Demand-Driven Integration and Divorcement Policy *

Patrick Legros†and Andrew F. Newman‡

May 14, 2016.

Abstract

Traditionally, vertical integration has concerned industrial economists only insofar as it affects market outcomes, particularly prices. This paper considers reverse causality, from prices – and more generally, from demand – to integration in a model of a dynamic oligopoly. If integration is costly but enhances productive efficiency, then a trend of rising prices and increasing integration could be due to growing demand, in which case a divorcement policy of forced divestiture may be counterproductive. Divorcement can only help consumers if it undermines collusion, but then there are dominating policies. We discuss well-known divorcement episodes in retail gasoline and British beer, as well as other evidence, in light of the model.

Keywords: theory of the firm, reverse causality, vertical integration, OIO, regulation, antitrust.

JEL codes: D23, D43, L2, L4, L5

---

*This paper has benefited from comments by seminar participants at Université libre de Bruxelles, CRESSE 2014, Tilburg university (TILEC) and Paris Dauphine. The authors are grateful to the Editor and the referees for their comments. Legros gratefully acknowledges support of the European Research Council (Advanced Grant 339950).

†Université libre de Bruxelles (ECARES), Northeastern University and CEPR.

‡Boston University and CEPR.
1 Introduction

A regulator observes that the firms in an industry he suspects of being imperfectly competitive have been vertically integrating over time. Armed with the traditional tools of industrial economics, he reckons that either integration is occurring because it enhances the productive or allocative efficiency of the firms in the industry, or because the firms are attempting to enhance their market power. Efficiency gains can arise because integration or vertical restraints help the vertical chain to internalize some externalities (e.g., double marginalization or free riding by distributors), in which case prices ought to fall. Market power enhancement could be due to foreclosure (increase rivals’ costs, refusal to supply) or to increased ability for vertical chains to collude; either way, prices should rise either at the wholesale or retail level.\(^1\) Hence, theory suggests that integration may lead to higher or lower prices depending on whether the dominant effect is foreclosure or efficiency. Telling the difference is straightforward: if prices are falling with integration, efficiency effects predominate. If they are rising, likely the firms are succeeding in enhancing their market power.

In the case of decreasing prices, the regulator, whose main constituency is consumers, has little reason to be concerned. In the other case, though, the regulator might be tempted to invoke a divorcement policy in order to limit the apparent effects of integration, either by intervening in the control structure of the production chain (for instance by ordering franchise gasoline retailers rather than their supplying refiners to make pricing decisions) or, more drastically, by ordering asset divestitures (as in the forced sale of pubs by the brewers that own and supply them).\(^2\) Being a practical person mainly interested in effective policy implementation, the regulator is not apt to ask the seemingly academic question of why integration has increased recently rather than some time in the distant past; the issue is how to act given the rise in prices. (In the case of falling prices, the regulator might take reasonable

---

\(^1\) See e.g., Lafontaine and Slade (2007); Rey and Tirole (1997); Riordan (2005). On balance, the empirical literature tends to provide support for the efficiency effect of vertical integration or vertical restraints (Cooper et al., 2005; Lafontaine and Slade, 2008).

\(^2\) The terms dissolution, divestiture and divorcement are often used interchangeably. The following excerpt from Oppenheim (1948) — cited in Adams (1951) — clarifies usage: “divestiture refers to situations where the defendants are required to divest themselves of property, securities or other assets. Divorcement is [...] used to indicate the effect of a decree where certain types of divestiture are ordered. It is especially applicable to cases where the purpose of the proceeding is to secure relief against anti-trust abuses flowing from [vertically] integrated ownership and control. The term ‘dissolution’ is generally used to refer to any situation where the dissolving of an allegedly illegal combination or association is involved, including the use of divestiture and divorcement as methods of achieving that end. While the foregoing definitions differentiate three aspects of remedies, the terms are frequently used interchangeably without any technical distinctions in meaning.” We thank Yossi Spiegel for suggesting this reference.
comfort in chalking it up to changes in the technology of production or distribution.)

But as is often the case, there are dangers in avoiding the academic questions. Indeed, in oft-studied cases in US retail gasoline and British beer, regulators imposed divorcement policies following long periods of increasing integration and rising prices. What ensued was a surprising continuation of rising prices instead of the expected fall. What is more, firms’ profitability fell, despite the price increases.

Standard industrial economic theories have a hard time explaining these episodes, but they make sense in the light of more recent developments in organizational economics, which has traditionally been concerned with the causes of integration more than its consequences (at least for the market). In a nutshell, the combination of rising prices, increasing integration, and reduced profits with continued rising prices post-divorcement can all be attributed to efficiency effects along with rising demand: in this view, causality runs from demand to integration, rather than from integration to market outcomes. The basis for this explanation is very simple. If integration indeed increases productive efficiency — a view that also has several, sometimes competing, sometimes complementary, foundations in organizational economics — then it follows from maximizing behavior that demand conditions must influence the integration decision. For if integration is costly, as it must be, else firms would always integrate to the maximum possible extent, then the productivity gain it offers is only worth the cost when the extra output produced is sufficiently valuable, namely with high demand. If demand is low, the cost of integrating outweighs the benefit, and the firm remains non-integrated.

The influence of demand on integration is at the heart of a recent paper (Legros and Newman, 2013), which considers the case of perfect competition, where the logic is most transparent. In this case, the role of demand is represented entirely by the price of the final product that a perfectly competitive supply chain faces. The gist of the argument can be made in the following reduced form model. Suppose that a chain’s technology is represented by the cost function

$$\phi(d)c(q) + h(d),$$

where $q$ is output and $c(q)$ is a standard cost function; we assume that there are eventually diminishing returns to scale so that this chain is not able to serve the entire market at constant marginal cost. The choice variable $d$ is the degree or depth of vertical integration, for instance, the number of units in the supply chain that belong to
a single firm with the rest remaining stand-alone firms. The function $\phi(d)$ represents how integration affects productive efficiency; $h(d)$ represents costs of integration. Examples of the former include improved coordination (Hart and Holmström 2010); better multitasking incentives (Holmström and Milgrom 1991); alignment of control and incentives (Grossman and Hart 1986; Hart and Moore 1990); or reductions in the costs of transactions, adaptation, or opportunism (Williamson 1971, 1975; Klein et al. 1978). In many cases, costs $h(d)$ can be generated by the same factors: incentives over multiple tasks are difficult to balance, and ceding control often means exchanging one incentive problem for another, resulting in decisions that are difficult for some parties to achieve given training, prior investments, or vision. Or they may result from maintaining a communication and monitoring infrastructure within the firm. Assuming that the chain chooses $d$ and $q$ to maximize its (joint) profits $Pq - \phi(d)c(q) - h(d)$ given market price $P$, a first observation is that $P$ affects the choice of integration level, just as it affects the choice of quantity produced, because it is a parameter of an optimization problem. To be more concrete, assume that $\phi(d)$ is decreasing and $h(d)$ increasing. Then profit is supermodular in the choices $(d, q)$ and has increasing differences in $(P, q)$. As a result, optimal $q$ and therefore optimal $d$ increase with $P$: when output is more valuable, and the chain therefore wants to produce more of it, it is worth investing more in the reduced costs of doing so.

Now consider the policy maker’s conundrum. If demand was increasing over time (and not compensated by entry), then price would be rising. This would induce firms to integrate more; their costs would be lower and profits (both net and gross of the integration cost) higher. Each firm would supply more (but not so much that the industry price would be reduced to its previous level, else firms would return to their previous integration and supply levels leading to excess demand). The new equilibrium price would be higher, but rather despite integration than because of it.

---

3This is a drastic simplification, since combinations of the members of the supply chain into several non-singleton firms, let alone recombinations across supply chains, are not allowed. But it is enough to make the point.

4Typically in organizational models, at least part of the costs or benefits of integrating are private, unobservable, and in any case non-contractible. Practically speaking this may mean that they will be difficult for the empirical investigator or policy maker to observe. In particular a distinction between gross (i.e. revenue minus costs of measured inputs) profitability and net profitability (gross profits minus integration costs) is worth bearing in mind.

5To be sure, in some models, particularly those in which incentives play a role, the extent of the efficiency gains, or the costs of integrating, may depend on other variables besides $d$, such as the price $P$ or the distribution of the profits among the various production units. For instance in Legros and Newman (2013), both the integration benefit and the integration cost display decreasing differences in $(d, P)$, but the net effect is that $d$ is always increasing in $P$. Other models of firms may have non-monotonic predictions; indeed, the differences across models could enable market data to serve as a proving ground for organization theory (Legros and Newman 2014).
For if the policy maker forced firms to reduce integration to some prior and lower level, their costs would rise, industry supply would be lower and the price even higher. This outcome is evocative of what happened in the gasoline and beer episodes that we document in section 4.

Of course, there are important differences, not least that neither of these industries appeared to be \textit{prima facie} competitive. Extending the perfectly competitive framework to an oligopolistic one is the task of this paper. We are not attempting any sort of generality here, only enough to highlight some of this issues. We consider a model of Cournot competition among supply chains that can choose the level of vertical integration, which reduces their marginal production cost. This is an appropriate setting to address another policy concern – expressed for example by policy makers in the British beer case — of whether and how integration facilitates collusion among these chains, as well how policies that regulate how integrated affect industry performance.

**Summary of Findings**  The first observation is that there is a conflict of interest between firms and consumers concerning the level of integration: as in other efficiency models, consumers would like there to be a high degree of integration, since that tends toward low costs and therefore low prices. But from the point of view of the firms in the market, there is too much integration: each firm confers a negative externality on its rivals when it integrates, because the cost reduction results in business stealing. In equilibrium, while measured profits may be substantial due to low marginal production costs, net profits that take account of the cost of integration (but at least in some of the interpretations alluded to above would be difficult to measure), will be low.

Second, as in the perfectly competitive case, demand plays a role in determining the integration decisions that firms make. Increasing demand always increases integration. But whether it is accompanied by rising or falling prices depends on which parameter of the (linear) demand is shifting. Consistent with the model, co-variation of integration and price is not always observed: indeed the beer and gasoline cases appear somewhat unusual in the trends that led to the policy responses.

Third, we can address the question of whether integration serves to facilitate collusion, which seems to have been a particular concern for regulators in the British beer case ([Spicer et al., 2012](#)). A first answer is not at all: as we have suggested, the industry not only has a collective motive to restrict output, but also to reduce integration levels. Indeed, if they are able to sustain collusion through repeated
interaction, then they will choose a *lower* integration level than they would in the non-collusive Cournot equilibrium. In this model, at least, high levels of integration serve as signs of low levels of collusion.\(^6\)

But there is a sense in which integration can support collusion. For the punishment inflicted on a deviator from a collusive strategy profile in which low levels of integration and output are being sustained is to revert to the higher integration and output levels of Cournot equilibrium. If integration were exogenous, or at least capped at a low level, this punishment would not be so severe, and collusion more difficult to sustain. By threatening the very low Cournot payoffs that integration affords, it is the possibility, rather than the reality, of more integration that helps sustain collusion.

Our fourth finding concerns the effects of divorcement policy, modeled as a binding cap on integration that is below the current level, and therefore requires divestment of assets. It follows from the over-investment result that if the industry is in the non-collusive equilibrium, then divorcement typically helps firms. Consumers are not helped by this, of course, because marginal costs and therefore prices rise.

If, however, the industry was colluding before the divorcement policy implementation, then two things can happen. Either collusion continues anyway, in which case consumers are harmed relative to the pre-divorcement outcome because marginal costs have increased, or it is undermined, because the Cournot payoff is now relatively high and does not constitute an adequate threat against deviation. This provides a potential benefit to consumers that would in general have to be weighed against the increase in marginal costs resulting from inefficient organization. However, as discussed in the final section of the paper, which considers the gasoline and beer divorcement episodes in more depth, it does not appear in either case that the trade-off between collusion and organizational inefficiency was managed to consumers’ benefit.

This is not to say that leaving industries unregulated is optimal. We discuss more general integration regulations that place caps on integration that need not force firms to divest. These may still destabilize collusion by increasing the Cournot payoff, but would have less detrimental effects on costs in case they don’t.\(^7\)

\(^6\)To be sure, a collusive industry that experiences rising demand would increase its level of integration, just as a monopolist would. But a non-collusive industry would do the same, and would always have higher levels of integration than the collusive one.

\(^7\)The possibility that regulation may enhance welfare by limiting collusion through restrictions that only bind off the equilibrium path has been observed in other contexts; see Chassang and Ortner (2015).
Links to the Literature  Our model is similar to models of investment in cost reduction, where the reduction in the marginal cost of production comes at the expense of a higher fixed cost (see Vives, 2008 for a recent survey), and in a static world, there would be indeed very little difference in interpreting \( d \) as an investment in process innovation or degree of integration. The difference of interpretation is important when we consider dynamics and collusion since we will assume that the cost reduction is temporary and therefore reversible: firms can choose the level of integration each period. This assumption of reversibility would be hard to rationalize in a model where the cost reduction is due to innovation. The assumption is natural, however, in an organizational context since firms can divest assets, or integrate more assets each period.\(^8\)

The choice of \( d \) affects the ability of firms to compete, since it modifies the marginal cost, but does not modify the demand for the product. If \( d \) could also modify the quality of the products, there would be a demand effect, as in Sutton (1991), with the important caveat that the cost \( h(d) \) is not sunk but fixed.

There is surprisingly little literature on tacit collusion when firms make investments that affect costs (or demand) before competition on the product market. An exception is Nocke (2007) which considers collusion in a sunk cost industry; investments permanently modify the demand for the products. This is not the case in our world since integration decisions are reversible. See also Schinkel and Spiegel (2016) in this volume.

Our results help us understand the role that organizational design can play in tacit collusion, and highlight that there may be costs and few benefits of forcing firms to divest even after observing covariation prices and integration. They should not, of course, be interpreted as implying that divestment and other forms of integration regulation are never desirable. For instance when vertical integration may lead to input foreclosure, divestitures may be pro-competitive (Sibley and Weisman, 1998; Vickers, 1985). But policy makers ought to be aware that upward trending prices and profits in the wake of integration need not be the result of foreclosure. As has been noted in other contexts, caution must be exercised in the design of policies that regulate vertical relationships (Cremer et al., 2007; Fiocco, 2011; Höfler and Kranz, 2011; Karl and Legros, 2015).

\(^8\)Clearly, transaction costs in the market for trading assets will limit reversibility; something we ignore here but would be relevant in a general analysis. The assumption we make is in line with organizational theories such as the “property rights” literature, which emphasize the private and incentive costs of relinquishing control rights under integration.
2 A Model of Integration

We consider an industry populated by vertical supply chains that are isolated from each other except at the final downstream stage, where they sell in a common market. Intermediate goods along the supply chain have no market. This portrait of the industry is mainly for simplicity, but it is also in the spirit of much of the organization literature that emphasizes “relationship specificity.” An example would be coal-fired electric generating station located next to a coal mine: coal is costly to transport and low in value, so the mine’s market is limited primarily to the power plant, but the electricity can be sold in a national market via the power grid. Our reduced form model is consistent with richer theories of integration, e.g., Legros and Newman (2013).

The timing of the model is as follows:

- There are $n$ downstream producers indexed by $i$ or $j$, and each makes a decision $d$ to integrate with suppliers (we do not consider horizontal integration).

- Integration decisions are observed and firms choose the quantity to produce.

- The product market clears, that is if $Q$ is the total quantity produced by the firms the price on the market is $P(Q) := a - Q$, the price equal to value of the inverse demand function at $Q$, where $a > 0$.

If the degree of integration in a firm is $d \in [0, \tilde{d}]$, and the quantity produced is $q$, the cost borne by the firm is $C(q; d)$, where to simplify the analysis we use the following specification:

$$C(q; d) := (c - d)q + d^2.$$ 

To insure an interior solution in the choice of integration in the Cournot game, we assume that $c$ is not too small with respect to $a$:

$$c < a < \frac{(n + 1)^2}{n} c. \quad (1)$$

The cost $d^2$ is best thought of as “fixed” and independent of output or price. Firms make a decision on the degree of integration $d$ and the higher $d$ is the larger are the fixed cost $d^2$ and the reduction of marginal cost $c - d$. The key feature of our specification for our results is that the cost function has negative cross partials in $d, q$, that is the marginal cost of production is a decreasing function of the degree of integration.
Cournot equilibrium

We consider subgame perfect equilibria in integration decisions \( \{d_i\} \), which are followed by integration-contingent quantity decisions \( \{q_i\} \). Output choices are contingent on the choices of integration by all firms in the industry because the marginal costs of firms are affected by these organizational choices.

For a given profile of integration decisions \( d := (d_1, \cdots, d_n) \), the continuation game is a standard Cournot game with marginal costs \( \{c - d_i, \ i = 1, \ldots, n\} \). Therefore the quantities, price, and profit levels are (sums are over \( j = 1, \ldots, n \) inclusive of \( i \) unless otherwise noted):

\[
q^*_i(d) = \frac{a - c - \sum_j d_j + (n + 1)d_i}{n + 1}, \quad Q^*(d) = \frac{n(a - c) + \sum_j d_j}{n + 1},
\]

\[
P(Q^*(d)) = \frac{a + nc - \sum_j d_j}{n + 1}.
\]

The equilibrium Cournot profit (gross profit less the fixed integration cost \( d^2 \)) is:

\[
\pi^*_i(d) = \frac{(a - c - \sum_j d_j + (n + 1)d_i)^2}{(n + 1)^2} - d^2_i. \tag{2}
\]

Note that the organizational choices are strategic substitutes in the first stage since \( \pi^*_i(d) \) has negative cross partials in \( (d_i, d_j) \). Hence if firm \( i \) expects other firms to integrate less, it will integrate more.

When firms choose their integration structure, they anticipate the equilibrium profit function (2). Because the marginal gross profit is \( \frac{2n}{(n+1)^2}(a - c - \sum_j d_j) + \frac{2nd_i}{n+1} \), as long as \( a - c \) is positive, some firms will choose positive integration in equilibrium because the marginal fixed cost is equal to zero when \( d = 0 \). Equilibrium is characterized by a non-singular linear system, for which the unique solution is symmetric:

\[
d^* = \frac{n(a - c)}{n^2 + n + 1}; \tag{3}
\]

assumption (1) ensures that \( d^* < c \).

**Lemma 1.** Under assumption (1), there exists a unique subgame perfect equilibrium. Each firm chooses a degree of integration

\[
d^* = \frac{n(a - c)}{n^2 + n + 1}.
\]
produces \( q^* = \frac{n+1}{n^2+n+1}(a - c) \) and the industry market prices is \( P^* = \frac{a+n(n+1)c}{n^2+n+1} \).

The integration decision by an individual firm imposes an externality on the others. To see this, consider \( \Pi^*(d) \), the per-firm profit when all firms use the same degree of integration \( d \) and then play Cournot:

\[
\Pi^*(d) = \frac{(a - c + d)^2}{(n + 1)^2} - d^2.
\]

A “Cournot planner” who chooses the (common) level of integration \( d^P \) for each chain, assuming they go on to play Cournot equilibrium in quantities, regards the marginal benefit of integration as \( \frac{2(a-c+d^P)}{(n+1)^2} \), and equates this to the marginal cost of integration \( d^P \):

\[
d^P = \frac{a - c}{n(n + 2)}.
\]

But from (2), each firm regards the marginal benefit as \( n \) times larger, because a cost reduction not only expands the market for the industry, but increases the market share of that firm. This business-stealing effect implies that Cournot competitors over-invest in integration (from their point of view, of course – consumers would not agree), which is an additional motive over and above the usual output restriction motive for collusion. The business-stealing effect is especially severe in our specification, as simple algebra shows that \( \Pi^*(0) > \Pi^*(d^*) \). Other properties of the function \( \Pi^*(d^*) \) are summarized here, as they will be useful later.

**Proposition 1.** Cournot-competing supply chains over-invest in integration in equilibrium: the net profit \( \Pi^*(d) \) they obtain when integrating at a common level \( d \) is a strictly concave function on \([0, d^*]\) with interior maximum at \( d^P < d^* \) and minimum at \( d^* \).

**Collusive Outcome**

We think of collusion as sustained in a repeated game with common discount factor \( \delta \), where firms choose integration as well as quantity decisions \((d_i, q_i)\) every period, since they are both reversible and costly for as long as the supply chain is operating. We will assume that collusion leads to the maximum per-firm profit absent side-payments. That is, we assume that firms collude on

\[
(d^M, q^M) := \arg \max_{d,q} (a - nq - c + d)q - d^2,
\]
The two first order conditions are $q = \frac{a-c+d}{2n}$ and $q = 2d$, implying that
\[
d^M = \frac{a-c}{4n-1}, \quad q^M = \frac{2(a-c)}{4n-1}, \quad P^M = \frac{2n-1}{4n-1}a + \frac{2n}{4n-1}c,
\]
and the maximum collusive profit is
\[
\Pi^M(d^M) := \frac{(a-c+d^M)^2}{4n} - (d^M)^2 = \frac{(a-c)^2}{4n-1}.
\]
Observe that $d^M > d^P$: the reason is that marginal returns to integrating under monopoly exceed those of the Cournot planner by a factor $\frac{(n+1)^2}{4n}$.

The repeated game in $(d^M, q^M)$ is somewhat non-standard, since the stage game is itself dynamic: firms choose integration first and then, upon observing the integration structure in the industry, choose their quantities. For this reason we are explicit in the way the trigger strategy is defined.

- At time 1,
  - Each firm $i$ chooses $d_i$.
  - If for each $i$, $d_i = d^M$, firms play $q_i = q^M$.
  - If for some $i$, $d_i \neq d^M$, firms play the Cournot equilibrium action $q_i(d)$.

- For time $t \geq 2$, define a history to be collusive if at each previous period, every firm chose $(d^M, q^M)$; otherwise the history is non-collusive.
  - If the history is collusive, play $d^M$. Once $d$ is observed play $q^M$ if $d_i = d^M$ for each $i$; otherwise play the Cournot quantity $q^*_i(d)$.
  - If the history is non-collusive, play $d_i = d^*$. For each observed $d$ play the Cournot quantity $q^*_i(d)$.

There are two incentive compatibility conditions: first, a firm must not want to change its organization (degree of integration) to $d_i \neq d^M$ and immediately face the Cournot profits $\pi^*_i(d^Me_{-i}, d_i)$ and, second, a firm must not want to change its output given $(d^Me, q^Me_{-i})$. Here $e$ is a vector of 1’s; the notation $xe$ ($xe_{-i}$) denotes that all firms (all firms but $i$) are playing the same action $x$.

Let $\Pi_{d}^{\text{dev}}(\hat{d}) := \max_d \pi_i(d e_{-i}, d) = \max_d \left\{ \frac{(a-c-(n-1)d+nd)^2}{(n+1)^2} - d^2 \right\}$ be the maximum profit within a period a firm could have by deviating to $d \neq \hat{d}$ and facing immediate Cournot competition when all other firms choose $\hat{d}$. From the remark following
Equation (2) about the strategic substitutability of the integration decisions, the optimal deviation if \( \hat{d} < d^* \) exceeds \( d^* \) and therefore \( \hat{d} \).

Let \( \Pi^\text{dev}_q(\hat{d}) := \max_q (a - c + \hat{d} - (n - 1)\hat{q} - q)q - \hat{d}^2 \) be the maximum profit within a period a firm can achieve by integrating to the level \( \hat{d} \) but subsequently deviating in quantity from the collusive quantity \( \hat{q} \) chosen by the other firms. (The notation reflects that the deviation profit only depends on \( \hat{d} \), because we always take \( \hat{q} \) to be the optimal monopoly quantity given the integration level \( \hat{d} \).)

When a firm deviates upward from \( \hat{d} = d^M \), it gains a competitive advantage due to a lower marginal cost, but this benefit is significantly reduced not only by the additional fixed cost, but also by an immediate change of conduct by the other firms: following this deviation in integration, firms will immediately shift to Cournot play, so that the cost reduction benefit is obtained only over the relatively small Cournot quantity. By contrast, by going along with the collusive integration decision \( d^M \), a firm can deviate in quantity, temporarily gaining a large share of the market, while facing only a delayed punishment by the other firms. Indeed, explicit computation of the values of the deviation profits defined above reveals that deviating in quantity while being obedient in integration brings a higher gain than deviating in integration:

\[
\Pi^\text{dev}_d(d^M) \leq \frac{9n^2}{(2n + 1)(4n - 1)^2} (a - c)^2
\]

\[
\Pi^\text{dev}_q(d^M) = \frac{n(n + 2)}{(4n - 1)^2} (a - c)^2.
\]

(In (6), the right hand side is computed for an interior solution, which is not guaranteed by (1); the inequality allows for the possibility that the optimal deviation is at a corner.) Because \( \Pi^\text{dev}_q(d^M) > \Pi^\text{dev}_d(d^M) \) and following either type of deviation firms play \((d^*, q^*)\) in all subsequent periods, the binding incentive constraint is the one for quantity deviation.

**Lemma 2.** The maximum profit a firm can obtain when deviating is \( \Pi^\text{dev}_q(d^M) \).

In equilibrium, firms produce \( q^M = \frac{a - c + d^M}{2n} \), and therefore the best deviation from this output is the one that achieves \( \Pi^\text{dev}_q(d^M) \). There is an incentive to deviate if \( \Pi^M(d^M) < (1 - \delta)\Pi^\text{dev}_q(d^M) + \delta \Pi^*(d^*) \), that is when:

\[
\delta < \delta^{\text{cutoff}}(d^M, d^*) := \frac{\Pi^\text{dev}_q(d^M) - \Pi^M(d^M)}{\Pi^\text{dev}_q(d^M) - \Pi^*(d^*)},
\]

where we make explicit the fact that the cutoff discount factor depends on the integration level chosen under collusion and under Cournot behavior. It is straightforward
to check that the critical discount factor $\delta^{no}(d^M, d)$ is decreasing in its second argument for $d > d^P$ (since $\Pi^*(d)$ is decreasing in that range). This property of the critical discount factor plays a key role in our analysis of policy in the next section.

3 Demand Changes and Divorcement Policy

Inspection of the outcomes under Cournot and collusion (Lemma 1 and expression (5)) shows that when firms do not expect to collude, they will integrate more: $d^* = \frac{n(a-c)}{n^2+n+1} > d^M = \frac{a-c}{4n-1}$ for $n > 1$. In addition to this level effect on integration, firms’ conduct modifies how sensitive integration is to demand shocks, and Cournot behavior leads to larger changes in integration than collusive behavior since $\frac{dd^*}{da} > \frac{dd^M}{da}$.

However, while increases in $a$ result in higher integration levels, they generate two opposite forces on prices. By themselves, these demand shifts would increase price. But there is a countervailing effect brought on by the induced reduction in marginal cost. Nevertheless, it is clear from the equilibrium values in Lemma 1 and expression (5) that the demand effect dominates, and the prices under Cournot competition and under collusion to monopoly are both increasing in $a$ — the variation is $\frac{1}{n^2+n+1}$ for Cournot and $\frac{2n-1}{4n-1}$ for collusion (thus, opposite integration, price is more responsive under collusion than under Cournot). Because integration also increases when $a$ increases, we have a co-variation between price level and integration. Notice that in the standard model with exogenous marginal cost, the variation in price due to $a$ would be larger than it is here: $\frac{1}{n+1}$ in the Cournot case, $\frac{1}{2}$ under collusion. Thus price increases in our model are not caused by integration at all, and indeed because of it are are smaller than they would otherwise be.

Proposition 2. (i) Rising demand in the form of increasing $a$ generates a higher degree of integration, both under Cournot and under collusion.

(ii) Price and integration co-vary in response to changes in $a$.

(iii) Integration is more responsive, and price less so, under Cournot than under collusion.

Note that in terms of elasticities rather than rates of change, price is more elastic under monopoly than Cournot, but integration is equally elastic in the two cases.

If the demand curve is $P = a - bQ$, the effects of changes in $a$ are as before. But increases in demand that take the form of reductions in the “market size” parameter
have somewhat different effects. On the one hand, integration increases, as long as interior solutions exist (this requires \( b > \frac{n}{(n+1)^2} \) in the case of Cournot and \( b > \frac{1}{4n} \) in the case of collusion). On the other hand, in neither case does price increase when \( b \) falls: if integration were held fixed, the fall in \( b \) would not lead to a price change (as in the standard model with linear demand and constant marginal costs), but since integration does increase, costs fall, and therefore both the monopoly and Cournot prices fall in equilibrium. Thus changing demand does not always generate co-variation in price and integration. Of course, neither is such co-variation universally observed. It is the possibility of such co-variation under demand-driven integration that we are pointing out here.

To simplify the remainder of the exposition, we assume that there is a one-time demand shift in the form of an increase in \( a \). We also assume there is no change in conduct following the demand increase. Let us assume that upon observing this joint increase in price and integration, the regulator puts more weight on the potential foreclosure effect of integration than on its efficiency benefits, and decides to regulate the industry by preventing any integration above some level \( d^r \) that we always take to be less than \( d^* \) (else the policy has no bite). In other words, the regulator believes that the causal relationship flows from integration to prices.

We will first consider how the regulation affects play assuming there is no change in conduct (i.e., collusion or Cournot). Using these observations, we derive how the critical discount factor \( \delta^{no}(\cdot, \cdot) \) is affected by policy, which allows us to determine whether there can be a change of conduct. Finally, we put these pieces together to assess the effects of divorcement and other policies that regulate the degree of integration.

Suppose first that firms play Cournot before and after the policy. Since \( d^r < d^* \), they are constrained to play \( d_i \leq d^r \), and will all choose \( d^r \) in equilibrium: since integration strategies are substitutes, when the other firms are constrained to choose \( d_j \leq d^r \), firm \( i \) will want to choose \( d_i > d^* \) but is also constrained and therefore chooses \( d^r \). The net Cournot profit is \( \Pi^*(d^r) \), which by Proposition 1 exceeds \( \Pi^*(d^*) \).

In effect, the regulation moves in the direction the Cournot planner would want by helping to limit the effects of the business-stealing externality.

---

9A full analysis of collusion under general shifts in demand, booms or busts, in the context of a two stage static game like ours is beyond the scope of this paper. See for instance Rotemberg and Saloner (1986) and Kandori (1991) for an analysis of repeated games with variable demand when there is a one-dimensional strategic variable.

10One can check that, as each of the three net profit expressions in equation (8) are expressible as functions of \( n \) times \((a - c)^2\), \( \delta^{no}(d^M, d^r) \) is independent of \( a \). Thus, increasing demand has no impact on the feasibility of collusion.
If firms are colluding before and after regulation, the policy has differing effects depending on whether $d^r$ is smaller than $d^M$. If $d^r \geq d^M$, the policy does not bind on what the firms do in equilibrium, but does bind on what they can achieve off the equilibrium path: the permanent Cournot punishment for quantity deviation now yields a payoff of $\Pi^*(d^r) > \Pi^*(d^r)$. However, colluding firms can still try to play $d^M$ in equilibrium, and the punishment within the period of playing Cournot at costs $c - d^M$ is unchanged, while the gain, which is now constrained by $d \leq d^r$, cannot be larger than before the regulation. On the other hand, the one-shot gain from deviating in quantity while integrating at $d^M$ is unaffected by the policy: Lemma 2 still holds in this case, and as before, the no-quantity deviation constraint is the one that binds.

If the policy is more severe, with $d^r < d^M$, then firms are constrained in the levels of integration they can maintain on path as well as off. Since the integration profit is concave in $d$, the cartel will wish to collude on $d^r$. The maximum profit level under collusion and the maximal quantity deviation profit are:

$$\Pi^M(d^r) = \frac{(a - c + d^r)^2}{4n} - (d^r)^2; \quad \Pi^{dev}_q(d^r) = \frac{(n + 1)^2}{16n^2}(a - c + d^r)^2 - (d^r)^2. \quad (9)$$

As in the unregulated case, there are two incentive compatibility conditions for stability, and again the condition with respect to deviations from $d^r$ can be sustained within a period by the threat of immediate reversion to Cournot play. As before, then, it is the quantity deviation incentive compatibility condition that binds.

To summarize:

**Lemma 3.** Suppose the regulator imposes $d^r < d^*$. If firms collude before and after the policy change, they play $d^M$ if $d^r \geq d^M$, and $d^r$ if $d^r < d^M$. The permanent Cournot punishment following any deviation leads to a per-period payoff of $\Pi^*(d^r)$.

To determine the effects of the policy on the ability of firms to collude, we need only check how it affects the critical discount factor $\delta^{no}(\cdot, \cdot)$. When $d^r \geq d^M$, the sustained profit is $\Pi^M(d^M)$ and deviation profit is $\Pi^{dev}_q(d^M)$. The Cournot profit is

\[\text{To see this, let } d^*(d^r) := \arg\max_d \pi_i(d^r e, d) \text{ denote the unconstrained optimal deviation for } i \text{ when the other firms play } d^r. \text{ Since } d^r < d^M, \text{ strategic substitutability of the objective implies } d^*(d^r) > d^*(d^M) > d^M, \text{ where the second inequality was shown earlier in the discussion leading to Lemma 2. Concavity of } \pi_i(d^r e, d) \text{ then implies that the objective is increasing on } [0, d^r(d^r)], \text{ so the solution to the constrained problem with } d \leq d^r \text{ is } d^r. \text{ Thus the firm can do no better by deviating than the Cournot planner’s payoff at } d^r, \text{ which is less then the monopoly payoff at } d^r \text{ that it can obtain by sticking to the collusive integration level.}
Π*(d′). Thus collusion is sustainable when
\[ \delta \geq \delta^{no}(d^M, d^r) = \frac{\Pi_q^{dev}(d^M) - \Pi^M(d^M)}{\Pi_q^{dev}(d^M) - \Pi^*(d^r)}, \]
which is a decreasing function of d′, since by Proposition 1, Π*(d) is decreasing on (dP, d′), which contains (dM, d∗). When d′ ≤ dM, collusive firms integrate to the level d′ and the critical discount factor to sustain collusion is
\[ \delta^{no}(d^r, d^r) = \frac{\Pi_q^{dev}(d^r) - \Pi^M(d^r)}{\Pi_q^{dev}(d^r) - \Pi^*(d^r)}. \]

From (9) and (4), this ratio is a constant, equal to δno(dM, dM) for any d′ ≤ dM; this is because firms always bear the same fixed cost d′2 whether they collude, deviate or play Cournot and therefore the numerator and the denominator are proportional to (a − c + d′)2. Thus the minimum discount factor that sustains collusion is a continuous, non-increasing function of d′, constant on [0, dM], and decreasing on (dM, d∗).

We are now in position to analyze the effects of policy. There are three cases, depending on the value of the discount factor δ. For simplicity, we assume that firms collude as long as their discount factor exceeds the critical level.

**Low discount factors**, \( \delta < \delta^{no}(d^M, d^r) \). In this case, firms are unable to sustain collusion before the policy, and since imposing the policy raises the critical discount factor, they will not collude after. The analysis of Cournot play applies, and the result is
- integration falls from d∗ to d′ and marginal costs rise;
- product price rises – consumers lose;
- gross profits (revenue less production costs) fall;
- net profit (gross profit less integration costs) increase – firms are better off.

**High discount factors**, \( \delta \geq \delta^{no}(d^M, d^M) \). Here firms are able to sustain collusion both before and after the policy, regardless of how severe it is (how low d′) and thus the policy itself would have no impact on conduct. A policy d′ ≥ dM would have no impact on market outcomes, but neither is it meaningfully a divorcement policy, since it does not force firms to divest any of the assets they own. A true divorcement policy has d′ < dM, in which case
• integration falls from $d^M$ to $d^r$ and marginal costs rise;
• product price rises – consumers lose;
• gross profits (revenue less production costs) fall;
• net profits (gross profit less integration costs) fall – firms are worse off.

**Moderate discount factors**, $\delta \in (\delta^{no}(d^M, d^r), \delta^{no}(d^M, d^M))$. The principle difference between this case and the two others is the possibility that the regulations engenders a change of conduct, by raising the critical discount factor above the existing one. Let $d^r(\delta) \in (d^M, d^r)$, solve $\delta^{no}(d^M, d^r) = \delta$. Given $\delta$, this is the minimum level of integration at which collusion is sustainable. If $d^r \geq d^r(\delta)$, then collusion is possible after as well as before the policy change, and the policy has no impact (again, since the integration ceiling is above the level that was being sustained in equilibrium, this policy is not really divorcement). But if $d^r < d^r(\delta)$, collusion is not possible under the policy. Integration now goes from $d^M$ to $d^r$, and the firms will deliver the Cournot quantity given $d^r$. Hence, if $d^r \in (d^M, d^r(\delta))$ the result is

• integration rises from $d^M$ to $d^r$ and marginal costs fall;
• product price falls – consumers gain
• gross profits (revenue less production costs) fall;
• net profits (gross profit less integration costs) fall – firms are worse off.

Notice that the change of conduct in this regime is generated by the effect of $d^r$ on off-path play: there is no constraint on on-path integration levels as long as $d^r \geq d^M$. Compliance with the policy then does not actually require firms to sell off assets.

A divorcement policy in this case would really amount to setting $d^r < d^M$, which is inefficient, because it does no more to change firms’ conduct than setting $d^r$ just under $d^r(\delta)$, but does force firms to engage in Cournot competition at higher-than-necessary costs. To be sure, it still benefits consumers relative to no regulation, because destabilizing collusion always lowers the market price in this model. Even if the regulator imposes the strongest divorcement policy by preventing any integration ($d^r = 0$), the monopoly price at $d^M$ exceeds Cournot price at $d^r = 0$. But he would do better to cap integration at above $d^M$ and below $d^r(\delta)$: it also destroys collusion but allows a greater degree of cost-reduction – and therefore a lower Cournot price.

Combining the three cases, we conclude:
Proposition 3. (i) Following divestiture, the price consumers pay decreases if, and only if, the firms’ conduct changes from collusion to Cournot competition.

(ii) Firms benefit from divorcement if and only if they are not colluding.

(iii) If firms collude on $d^M$ and the conduct of firms can change with an integration ceiling, a policy that dominates divorcement limits integration to just under $d^r(\delta) \in (d^M, d^*)$.

The policy described in Proposition 3(iii) is near-optimal in $d^r$ (the change of conduct at $d^r(\delta)$ introduces a discontinuity, so an optimum does not exist). It would result in a shift in conduct from collusion to Cournot and consequently an increase in the level of integration from $d^M$ to $d^r$. We are not aware of any policy episodes resembling this possibility; note it requires information on the discount factor $\delta$, which is unlikely to be available. There is, however, a policy that economizes on information for the regulator and will dominate divorcement. This soft policy simply caps integration at the existing level rather than forcing divestment to some level below it: it will break collusion whenever a stricter divorcement policy would, is neutral with respect to prices and integration whenever it does not, but allows larger cost reductions in case firms do switch to Cournot.

One thing we have not considered so far is the effects of divorcement on exit, and while a full treatment is beyond the scope of this paper, we make some observations. If firms incur an additional fixed cost to operate, then a fall in net profits following divorcement may lead some to exit. There is evidence this happened in the British beer case. Exit may not be harmful of course if firms were making monopoly profits prior to the regulation. But if they were not colluding, the regulation will force costs to be higher, and after exit, the smaller number of firms combined with higher costs will drive prices up even further. What is more, with fewer firms left, collusion may be more sustainable than before, raising the possibility that divorcement might facilitate collusion.

4 Some Evidence for Demand-Driven Integration

Evidence for the effects of demand on integration is starting to be collected. Some suggestive evidence comes from some single-industry studies (e.g. Forbes and Leder-
man (2010), which shows that airlines are more apt to integrate with regional carriers on more “valuable” routes (specifically those where failures are more costly); Forbes and Lederman (2009) also show that integrated relationships are more productive, which bolsters the key assumption in this paper.

A few papers try systematically to test for demand effects on integration by focusing on ostensibly competitive industries wherein the influence of demand would manifest itself through the price level. An empirical challenge is to find exogenous sources of price variation and look for correlation with integration. One approach is provided in Alfaro et al. (2016), which uses variation in the Most-Favored-Nation (MFN) tariffs applied by GATT/WTO members as a proxy for price variation. The idea is that tariffs affect prices, and through that integration, but (vertical) integration is unlikely to affect tariffs. The argument for exogeneity of MFN tariffs comes partly from the institutional structure by which they are set: long rounds of multilateral bargaining and a non-discrimination principle that forces uniform application of tariffs to all trading partners makes the MFN tariffs much more resistant to lobbying than other forms of trade barriers. Because tariffs raise product prices in the domestic market (as compared to the world price), reverse causality suggests that they should lead to more integration among firms selling in that market.

Alfaro et al. (2016) defines the degree of integration to be the fraction of inputs (in value-added terms) that are produced within the firm (this measure is due to Fan and Lang (2000)); in the data, the average is about 6%, though there is considerable variation across as well as within (4-digit SIC) industries. Tariffs should have stronger effects on firms that do not sell abroad, since exporters face the world price, not the just the domestic one. Focusing on this differential effect of tariffs on domestic firms and exporters, and using country-sector fixed effects to control for possible omitted variables that might be driving integration and tariffs, the paper reports strong effect of tariffs on the degree of integration. The estimated tariff elasticity of vertical integration is in the range 0.02 – 0.09, which, since tariffs average around 5%, translates into a price elasticity in the range 0.4 – 2.

An example of a single-industry study that tries to identify price effects on vertical integration is McGowan (2015), which looks at changes in the vertical structure of U.S. coal processing plants and the mines that supply them. The Staggers Railroad Act of 1980 deregulated railroad pricing; greater competition among railroads led to falling shipping costs, which enabled electric power plants, particularly in the East, to profitably source coal from anywhere in the country. Cheap coal from the Powder River Basin in Montana and Wyoming was now able to compete with Eastern coal
among Eastern power plants, reducing coal prices there. However, in the West, there was little change in shipping costs, mainly due to the relative sparseness of rail networks. As our model hypothesizes, the data reveal a positive correlation between processing plant productivity and vertical integration. And consistent with its predictions, by the mid 1980s, vertical integration (the fraction of mines owned by processing plants) in the East had fallen 28% relative to the West.

In both the British beer and US gasoline divestment episodes, the regulators forced divestitures in vertical chains, admittedly because there was an increasing trend in prices and the fear was that this was due to foreclosure effects facilitated by vertical integration.

The econometric challenge to identify the effect of divestment on prices is to control for the possibility that other factors, such as changes in accounting conventions or tax rules for the beer industry, contributed to the increase in prices. If the foreclosure story was the right one, divestment should have led to a decrease in prices. If divestment led to an increase in prices, there is support for an efficiency view of integration.

Slade (1998a) documents the effects of the Beer Orders of the 1980s, a UK Monopolies and Mergers Commission decision to force divestiture of 14,000 public houses. She contrasts four types of organization in the vertical chain brewer-pub: company owned, franchised pubs with or without fixed fees, and arms-length relationship. Company owned is akin to integration, while the two other forms are weak and strong forms of non-integration.

The upstream segment of the industry was relatively concentrated at the time of the decision, with seven large national brewers and a larger number of microbreweries. The divestment effectively put ceilings on the number of pubs (licenses) that a brewer could have. The divestitures led to the emergence of public-house chains, which have long-term contractual agreements with national brewers, as well as decisions by some brewers to stop production and shift to retailing. The main finding was that following the decision, retail prices increased (in the houses closely tied to the brewers, but not in the free houses), and profits of brewers decreased. Consistent with our results, following the policy, some brewers exited the industry.

Interestingly, some commentators on the beer case (e.g. Spicer et al., 2012) note that demand for pub beer was growing over the decade leading to Beer Orders, due to income growth and greater leisure time. They also document the strong industry opposition to the policy. In terms of our model, the evidence fits the scenario in which collusion was sustained both before and after the intervention: this accounts
for the increased price, the fallen profits, and the industry opposition. If collusion had been stopped by the policy, most likely prices would have fallen, at least for a while, as suggested in the moderate-discount-factor case in Section 3. If there had been no collusion, brewers should have welcomed some version of the policy as a check on their over-investment.

Barron and Umbeck (1984) studies the effects of the divorcement law enacted in 1974 in Maryland that prohibited refiners’ control of gasoline stations, and re-allocated control rights for hours of operation and retail pricing to the stations. Contrary to the beer example, here the refiners were not obliged to divest their assets but to move to a franchising system where the gas station franchisees would have control over operation decisions, including the retail price. As in the previous example, the effect of the divorcement has been an increase in retail prices. Barron and Umbeck (1984) cite evidence that the supporters of the legislation included owners of independent gasoline stations, who are indeed likely to gain from the divorcement since price competition will be soften at the retail level, while opponents to the legislation included, obviously, refiners affected by the divorcement policy but also consumers.

These results are consistent with an efficiency view of integration, a point that has been made in many other empirical studies of vertical relationships (Lafontaine and Slade, 2007), but are especially pertinent for our discussion. A regulator stepped in and forced divorcement following an upward trend in both retail prices and vertical integration, yet prices continued to rise. This is consistent with the view that the pre-divorcement upward trend in prices was due primarily not to integration but to changes in demand that were driving both integration and price.

5 Conclusion

One of the challenging tasks in evaluating mergers or the performance of integrated firms is to disentangle the efficiency and market power effects of integration. Our understanding of the causal relationship between price levels and degrees of integration not only guides econometric efforts to separate these effects but also influences policy decisions. In the divorcement episodes discussed here, making allowance in the policy discussion for the possibility of demand-driven integration may have led to more satisfactory outcomes. In terms of our model, a simple cap on integration at then-current levels would have had the same potential benefit as divorcement.

13See also Barron et al. (1985); Slade (1998b); Vita (2000) and Blass and Carlton (2001).
terms of destabilizing collusion, but without the damage to firms’ cost structures that ultimately kept prices high.

References


