# Limiting and Taxing Competition Through Vertical Contracts: Anticompetitive Vertical Restraints Made Simple[†]

## Rev 8 Sept 2019


## Joseph Farrell[*]


"Sometimes things are just what they seem to be…"—Charles Bukowski

## Table of Contents



---


[†] This paper was begun in 2012 and has been extraordinarily slow (even for me) in the writing, starting as a more academic echo of my analyses of vertical concerns while I served at the FTC in 2009-2012. I learned a lot through discussions with FTC economists Pat DeGraba, Dan O'Brien, and especially Alison Oldale, as well as with academic colleagues, notably Louis Kaplow, Carl Shapiro, and Ralph Winter. My thinking has also benefited greatly from seminar presentations at Berkeley, Northwestern/Searle (2013), DG-COMP, OECD, the UK CMA, Oxford, and UEA Norwich. I thank Zarek Brot-Goldberg, Leo Chang, and Sean Wu for excellent research assistance. But none of them are responsible for my views.



[*] Professor of Economics, University of California at Berkeley, joseph.farrell@berkeley.edu, and Partner at Bates White LLC.

## 1. Introduction

The most profitable number of monopoly markups is 1.0.

This principle is often presented to explain that if a vertical value chain involves a good deal of market power at more than one level—perhaps a pure monopoly complemented by another monopoly, or by a layer with some market power—then the firms have a collective incentive to negotiate sophisticated contracts to bring the number of monopoly markups down towards 1. This is often called "elimination of double marginalization."

But of course the principle operates in the other direction too. If a monopoly is undermined by (surprisingly) self-destructive undercutting temptations in its dealings with a competitive complementary layer, there is an incentive to structure its relationships and contracts—for instance, using MFNs or exclusivity—so as to bring the number of monopoly markups back up to 1.[1] This in part frames the literature on vertical restraints in the "triangular" case.[2]

And if, absent e.g. a class of restraints, there is strong enough competition at each level so that the final prices to consumers are below the joint monopoly level—so that in total there is strictly less than one monopoly rent being extracted—then again the Coasean joint incentive (for the firms, not of course including consumers) is to bring the effective amount of monopoly up towards 1.0. Antitrust should of course seek to prevent that.[3]

This paper studies two models in which firms can use nonlinear contracts that reference rivals so as to raise prices above the imperfectly competitive level that would otherwise prevail. The models evaluate two widely discussed limits on anticompetitive vertical contracting. First, there is the widely studied commitment problem mentioned above. Second, there is a holdout problem (also well-known): a price-increasing deal with one downstream firm raises the reservation payoff of other downstream firms, creating a "positive contracting externality on the non-signer" (Segal, 1999), which discourages such deals. I show that each of these limits does indeed prevent profit-maximizing firms from contracts that would sustain the joint monopoly price—but does not prevent them from contracts that substantially raise the price above the imperfectly competitive "but-for" level.

Such partial—but not full—monopolization fits with my experience that enforcement agencies often encounter worrisome vertical contracts that reference rivals and thereby hinder trade with, but do not totally exclude, rivals. I use a "tax" metaphor—inspired by some instances in which it is more than a metaphor, but I hope more broadly illuminating—to calibrate such partial exclusion.

In what follows, a manufacturer M, facing rivals R, sells a wholesale product to distributors or retailers D1 and D2, who then compete to sell a retail product to consumers. Below, Section 2

---

[1] See for example Whinston (2006, p.160), Tirole (2015, section I.A).

[2] See for example Whinston (2006, p.176).

[3] Tirole (2015) and others have suggested attention to blocking firms' attempts to protect against a monopoly's self-destructive self-undermining temptations—to protect the Coase conjecture as a force protecting consumer welfare, one might say, in the triangle case. While that can have some policy appeal, personally I think it is more central to the antitrust agenda to block firms' attempts to limit competition in the much less studied "quadrilateral case."

discusses how vertical contracts that reference rivals can often discourage—or (at least metaphorically) tax—incremental trades between a dominant firm M's trading partners (here distributors, or D's) and M's rivals R. Such incremental trades, and especially those prompted by failure of M to fully serve consumers' needs, are essential to competition. Discouraging expansion of trade between D and R is complementary—for M's profits or for industry profits—with raising M's price. Thus, as I discussed at more length in an earlier version of this project, one should be very concerned with such tax effects. A possible Chicago riposte is that transaction costs—meaning difficulties reaching agreement on the bad contracts—will prevent harm, despite the prospect of higher aggregate profits. It is this riposte that I evaluate using the models below.

Section 3 flags (some) key questions that I'll seek to address.

Section 4 introduces a simple static model, and a linear-demand special case. In that model (in particular in its special case) it evaluates the two contracting challenges identified above, and shows that anticompetitive vertical restraints can readily emerge, although in the linear case full monopoly pricing will not—showing that it's analytically important to consider effects short of full monopoly.

However, it is difficult to discuss the static model seriously without yearning for some dynamics, so following Maskin and Tirole, section 5 introduces an infinite-horizon alternating-move dynamic model. I do not derive a full solution, but by leveraging off results in the static model, I show that anticompetitive contracts arise in dynamic equilibrium.

Section 6 is a mix of somewhat quirky literature review and comments on the need for robust or information-light guidance for enforcers in this very challenging area.

## 2. Limiting and Taxing Rivals' Trades with Customers

Limiting rivals' trades with customers is a central issue in antitrust.[4] It is conventionally discussed mostly in terms of one firm M discouraging (as e.g. in collusion) or preventing (as in raising rivals' costs or exclusion) a rival seller R from trading with M's customers D. But trade is two-sided, and there is a fundamental economic equivalence between discouraging M from trading with D and discouraging D from trading with M. This is often taught in public finance: it doesn't (generically) matter whether the buyer or the seller nominally pays a tax.[5]

A variety of vertical restraints, once agreed to, limit the degree to which the buyer can turn to the seller's rivals if the seller doesn't entirely satisfy the buyer's needs. Of course, to a degree the buyer will therefore insist on getting its needs met in the contract or via other mechanisms—but does that degree of insistence suffice? That's one way to formulate the basic question in this literature and in this paper.

Of course the starkest of these restraints, and probably the most studied, is fully exclusive dealing. Exclusive dealing can also be softened in various dimensions (in addition to duration).

[4] Edlin and Farrell (2015) suggest this as a central perspective, but I think most antitrust scholars would agree that it is at least a significant issue.

[5] There can be wrinkles, even important ones, in particular contexts (the interchange fee literature comes to mind).

For instance, the buyer may be allowed to buy a certain amount, or a certain proportion of its purchases, from other sellers. But if the limits on such behavior are binding, they limit the buyer's safety valve and the seller's penalty for failing to meet the buyer's needs (and if the limits aren't binding, it is not clear why the contract is negotiated).

Rather than a hard limit, a contract may specify pricing incentives for (partial) exclusivity. For example, there might be an exclusive price and a (higher) non-exclusive "list" price. Or the unit price might increase with the amount or proportion bought from the rival—which can amount to a tax, as Inderst and Shaffer (2010) and Farrell et al. (2010) analyzed. Explicit charges for such purchases arose for instance in Microsoft's "per processor" licenses to Windows, and it is also one (not the only) perspective on ECPR charges. Less explicit penalties for disloyalty have been alleged in various cases, notably including the FTC's settlement (2010) with Intel.

Other vertical restraints (can) also blunt the trade-expanding response by D and R to a worsening of the marginal deal that M offers D. Those include price-parity provisions and their non-price equivalents (better deals can't get preferential marketing by D to end-customers).

Some of these, such as per-processor fees, or ECPR, or purchase-share discounts, quite concretely amount to a (negotiated) tax on purchases from rivals.[6] But, ambitiously, I also view the tax idea as a helpful metaphor for negotiated disincentives for trades with rivals more broadly. At one level the idea is very simple: we are used to intuitions about how a tax on an activity operates as a (typically not absolute but often significant) disincentive for it. More technically, if we consider the choice of the buyer's trades, y, with the rival R, as providing them net benefit u(y), and if restraints cause a choice y=y* that is below what would otherwise be chosen, then $u'(y^*) > 0$, a disequilibrium choice if it were not for the restraints; the choice can be modeled as a response to a per-unit tax of $u'(y^*)$ combined with a lump-sum rebate of $y^*u'(y^*)$.

This is of course an imperfect model or metaphor: for instance (just to illustrate), with some of the restraints mentioned above, an increase in M's wholesale price to D automatically causes either a parallel increase in the tax (ECPR and often loyalty discounts) or a parallel increase in the retail price that D sets for R's products (price parity agreements). So I certainly don't claim that the analysis below literally covers those, or a fortiori all, restraints. But I do aspire to be learning something general about vertical restraints that choke off D's and R's competitive response to M's prices—not only (though that would be worthwhile too) something about those cases where the tax metaphor is especially tight. And after all Chicago thrived on making "price" metonymic for all aspects of a deal.

Because of this ambition and the gaps in the metaphor, I think what follows is more powerful as exemplifying theory (X can happen) more than as generalizing (Y must happen) or classifying (Z happens if and only if A) theory. In particular, my takeaway—perhaps not startling to this audience—is that the identified barriers to negotiating anticompetitive increases in prices through tax-like vertical restraints can be—I would say can easily be—pretty porous.

[6] As discussed below, this differs sharply from a quantity discount on purchases from this seller, because of how they respond to increases in total purchases.

### *A. Limiting Rivals' Trades with Customers, or Vice Versa*

## 3. Key Questions

Can such restraints plausibly (profitably) be negotiated, if their only effect is anticompetitive?

One well-known challenge (e.g. McAfee and Schwartz 1994) is commitment: when M negotiates with "the last" D in a series of negotiations, it has an incentive to undermine what has come before (and anticipation of this in itself undermines earlier negotiations). This of course depends on a view about whether M and other Ds can "respond" to "the last" negotiation. In a rigidly interpreted static model, by definition they can't. But even in the context of a static model, one can more intelligently and flexibly discuss the issue. And in an infinite-horizon model in which decisions matter but are not permanent, inspired by 1980s work of Maskin and Tirole, one can analyze the issue more quantitatively.

The second challenge is what Segal (1999) identified as the "positive contracting externality on non-participants" ("PCE") and that is sometimes called "holdout." When M and D1 sign a deal that raises D1's marginal costs, D2 benefits even (or especially) if it does not sign a deal with M, and this improvement in D2's reservation payoff discourages M and D1 from such deals. Another way to put this is that if a web of deals among M and the D's raises downstream prices and thereby increases their profits, a more than proportionate share of the anticompetitive gains must go to the D's, putting in question whether M will want to bother or can even afford it. Again, one can see the bones of this issue in a static model, but it is also helpful to frame the issue with explicit dynamics.

The next two sections, therefore, introduce first a static model and then an alternating-move infinite-horizon dynamic model. In each, I comment on those two challenges to the successful negotiation of M-D deals that yield anticompetitive profits by raising downstream final prices, despite the bulletproof presence of a robust fringe offering intermediate products at M's cost. In each, I find that the challenges do constrain the industry's ability to realize full joint monopoly outcomes, but do not constrain it to behave competitively.

## 4. Static Model

Manufacturer M and a competitive fringe of its rivals R, operating under constant returns to scale, compete to supply a homogeneous intermediate good to D1 and D2, who in turn compete to sell their symmetrically horizontally differentiated final products to consumers. M's unit cost is (for simplicity) 0; R's is $h$, which I will also take to be 0 for further simplicity.[7] Each D uses one unit of M's/R's product for each unit of output; again for simplicity, I assume the D's have no other marginal costs.

[7] I conjecture that when M is more efficient than R (h>0), it is easier for M to negotiate restrictive deals of the kind considered here, because D's reservation payoff will be lower. At the same time, many commenters would view this as problematic only if M's negotiated marginal wholesale prices exceed h. Extending my analysis to the case h>0 is high on my priority list.

M can offer contracts to the Ds; those contracts can be nonlinear and can reference rivals. A contract between M and D can specify (i) a per-unit price, denoted s or t below; (ii) a fixed fee, $F$, paid by M to D at the time the contract is signed; and (iii) a per-unit tax on purchases from R. In the simple model that tax will always be set equal to or higher than the per-unit price s or t, so for simplicity I do not define separate notation for it (see the final paragraph of this subsection).

If the Ds have constant unit (marginal and average) costs $s$ and $t$ respectively, write $\pi(s,t)$ for the first D's profits, and correspondingly $\pi(t,s)$ is the second D's profit (that is, the first argument describes "own" costs, and the second argument is "rival's cost").[8] Similarly, write $x(s,t)$ for D's output, and $p(s,t)$ for its price. Sometimes I will write $q(s,t)$ for the price charged by the rival of a firm, so $q(s,t) \equiv p(t,s)$.

Note for clarity that even limiting attention to what they pay for the input supplied by M or R, each D may not (in equilibrium often will not) have constant average costs, because of $F$; we keep that separate from $\pi$, so the profit measure for D is $\pi(s,t) + F$.[9]

If contract negotiations succeed, the agreed per-unit price $s$ or $t$ becomes common knowledge among D1, D2, and M. D1 and D2 then set retail prices p and q for their goods, then consumers observe those retail prices and choose quantities. If negotiations fail, the lack of an agreement also becomes common knowledge, as is the fact that the no-deal D will buy from R at cost h=0.

These private marginal costs (wholesale prices) for the D's fully determine their retail prices and outputs, which in turn determine consumer welfare, total welfare, and total profits, so they are the focus of our analysis. The fixed fees $F$ matter through affecting agreements on those marginal wholesale prices, and also affect the distribution of profits among the firms.

We investigate whether M can profitably negotiate a deal with each D that raises D's marginal cost, anticompetitively raising retail prices. The analysis ignores both efficiencies and exclusion; it's all about this particular mechanism for softening competition.

There is an expositional paradox in the way this paper has grown; I think it is only superficial but it's worth surfacing. I'm interested in the role of (literal or metaphorical) taxes on D's purchases from R, negotiated between D and M; of course such taxes can let M make more of the sales to D at a higher price. In pursuit of a tractable model, I've ended up with a stark model in which M makes all the wholesale sales, limit-pricing against the tax-disadvantaged R; and this is formally equivalent to a pure exclusive deal with a wholesale price negotiated at the same time. I certainly think of it much more in the vein of a tax on R that would more generally display M competing at a tax-based advantage, and the limit pricing in the model as only a stark and simplified version of that. But model is the model, and the reader can of course interpret it in either way.

[8] Remember that s and t are not the retail prices set by D1 and D2 but are the (marginal) wholesale prices that they set, so this is an "indirect" notation.

[9] Of course D could have other costs (or revenues) that can in principle be built in to the function $\pi$, as long as they don't vary with the variables studied here and as long as D remains active.

### A. *Two Negative Results*

I start by noting two negative results: circumstances in which M cannot raise D's marginal cost above R's offer price, h.[10]

#### 1. Contracts that don't reference rivals

First, when M's contracts cannot reference rivals, R's offer remains available to D. Of course M could raise its own price, but D1 and D2 would simply turn to R; they, and downstream consumers, would be unharmed. Because M's and R's wholesale products are undifferentiated substitutes, D's indirect cost function is the minimum of M's and R's prices.

More precisely, suppose that M offers D a general nonlinear tariff under which D can choose any quantity $z \in Z$ (which may be a limited or unlimited choice set) from M and pay $T(z)$ for it; then D's total cost for output $x$ is $C(x) = \min[h.x, \inf[T(z) + h.(x-z): z \in Z]$. Since one of D's options for expanding output from say x to $x'$ is to hold z fixed at the level chosen for output x (or to refrain from trading with M if that is what D chose for output x), and buy all the additional quantity from R, we get the inequality $C(x') \leq C(x) + h(x'-x)$, so D's per-unit incremental cost on that increment cannot exceed h.

Thus irrespective of the set of offered options Z and the tariff T, M cannot raise D's incremental (e.g. marginal) cost above R's offer price, h. Raising D's marginal cost above h requires violating the inequality $C(x') \leq C(x) + h(x'-x)$, which requires prohibiting or worsening at least some of D's options for incremental trade with R.

A couple of further comments on that:

a) Losing a Quantity Discount via Diversion

One popular objection to that reasoning is the following. Consider D's incentives when its contract with M does not reference rivals but does have quantity discounts, so $B(x,y) = b(x)$ with $b(x)/x$ decreasing in x. Typically if D sells one more unit of R's product y, it will sell some number (say e) fewer units of x; e is often called D's diversion ratio between M and R. If that reduction in x causes D to lose a quantity discount from M, why isn't that lost discount equivalent to a tax on the incremental unit of R?

I give two related answers.

When D *could* equally choose to maintain x (its quantity of M sold), even if it would choose not to, the envelope theorem applies: since D also optimizes x, one can calculate the profit impact for D of one more unit of y as if D held x fixed, and thus independent of the shape of the wholesale-price schedule for x. Thus D will ensure that its downstream marginal revenue from selling R's product (net of other marginal costs) is equal to, and in particular does not exceed, h, and this constrains the degree to which consumers can be overcharged in general.

Thus define D's profit as a function of its quantities, x and y: $R(x,y) - B(x,y) - hy$,

[10] I am generally assuming $h = 0$ for simplicity, but it seems clearer here to retain notation for R's offer price.

where R(x,y) measures D's profit gross of wholesale pricing. If $B(x,y) \equiv b(x)$, the first-order conditions from varying x and y separately are: $\partial R / \partial x = b'(x)$ and $\frac{\partial R}{\partial y} = h$. That last equation is the key point: M's own-quantity price schedule b(.) does not affect $\partial R/\partial y$.

Another approach is as follows. In the thought-experiment "sell another unit of y and re-optimize x," which is to say decrease x by e, the first-order condition is $\frac{\partial R}{\partial y} - h - e\left[\frac{\partial R}{\partial x} - b'(x)\right] = 0$. Since $eb'(x)$ appears in that equation, one might think that it would influence the equilibrium value of $\partial R / \partial y$, loosely consistent with the shape of b(.) affecting things much as a tax on y would, and I think this is the intuition we're considering.

But another possible deviation for D would be to increase x by 1 and (continuing with the same intuitive approach) then decrease y by the relevant diversion ratio, k ("sell another unit of x and re-optimize y"). That yields another first-order condition: $\frac{\partial R}{\partial x} - b'(x) - k\left[\frac{\partial R}{\partial y} - h\right] = 0$.

Unless $ek = 1$, that pair of equations (viewed as linear equations in the variables $\partial R / \partial x - b'(x)$ and $\frac{\partial R}{\partial y} - h$ ) is non-singular and implies $\partial R / \partial x = b'(x)$ and $\frac{\partial R}{\partial y} = h$, whatever the shape of $b(\cdot)$.[11] The correct intuition is that when M's contract doesn't reference rivals, D can buy y at marginal price h and nonlinearity of b doesn't create a tax.

This seems to me an important reason to focus, as Scott Morton (2013) vigorously helped lead us toward, on contracts that reference rivals: they have the potential to raise D's cost of buying "more input" in a way that other vertical contracts, including steep discount schedules, don't. But one may have to interpret this with care:

b) Ratchet Effects with Individually Tailored Quantity Discounts

In a dynamic setting, pricing may in practice reference rivals without explicitly doing so. This could arise through a conscious intent to reward or punish loyal or disloyal Ds "next time," but it can also arise without such intent. An example is the following. M offers D a nonlinear price structure b(x) that does not reference rivals. But M's choice of that structure (e.g., where quantity discounts kick in) depends on its beliefs about D's likely scale; and D's previous purchases from R (as well as from M) affect those beliefs. For example, M might offer D a quantity discount for D's purchases from M that are at least say 90% of the previous period's total (M plus R) purchases by D.

Now when D buys an additional unit of y, one consequence is that next period it will face a more demanding nonlinear pricing schedule. Thus holding fixed next period's planned quantities (in an envelope-theorem way), it would have to pay more for those purchases, the more it buys from R today. Then the present value to D of its bills payable to M does depend on the y's and not only on the x's. Each contract by itself doesn't reference rivals, but the pattern of contracts does.

---

[11] If each D faces a *constraint* that the total quantity (x+y) is constant, the envelope theorem would not apply, we would have ek=1, and the concern described in the text could have force. I believe this suggests that one should avoid simplified models with that total-quantity constraint. (Indeed, they deny the possibility of competition among the D's taking business from one another.)

### 2. Restriction to linear pricing

As a second "negative" result, M is also powerless to raise D's marginal cost if it can offer only linear contracts, i.e. if $F \equiv 0$. A linear contract with s>h would raise D's marginal cost but also raise D's average cost, so especially if it also limited D's use of R, it would be rejected. This idea is the key to a passage by Bork, discussing exclusive dealing by Standard Fashions:

> Exclusivity has necessarily been purchased from [the store, or D], which means that the store has balanced the inducement offered by Standard [=M], presumably in the form of lower prices, against the disadvantage of handling only Standard's patterns and found the inducement greater. The advantage to the store is also an advantage to consumers because monopolists…sell more at lower prices when their costs are lowered.[12]

In the present model, with the additional constraint/assumption $F \equiv 0$, this argument works; without that condition it fails because it is not at all "presumabl[e]" that the inducement takes the form of a lower marginal wholesale price—in fact, it won't. As with pay-for-delay, the fixed fee is key.[13]

## *B. Linear differentiated Bertrand duopoly downstream*

In parts of what follows, I specialize to differentiated Bertrand duopoly downstream with linear final demand. If D1 and D2 set prices *p* and *q* respectively, they get demand

$$x = 1 - p + dq$$
$$y = 1 - q + dp$$

Here *d* is the diversion ratio (symmetric). We can calculate the Ds' prices, outputs, and profit function, $\pi(\cdot,\cdot)$, in terms of these demand fundamentals. For convenience, define $K \equiv 4 - d^2 = (2+d)(2-d)$. Then we find:

$$Kp(a,b) = 2 + d + 2a + db$$
$$Kq(a,b) \equiv Kp(b,a) = 2 + d + 2b + da$$

Hence we have gross margins:

$$K(p-a) = 2 + d - (2 - d^2)a + db$$
$$K(q-b) = 2 + d - (2 - d^2)b + da$$

As is common in these standardized linear-demand frameworks, formulae for quantities echo those for gross margins:

$$Kx(a,b) = 2 + d - (2 - d^2)a + db$$
$$Ky(a,b) \equiv Kx(b,a) = 2 + d - (2 - d^2)b + da$$

Hence, again familiarly, the profit function is a perfect square:

$$\boldsymbol{K^2\pi(a,b) = (2 + d - (2 - d^2)a + db)^2}$$

---

[12] Bork, *The Antitrust Paradox* (1978), p.307.

[13] Like Bork, Klein and Murphy (2008) state that (in their model) linear pricing is an innocuous assumption; that is not the case in the present model.

[We can use the fact that in this special model, $x(a,b) = p(a,b) - a$.]

Notice that this is not the (more usual) profit function of the choice variables $p$ and $q$. In particular, while the cross-derivative of a firm's profit with respect to $p$ and $q$ is positive (strategic complements), we will use below the fact that:

$$\boldsymbol{K^2 \frac{\partial^2}{\partial a \partial b} \pi(a,b) = -2d(2 - d^2) < 0}$$

An integrated monopolist would set the (symmetric) retail price $p = q = {}^1\!/_{2(1-d)}$, and an unintegrated monopolist M setting wholesale price $t$ but not directly controlling p and q, would induce downstream monopoly retail pricing by setting t at the joint-monopoly wholesale price

$t = t^m = {}^d\!/_{[2(1-d)]}$

### *C. Commitment*

In a static model with unobservable contracts, if M has signed (or expects to sign) a deal with D1 at marginal wholesale price *t*, then M and D2 jointly maximize their surplus by setting a marginal wholesale price *dt*, as that is then M's private marginal (opportunity) cost of a marginal sale to D2. When *d*<1, if we envision M and D2 "best-responding to" the deal struck (or expected) between M and D1, this "unravels" and the only static equilibrium is *t*=0. This is a familiar argument in the literature, and in itself does not depend on the existence of R.

A natural response is that, if *d* is relatively large (close to 1), departing from that ex post bilaterally optimal value of *dt* and (for instance) charging *t* instead sacrifices only a second-order amount for M and D2. MFN provisions and simple resistance to temptation might well enable M to refrain from undermining its own power—humans and organizations often give in to self-destructive temptation, but not always. Indeed, although the argument is familiar in the literature, I don't think most antitrust enforcement would take the view—nor should it—that a monopoly selling through multiple outlets (even) without restrictive contracts is harmless or powerless.

In our model, there is a different response: contracts are observable before retail pricing occurs, and this makes the best-response function quite different. I analyze this in the linear model. Contracts (or the per-unit element, *s*, which is what counts) might in reality be promptly observable via mechanisms to enforce MFNs, for instance (especially where such mechanisms make higher profits more sustainable), or might be leaked, or each D might be able over time to infer the other D's private marginal cost from its market conduct. In any case, I don't claim that contracts are necessarily observed, but it seems too strong to claim that they can't be—thus we should explore this branch.

#### 1. Commitment in the Linear Model

Following the literature, we consider a game played between (a) M and D1, negotiating a deal in the absence of D2, and (b) M and D2, negotiating a deal in the absence of D1. It's a somewhat odd kind of game in that M is on both teams, but the idea is that each team maximizes its joint payoff, by choosing its agreed-on marginal wholesale price, given what it expects the other team to be doing in that respect. (Each team also determines a value of *F*, but we can keep that in the background.) In negotiating the value of t to charge D1, M takes into account the effect on its revenues from D2; but a surprise (off-equilibrium-path) outcome of the M-D2 negotiation would not factor into the M-D1 negotiation.[14]

Suppose that M and (say) D2 know or believe that M and D1 will trade at marginal price s. Given s, their optimal response $t$=$\rho(s)$ maximizes their joint payoff

$$\boldsymbol{J(s,t) - \pi(s,t) = \pi(t,s) + sx(s,t) + tx(t,s)},$$

where the function *J* represents total three-way profits, from which we subtract D1's product-market profits—this way to express it turns out to be useful below.

Importantly, in this formulation, M and D2 effectively assume that D1 has irrevocably agreed (or will inevitably agree) to pay some *s* that M-D2 now can't affect. This analysis therefore focuses exclusively on the commitment issue, omitting the PCE issue: PCE concerns how a higher-price M-D2 deal worsens M's bargaining position in an upcoming negotiation with D1. (Next, I will turn to that issue separately in the static model, and then comment on both a dynamic model.)

In our linear case, it is convenient to multiply by $K^2$, and we have

$$\begin{aligned}\boldsymbol{K^2[J(s,t) - \pi(s,t)]}& \\ &\boldsymbol{= (2+d+ds-(2-d^2)t)^2 + Kt(2+d+ds-(2-d^2)t) + Ks(2+d+dt} \\ &\boldsymbol{-(2-d^2)s)}\end{aligned}$$

The first-order condition with respect to *t* is

$$\begin{aligned}0 = 2(2+d+ds-(2-d^2)t)[-(2-d^2)] + K(2+d+ds-(2-d^2)t) - Kt(2-d^2) \\ + Ksd\end{aligned}$$

so *t* will be set at

$$\boldsymbol{\rho(s) = \frac{d^2+2d+4s}{4(2-d^2)}d.}$$

This reaction function is upward-sloping and has a strictly positive intercept. The upward-sloping reaction function is significant in the dynamic model later, so we give this feature of the market environment a name: strategic complements in wholesale prices (SCWP).[15]

To find the symmetric equilibrium wholesale price, set $\rho(s) = s$ and solve, giving:

$$\boldsymbol{\sigma = \frac{d^2}{4(1-d)}}$$

[14] In a sense this is like the "separate negotiating teams" in the FTC's remedy in *Evanston Northern*. Many economists have commented that, even if fully enforced, that remedy would not resolve the anticompetitive effects of the merger; in a sense that represents a rejection of the argument that the commitment problem here would prevent M from exercising any power.

[15] I derived it in this very specific model, but don't know how widely it's likely to hold. The well-known model of private offers, with its best-response function $t = ds$, also satisfies SCWP (although it differs fundamentally, so I don't claim the same analysis applies).

This is ($d$/2) times the monopoly-inducing wholesale price $t^m$. In this model the commitment issue, interpreted as symmetric Nash equilibrium, holds the wholesale price below half of the monopoly-inducing wholesale price, but for large values of the diversion ratio, it approaches that bound.

For $s < \sigma$, $\rho(s) > s$, which means that increasing $t$ slightly above $s$ raises three-way total profits $J(s,t)$ by more than it raises $\pi(s,t)$. Using subscripts for partial derivatives, $J_2(s,s) > \pi_2(s,s)$ for $s < \sigma$. We will use this result in the dynamic model below.

### *D. Overcoming Holdout Incentive (PCE)*

We now separately address the holdout or PCE issue. Separation is achieved by allowing M to make simultaneous (symmetric) public offers to the two Ds, who must simultaneously accept or reject; thus we set aside whether the incentives for higher wholesale prices would be undermined ($\rho(s) < s$) or emboldened ($\rho(s) > s$) by an ability of M and one D to "respond" to the deal struck with the other D.

#### 1. General treatment of small t

Suppose that M offers each D the contract: "pay me t≥0 per unit, and accept lump sum F in return."

If accepted by both Ds, this scheme is profitable for M (relative to limit pricing at $h$=0) if and only if $tx(t,t) \geq F$ (here and below I am indeed using the simplifying assumption $h$=0). If each D expects the other to accept, and expects the other's deal to stay in place even if it itself declines the offer, then the condition for D to accept is $\pi(t,t) + F \geq \pi(0,t)$. The PCE is the fact that $\pi(0,t)$ increases with $t$.

Define $G(t) \equiv \pi(t,t) + tx(t,t) - \pi(0,t)$. Thus there exists $F$ strictly satisfying both the acceptance and the profitability (for M) conditions, if and only if $G(t) > 0$.

Since $G(0) = 0$, a sufficient condition for this to hold for some (indeed, all sufficiently small) $t$>0 is that $G'(0) > 0$.[16] Using subscripts for partial derivatives,

$$\boldsymbol{G'(0) = x(0,0) + \pi_1(0,0)}$$

Observe that $2G(t) \equiv J(t,t) - 2\pi(0,t)$. Moreover, starting at symmetry, $(t,t)$, the effect on three-way total profits of raising one wholesale price is the same as the effect of raising the other wholesale price. Consequently, $2J_1(t,t) = 2J_2(t,t) = J_1(t,t) + J_2(t,t)$. The condition $G'(t) > 0$ thus means that slightly increasing just one $t$ raises three-way total profits $J(t,t)$ by more than it raises $\pi(0,t)$—in other words, that

$J_1(t,t) = J_2(t,t) > \pi_2(0,t)$.

If the Ds do not compete, or ignoring "strategic" effects, the impact of a small change in unit cost on D's profits is of course equal to its output, so $\pi_1 = -x$ and $G'(0) = 0$: our sufficient condition (barely) fails. In fact (with downward-sloping demand) without strategic competition

[16] This sufficient condition is a necessary condition if G is concave. In the linear example, the maximands I deal with are linear-quadratic in t (not in d). Outside the linear example, I assume that maximands are single-peaked where I need to.

between the Ds, there is no profitably sustainable $t$>0. This is a version of the "one monopoly rent theorem"—M and D jointly would like D to exercise whatever pricing power they jointly possess without its choices being distorted by a wholesale price above marginal cost.

But the narrowness of the sufficient condition's failure is an indication that it doesn't take much to reverse it. In particular, when the D's compete as strategic complements (as they do in our linear example), $G'(0) > 0$, due to the "strategic" positive value of having a higher marginal cost: one's rival prices higher as a result.[17] That is, M and D1 are jointly first-order indifferent—though second-order averse—to raising $t$ just above zero, until they recognize that when D2 sees them doing so, D2's price will rise, benefiting M and D1 jointly to first order.[18] (By the envelope theorem applied to D2's pricing, D2 is first-order indifferent to that price increase in itself, so M and D1 jointly gain whenever the price increase raises three-way (industry) joint profit.)

### 2. Versions of D's reservation payoff (intuiting reduced-form dynamics)

The above assumes a stringent version of the condition for M and D1 to find gains from trade in higher values of $t$, because it depicts the outcome as gifting D2 an improved reservation value of $\pi(0,t) > \pi(0,0)$. This is the "positive contracting externality" (on the non-participant D2), as in Segal (1999). Indeed this seems to follow if we take literally that M makes simultaneous offers to both Ds and that there is no further action after D1 and D2 simultaneously accept or reject. Still, one should recognize that the timing in this static model is a simplification, and in reality a response would make sense—if there were any opportunity to respond, the assumption of no response would be irrational ex post and would leave M and D1 exceptionally vulnerable ex ante. Sticking rigidly to the static model is like—in fact, precisely like—asking whether a horizontal cartel can operate in a static model, where each participant simultaneously chooses its price or output taking as given others' choices, and with no chance for responses.

But what response? And what about responses to responses…? That takes us toward a fully dynamic model, which is more satisfactory in principle but much harder to solve (see below). Before going there, let me suggest that at least three formulae for D2's reservation payoff seem simple but reasonable, outside the strict model but as it were not far outside it. First, as above, we might assume that M and D1 are caught flat-footed; as above, this yields D2 a reservation payoff $\pi(0,t)$. Second, M and D1 might quickly renegotiate, or to have written into their contract an ex post optimal choice of what to do if D2 rejects; writing $\rho$ for the best-response function, that would yield D2 a reservation payoff $\pi(0,\rho(0))$.[19] Third, M and D1 might abandon their attempt to raise final prices if D2 doesn't sign on, in which case D2's reservation payoff would be $\pi(0,0)$.

---

[17] The calculation is taken from the appendix of Farrell and Shapiro (2008), where it was applied to the licensing of a probabilistic patent. In both cases, the issue is a downstream firm's willingness to pay an elevated marginal cost in return for a reverse fixed fee together with an expectation that rivals will also pay the elevated marginal cost.

[18] By the same token, anything that pushes in the opposite direction can boost the failure from borderline to strict. In particular preliminary calculations with a linear Cournot model downstream seem to support that with strategic substitutes, the PCE is much more constraining.

[19] Because I assume that wholesale prices become common knowledge before retail prices and outputs are set, this best-response function is not simply $\rho(t) = dt$, as would be the case with private offers.

## 3. PCE in the linear model

In our linear model, multiplying by $K^2 = (4-d^2)^2$ for convenience, we have

$$\begin{aligned} K^2G(t) &= K^2[p(t,t)x(t,t) - \pi(0,t)] \\ &= (2+d+2t+dt)(2+d-(2-d^2)t+dt) - (2+d+dt)^2. \end{aligned}$$

This represents (a constant times) M's profit from reaching symmetric agreement on t if it must give each D a payoff equal to D's reservation value $\pi(0,t)$. Maximizing this with respect to t will display in pure form the PCE effect on profits and joint profit maximization. The first-order condition is:

$$\begin{aligned} 0 = K^2G'(t) = (2+d)(2+d-(2-d^2)t+dt) - (2-d^2-d)(2+d+2t+dt) \\ - 2d(2+d+dt). \end{aligned}$$

Collecting terms with and without *t*, this can be written

$$(2+d)[2+d-(2-d^2-d)-2d] = [2(2+d)(2-d^2-d)+2d]t,$$

yielding:

$$\tau = \frac{1}{2}\frac{(2+d)d^2}{(4-2d^2-d^3)}$$

It is also straightforward to calculate that $G(0) \geq 0$ whenever *t* is below twice $\tau$:

$$\boldsymbol{t \leq 2\tau = \frac{d^2(2+d)}{4-2d^2-d^3}}$$

At d=0 this constraint amounts to t≤0. This simply re-derives the Bork result that with non-competing direct buyers D, the scheme with t>0 is a loser. (Even though it doesn't have to overcome a positive contracting externality, as there are no market spillovers between the Ds, there is simply no joint gain available to M and any D, or to M and the Ds collectively, let alone to them and R collectively). And as the $d^2$ in the numerator suggests, for fairly small values of d, the problem will be relatively mild. For example, if $d = 0.2$, the joint-monopoly wholesale price is $\frac{d}{[2(1-d)]} = 0.125$, while $\tau$ is about .01.

But $\tau$ and the upper bound on *t* are strictly increasing in *d* (between 0 and 1), so when the Ds actually compete with one another (*d*>0), for every sufficiently small *t*>0 the scheme overcomes even the assumed very demanding version of the PCE holdout condition. And "sufficiently small" needn't mean economically negligible: *t* can be a substantial fraction of the monopoly-inducing level, capping out at a bit over 40% (at around *d*=0.7). See Figure 1.

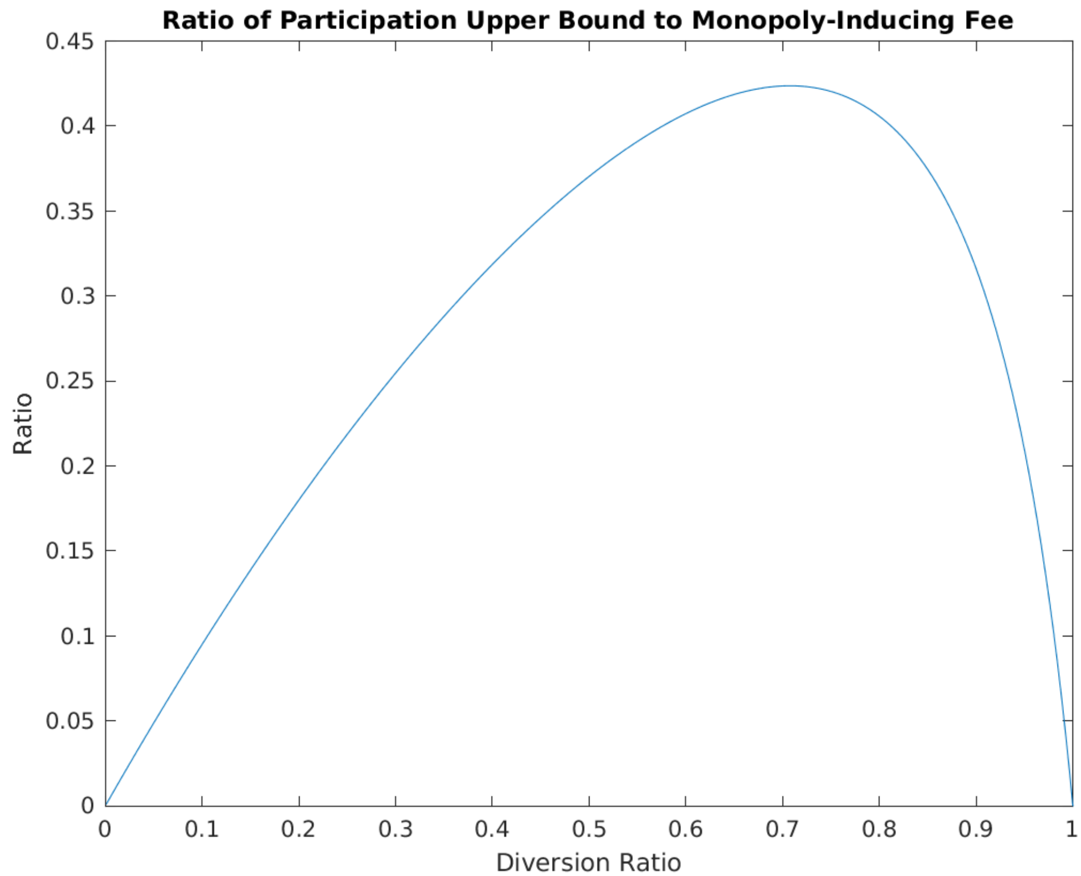


Figure 1.

Observe that:

**Proposition: If $\pi_{12} < 0$, as is the case in the linear example, the static Nash marginal wholesale price, $\sigma$, is at least as great as the value $\tau$ that maximizes $G(\cdot)$. That is, $\sigma \geq \tau$. If** $\pi_{12} > 0$ then $\sigma \leq \tau$.

We can verify $\sigma \geq \tau$ by direct calculation in the linear example, but a more general and perhaps more illuminating calculation is as follows. Recall that $\tau$ is the value of $t$ that maximizes $G(t) = J(t,t) - 2\pi(0,t)$, so the first-order condition is $0 = G'(t) = J_1(t,t) + J_2(t,t) - 2\pi_2(0,t) = 2[J_2(t,t) - \pi_2(0,t)]$, using symmetry of $J$. Meanwhile $\sigma$ is the value of $t$ that maximizes $J(\sigma,t) - \pi(\sigma,t)$, so the first-order condition is $0 = J_2(\sigma,\sigma) - \pi_2(\sigma,\sigma)$. Thus $G'(\sigma) = 2[J_2(\sigma,\sigma) - \pi_2(0,\sigma)] = 2[J_2(\sigma,\sigma) - \pi_2(\sigma,\sigma) + \pi_2(\sigma,\sigma) - \pi_2(0,\sigma)] = 2[\pi_2(\sigma,\sigma) - \pi_2(0,\sigma)]$ (using the first-order condition for $\sigma$), and the cross-derivative condition on $\pi$ implies that this is negative, so that $\sigma$ is too large to maximize $G$.

In this model, while substantial anticompetitive effects are possible, full monopoly pricing is not sustainable for any value of d. If exclusive dealing is a clumsy instrument that can't include a commitment to stop short of monopoly pricing (in game-theory terms, exclusives are signed before prices are set—perhaps because of cost uncertainty or the like), substantial anticompetitive joint gains can only be realized through the use of a more nuanced mechanism, such as a limited (negotiated) tax on purchases from R. So if an economist were to simply check whether or not M can profitably convince Ds to accept full exclusive dealing (with monopoly pricing, as M can't commit otherwise), she would find that, Chicago-style, the PCE makes such contracts unprofitable. But it would be wrong to interpret that as showing that exclusive-dealing concerns are absent here: partial exclusive dealing of the kind we're considering can profitably arise despite the PCE. Dissenting appeals Judge Greenberg in *Eaton-Meritor* suggested that "de facto partial exclusive dealing" was oxymoronic (or at least neologistic in a bad way), but this analysis shows that such contracts may be profitable and quite harmful. The tax analysis enables us to see these risks, which would be missed or obscured if we focused on pure exclusive dealing with concomitant monopoly wholesale pricing.

## 5. Alternating-Move Model

To further study the issues of commitment, reaction and holdout, I use an alternating-move Markov model inspired by Maskin and Tirole (1982 et seq.). As Maskin and Tirole argued, this helps us formulate some of these questions more rigorously than trying to shoehorn them into what is formally a static model, and more responsively than treating the static model literally.

In each of infinitely many discrete periods, M and one D or the other have an opportunity to negotiate a contract that will prevail for two periods. In odd-numbered periods, M and D1 can contract; in even-numbered periods, M and D2 can contract. Within each period, contract negotiations happen first and the outcome is revealed, followed by retail pricing and the determination of outputs.

Thus in an even-numbered period, D1 either (a) is in the second period of a two-period contract negotiated at the beginning of last period, in which case it still has constant unit cost equal to s, the per-unit price that it agreed to pay M in that negotiation; or (b) failed to reach agreement with M last period and thus has constant unit cost 0.[20] It cannot renegotiate now but must wait for next period. All those facts are common knowledge. In that environment, M and D2 now negotiate a per-unit price t and a fixed fee (lump-sum transfer) F, each firm seeking to maximize its present discounted value of profits, given s and assuming that both parties, as well as D1, foresee how the game will play out. Unlike the artificial "Evanston Northern" assumption in the static model, M (as well as the Ds) observes and can (but only later!) react to any off-equilibrium deal struck by M and one of the Ds.

I assume that only M can make an offer, so it sets $t$ to maximize the sum of its and D2's present-value profits, given the deal in place with D1 ("$s$") and the prospects for how things will evolve over time; it also sets $F$ to hold D2 to the present value of profits that D2 could expect if there were no agreement this period; and in equilibrium D2 (barely) accepts.

[20] Of course, in an odd-numbered period, things are the same except that the roles of D1 and D2 are inverted.

I write $t = r(s)$ for the dynamic reaction function; in Markov perfect equilibrium, *t* does not otherwise depend on history, nor does it depend on any $F$. As discussed by Maskin and Tirole, *r* is a true "reaction function," in that M and D2 react to *s* according to it.[21]

Observe that D2 and M jointly face exactly the same problem if D1 failed to agree with M last time as they do if D1 agreed on s=0 last time; thus they will agree on $t = r(0)$ in either case.

With *s* still in force from the previous period, and having just negotiated *t* (or failed to), the intra-period product market phase is as in the static model. That is, downstream duopoly competition between D1 and D2 yields D1's intra-period price $p(s,t)$ and output $x(s,t)$, with resulting quasi-profits $\pi(s,t)$; and symmetrically yields D2 quasi-profits equal to $\pi(t,s)$, with price $q(s,t) \equiv p(t,s)$ and output $x(t,s)$. M's intra-period quasi-profits are equal to $sx(s,t) + tx(t,s)$, since *s* and *t*, which are costs to the D's, are revenues for M. Three-way joint intra-period profit is $J(s,t) = p(s,t)x(s,t) + p(t,s)x(t,s)$, and the omission of $F$ doesn't matter for that calculation since it is just a transfer. In some but not all of what is below, I will specialize to the linear differentiated-product special case treated above.

It is helpful to define notation for each party's present value—now including $F$ in each case—as evaluated at the outset of a period's negotiation phase. Continue to write *s* for the carry-forward unit price, committed to last period by say D1. Then each firm's present value depends on s and is otherwise independent of history, but also depends on the deal $(t, F)$ that M and D2 are about to strike—for some purposes we need to keep $(t, F)$ as separate variables, while recognizing that they themselves are predictable functions of *s*.

Write $U(s)$ for D1's present value at the beginning of such a period. Then as an accounting matter we have:[22]

$$U(s) = \pi(s,t) + \delta V(t)$$

Similarly, writing $V(s)$ for D2's present value, we have

$$V(s) = F + \pi(t,s) + \delta U(t)$$

But we also know that D2 will be held to its reservation present-value payoff. If D2 failed to reach agreement with M, it would trade with R instead, at cost h=0; in the current period that would be competing against the established M-D1 wholesale price, *s*, but next period M and D1 would renegotiate their marginal wholesale price to $r(0)$. Thus:

$$V(s) = \pi(0,s) + \delta\pi\big(0, r(0)\big) + \delta^2 V(r(0))$$

from which we get both that $V'(s) = \pi_2(0,s)$, so $V$ is increasing, and also (setting $s = r(0)$) that $(1-\delta^2)V\big(r(0)\big) = (1+\delta)\pi(0, r(0))$, giving us the solution for $V$ if we knew $r(0)$, viz.:

[21] The quite distinct one-shot Stackelberg reaction function (sometimes confusingly called the Nash best-response function), which I denote $t = \rho(s)$, was discussed above and solved for in the linear case.

[22] Below, I write out some equations partly in order to clarify what the model is, partly to record my progress towards solving it (which I haven't done but hope to do), and partly because I will be able to make use of one or two key equations.

$$V(s) = \frac{\pi(0, r(0))}{1-\delta} + \int_{r(0)}^{s} \pi_2(0, z)\mathrm{dz} = \pi(0, s) + \frac{\delta}{1-\delta}\pi\big(0, r(0)\big)$$

(One could also see this more directly: because D is indifferent to reaching agreements at each negotiation node, it is also indifferent to never reaching agreement at any of those nodes, in which case after the initial period at *s*, its rival will always be at $r(0)$.)

We can also write *F* as a function of $(s, t)$, and because *F* is set at a constraint rather than optimizing a smooth function, it is easier to interpret it as that function rather than as a choice variable:

$$F = V(s) - \pi(t, s) - \delta U(t).$$

We also have M's present value, $W(s)$:

$$W(s) = -F + sx(s, t) + tx(t, s) + \delta W(t)$$

Finally we have the joint optimization of *t* by M and D2 to maximize the sum of their present values (of course *F* drops out):

$$\boldsymbol{t = r(s)\mathbf{max}[J(t, s) - \pi(s, t) + \delta U(t) + \delta W(t)]} \tag{1}$$

In this joint maximization, the first two terms constitute M's and D2's joint intra-period payoff, as studied in the static model above. In the two final ($\delta$-discounted) terms, only t appears, so the cross-derivative vanishes (in economic terms, by the time next period rolls around, s no longer applies). Thus the cross-derivative with respect to s and t is equal to—and *a fortiori* therefore has the same sign as—the cross-derivative of the intra-period joint payoff. Consequently, although solving for the function $r(\cdot)$ is difficult, we know that it is increasing/decreasing if and only if the cross-derivative of the intra-period joint M-D2 payoff, $J(t, s) - \pi(s, t)$, is positive/negative.[23] But we calculated that the linear static duopoly model has the feature of strategic complements in wholesale prices (SCWP)—in the M-D1/M-D2 game, the Stackelberg reaction function is increasing (even though $\pi_{12} < 0$). Thus the dynamic reaction function $r(\cdot)$ is also increasing.

Could there be a steady-state Markov equilibrium at s=0? If there were, starting at $(0,0)$, at the beginning of a period in which M and D1 are negotiating, consider the effects of their agreeing to a slightly higher marginal wholesale price, d*w*, for one cycle (two periods) only, followed by reversion to the zero steady state. To first order, this changes total three-way first-period joint profits by d*w* times $J_1(0,0)$. In the second period, D1 continues to pay d*w* at the margin, and D2 will negotiate with M a marginal price equal to $r'(0)dw$, so the impact on second-period three-way joint profits is dw times $J_1(0,0) + r'(0)J_2(0,0)$. In the third period, M and D1 revert to the but-for path, but D2 is still paying $r'(0)dw$, so three-way joint profits are $r'(0)J_2(0,0)dw$ above steady-state. In the fourth period and thereafter, we are back on steady-state conduct and outcomes. The effect on three-way joint present discounted profits, evaluated at the beginning, is therefore d*w* times

$$\boldsymbol{J_1(0, 0) + \delta[J_1(0, 0) + r'(0)J_2(0, 0)] + \delta^2 r'(0)J_2(0, 0).}$$

[23] This is a quicker and more intuitive proof and perhaps a modest extension of a lemma in Maskin and Tirole.

How much of this does D2 collect? In the first period D2 gets a product-market windfall of $\pi_2(0,0)\mathrm{d}w$. At the beginning of the second period, $F$ adjusts, to continue to hold D2 (in present-value terms) to its reservation value. Consequently the profit effect on D2 for the second and all subsequent periods combined, is simply the change in its reservation value at that node, which from above is simply $\pi_2(0,0)\mathrm{d}w$. Thus, of the three-way joint present value profit impact, D2 collects $(1+\delta)\pi_2(0,0)\mathrm{d}w$. That leaves the change in M's and D1's combined present values as:

$$J_1(0,0) + \delta[J_1(0,0) + r'(0)J_2(0,0)] + \delta^2 r'(0)J_2(0,0) - (1+\delta)\pi_2(0,0)$$

(all times d$w$). Given that $r'(0) \geq 0$, we can sign this as positive when the intra-period game involves strategic complements, because we know from the static analysis that $G'(0) > 0$, i.e. $J_1(0,0) = J_2(0,0) > \pi_2(0,0)$.

This tells us that the benign outcome in which M and R price at 0 is not an equilibrium, let alone a unique one. Higher-priced outcomes arise out of natural Markov reactions; there is no need (although of course it can happen) for the parties to view that as a strategic goal, and competitive (here zero) wholesale prices will not survive—if the nonlinear contracts and restrictions on rivals are available.

Similar logic in fact gives us a strictly positive lower bound on any steady-state value of $s$, at least for the linear example. Consider any $s < \min[\tau, \sigma]$, where recall that $\sigma$ is the "Nash" value in the static analysis, equal to $\frac{d^2}{[4(1-d)]} > 0$ in that example, and $\tau$ maximizes $G(t)$, equal to $\frac{1}{2}\frac{(2+d)d^2}{(4-2d^2-d^3)} > 0$ in that example. Recall also that in the linear example, or more generally with $\pi_{12} < 0$, we have $\sigma \geq \tau$ so $\min[\tau, \sigma] = \tau$.

Consider M and D1 negotiating in a state where, last period, D2 had agreed to a marginal wholesale price $s < \min[\tau, \sigma]$. I will claim that $r(s) > s$. If a hypothetical steady-state Markov equilibrium at $s$ were expected, there would be an incentive to deviate upwards.[24]

To see this, consider the effects of M and D1 setting $t$ slightly above $s$ now, and then in their next negotiation (two periods from now) reverting to $s$. The effect on three-way joint present value profit is, much as above (but starting at $s$, not 0):

$$J_1(s,s) + \delta[J_1(s,s) + r'(s)J_2(s,s)] + \delta^2 r'(s)J_2(s,s)$$

The effect on D2's present value profit is, much as above, $\pi_2(s,s) + \delta\pi_2(0,s)$. But because $s < \sigma$, $J_1(s,s) > \pi_2(s,s)$. And because $s < \tau$, $J_1(s,s) = J_2(s,s) > \pi_2(0,s) \geq 0$. Since we have $r'(s) \geq 0$, it follows that the change in D2's present value profits is less than the change in three-way joint profits—in other words, that M and D1 jointly gain from their deviation. Thus we have:

**Proposition. There is no steady-state Markov perfect equilibrium at $s < min[\sigma, \tau]$. Starting from such s, there is an incentive for upward deviation: $r(s) > s$.**

---

[24] Narrowly the proof shows no steady-state Markov equilibrium at such small values of s. I believe that with reasonable assumptions (true in the linear example) about single-peaked objective functions it can be expressed this bit more concisely.

When $r$ has slope less than 1, we can expect there to be a symmetric steady-state Markov equilibrium, so the finding that there are no such equilibria at zero or low wholesale prices implies that there is one at higher prices. If the slope of $r$ is above 1, a steady state would be unstable, but I think informally we can observe that, starting from low prices, the incentive for M and D1 to raise the wholesale price to be paid by D1 would be strong because it would strongly raise the price that D2 would agree to, increasing joint profits (it doesn't matter that those profits would go in part to D1 via an increase in next period's $F$).[25]

A very similar analysis shows that it cannot be a steady-state Markov perfect equilibrium to charge wholesale prices $t^m$ and thereby support the full joint monopoly outcome: the partial derivatives of J are zero at $(t^m, t^m)$ and therefore since $\pi_2 > 0$ there is an incentive to deviate downwards.[26]

## 6. Literature (written and unwritten)

Two literature reviews are needed here: a review of the analytical academic literature, and a review of what Keynes' "practical men… usually the slaves of some defunct economist" tend to think. I start with the latter, with apologies for inevitably (perhaps badly) oversimplifying and underestimating. And this is partly because when I was a "practical man" in Washington, I felt that the literature ought to give us more robust guidance. One respect in which it fell short was in the degree to which it focused on what Whinston (2006) calls the "triangle case," in which even absent restraints there is already a monopoly at one level. While there can of course be challenging issues in that context, it does not "really" get to grips with whether restraints move an industry toward market power that would not exist without the restraints. Secondly, to the extent that the literature did address that (analytically difficult) "quadrilateral case," there was a need for a way of stepping beyond very specific models. Each vertical restraint tends to be different, and the post-Chicago literature tends to stress that details matter. Is that the stubborn truth about the world, or are there opportunities for general guidance?

This paper doesn't address that overarching question (yes I think there are), but it does try to explore an important aspect of it. There are multitudinous ways in which M might ask D to hinder, limit, burden, and frictionalize D's trades with R. It's not surprising that M unilaterally would want such limits; the great contribution of the Chicago school here is to insist that it's much less clear that M "can afford" to persuade D to agree, and that "can afford" is not about M's wealth but is about an agreement's impact on joint value. Thus this paper considers a negotiated increase in the wholesale price that D will pay, which I mean as a catch-all metaphor for lots of different possible limits on D-R trade, rather than as a literal contract provision; and asks whether such limits might increase joint value.

[25] Caution is appropriate, however, because these models can be quite confusing!

[26] The equations here for the gain from deviation make me intuit there is "almost" a result ruling out values of s above $\max[\sigma, \tau]$, but I am not seeing how we would deal with the term in $\delta^2$. There is some "slack" in the arguments in that some of the terms are signed by the comparison between s and $\sigma$, and others by the comparison with $\tau$, and I conjecture that a result is to be found by "using" that slack to address the $\delta^2$ term.

Of course this question was not new, and Bernheim and Whinston (1998) had given an impressively general treatment of exclusive dealing, but it may have been too intimidating for full policy impact. Many "practical men" simplify this down to looking for whether an exclusive deal between M and D1 gave M additional power over "other" buyers D2 *through denial of scale* to entrants.

This denial-of-scale or divide-and-conquer theory of harm, associated with Rasmusen et al. (1991) and Segal and Whinston (2000), pictures M as harming its direct trading partners or buyers—in the simplest forms of the theory, when M negotiates with any one buyer B, M's deals with other buyers have already foreclosed (or will inevitably foreclose) B's option to trade with M's rival R. While this is obviously an important theory (and analyzed with impressive elegance and rigor), I was surprised to find that it was widely assumed to be *the* intellectually coherent theory of harm.[27] What about just agreeing to limit competition so that they could raise prices and extract more money from consumers? Of course, D could raise prices anyway, so the point would have to be that multiple otherwise-competing Ds are involved. And it requires that M's direct trading partners D are not themselves the (same) final consumers. But in themselves these are frequent circumstances. Inspired by some of the same issues as I was, DeGraba (2013) developed these ideas at the FTC (where John Simpson also worked at the time).

Some literature addressed this idea primarily as a modification of the divide-and-conquer theory: Fumagalli and Motta (2006, 2008), Simpson and Wickelgren (2007), and Abito and Wright (2008). Simpson and Wickelgren explicitly discuss the case of no economies of scale, but in my experience denial of scale to rivals largely remained (probably remains) the most popular go-to theory. This was a bit surprising, and may have undermined the influence of concerns about vertical restraints, because denial-of-scale theories are in a sense isomorphic to predatory pricing concerns. In each case, there is an economy of scope between sales to one group of buyers and sales to another group, and a firm can monopolize sales to one group by taking all, or "too much," of the sales to the other.[28] Of course, being isomorphic to predatory pricing concerns is not a recipe for a competitive concern to be hospitably received by most courts. In contrast to those purely exclusionary concerns, this paper joins some of those above in focusing on the creation of joint value for participating firms—which in my model includes not only M but the Ds, and can also be made to include fringe firms R, making them less apt to complain or resist with full vigor—at the expense of downstream consumers. In a related vein, Asker and Bar-Isaac (2014) studied the possibility that retailers who are dealt in to a profitable monopoly may be reluctant to upset the apple cart by cooperating with a disruptive rival entrant manufacturer.

Perhaps the work closest to this is Justin Johnson (2014). Johnson considers the formation of (in general) several networks of upstream and downstream firms linked by exclusive dealing contracts and internalizing competition among downstream firms within each network. In his model, once one such network becomes large, the PCE drives other firms to other networks, limiting the scope of internalizing competition (i.e. the set of firms that join any one network), whereas I model it as limiting the degree of internalizing competition while (in equilibrium) all

---

[27] This work may be so influential partly because its elegance makes it a natural pick for a graduate school reading list. It's also an inspiring reading because (I think) Segal himself was a graduate student when it was written.

[28] In each case it is possible that the fierce competition for the pivotal sales never quite happens: see Segal and Whinston (2000) and Edlin (2002).

(both) Ds join the (only) network. Johnson also appears to build in a form of simultaneity that (I believe) removes the commitment issue from his model.

Several of these papers, notably Fumagalli and Motta and Simpson and Wickelgren, explore various positions on whether and how (in my language) M and D1 can "respond" if D2 fails to agree to M's offer. Those assumptions, which are inevitably somewhat ad hoc in models without explicit dynamics, can matter a lot for the model's results, and for example may largely account for the apparent contradictions between those two papers' conclusions.[29] The alternating-move dynamics in this paper are my attempt to go beyond competing assumptions in that particular respect; technically this approach is of course closely inspired by Maskin and Tirole, who used the framework to study horizontal duopoly; as far as I am aware the approach has not previously been used in studying vertical restraints.

## 7. Conclusion

In what I think are some natural models of limit pricing by M under the umbrella of tax-like vertical restraints—or, if the reader prefers, models of exclusive dealing with the wholesale price negotiated alongside the exclusivity—I evaluated the "commitment" and "positive contracting externality" limits on the creation of market power through vertical restraints. In the models, those limits keep firms away from the full monopoly outcome, but allow quite substantial taxes and correspondingly prices (wholesale and retail) well above competitive levels. In the models, the protagonist M pays downstream firms lump sums to accept [taxes on purchases from M's rivals that support] supracompetitive wholesale prices. The supply (offer) price from M's rivals does not budge—there is no exclusion in that sense—but the deals harm consumers relative to a world in which the restraints in question are prohibited. If, as in the models, only M has the ability to develop such a web of restraints, consumers are harmed relative to a world in which M does not exist.

It should go without saying that this doesn't claim to be an exhaustive analysis of vertical restraints. Subject as always to more analysis and evidence, however, it has convinced me that we should not take very much comfort in these two often cited limits on voluntary negotiation of collusive vertical restraints that reference rivals.

## 8. References

Abito, Jose and Julian Wright (2008), "Exclusive Dealing with Imperfect Downstream Competition," IJIO

Asker, John, and Heski Bar-Isaac (2014), "Raising Retailers' Profits: On Vertical Practices and the Exclusion of Rivals," AER.

Bernheim, Douglas and Randal Heeb (2015), "A Framework for the Economic Analysis of Exclusionary Conduct," in Blair and Sokol, Oxford Handbook of International Antitrust Economics

[29] See e.g. the discussion in Farrell (2005).

Bork, Robert (1978), The Antitrust Paradox

DeGraba, Patrick (2013), "Naked Exclusion by a Dominant Input Supplier: Exclusive Contracting and Loyalty Discounts," IJIO

Edlin, Aaron (2002), "Stopping Above-Cost Predatory Pricing," Yale Law Journal.

Edlin, Aaron and Joseph Farrell (2015), "Freedom to Trade," in Blair and Sokol, Oxford Handbook of International Antitrust Economics

Farrell, Joseph (2005), "Deconstructing Chicago on Exclusive Dealing," Antitrust Bulletin

Farrell, Joseph, Janis Pappalardo, and Howard Shelanski (2010), "Economics at the FTC"

Farrell, Joseph and Carl Shapiro (2008), "How Strong are Weak Patents?" AER

Fumagalli, Chiara and Massimo Motta (2006), "Exclusive Dealing and Entry, when Buyers Compete," AER

Inderst, Roman, and Greg Shaffer (2010), "Market-Share Contracts as Facilitating Practices," RAND Journal

Johnson, Justin (2014), "Anticompetitive Exclusive Dealing Without Exclusion," working paper.

Klein, Benjamin and Kevin Murphy (2008), "Exclusive Dealing Intensifies Competition for Distribution," Antitrust Law Journal.

Maskin, Eric and Jean Tirole (1982, 1985), "A Theory of Dynamic Oligopoly, I and II" Econometrica

Maskin, Eric and Jean Tirole (1987), "A Theory of Dynamic Oligopoly, III: Cournot Competition," EER

McAfee, Preston and Marius Schwartz (1994), "Opportunism in multilateral vertical contracting," AER

Rasmusen, Eric, Mark Ramseyer, and John Wiley (1991), "Naked Exclusion," AER

Scott Morton, Fiona (2013), "Contracts that Reference Rivals," Antitrust

Segal, Ilya (1999), "Contracting with Externalities," QJE

Segal, Ilya and Michael Whinston (2000), "Naked Exclusion: Comment," AER

Simpson, John and Abraham Wickelgren (2007), "Naked Exclusion, Efficient Breach, and Downstream Competition"AER

Tirole, Jean (2015), "Market Failures and Public Policy," Nobel Lecture, AER

Whinston, Michael (2006), Lectures on Antitrust Economics.

Zenger, Hans (2010), "When Does Exclusive Dealing Intensify Competition for Distribution?" Antitrust Law Journal.