this by setting $Q_i^H = Q^H$ (the solution is unique when $\Delta_i > \Delta_j$).[18] In this case, firm j faces the largest negative revenue shock from facing a low-cost competitor, and so should not produce *any* high-cost output. Without loss of generality, then, let $Q_1^H = Q^H$. Replacing into (10) and

[16] Kaufmann and Lafontaine (1994) and Michael and Moore (1995) use franchise financial return data to show that franchisors leave ex ante rents to franchisees, and Kaufmann and Lafontaine (1994) explicitly hypothesize an explanation based on limited liability.

[17] In fact, the constraint on $T_i(0)$ is that it both satisfies incentive compatibility for low-cost firm i, and ensures non-negative profit when facing another low-cost firm. See appendix for further details.

[18] One can see this by differentiating (10) with respect to Q_i^H for a fixed Q^H using the substitution $Q_j^H = Q^H - Q_i^H$.

rearranging gives

$$
\begin{aligned}
& r^2\left(Q^H+\Delta\right) P\left(Q^H+\Delta\right)-r(1-r) Q^H P\left(Q^H\right)-c Q^H+ \\
& r(1-r)\left[\left(Q^H+\Delta_1\right) P\left(Q^H+\Delta_1\right)+\left(Q^H+\Delta-\Delta_1\right) P\left(Q^H+\Delta-\Delta_1\right)\right]+ \\
& (1-r)\left[Q^H P\left(Q^H+\Delta-\Delta_1\right)\right]. \qquad (11)
\end{aligned}
$$

This objective function cleanly combines the trade-offs we have highlighted throughout the paper. The second line is identical to the term in the objective function with risk neutrality (5) that reflects the incentive to smooth aggregate output. To maximize it, the upstream firm sets $\Delta_1=\Delta_2=\Delta/2$. The third line reflects the surplus lost due to competition, and is in fact equivalent to the final line in the objective function with infinite risk aversion (7) following the substitution $Q_1^H=Q^H$. The upstream firm maximizes it by choosing $\Delta_1=\Delta$.

The first order condition for the optimal Δ_1 for a fixed Δ is

$$
\mathrm{MR}\left(Q^H+\Delta_1^*\right)=\mathrm{MR}\left(Q^H+\Delta-\Delta_1^*\right)+\frac{Q^H P'\left(Q^H+\Delta-\Delta_1^*\right)}{r}. \qquad (12)
$$

This is identical to the first-order condition in the baseline case (6) except for the final term on the right-hand side. Without this, the upstream firm would select a symmetric outcome $\Delta_1=\Delta/2$. But with limited liability, there is an extra benefit to increasing Δ_1. Since high-cost firm 1 makes zero profit, low-cost firm 1 must make positive profit when facing an efficient firm 2. By allocating more low-cost production to firm 1 relative to firm 2, i.e. increasing Δ_1 for a fixed Δ, the upstream firm reduces this surplus.

Figure 3 illustrates the argument. The intersection of marginal revenue curves gives the optimum without limited liability. However, with limited liability, the term $Q^H P'(\cdot)/r$ (which is negative as demand is decreasing) creates an additional marginal benefit to increasing Δ_1. This acts to push the optimal Δ_1 above $\Delta/2$.

As one can observe, the lower is r, the greater is the effect of limited liability. The gains to output smoothing are higher when low-cost firms are more likely since without them output is constant in any case. As such, when r is sufficiently low the incentive to increase surplus extraction at the expense of output smoothing becomes more pronounced, and the upstream firm chooses full exclusion.

Proposition 4 *There exists an $r^*>0$ such that full exclusion is optimal for all $r<r^*$.*

The main intuition comes from observing figure 3. When r is small, the wedge $Q^H P'(\cdot)/r$ becomes so large that the intersection of the two sides of (12) occurs at $\Delta_1=\Delta$.

We find a remarkable similarity between high risk aversion and limited liability. In both versions of the model, if the probability of finding an efficient firm is sufficiently

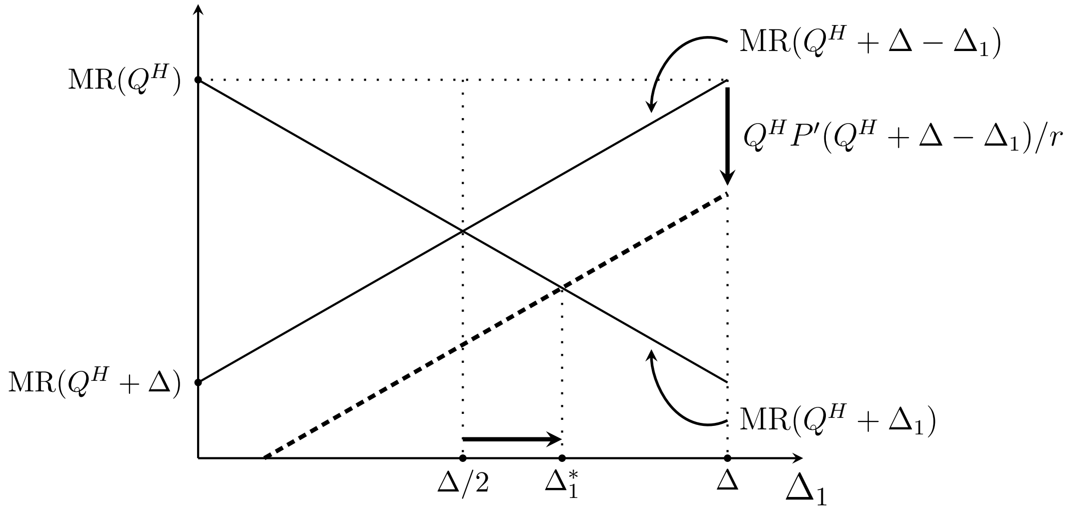


Figure 3: The effect of limited liability on the optimal Δ_1

This figure extends figure 1 to show the effect of limited liability on Δ_1^*. Whereas the intersection of the marginal revenue curves at $\Delta_1 = \Delta/2$ is optimal without limited liability, Δ_1^* is higher than $\Delta/2$ with limited liability. The linearity of the curves is for concreteness and is not assumed in the model.

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

5 Discussion and Conclusion

This paper identifies a new rationale for using exclusivity provisions: when firms compete downstream, and do not perfectly observe one another's productivity shocks, competition generates uncertainty, leading risk-averse (limited-liable) agents to require a risk premium (additional surplus). To save on these costs, the upstream firm sometimes prefers to deal exclusively with one firm, and more generally offers asymmetric contracts in many cases.

As mentioned in the introduction, an alternative story for observing exclusion is that upstream firms suffer from commitment problems. In franchise networks, the main commitment problem is *encroachment* whereby franchisors allow new franchisees to open outlets in areas previously successfully developed by established franchisees. Exclusive territories are often cited as a means of reassuring franchisees that such encroachment will not take place. Blair and Lafontaine (2011) argue that encroachment should be more problematic the larger the network, and then contrast this with the empirical evidence. One study found that only 26 of the largest 50 restaurant franchisors in the US offered an exclusive territory, compared with an overall incidence for restaurant franchisors of around 75%. On the other hand, Azoulay and Shane (2001) collected a dataset of newly founded franchises across a variety of industries, and report that 84% offered exclusive territories, whereas the cross-industry incidence in the US is around 73%. Hence in the networks where encroachment should be seemingly less of a problem, exclusion is *more* likely to be observed. An explanation of this observation is that franchisees in new networks face greater uncertainty than in existing networks, and that exclusion protects them in part against the concomitant risk.

In Hansen and Motta (2012), we provide several robustness checks on the basic results of this paper. The most notable one is the introduction of ex ante participation constraints in place of interim ones, as well as correlation between shock realizations. This brings the model very close to that of Rey and Tirole (1986), who adopt a setup with perfect correlation. It is also compatible with downstream firms' facing a common shock in addition to an idiosyncratic one. We show that *any* imperfect correlation leads to the optimality of full exclusion with infinite risk aversion, in direct contrast to Rey and Tirole (1986). This underscores the point that even small amounts of uncertainty can have dramatic impacts on the distribution of output across firms.

One important issue is why the upstream firm does not offer more insurance to downstream firms, as this would clearly improve its payoff. For example, one possibility would be to pay downstream firms a fixed amount to produce, but then collect the revenue itself. However, one can show that in this case total downstream profits actually decrease since there is no longer a monopoly producer. Moreover, one can construct examples in which producer surplus (the sum of upstream and downstream profits) is higher with exclusive contracts than with revenue-sharing contracts. Hence if there is a mechanism through which the upstream firm internalizes the effect of providing insurance to downstream firms, such as ex ante bargaining, it will not necessarily choose to do so.

More broadly speaking, another relevant concern for upstream firms is moral hazard. There is a large literature, summarized in Blair and Lafontaine (2011), that models upstream firms as risk-neutral principals and downstream retailers as risk-averse agents who make investments in quality. As is well known in the literature (Holmström 1979), insurance comes at the cost of incentives and would be expected to reduce the amount of investment. Seen this way, our model represents a means of endogenizing the noise that downstream firms face in vertical markets, whereas many papers in the moral hazard literature simply take as given stochastic downstream payoffs (examples include Brickley and Dark 1987 and Lafontaine 1992).

Finally, as mentioned in the introduction, one restriction in the model is that the upstream firm does not react to the messages it receives from downstream firms. Dequiedt and Martimort (2015) show that when it can react, the equilibrium outcome is akin to an all-pay auction in which the lowest-cost firms produce all the output. An initial observation is that, in many vertical markets, upstream firms do not create mechanisms to allocate all production to the most efficient firm. Blair and Lafontaine (2011, section 3.1) describe a norm in which franchisors offer all potential franchisees identical terms-of-trade rather than trying to tailor contracts to firm-specific characteristics. At the same time, there are clearly many idiosyncratic, transient productivity shocks that can hit competing downstream producers (e.g. equipment failure, personnel conflict, and so on) that generate output fluctuations. Our model captures this type of uncertainty.

Of course, there are certainly situations in which selection is crucial and our restricting the upstream firm's ability to react to type reports is not appropriate. In these cases, competition is shifted from the production stage to the bidding stage in the auction-like mechanism. The analogue of the question we ask in this paper would be how many bidders the upstream firm would want to participate in the auction. We conjecture that risk aversion might provide a bound on this number since each new participant imposes an uncertain negative externality on existing ones. In fact, our basic logic 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 types 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.

Banal-Estañol, A., and M. Ottaviani (2006): "Mergers with Product Market Risk," *Journal of Economics & Management Strategy*, 15(3), 577–608.

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.

Dasgupta, S., and D. F. Spulber (1989): "Managing procurement auctions," *Information Economics and Policy*, 4(1), 5–29.

Demougin, D. M., and D. A. Garvie (1991): "Contractual Design with Correlated Information under Limited Liability," *RAND Journal of Economics*, 22(4), 477–489.

Dequiedt, V., and D. Martimort (2015): "Vertical Contracting with Informational Opportunism," *American Economic Review*, 105(7), 2141–82.

Hansen, S., and M. Motta (2012): "Vertical Exclusion with Endogenous Competiton Externalities," CEPR Discussion Papers 8982, C.E.P.R. Discussion Papers.

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

Holmström, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10(1), 74–91.

Kaufmann, P. J., and F. Lafontaine (1994): "Costs of Control: The Source of Economic Rents for McDonald's Franchisees," *Journal of Law and Economics*, 37(2), 417–53.

Laffont, J.-J., and J. Tirole (1987): "Auctioning Incentive Contracts," *Journal of Political Economy*, 95(5), 921–37.

Lafontaine, F. (1992): "Agency Theory and Franchising: Some Empirical Results," *RAND Journal of Economics*, 23(2), 263–283.

Lafontaine, F., and M. Slade (2008): "Empirical Assessment of Exclusive Contracts," in *Handbook of Antitrust Economics*, ed. by P. Buccirossi. MIT Press, Cambdrige.

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.

McAfee, R. P., and M. Schwartz (1994): "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *American Economic Review*, 84(1), 210–30.

Michael, S. C., and H. J. Moore (1995): "Returns to franchising," *Journal of Corporate Finance*, 2(1-2), 133–155.

Nocke, V., and J. Thanassoulis (2014): "Vertical Relations Under Credit Constraints," *Journal of the European Economic Association*, 12(2), 337–367.

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.

Scott Morton, F. (2012): "Contracts that Reference Rivals," Speech delivered at Georgetown University Law Center, April 5, 2012. Available at http://www.justice.gov/atr/file/518971/download.

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

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

A Proofs

A.1 Proof of Lemma 1 / Upstream Objective Function

Proof. Suppose an agent faces the lottery $[(w, w-L), (1-r, r)]$ and has CARA utility. The certainty equivalent C is defined by

$$\exp(-aC) = (1-r)\exp(-aw) + r\exp(-a(w-L))$$

which after algebraic manipulations gives $C = w - \Gamma(L, a, r)$ where $\Gamma(L, a, r) \equiv \frac{\ln[(1-r)+r\exp(aL)]}{a}$. Applying this expression to the lotteries faced by downstream firms in the model gives

$$\text{CertRev}_i(\hat{c}_i) = \\ Q_i(\hat{c}_i)P[Q_i(\hat{c}_i) + Q_j(c)] - \Gamma\left(Q_i(\hat{c}_i)\left\{P[Q_i(\hat{c}_i) + Q_j(c)] - P[Q_i(\hat{c}_i) + Q_j(0)]\right\}, a, r\right). \tag{A.1}$$

Incentive compatibility for the low-cost firm and participation of the high-cost firm imply participation of the low-cost firm since

$$\text{CertRev}_i(0) - T_i(0) \geq \text{CertRev}_i(c) - T_i(c) > \text{CertRev}_i(c) - Q_i(c)c - T_i(c) \geq 0.$$

So the incentive compatibility constraint for the low-cost firm must bind, since otherwise the upstream firm could increase $T_i(0)$. Incentive compatibility for the high-cost firm can be written as[19]

$$c[Q_i(0) - Q_i(c)] \geq [T_i(c) - T_i(0)] + [\text{CertRev}_i(0) - \text{CertRev}_i(c)] \geq 0.$$

So $Q_i(0) \geq Q_i(c)$ is necessary for incentive compatibility, and, under this condition, incentive compatibility for the low-cost firm implies incentive-compatibility for the high-cost firm. This leaves the participation constraint for the high-cost firm, 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 $Q_i(0) \geq Q_i(c)$. ■

We now derive the analytical expression for the upstream firm's objective function we use in the remaining formal results. Let $Q_i^H \equiv Q_i(c)$, $\Delta_i \equiv Q_i(0) - Q_i^H$, $Q^H \equiv Q_1^H + Q_2^H$, and $\Delta \equiv \Delta_1 + \Delta_2$. We can then express the choice variables in the maximization problem in terms of the vector $\mathbf{S} = (Q_1^H, Q^H, \Delta_1, \Delta)$, which must satisfy the constraints $Q^H \geq Q_1^H \geq 0$ and $\Delta \geq \Delta_1 \geq 0$. Plugging into (A.1), we obtain

$$\begin{aligned} &\text{CertRev}_i(c) = Q_i^H P(Q^H) - \Gamma\left\{Q_i^H\left[P(Q^H) - P(Q^H + \Delta_j)\right], a, r\right\} \text{ and} \\ &\text{CertRev}_i(0) = (Q_i^H + \Delta_i)P(Q^H + \Delta_i) - \Gamma\left\{(Q_i^H + \Delta_i)\left[P(Q^H + \Delta_i) - P(Q^H + \Delta)\right], a, r\right\} \end{aligned}$$

After applying the substitutions $Q_2^H = Q^H - Q_1^H$ and $\Delta_2 = \Delta - \Delta_1$, the upstream firm's

[19] The last inequality follows from low-cost incentive compatibility.

objective function becomes[20]

$$\pi(\mathbf{S},a,r,c) =$$
$$(1-r)Q^H P\left(Q^H\right) + r\left[\left(Q_1^H+\Delta_1\right)P\left(Q^H+\Delta_1\right) + \left(Q_2^H+\Delta-\Delta_1\right)P\left(Q^H+\Delta-\Delta_1\right)\right] -$$
$$(1-r)\Gamma\left\{Q_1^H\left[P\left(Q^H\right)-P\left(Q^H+\Delta_2\right)\right]\right\} - (1-r)\Gamma\left\{\left(Q^H-Q_1^H\right)\left[P\left(Q^H\right)-P\left(Q^H+\Delta_1\right)\right]\right\} -$$
$$r\Gamma\left\{\left(Q_1^H+\Delta_1\right)\left[P\left(Q^H+\Delta_1\right)-P\left(Q^H+\Delta\right)\right]\right\} -$$
$$r\Gamma\left\{\left(Q^H-Q_1^H+\Delta-\Delta_1\right)\left[P\left(Q^H+\Delta-\Delta_1\right)-P\left(Q^H+\Delta\right)\right]\right\} - cQ^H. \tag{A.2}$$

By l'Hôpital's Rule we obtain the relationships

$$\lim_{a\to 0}\Gamma(L,a,r) = \lim_{a\to 0}\frac{rL\exp(aL)}{r\exp(aL)+1-r} = rL$$
$$\lim_{a\to\infty}\Gamma(L,a,r) = \lim_{a\to\infty}\frac{rL\exp(aL)}{r\exp(aL)+1-r} = \lim_{a\to\infty}\frac{L}{1+\frac{1-r}{r}\exp(-aL)} = L.$$

In other words, the limit of certainty equivalent revenue as $a \to 0$ corresponds to risk neutrality, as certainty equivalent income is expected income. The limit as $a \to \infty$ corresponds to infinite risk aversion, as the payoff from a lottery is its worst realization. We extend the definition of $\pi(\mathbf{S},a,r,c)$ to $a = 0$ with expression (5) in the main text, and to $a = \infty$ with (7). Given the limit results, $\pi(\mathbf{S},a,r,c)$ defined in this way in continuous in $a \in \mathbb{R}_+ \cup \infty$.

A.2 Proof of Proposition 1

Proof. It remains to be shown that the optimal value of Δ is positive. The only other possibility is that $\Delta = 0$ since $\Delta < 0$ would violate incentive compatibility. Suppose that $\Delta = 0$, and let the optimal value of Q^H under this restriction be $Q^{H\prime}$. The total profit of the upstream firm from this solution is $Q^{H\prime}\left[P\left(Q^{H\prime}\right)-c\right]$. Note that this payoff can be obtained with the exclusive contract $Q_1^H = Q^{H\prime}$ and $Q_2^H = 0$. (By assumption $\Delta_1 = \Delta_2 = 0$).

Now consider the upstream firm's choice of the optimal exclusive contract, i.e. a contract in which, without loss of generality, $Q_1^H = Q^H$ and $\Delta_1 = \Delta$. The relevant program is

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

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

$$\mathrm{MR}\left(Q^{H*}+\Delta^*\right) \leq 0 \tag{A.4}$$
$$r\mathrm{MR}\left(Q^{H*}+\Delta^*\right) + (1-r)\mathrm{MR}\left(Q^{H*}\right) - c \leq 0 \tag{A.5}$$

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

[20]Here for notational compactness we have dropped the a and r parameters from the Γ functions.

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, their optimality is contradicted. ■

A.3 Proof of Proposition 2

Proof. The proof relies on continuity in the upstream firm's objective function and its derivatives. We first establish some relevant notation and properties of the solutions.

First, let $\mathbf{S}^*(a, r, c) \subset \mathbb{R}^4_+$ denote the set of solutions to program (3) given a. By arguments in the main text, $S^*(\infty, r, c) = \{(Q^{H*}, Q^{H*}, \Delta^*, \Delta^*), (0, Q^{H*}, 0, \Delta^*)\}$ whenever $Q^{H*} > 0$. Moreover, the optimal values for Q^H and Δ are derived in the proof of proposition 1 in which we solve for the optimal exclusive contract. From (A.4) and (A.5), $Q^{H*} > 0$ if and only if $r < 1 - \frac{c}{MR(0)}$. Moreover, $\Delta^* > 0$ as argued in the proof of proposition 1.

Second, from expression (7)

$$\frac{\partial \pi(\mathbf{S}, \infty, r, c)}{\partial Q_1^H} = (1 - r)[P(Q^H + \Delta - \Delta_1) - P(Q^H + \Delta_1)] \gtreqless 0 \Leftrightarrow \Delta_1 \gtreqless \frac{\Delta}{2}.$$

Moreover,

$$\frac{\partial \pi(\mathbf{S}, \infty, r, c)}{\partial \Delta_1} \propto Q^H P'(Q^H + \Delta_1) - Q_1^H[P'(Q^H + \Delta - \Delta_1) + P'(Q^H + \Delta_1)].$$

This is strictly increasing in Q_1^H since $P' < 0$. Moreover, this expression is negative when $Q_1^H = 0$ and positive when $Q_1^H = Q^H$. So, we conclude there exists some $\Theta(\Delta_1) \in (0, Q^H)$ such that $\frac{\partial \pi(\mathbf{S},\infty,r,c)}{\partial \Delta_1} \gtreqless 0 \Leftrightarrow Q_1^H \gtreqless \Theta(\Delta_1)$. The signs of these derivatives are plotted in figure 2b in the main text.

Third, note that (A.2) is bounded above by

$$(1 - r)Q^H P\left(Q^H\right) + r\left[\left(Q_1^H + \Delta_1\right) P\left(Q^H + \Delta_1\right) + \left(Q^H - Q_1^H + \Delta - \Delta_1\right) P\left(Q^H + \Delta - \Delta_1\right)\right].$$

By assumption the first term is unbounded below as $Q^H \to \infty$ and, similarly, at least one of the two terms in square brackets is unbounded below as $\Delta \to \infty$. Also note that the upstream firm can always guarantee itself a payoff of zero by choosing $Q^H = \Delta = 0$. We can therefore restrict attention to the upstream firm's choosing Q^H and Δ in a closed, bounded set D.

Since $\pi(\mathbf{S}, a, r, c)$ is continuous and the constrained set of parameters is compact, $\mathbf{S}^*(a, r, c)$ is upper-semicontinuous by the Maximum Theorem (see Sundaram (1996) theorem 9.14 for details). Let V be an open set such that $\mathbf{S}^*(\infty, r, c) \subset V$ and for which, for all $\mathbf{S} \in V \cap D$ either $\frac{\partial \pi(\mathbf{S},\infty,r,c)}{\partial Q_1^H}, \frac{\partial \pi(\mathbf{S},\infty,r,c)}{\partial \Delta_1} > 0$ or $\frac{\partial \pi(\mathbf{S},\infty,r,c)}{\partial Q_1^H}, \frac{\partial \pi(\mathbf{S},\infty,r,c)}{\partial \Delta_1} < 0$. By upper-semicontinuity there exists some a^1 such that $\mathbf{S}^*(a, r, c) \subset V$ for all $a > a^1$. Moreover, since $\pi(\mathbf{S}, a, r, c)$ has continuous derivatives, for any $\mathbf{S} \in V \cap D$ there exists some $a^2(\mathbf{S})$ such that $\frac{\partial \pi(\mathbf{S},a,r,c)}{\partial Q_1^H}, \frac{\partial \pi(\mathbf{S},a,r,c)}{\partial \Delta_1} > 0$ or $\frac{\partial \pi(\mathbf{S},a,r,c)}{\partial Q_1^H}, \frac{\partial \pi(\mathbf{S},a,r,c)}{\partial \Delta_1} < 0$ for all $a > a^2(\mathbf{S})$. Hence $\mathbf{S}$ cannot maximize profit unless it is fully exclusive. The proof is completed by taking $\overline{a} \equiv \max\{a^1, \max_{\mathbf{S}} a^2(\mathbf{S})\}$. ■

A.4 Proof of Lemma 2

Proof. The proof for why $\Delta^* > 0$ proceeds exactly as in the proof of proposition 1. In particular, note that when $\Delta = 0$ upstream profits from (A.2) become $Q^H \left[P\left(Q^H\right) - c\right]$, which can be obtained with an exclusive contract. But the optimal exclusive contract must have $\Delta > 0$.

Now consider some $\mathbf{S}' = (0, 0, \Delta_1, \Delta)$ for any $0 \leq \Delta_1 \leq \Delta$. From (A.2) it is easy to check that $\lim_{r\to 0, c\to 0} \pi(\mathbf{S}', a, r, c) = 0$. On the other hand, let $Q^{H\prime} \equiv \arg\max_{Q^H} Q^H \left[P\left(Q^H\right) - c\right]$. Let $\mathbf{S}' = (Q^{H\prime}, Q^{H\prime}, 0, 0)$. Clearly $\lim_{r\to 0, c\to 0} \pi(\mathbf{S}'', a, r, c)$ is positive, which implies there exists some r^* and c^* such that $\mathbf{S}''$ produces higher profit that $\mathbf{S}'$, meaning that $\mathbf{S}'$ cannot be optimal. ■

A.5 Proof of Proposition 3

Proof. The strategy of the proof is as follows. We fix some $Q^H > 0$ and $\Delta > 0$, and then consider the problem of maximizing $\pi(Q_1^H, Q^H, \Delta_1, \Delta, a, r, c)$ with respect to Q_1^H and Δ_1 treating Q^H and Δ as fixed, positive parameters. Denote the derived function as $\pi(Q_1^H, \Delta_1)$. Without loss of generality, we consider contracts in which $Q_1^H \geq Q^H/2$. We show that $\Delta_1 > \Delta/2$ and $Q_1^H > Q^H/2$ is optimal within this class. For notational compactness, we express $\Delta - \Delta_1$ as Δ_2 and $Q^H - Q_1^H = Q_2^H$ in some expressions below in line with the definitions in the main text.

Several properties of the Γ function defined in the proof of lemma 1 are useful in the proof:

1. $$\frac{\partial \Gamma(L, a, r)}{\partial L} = \frac{r \exp(aL)}{1 - r + r\exp(aL)} > 0.$$

2. From the above expression, we clearly have $\frac{\partial \Gamma(L,a,r)}{\partial L} \in (r, 1)$.

3. $$\frac{\partial^2 \Gamma(L, a, r)}{\partial L^2} = \frac{(1-r)ar\exp(aL)}{[1 - r + r\exp(aL)]^2} > 0.$$

The first step in the proof is to compute and sign the partial derivatives of $\pi(Q_1^H, \Delta_1)$. From (A.2), we can compute

$$\begin{gathered}
\frac{\partial \pi}{\partial Q_1^H} = r\left[P\left(Q^H + \Delta_1\right) - P\left(Q^H + \Delta_2\right)\right] - \\
r\left[P\left(Q^H + \Delta_1\right) - P\left(Q^H + \Delta\right)\right]\Gamma'\left\{\left(Q_1^H + \Delta_1\right)\left[P\left(Q^H + \Delta_1\right) - P\left(Q^H + \Delta\right)\right]\right\} + \\
r\left[P\left(Q^H + \Delta_2\right) - P\left(Q^H + \Delta\right)\right]\Gamma'\left\{\left(Q^H - Q_1^H + \Delta_2\right)\left[P\left(Q^H + \Delta_2\right) - P\left(Q^H + \Delta\right)\right]\right\} - \\
(1-r)\left[P\left(Q^H\right) - P\left(Q^H + \Delta_2\right)\right]\Gamma'\left\{Q_1^H\left[P\left(Q^H\right) - P\left(Q^H + \Delta_2\right)\right]\right\} + \\
(1-r)\left[P\left(Q^H\right) - P\left(Q^H + \Delta_1\right)\right]\Gamma'\left\{\left(Q^H - Q_1^H\right)\left[P\left(Q^H\right) - P\left(Q^H + \Delta_1\right)\right]\right\}
\end{gathered}$$

In this expression and those that follow in the proof, $\Gamma'\{X\}$ should be understood as $\frac{\partial \Gamma(L,a,r)}{\partial L}$ evaluated at $L = X$. Notice that since $\Gamma'\{X\}$ is increasing in X, the above expression is monotonically decreasing in Q_1^H. So for each Δ_1 there is a unique optimal value for Q_1^H.

Evaluating this expression with linear demand and simplifying gives

$$\begin{aligned}\Delta_1 \left\{r\Gamma'[(Q^H - Q_1^H + \Delta_2)\Delta_1] + (1-r)\Gamma'[(Q^H - Q_1^H)\Delta_1] - r\right\} - \\ \Delta_2 \left\{r\Gamma'[(Q_1^H + \Delta_1)\Delta_2] + (1-r)\Gamma'[Q_1^H \Delta_2] - r\right\}.\end{aligned} \tag{A.6}$$

First note that when $\Delta_1 = \Delta_2$ this derivative is zero at $Q_1^H = Q^H/2$.

We now show that whenever $\Delta_1 > \Delta_2$, (A.6) is positive when evaluated at $Q_1^H = Q^H/2$. This implies that the optimal value of Q_1^H when $\Delta_1 > \Delta/2$ is greater than $Q^H/2$. The proof proceeds by the construction of lower bounds. First, by the assumption that $\Delta_1 > \Delta_2$ and monotonicity of Γ', we know that (A.6) is larger than

$$\begin{aligned}\Delta_1 \left\{r\Gamma'\left[\left(\frac{Q^H}{2} + \Delta_1\right)\Delta_2\right] + (1-r)\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] - r\right\} - \\ \Delta_2 \left\{r\Gamma'\left[\left(\frac{Q^H}{2} + \Delta_1\right)\Delta_2\right] + (1-r)\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] - r\right\}\end{aligned}$$

which itself is larger than (again by monotonicity of Γ')

$$\begin{aligned}\Delta_1 \left\{r\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] + (1-r)\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] - r\right\} - \Delta_2 \left\{r\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] + (1-r)\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] - r\right\} > \\ \Delta_1 \left\{\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] - r\right\} - \Delta_2 \left\{\Gamma'\left[\frac{Q^H}{2}\Delta_2\right] - r\right\}.\end{aligned}$$

Finally, we know that the final expression above is positive since $\Delta_1 > \Delta_2$ and $\Gamma' \in (r, 1)$. This implies that whenever $\Delta_1 > \Delta_2$, $Q_1^{H*} > Q^H/2$.

With linear demand, the part of the objective function (A.2) that depends on Δ_1 becomes

$$\begin{aligned}r[(Q_1^H + \Delta_1)(1 - Q^H - \Delta_1) + (Q_2^H + \Delta - \Delta_1)(1 - Q^H - \Delta + \Delta_1)] - \\ r\Gamma[(Q_1^H + \Delta_1)(\Delta - \Delta_1)] - r\Gamma[(Q_2^H + \Delta - \Delta_1)\Delta_1] - (1-r)\Gamma[Q_1^H(\Delta - \Delta_1)] - (1-r)\Gamma[Q_2^H \Delta_1].\end{aligned}$$

The derivative with respect to Δ_1 is

$$\begin{aligned}r(1 - Q^H - Q_1^H - 2\Delta_1) - r(1 - Q^H - Q_2^H - 2(\Delta - \Delta_1)) - \\ r(\Delta - 2\Delta_1 - Q_1^H)\Gamma'[(Q_1^H + \Delta_1)(\Delta - \Delta_1)] - r(\Delta - 2\Delta_1 + Q_2^H)\Gamma'[(Q_2^H + \Delta - \Delta_1)\Delta_1] - \\ (1-r)(-Q_1^H)\Gamma'[Q_1^H(\Delta - \Delta_1)] - (1-r)Q_2^H \Gamma'[Q_2^H \Delta_1].\end{aligned}$$

which can be re-written as

$$\begin{aligned}r(\Delta - 2\Delta_1)\left\{2 - \Gamma'[(Q_1^H + \Delta_1)(\Delta - \Delta_1)] - \Gamma'[(Q_2^H + \Delta - \Delta_1)\Delta_1]\right\} + \\ Q_1^H \left\{r\Gamma'[(Q_1^H + \Delta_1)(\Delta - \Delta_1)] + (1-r)\Gamma'[Q_1^H(\Delta - \Delta_1)] - r\right\} - \\ Q_2^H \left\{r\Gamma'[(Q_2^H + \Delta - \Delta_1)\Delta_1] + (1-r)\Gamma'[Q_2^H \Delta_1] - r\right\}.\end{aligned} \tag{A.7}$$

When $Q_1^H = Q_2^H = Q^H/2$, (A.7) is zero when $\Delta_1 = \Delta_2 = \Delta/2$. Now instead suppose that $Q_1^H > Q_2^H$. One can easily show the derivative evaluated at $\Delta_1 = \Delta/2$ is positive: the first line of (A.7) is zero, and the last two lines can be bounded below by a positive number in a manner similar to the argument above for the derivative in Q_1^H.

Moreover, one can also argue that the derivative is positive at $\Delta_1 < \Delta/2$ when $Q_1^H > Q_2^H$. First, the first line of (A.7) is positive since $\Gamma' < 1$. Second, note that $(Q_1^H + \Delta_1)(\Delta - \Delta_1) > (Q_2^H + \Delta - \Delta_1)\Delta_1$ whenever $\Delta_1 < \frac{Q_1^H}{Q^H}\Delta$, which is greater than $\Delta/2$ by the assumption $Q_1^H > Q_2^H$. Also $Q_1^H(\Delta - \Delta_1) > Q_2^H \Delta_1$ by assumption. So by monotonicity of Γ', we obtain a positive derivative at $\Delta_1 < \Delta/2$.

Thus the maximizers of $\pi(Q_1^H, \Delta_1)$ either satisfy $Q_1^{H*} = Q^H/2$, $\Delta_1^* = \Delta/2$ or $Q_1^{H*} > Q^H/2$, $\Delta_1^* > \Delta/2$. By the arguments above there exists a set $X = \{(Q_1^H, \Delta_1) \mid Q^H/2 + \varepsilon \geq Q_1^H \geq Q^H/2, \Delta/2 + \varepsilon \geq \Delta_1 \geq \Delta/2\} \backslash \{(Q^H/2, \Delta/2)\}$ for some $\varepsilon > 0$ such that the gradient of $\pi(Q_1^H, \Delta_1)$ is positive for all $x \in X$. Hence $Q_1^{H*} = Q^H/2$, $\Delta_1^* = \Delta/2$ cannot be a solution. ■

A.6 Proof of Proposition 4

Proof. We first analyze the program

$$\max_{\{Q_i(\widehat{c}_i), T_i(\widehat{c}_i)\}_{i=1}^2} \sum_{i=1}^{2} rT_i(0) + (1-r)T_i(c) \quad \text{such that} \quad \forall c_i \tag{A.8}$$

$$\min\{\pi_i(c_i, 0, c_i), \pi_i(c_i, c, c_i)\} \geq 0 \tag{LL}$$

$$U\left[L_i(c_i \mid c_i)\right] \geq U\left[L_i(\hat{c}_i \mid c_i)\right] \quad \text{for } \hat{c}_i \neq c_i \tag{IC}$$

$$Q_i(c_i) \geq 0. \tag{NN}$$

Following exactly the same arguments as in the proof of lemma 1, we can show that $\Delta_i \geq 0$ is necessary and sufficient for incentive compatibility, and that one can ignore the IC constraint for the high-cost firm. $\Delta_i \geq 0$ implies one only needs to consider the LL constraints corresponding to meeting an efficient competitor. Clearly then it is optimal to choose

$$T_i(c) \;=\; Q_i^H[P(Q^H + \Delta_j) - c].$$

The situation for the low-cost firm is less clear. The optimal contract must satisfy both the LL and IC constraints, which means the optimal transfer satisfies

$$T_i(0) = \min\left\{\begin{array}{c} (Q_i^H + \Delta_i)P(Q^H + \Delta), \\ (Q_i^H + \Delta_i)[rP(Q^H + \Delta) + (1-r)P(Q^H + \Delta_i)] - \\ Q_i^H[rP(Q^H + \Delta_j) + (1-r)P(Q^H)] + T_i(c) \end{array}\right\}.$$

Rather than determine which constraint binds in general, we focus on a binding IC constraint, and provide conditions under which a fully exclusive contract maximizes the resulting objective function. Fully exclusive contracts that are incentive compatible also satisfy the LL constraint. Clearly this is true for the firm that produces nothing for any cost type. For the firm that does produce, $T_i(c) = Q^H[P(Q^H) - c]$ which when plugged into the IC constraint gives

$$T_i(0) = (Q^H + \Delta)[rP(Q^H + \Delta) + (1-r)P(Q^H + \Delta)] - Q^H P(Q^H) + Q^H P(Q^H) - cQ^H = (Q^H + \Delta)P(Q^H + \Delta) - cQ^H$$

which is less than $(Q^H + \Delta)P(Q^H + \Delta)$ since $Q^H \geq 0$. Thus the same conditions that lead to full exclusion while ignoring the LL constraint also correspond to conditions under which full exclusion arises when $T_i(0)$ must satisfy limited liability constraints.

The upstream objective from plugging $T_i(c)$ and $T_i(0)$ given by the IC constraint into the objective function becomes (following arguments in the main text)

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

where $R(Q) \equiv QP(Q)$ is revenue. Recall that this expression is valid under the normalization $\Delta/2 \geq \Delta_2$, in which case we know that $Q_1^H = Q^H$ is optimal. The solutions we study below will all satisfy this normalization. The partial derivatives of $\pi(Q^H, \Delta, \Delta_2)$ are

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial Q^H} = \left\{ \begin{array}{c} r^2 \mathrm{MR}\left(Q^H + \Delta\right) - r(1-r)\mathrm{MR}\left(Q^H\right) + \\ r(1-r)\left[\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) + \mathrm{MR}\left(Q^H + \Delta_2\right)\right] + \\ (1-r)\left[P\left(Q^H + \Delta_2\right) + Q^H P'\left(Q^H + \Delta_2\right)\right] \end{array} \right\} - c \tag{A.9}$$

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial \Delta} = r^2 \mathrm{MR}\left(Q^H + \Delta\right) + r(1-r)\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) \tag{A.10}$$

$$\frac{\partial \pi(Q^H, \Delta, \Delta_2)}{\partial \Delta_2} = \left\{ \begin{array}{c} r(1-r)\left[-\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) + \mathrm{MR}\left(Q^H + \Delta_2\right)\right] + \\ (1-r)Q^H P'\left(Q^H + \Delta_2\right) \end{array} \right\}. \tag{A.11}$$

Solution I: $Q^H = 0$, $\Delta_2 = \frac{\Delta}{2}$. This solution exists only if there is some $\Delta > 0$ such that the partial derivatives evaluated at $(Q^H, \Delta, \Delta_2) = (0, \Delta, \Delta/2)$ satisfy $\frac{\partial \pi(Q^H,\Delta,\Delta_2)}{\partial Q^H} < 0$, and $\frac{\partial \pi(Q^H,\Delta,\Delta_2)}{\partial \Delta} = \frac{\partial \pi(Q^H,\Delta,\Delta_2)}{\partial \Delta_2} = 0$. (A.11) is clearly satisfied, while (A.9) and (A.10) rewrite as[21]

$$-r(1-r)\mathrm{MR}(0) + r(1-r)\mathrm{MR}\left(\frac{\Delta}{2}\right) + (1-r)P\left(\frac{\Delta}{2}\right) < c. \tag{A.12}$$

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

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.

Solution II: $Q^H > 0$, $0 < \Delta_2 \leq \frac{\Delta}{2}$. This solutions exists only if the partial derivatives are all

[21] Here we have also plugged (A.9) into (A.10).

0. The resulting system of equations simplifies to

$$\left\{ \begin{array}{c} -r(1-r)\mathrm{MR}\left(Q^H\right) + r(1-r)\mathrm{MR}\left(Q^H + \Delta_2\right) + \\ (1-r)\left[P\left(Q^H + \Delta_2\right) + Q^H P'\left(Q^H + \Delta_2\right)\right] \end{array} \right\} = c \tag{A.14}$$

$$r\mathrm{MR}\left(Q^H + \Delta\right) + (1-r)\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) = 0 \tag{A.15}$$

$$r(1-r)\left[-\mathrm{MR}\left(Q^H + \Delta - \Delta_2\right) + \mathrm{MR}\left(Q^H + \Delta_2\right)\right] + (1-r)Q^H P'\left(Q^H + \Delta_2\right) = 0. \tag{A.16}$$

Since $P' < 0$, (A.16) implies that $\mathrm{MR}\left(Q^H + \Delta_2\right) > \mathrm{MR}\left(Q^H + \Delta - \Delta_2\right)$ 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.16) must be strictly negative. So this solution cannot exist for r sufficiently small.

Solution III: Exclusive contract with $Q^H > 0$, $\Delta > 0$, $\Delta_2 = 0$. This solution exists only if

$$r\mathrm{MR}\left(Q^H + \Delta\right) + (1-r)\mathrm{MR}\left(Q^H\right) = c \tag{A.17}$$

$$r\mathrm{MR}\left(Q^H + \Delta\right) = 0 \tag{A.18}$$

$$r(1-r)\left[-\mathrm{MR}\left(Q^H + \Delta\right) + \mathrm{MR}\left(Q^H\right)\right] + (1-r)Q^H P'\left(Q^H\right) < 0. \tag{A.19}$$

which simplifies to

$$(1-r)\mathrm{MR}\left(Q^H\right) = c \tag{A.20}$$

$$\mathrm{MR}\left(Q^H + \Delta\right) = 0 \tag{A.21}$$

$$r\mathrm{MR}\left(Q^H\right) + Q^H P'\left(Q^H\right) < 0. \tag{A.22}$$

This solution clearly exists when r is small.

Note that we have not considered a solution in which $\Delta = 0$. To rule this out, one can follow the exact same logic as in the proof of proposition 1. ■