Vertical Integration and Market Foreclosure

Few people would disagree that horizontal mergers have the potential to restrict output and raise consumer prices, but there is much less agreement about the anticompetitive effects of vertical mergers. Some commentators have argued that a purely vertical merger will not affect a firm's monopoly power, because the merger of an upstream and a downstream firm, each of which controls, say, 10 percent of its market, does not change market shares: other firms continue to possess 90 percent of each market. Other commentators have responded by developing models in which vertical integration can lead to the foreclosure of competition in upstream or downstream markets. These models, however, rely on particular assumptions about contractual arrangements between nonintegrated firms (for example, that pricing must be linear) or about the ability of integrated firms to make commitments (for example, that an integrated supplier can commit not to undercut a rival). Thus at this stage the debate about the conditions under which vertical mergers are anticompetitive is far from settled.

The purpose of this paper is to develop a theoretical model showing how vertical integration changes the nature of competition in upstream and downstream markets and identifying conditions under which market...
foreclosure will be a consequence or a purpose, or both, of such integration. In contrast to much of the literature, the paper does not restrict upstream and downstream firms to particular contractual arrangements, but instead allows them to choose from a full set of arrangements both when they are integrated and when they are not (so, for example, two-part tariffs are permitted). We also allow nonintegrated firms to respond optimally to the integration decisions of other firms, either by remaining nonintegrated, exiting the industry, or themselves integrating. We use the analysis to shed some light on prominent vertical mergers involving the cement industry, computer reservation systems for airlines, and the St. Louis Terminal Railroad.

The paper follows the recent literature on ownership and residual control rights in the way vertical integration is formalized. We assume that the upstream and downstream firms do not know ex ante which type of intermediate good will be the appropriate one to trade and that the large number of potential types makes it too costly to write contingent forward contracts. As a result, the only way to influence ex post behavior is through the allocation of residual rights of control over assets. Moreover, we take the point of view that the shift in residual control rights that occurs under integration permits profit sharing between upstream and downstream units. As a consequence, all conflicts of interest about prices and trading policies are removed. In this respect, vertical integration does not differ formally from a profit-sharing scheme between independent contractors. Profit sharing may be difficult to implement in the absence of integration, however, because independent units can divert money and misrepresent profits. In contrast, the owner of a subordinate unit, because he or she has residual rights of control over the unit’s assets, may be able to prevent diversion and enforce profit sharing.

2. This means that the elimination of the double marginalization of prices is not a motive for integration in our model. For a discussion of this issue, see Tirole (1988).
4. On this, see Williamson (1985) and, for formal models, Hart (1988), Holmstrom and Tirole (1989), and Riordan (1989). As an extreme example, consider an independent unit, A, that has signed a profit-sharing agreement with firm B. One way A can misrepresent and divert its profits is by purchasing an input at an inflated price from another company in which A’s owners have an interest. It may be hard for B to write an enforceable contract ex ante to prevent such a diversion, even though B may be well aware of the practice ex
Although integration removes conflicts of interest about pricing and trading policies, it is accompanied by costs. First, after integration, a subordinate manager may have lower incentives to come up with good ideas to reduce production costs or to raise quality because this investment is expropriated by the owner of the firm.\footnote{See Grossman and Hart (1986) or Hart and Moore (1988). We assume that effort costs cannot be reimbursed as part of a profit-sharing scheme.} Second, there may be a loss in information about the subordinate’s performance, and therefore less incentive to make improvements, because vertical integration reduces or eliminates the fluidity of the market for the stock of the newly subordinate unit.\footnote{See Holmström and Tirole (1990).} Third, there may be legal costs associated with the merger. We do not explicitly formalize these costs of integration, although it is easy to do so. Instead, it will be enough to represent them by a fixed amount $E$.

\textit{Description of the Model}

The basic model consists of two potential suppliers or upstream firms, $U_1$ and $U_2$, and two potential buyers or downstream firms, $D_1$ and $D_2$.\footnote{The model could easily be generalized to the case of more than two upstream or downstream firms, however.} The downstream firms compete on the product market and sell perfect

post (the information that the input is overpriced is observable but not verifiable). On the other hand, if $A$ and $B$ are integrated, $B$ can refuse ex post $A$’s manager’s request to spend company resources on the expensive input, thus effectively blocking the transaction. This is because $B$ now possesses residual rights of control over company $A$’s resources by virtue of integration.

Of course, diversion problems are not completely eliminated by integration. In particular, if $B$ owns $A$, $B$ can use its residual control rights to divert money from $A$. However, as long as $B$ diverts on a proportionate basis from both units $A$ and $B$—and as long as this diversion is less than 100 percent—$A$’s subordinate manager can be given a compensation package that is some fraction of $A$’s and $B$’s joint profit. Given this, $A$’s subordinate manager will have an incentive to choose pricing and trading policies that are in the interest of the company as a whole.

Another argument can be given as to why a merger reduces conflicts of interest over prices and trading policies. Under integration, a subordinate manager will act in the interest of the parent company, since otherwise he or she will be dismissed. But the pressure on the manager of an independent unit to act in the interest of another independent contractor is less because the only sanction available to the independent contractor is to sever the whole relationship with the unit (the contractor cannot fire the unit’s manager alone). On this, see Hart and Moore (1988).
substitutes. The upstream firms produce the same intermediate good at constant, although perhaps different, marginal costs, \( c_1 \) and \( c_2 \), subject possibly to a capacity constraint.

Three variants of the basic model are developed, each of which illustrates a different motive for integration. Variant 1, called \textit{ex post monopolization}, focuses on the incentive for a relatively efficient upstream firm to merge with a downstream firm to restrict output in the downstream market. To understand the idea, consider the special case in which one of the upstream firms, \( U_2 \), say, has infinite marginal cost. It is sometimes claimed that in this case \( U_1 \) would never have an incentive to merge with a downstream firm, \( D_1 \) say, because \( U_1 \) is already a monopolist in the upstream market.\(^8\) This claim is false unless enforceable exclusive-dealing contracts are feasible, or unless the offers of \( U_1 \) to \( D_1 \) and \( D_2 \) are public. In particular, in the absence of exclusive-dealing contracts, \( U_1 \) has an incentive to supply both \( D_1 \) and \( D_2 \) and, in so doing, to produce more than the monopoly output level. For example, suppose \( U_1 \) tries to monopolize the downstream market by selling the monopoly output \( (q^m) \) to \( D_1 \) for a lump-sum fee equal to monopoly profit. It is not an equilibrium for \( D_1 \) to accept such an offer because the firm knows that \( U_1 \) has an incentive to sell an additional amount to \( D_2 \), thus causing \( D_1 \) to make a loss. On the other hand, suppose \( U_1 \) tries to monopolize the downstream market by offering \( \frac{1}{2} q^m \) to each of \( D_1 \) and \( D_2 \) at a fee equal to half the monopoly profit. It is not an equilibrium for \( U_1 \) to make and \( D_1 \) and \( D_2 \) to accept these offers either, because if \( D_1 \), say, is expected to accept, \( U_1 \) has an incentive to increase its supply to \( D_2 \) above \( \frac{1}{2} q^m \), and \( D_1 \) again makes a loss.

Integration can be a way around the inability of \( U_1 \) to restrict output. If \( U_1 \) and \( D_1 \) merge, \( U_1 \) has no incentive to supply \( D_2 \). The reason is that under integration the profits of \( U_1 \) and \( D_1 \) are shared, and every unit sold to \( D_2 \) reduces the combined profit of \( U_1-D_1 \) by depressing price. Thus the unique equilibrium now is for \( U_1 \) to supply \( q^m \) to \( D_1 \) and nothing to \( D_2 \).

Why could \( U_1 \) not achieve the same outcome by writing an exclusive-dealing contract with \( D_1 \)? There are several answers to this. First,\(^8\) For example, as Posner and Easterbrook (1981, p. 870) have written. "there is only one monopoly profit to be made in a chain of production."

\(^8\)
exclusive dealing may be unenforceable for informational reasons. In particular, it may be difficult for $D_1$ to monitor or control shipments by $U_1$ to other parties without having residual rights of control over the assets of $U_1$, including buildings, trucks, and inventories. And even if shipments can be monitored, if there are third parties outside the industry with whom $U_1$ can realize gains from trade and who could bootleg its product to $D_2$, a strict enforcement of exclusive dealing requires not trading with these third parties, which may prove costly. Second, exclusive dealing may be unenforceable for legal reasons: the courts have taken a harsh stance on those exclusive-dealing contracts they think may result in foreclosure.

In addition, exclusive dealing, even if it is feasible, is not generally a perfect substitute for integration. In particular, if supply costs of $U_2$ are finite rather than infinite, then it is no longer optimal for an integrated $U_1-D_1$ pair not to supply $D_2$ at all. Instead $U_1-D_1$ will want to offer $D_2$ the same amount that $U_2$ would offer $D_2$, but at a slightly lower price. An exclusive-dealing contract will not achieve this. Moreover, a contract that limits the amount that $U_1$ can sell $D_2$ may be very difficult to enforce: given that $U_1$ is supplying $D_2$ anyway, it may be hard for $D_1$ to verify that supplies equal 100, say, rather than 200. Integration avoids this problem: profit sharing between $U_1$ and $D_1$ means that $U_1$ automatically finds it in its interest to supply the profit-maximizing level and quality of service to $D_2$.

In extensions of this first variant, we consider the possibility that it may not be known in advance whether $U_1$ or $U_2$ is the more efficient supplier and that the upstream and downstream firms must make ex ante industry-specific investments before trading ex post. We show that the more efficient (in a stochastic sense) upstream firm will have a greater incentive to merge to monopolize the market ex post. Also, if $U_1$ and $D_1$ merge, the profits of $D_2$ will typically fall, because if $U_1$ turns out to be the more efficient firm ex post, it will channel supplies toward $D_1$ at the expense of $D_2$. This fall in the profits of $D_2$ may cause it to stop investing or to exit the industry. To the extent that exit by $D_2$ reduces the profits of $U_2$ by lowering the total demand for its product,

9. An analysis of exclusive-dealing contracts is contained in appendix C.
10. The enforcement problem becomes even greater if $U_1$ wants to commit itself not to supply $D_2$ with quality of service above that provided by $U_2$. 
$U_2$ may have an incentive to rescue $D_2$ by merging with it and paying part of its investment cost (via profit sharing). In other words, bandwagoning may occur.

This first variant assumes that upstream firms engage in Bertrand competition in the price and quantity offers they make to downstream firms. The second variant, called \textit{scarce needs}, supposes instead that upstream and downstream firms bargain over the gains from trade in such a way that each upstream firm obtains on average a positive share of these gains. In addition we now assume for simplicity that $c_1 = c_2$: the upstream firms are equally efficient.

Under these conditions, there is a new motive for integration. An upstream firm may merge with a downstream firm to ensure that the downstream firm purchases its supplies from this upstream firm rather than from others. In particular, if $U_1$ and $D_1$ merge, then, rather than sometimes buying input from $U_1$ and sometimes from $U_2$ as under nonintegration, $D_1$ will now buy all its input all the time from $U_1$. Thus $U_1$ gains a valuable trading opportunity and $U_2$ loses one. (Scarce needs refers to the fact that $D_1$ and $D_2$ have limited input requirements.)

If $U_2$ remains in the industry (continues to invest), the only effect of the merger is to increase the $U_1-D_1$ share of industry profit and reduce the $U_2-D_2$ share. In particular, there is no ex post monopolization effect in this second variant: given that $U_2$ is as efficient as $U_1$, there is no reason for $U_1$ to restrict its supplies to $D_2$, because $U_2$ will make up the difference anyway. However, if the reduction in the profits of $U_2$ causes it to quit the industry, $U_1$ is left as the only supplier (we refer to this as ex ante monopolization) and, given that it is merged with $D_1$, it will be able to use this power to completely monopolize the market ex post (as part of a merged firm, it has no incentive to supply $D_2$). Thus total quantity supplied will fall and the price consumers pay will rise.

Bandwagoning does not occur in equilibrium in this second variant. However, $U_2-D_2$ may try to preempt $U_1-D_1$ by merging first. In real time, the upstream firm with lower investment costs will win this preemption game by merging early.

The third variant, \textit{scarce supplies}, reverses the role of upstream and downstream firms. Suppose that the upstream firms are capacity-constrained relative to downstream firms’ needs, with upstream and downstream firms again bargaining over the terms of trade. Under these
conditions, a third incentive to integrate arises: a downstream firm and an upstream firm may merge to ensure that the upstream firm channels its scarce supplies to its downstream partner rather than to other down-
stream firms.

If $U_1$ and $D_1$ merge, $D_2$ suffers because under nonintegration $D_2$ obtains some profit from being able to purchase supplies from $U_1$, whereas under integration $U_1$ channels all its supplies to $D_1$. The decline in its profits may cause $D_2$ to quit the industry. In this case $U_2$’s profits will also fall because it faces only one purchaser for its output: $D_1$. Hence $U_2$ may cease to invest. If this happens, capacity will be eliminated from the market, consumer price will rise, and the effect of the $U_1$-$D_1$ merger will have been to monopolize the market ex ante.

To avoid exit by $D_2$, firm $U_2$ may merge with it. Thus, as in the first variant, bandwagoning is a possible outcome. Also $U_2$ and $D_2$ may try to preempt a $U_1$-$D_1$ merger by merging first. The preemption game will lead to premature merger by $U_1$-$D_1$ or $U_2$-$D_2$.

Table 1 summarizes the three variants.

Welfare Analysis of Vertical Mergers

Our theory has a number of implications for the welfare analysis of vertical mergers. The model shows three sources of social loss from mergers and two sources of social gain. First, in variant 1 a merger of $U_1$ and $D_1$ raises consumer prices to the extent that it allows $U_1$-$D_1$ to monopolize the market ex post. This reduces the sum of consumer and producer surplus for the usual reasons. Second, in all three variants of the model, a merger of $U_1$ and $D_1$ may cause exit by $U_2$ or $D_2$ or both. This ex ante monopolization effect again gives $U_1$-$D_1$ greater market power ex post, causing consumer prices to rise and consumer plus producer surplus to fall. Third, mergers involve incentive and legal costs, which we have represented by a fixed amount $E$.

Offsetting these losses are two potential gains from mergers. First, a merger of $U_1$ and $D_1$ that causes exit by $U_2$ or $D_2$ or both leads to a saving in investment costs. To the extent that these costs were incurred by $U_2$ and $D_2$ to increase their aggregate profit at the expense of $U_1$-$D_1$, with no price effects, this represents a social gain. In other words, a merger-induced exit can be beneficial to the extent that it leads to a reduction in rent-seeking behavior.
Table 1. Summary of Three Variants

<table>
<thead>
<tr>
<th>Item</th>
<th>Variant 1: ex post monopolization</th>
<th>Variant 2: scarce needs</th>
<th>Variant 3: scarce supplies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output contraction</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bargaining effect</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Possible circumstances</td>
<td>No capacity constraints upstream</td>
<td>Downturn in D industry, or excess capacity in U industry</td>
<td>Downturn in U industry, or excess capacity in D industry</td>
</tr>
<tr>
<td></td>
<td>and downstream</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct victim of vertical integration</td>
<td>Nonintegrated D</td>
<td>Nonintegrated U</td>
<td>Nonintegrated D</td>
</tr>
<tr>
<td>Indirect victim (if direct victim exits)</td>
<td>Nonintegrated U (under certain conditions)</td>
<td>Nonintegrated D</td>
<td>Nonintegrated U</td>
</tr>
<tr>
<td>Trade between integrated unit and nonintegrated direct victim</td>
<td>Yes (but price squeeze)</td>
<td>No&lt;sup&gt;a&lt;/sup&gt;</td>
<td>No&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Incentive to integrate larger for</td>
<td>More efficient U firm</td>
<td>More efficient D firm</td>
<td>Larger U firm</td>
</tr>
<tr>
<td>Possible industry structures</td>
<td>Nonintegration; partial integration; bandwagon; integration and exit (downstream or downstream and upstream)</td>
<td>Nonintegration; integration and exit (upstream or downstream)</td>
<td>Nonintegration; partial integration; bandwagon; integration and exit (downstream or downstream and upstream)</td>
</tr>
</tbody>
</table>

<sup>a</sup> As long as integrated U does not operate at full capacity. Otherwise the integrated D may still buy some supplies from nonintegrated U.<br>
<sup>b</sup> As long as integrated D does not operate at full capacity. Otherwise, the integrated U may sell some of its supplies to a nonintegrated D.<br>
<sup>c</sup> If the downstream firms have the same demands. If they have different demands, say, because they have different storage or marketing facilities, then the same industry structures as in the scarce supplies case may emerge.

Second, there may be pure efficiency gains from mergers. In all three variants of the model, upstream and downstream firms make ex ante investments. Although these investments are taken to be industry-specific, given that the industry is imperfectly competitive, they have many of the characteristics of the relationship-specific investments emphasized by Williamson and Klein, Crawford, and Alchian.<sup>11</sup> In par-

---

ticular, an upstream firm, say, might be unwilling to invest, given that the absence of a perfectly competitive market for its product can cause it to be held up. Thus one motive for a merger between an upstream and downstream firm may be to encourage investments by reducing holdup problems. A merger carried out for these reasons will increase competition and reduce consumer prices. For simplicity, the formal model supposes that firms are prepared to invest under nonintegration and so holdup problems are not a motive for merger; it would be easy to relax this assumption, however.

Given these conflicting effects, it is hard to deliver clear-cut prescriptions for antitrust policy on vertical mergers. Any industry in which investments are industry-specific rather than relationship-specific (the particular cases we consider later all fit into this category) is either competitive, in which case neither holdup nor foreclosure effects should be important and vertical mergers should be irrelevant, or imperfectly competitive, in which case both holdup and foreclosure effects are potentially important and it is hard to distinguish between them. The theory can, however, give some guidance as to when the foreclosure effects are likely to be significant, so that the onus might be on the merging firms to show that there are substantial efficiency gains offsetting the anticompetitive effects. According to our variants, restriction of competition is most likely to be a factor when the merging firms are efficient (have low marginal costs or investment costs) or are large (have high capacities) relative to nonmerging firms. Because there is no strong reason to think that holdup problems will be more serious for efficient or large firms, the theory suggests that vertical mergers involving efficient or large firms should be subject to particular scrutiny by the antitrust authorities. The model also suggests that the antitrust authorities should only be suspicious of vertical mergers that significantly harm rivals. Thus a merger between an upstream and a downstream firm that have had substantial dealings with outside firms is potentially more damaging than one between those that have primarily traded with each other and where the foreclosure effect on rivals will be small.

The paper is organized as follows. The next section describes the model. The first variant is explored in the sections titled "Ex Post Monopolization: The Case of Perfect Certainty and No Investment" and "Ex Post Monopolization: Uncertainty and Positive Investments." The second variant is discussed in the section "Bargaining Effects: Scarce Needs" and the third variant in "Bargaining Effects: Scarce
Supplies.” (The section called “Ex Post Monopolization: Uncertainty and Positive Investments” is considerably more involved than the others, and the reader may wish to skip it on first reading.) These sections are followed by “Extensions”; “Applications,” which applies the analysis to various industries; and “Review of the Literature,” which puts this paper in the context of previous research. Finally, the appendixes contain technical material and an analysis of exclusive-dealing contracts.

The Model

There are two potential suppliers or upstream firms, $U_1$ and $U_2$, and two potential buyers or downstream firms, $D_1$ and $D_2$. The downstream firms compete on the product market and sell perfect substitutes. The demand function for the final good is $Q = D(p)$ with concave inverse demand $p = P(Q)$. The upstream firms produce the same intermediate good at constant marginal cost, $c_i$ ($i = 1, 2$). The intermediate good is transformed into the final good by the downstream firms on a one-for-one basis at zero marginal cost (the downstream firms are thus symmetric).

It is assumed that the upstream marginal costs, $c_i$, are sufficiently high relative to the downstream marginal cost (zero) that if the downstream firms, $D_1$ and $D_2$, have purchased quantities $Q_1$ and $Q_2$ in the “viable range,” the Nash equilibrium in prices in the downstream market has both firms charge the market-clearing price $P(Q)$, where $Q = Q_1 + Q_2$. For this reason the Cournot revenue functions, profit functions, and reaction curves are relevant. Define

$$r(q, \hat{q}) = P(q + \hat{q})q,$$

$$\pi^t(q, \hat{q}) = [P(q + \hat{q}) - c_i]q,$$

and

$$R_i(\hat{q}) = \arg \max_q \pi^t(q, \hat{q}).$$

12. See Tirole (1988, chap. 5) for more detail.
Assume that $\pi^i$ is strictly concave in $q$ and twice differentiable. $R_i(q)$ is then unique and differentiable. As is well known, the slope of a reaction curve is between $-1$ and $0$: $-1 < dR_i/dq < 0$.

We assume that for any costs $(c_1, c_2)$, the reaction curves $R_1$ and $R_2$ have a unique intersection $[q_1^*(c_1, c_2), q_2^*(c_1, c_2)]$; that is, the Cournot equilibrium is unique. We also introduce the monopoly output $q^m(c)$ and monopoly profit

$$
\pi^m(c) = \max_{Q} \{[P(Q) - c]Q\} = \{P[q^m(c)] - c\}q^m(c)
$$

at cost $c$. Last, for technical convenience, assume that firm $i$’s marginal revenue is convex in firm $j$’s output (as is the case, for instance, for linear demand curves). This assumption is needed only in the first variant and is a sufficient condition for contracts that induce random behavior by downstream firms not to be optimal for upstream firms.

The industry evolves in two stages: ex ante and ex post. The ex ante stage includes the decisions before uncertainty is resolved: vertical integration and industry-specific investments. The uncertainty is two-dimensional. First, the firms do not know ex ante which intermediate good will be the appropriate one to trade ex post. We adopt the Grossman and Hart (1986) methodology of presuming that the large number of potential technologies or products ex ante makes it too costly to write complete contracts and that the only way to influence ex post behavior is through the allocation of residual rights of control over assets. Second, the firms may not know which marginal cost structure $(c_1, c_2)$ for the relevant product will prevail. Rather they have prior cumulative distribution functions $F_1(c_1)$ and $F_2(c_2)$ on $[\underline{c}, \overline{c}]$; for simplicity $c_1$ and $c_2$ are drawn from independent distributions.

The timing is as follows:

**Ex Ante Stage**

**Step 1 (vertical integration).** First, firms decide whether to integrate vertically. Antitrust statutes prevent any merger with a horizontal element. They thus allow only mergers between a $U$ and a $D$, because a firm cannot include the two upstream units or the two downstream units. Assuming that the four parties are still active after the investment or exit stage (see step 2), four industry structures may emerge:

—NI (nonintegration). All four parties are separately run.
—PI1 (U1-D1 integrated). Firms U1 and D1 have merged, firms U2 and D2 remain independent (without loss of generality one can assume that U_i merges with D_i, since the two downstream firms are symmetric).

—PI2 (U2-D2 integrated). Only firms U2 and D2 have merged.

—FI (full integration). U1 and D1 have merged and so have U2 and D2. The industry has experienced bandwagon.

We also want to study the possibility of ex ante monopolization, in which vertical integration by a U and a D triggers exit by the other D, the other U, or both. We will denote these industry structures by M'_{U_i}, M'_{U_d}, and M'_{U_d} respectively; for instance, M'_{D} means that the integration of U_i and D_j has triggered exit of D_j and thus the ex ante monopolization of the downstream market (but not of the upstream market).

**Step 2 (investment or exit).** After choosing whether to integrate, the U and D units commit industry-specific investments: 0 or I for upstream units, 0 or J for downstream units. Investing 0 implies that the unit is not able to trade in the ex post stage and thus exits. A unit that invests is able to trade ex post. Investments are noncontractible and are thus private costs to the parties that commit them, in the tradition of the bilateral monopoly paradigms of Williamson (1975, 1985) and Grossman and Hart (1986), with the particularity that investments are industry-specific rather than firm-specific. Under integration, however, an implication of the profit-sharing assumption 1 below is that these investment costs can be internalized between the merging parties. At the end of this step, the industry structure is one of NI, PI1, PI2, or FI if all units have invested, or M'_{U_i}, M'_{D_j}, or M'_{U_d} if integration between U_i and D_i has triggered exit of U_j, D_j or both. The other configurations will be irrelevant under our assumptions.

**Ex Post Stage**

**Step 3 (resolution of uncertainty).** At the beginning of the ex post stage, all parties learn the relevant product to trade. They also learn the upstream marginal costs (c_1, c_2) to produce this product. There is no asymmetry of information (all parties know the marginal costs as well as the demand curve).
Steps 4 and 5 (contract offers and acceptances). The upstream and downstream firms contract about how much of the intermediate good to trade. Variants discussed later differ in the nature of competition between $U_1$ and $U_2$. The first variant presumes Bertrand competition, while the other two allow a more even distribution of bargaining power between the upstream and downstream firms.

Step 6 (production and payments). Outputs of intermediate good specified by contracts and internal orders are produced and delivered. Payments are made by the downstream firms to the upstream firms.

Step 7 (final product market competition). $D_1$ and $D_2$ transform the intermediate good into final product (at zero marginal cost) and sell their outputs $Q_1$ and $Q_2$ at price $P(Q_1 + Q_2)$. As noted above, it is optimal for them to do so, assuming that they learn each other’s output before choosing their prices and that $c_1$ and $c_2$ are sufficiently large.

Returning to the ex ante stage, we make the following assumptions about the consequences of vertical integration, a justification for which was given in the introduction.

Assumption 1: Integration between a $U$ and a $D$ results in their sharing profits ex post. This is the benefit of integration. This leads to the removal of all conflicts of interest about prices and trading policies, although conflicts over effort may remain.

13. A subtlety implicit in assumption 1 should be noted. What is actually being assumed is that under integration, profits of the parent and subsidiary are commingled in such a way that profit sharing is inevitable. In other words, the previous arrangement, whereby the manager of the subsidiary is paid according to the subsidiary’s profit, is no longer feasible. Assumption 1 is, of course, extreme, but it does seem reasonable to suppose that it is harder to identify the performances of the parent and subsidiary under integration than under nonintegration. Most of our results seem likely to generalize to the case where conflicts of interest over prices and trading policies are reduced even if not eliminated under integration.

An implication of assumption 1 is that it does not matter which is the parent company and which is the subordinate company in a merger, that is, it does not matter whether the upstream firm buys the downstream firm or vice versa. This simple view of mergers suffices for the analysis presented here, but the identity of the owning party does matter under more general conditions. See Grossman and Hart (1986) or Hart and Moore (1988) for a discussion.
**Assumption 2:** Integration between a U and a D involves a loss in efficiency equal to a fixed number, $E \geq 0$. This is the cost of integration.\(^{14}\)

We also make the following assumptions on the merger game.

**Assumption 3:** $U_i$ can merge with $D_i$ only.

This assumption is made for convenience. For example, allowing an upstream firm, say, to bargain with several downstream firms raises some thorny issues related to antitrust. What would happen under the antitrust statutes if both downstream firms agreed to merge with the same upstream firm? If we assume that an upstream firm can negotiate with a single downstream firm, assumption 3 involves no loss of generality because the downstream firms are symmetric.\(^{15}\) We will further assume that $U_i$ and $D_i$ make the optimal merger decision for them. The distribution of the gains from merging between them depends on their relative bargaining power and will not be investigated here because it does not affect industry structure and performance.\(^{16}\)

**Assumption 4:** Integration is irreversible.

Divestiture is ruled out by assumption 4. In practice, divestiture is costly because some of the integration costs are sunk and because new

14. As noted in the introduction, one component of the cost of integration is the loss caused by a subordinate manager’s dulled incentives. One case consistent with our hypothesis that $E$ is a fixed number independent of the rest of the model is that in which the subordinate’s dulled incentives concern activities having to do with the reduction of fixed (as opposed to marginal) production costs and the supply of goods to third parties (firms outside the industry).

15. Assumption 3 does have one important implication, however; it rules out the possibility of extortion by the upstream firms. For instance, it might be the case that the sum of the profits of $U_1$ and $D_1$ falls if they integrate, and yet $D_1$ accepts a low offer from $U_1$ to merge because of $U_1$’s threat to merge with $D_2$ and foreclose $D_1$ at the ex post stage.

16. As we shall see, a merger between $U_i$ and $D_i$ will often hurt $U_j$ or $D_j$ or both. One possibility we do not allow is that $U_j$ or $D_j$ bribes $U_i$ or $D_i$ not to merge. There are two justifications for this. First, such a bribe might be viewed with suspicion by the antitrust authorities. Second, there may be roundabout ways in which $U_i$ and $D_i$ can merge (for example, by forming a holding company that owns both $U_i$ and $D_i$) so as to evade a contract committing them not to combine. Note that this position is not inconsistent with the view that the antitrust authorities can prohibit mergers. There might be enough evidence that the formation of a holding company amounted to a merger for a court to rule against such a holding company in an antitrust case, but not enough evidence for a court to make the same ruling in a breach of contract case.
costs are incurred. However, assumption 4 would be unduly restrictive in industries in which demand and cost conditions change dramatically over time. Studying the cyclical integration and disintegration decisions of firms is an important item on the research agenda, and one to which our model is amenable, but it is outside the scope of this paper.

**Assumption 5:** If $U_i$ and $D_i$ integrate, $U_j$ and $D_j$ can follow suit before step 2 (immediate response).

This assumption deserves some clarification. It states that firms can react quickly to their rivals' integration decision. Formally, it corresponds to the following "reduced-form merger game" within step 1. First, the $U$ firms simultaneously decide whether to integrate. Second, if $U_i$ has integrated and $U_j$ has not, $U_j$ gets a chance to respond (but the firms cannot integrate in this "second period of step 1" if none has integrated in the "first period").

The reduced-form merger game is not rich enough to depict some interesting situations. Suppose for instance that if one of the $U$ merges, the unintegrated $D$ exits; it may be the case that the reduced-form merger game has two equilibria: "$U_1$ integrates, $U_2$ does not" and "$U_2$ integrates, $U_1$ does not." To select between the two equilibria and to give a more realistic picture of merger dynamics, we also develop a continuous-time version of the merger game. Suppose that time is continuous and that at each instant there is a new trading dimension ("product" in our model) on which to contract. Contracting must be done just before trading. Similarly, investment must be committed continuously for the firms to keep abreast of industry developments (that is, to avoid exit: we suppose that once a unit has stopped investing it cannot come back). The profits mentioned in the paper are then flow profits; $E$ is the present discounted value of the integration cost—it can be thought of as being equal to $E_0 + (E_1/r)$, where $E_0$ is the upfront integration cost, such as legal fees, $E_1$ is the flow loss of incentives, and $r$ is the rate of interest. In this continuous-time framework, the strategic variable is the date of integration. The loss for $U_i$ to integrating just after $U_j$, compared with integrating simultaneously, is negligible because the loss in flow profit is infinitesimal relative to present values of profits. We adopt the convention that the market "opens" at date 0. That is, the flow investment is incurred and the flow profits are received at each
instant from date 0 on. However, we let firms incur the integration cost before date 0 if they so wish in order to allow preemption.

Besides giving an interpretation of the immediate-response postulate of assumption 5, this continuous-time model selects among multiple equilibria and yields the date at which integration occurs. In those cases in which the reduced-form game has a unique equilibrium, the continuous-time model predicts the same integration pattern, which then occurs at date 0.

**Ex Post Monopolization: The Case of Perfect Certainty and No Investment**

We now develop the first variant of our model in which the upstream firms compete à la Bertrand in step 4. We consider first the case in which the firms' marginal costs are certain and investment costs are zero. Later we extend the analysis to uncertain marginal costs and positive investments.

**Step 4: Contract Offers under Bertrand Competition**

Both upstream firms make simultaneous and secret contract offers to each unintegrated D. In a vertically integrated firm, given the profit-sharing assumption, this offer is a willingness to supply any level of output at an internal marginal transfer price equal to the marginal cost, ci, of the upstream unit.

We will not put any restriction on the contracts that can be signed between a U and a D, given the information structure. A simple

---

17. The secrecy assumption reflects the possibility of hidden or side contracting. It allows us to abstract from the possibility of contracts committing the downstream firms to adopt certain behaviors in the final product market; see Fershtman and Judd (1986) and Katz (1987). In addition it rules out the possibility that an upstream firm can commit itself to limit its sales to some downstream firm by making an appropriate public offer to that firm.

18. Unlike most papers in this literature, our paper does not confer an exogenous advantage to the integrated firms by having the internal transfer price be equal to marginal cost while external transfer prices differ from marginal cost because two-part tariffs are ruled out. We will allow general contracts, including two-part tariffs, for external transactions.
contract between $U_i$ and $D_j$ specifies a transfer, $t_{ij}$, from $D_j$ to $U_i$ that depends on the quantity purchased by $D_j$ from $U_i$, which is $t_{ij}(q_{ij})$. (For instance, a two-part tariff is an affine function of $q_{ij}$.) We will actually allow a finer information structure and accordingly a larger class of feasible contracts. We suppose that $D_j$ can show to $U_i$ any amount of the good, or exhibit receipts for the sales on the final good market, as long as it does not exceed the total amount of the good bought by $D_j$ from $U_i$ and $U_j$. Thus if $Q_j = q_{ij} + q_{2j}$ is the quantity purchased by $D_j$, the firm can demonstrate any $\hat{Q}_{ij} \leq Q_j$ to $U_i$. Accordingly we allow conditional contracts, $t_{ij}(q_{ij}, \hat{Q}_{ij})$.19

**Step 5: Acceptance and Rejection of Contracts under Bertrand Competition**

The unintegrated downstream firms simultaneously accept or reject the contracts offered in step 4. If $D_j$ accepts $U_i$’s offer, it selects an input level, $q_{ij}$ and, in the case of a conditional contract, announces a quantity $\hat{Q}_{ij}$ to be exhibited later to $U_i$, such that $\hat{Q}_{ij} \leq Q_j = q_{ij} + q_{2j}$.

Assume, without loss of generality, that $c_1 \leq c_2$. We describe an equilibrium in the four industry structures that are possible, given that no firm exits, and relegate the study of uniqueness to appendix B.

—Nonintegration NI. The outcome under nonintegration is given in proposition 1.

**PROPOSITION 1:** Assume $c_1 \leq c_2$. Under nonintegration, $D_1$ and $D_2$ each buy $q^* = q^*(c_1)$ from $U_1$ and 0 from $U_2$, where $q^*$ is the Cournot level corresponding to marginal cost $c_1$: $q^* = R_1(q^*)$. They each pay a transfer $t^*$ to $U_1$ and 0 to $U_2$, where

$$r(q^*, q^*) - t^* = r[R_2(q^*), q^*] - c_2R_2(q^*)$$

19. The reason for introducing conditional contracts is technical. Conditional contracts turn out to be irrelevant in six of the seven possible industry structures, and the reader might as well think in terms of simple contracts. In the seventh industry structure, partial integration in which the higher-cost upstream firm is integrated, no equilibrium exists that involves simple contracts only, unless $c_1 = c_2$ or $|c_2 - c_1|$ is large. There exists an equilibrium in conditional contract offers in which the downstream firms end up choosing simple contracts, so that conditional clauses, although offered, are not selected on the equilibrium path. Furthermore, this equilibrium yields the reasonable outcome of a richer contract-offer game in which only simple contracts are enforceable: see note 20.
Total output is $2q^*$ and profits are

$$
U_1: \ U^{NI}(c_1, c_2) = 2[r(q^*, q^*) - c_1q^*]$$

$$- 2[r(R_2(q^*), q^*) - c_2R_2(q^*)]$$

$$U_2: \ U^{NI}(c_2, c_1) = 0$$

$$D_j: \ D^{NI}(c_1, c_2) = r[R_2(q^*), q^*]$$

$$- c_2R_2(q^*) \text{ for } j = 1, 2.$$ 

The intuition behind proposition 1 is as follows. In equilibrium each $D$ anticipates that its rival buys the Cournot output from the low-cost firm. Given this, it can do no better than buying $q^*$ from the low-cost firm too. The transfer price given by equation 1 is such that each $D$ is indifferent between accepting $U_1$’s offer to sell $q^*$ at $t^*$ and buying the best reaction to $q^*$ at a cost of $c_2$ per unit (from $U_2$). $U_1$’s profit is equal to industry profit minus the downstream firms’ profit. Note that, from Bertrand competition, $U^{NI}(c, c) = 0$ for all $c$.

The proof of proposition 1, as well as of other propositions in this section, is to be found in appendix A.

—Partial integration $PI_1$. Suppose now that $U_1$ and $D_1$ are integrated and $U_2$ and $D_2$ have remained independent. We index profits by $PI$. In particular, $D^{PI}(c, c')$ denotes the nonintegrated downstream firm’s profit when the integrated supplier has cost $c$ and the nonintegrated one has cost $c'$.

**Proposition 2:** Assume $c_1 \leq c_2$. Let $(q_1^*, q_2^*) = [q_1^*(c_1, c_2), q_2^*(c_1, c_2)]$ be given by $q_1^* = R_1(q_2^*)$ and $q_2^* = R_2(q_1^*)$. Thus $q_1^* \geq q^*(c_1) \geq q_2^*$ and $q_1^* + q_2^* \leq 2q^*(c_1)$. Under $PI_1$, $U_1$ produces $q_1^*$ for the internal buyer $D_1$ and sells $q_2^*$ at price $t_2^*$ to $D_2$ where

$$t_2^* = c_2q_2^*.$$ 

$U_2$ does not sell. Total industry output is $(q_1^* + q_2^*)$, and profits are

$$U_1D_1: \ V^{PI}(c_1, c_2) = E, \text{ where}$$

$$V^{PI}(c_1, c_2) = r(q_1^*, q_2^*) - c_1q_1^* + (c_2 - c_1)q_2^*$$

$$\geq U^{NI}(c_1, c_2) + D^{NI}(c_1, c_2)$$
\[ q_2 = R_1(q)_2 \]
\[ q_1 + q_2 = 2q^* \]
\[ q_1 = R_1(q)_1 \]

All inequalities in this proposition are strict if and only if \( c_1 < c_2 \).

In words, the equilibrium is the Cournot equilibrium between two firms with marginal costs \( c_1 \) and \( c_2 \), except that production efficiency holds. The low-cost integrated upstream firm supplies \( q_2^* \) to the external buyer at profit \( (c_2 - c_1)q_2^* \). The comparison with the nonintegrated case is depicted in Figure 1.

The difference from nonintegration stems from the fact that, because of profit sharing, an integrated \( U_1-D_1 \) has an incentive to restrict supplies to \( D_2 \) as much as possible. However, since it cannot stop \( U_2 \) from
supplying \( R_2(q_1^*) \), its best strategy is to undercut \( U_2 \) slightly and supply \( R_2(q_2^*) \) itself. Firm \( D_2 \) is partially foreclosed and is hurt by vertical integration, while the profit of the integrated firm rises. Ex post monopolization \((q_1^* + q_2^* < 2q^* \text{ if } c_1 < c_2)\) results because \(-1 < \frac{dR_1}{dq_2} < 0\) and \((q_1^*, q_2^*) \) and \((q^*, q^*)\) are both on the \( q_1 = R_1(q_2) \) reaction curve. Social welfare is reduced and, gross of the integration cost \( E \), industry profit has increased.

—Full integration \( FI \). Suppose now that \( U_1-D_1 \) and \( U_2-D_2 \) are integrated.

**Proposition 3**: Under full integration and \( c_1 \leq c_2 \), the allocation is the same as under \( PI_1 \), except that the integrated firm, \( U_2-D_2 \), also incurs efficiency loss \( E \). That is, \( U_1 \) supplies \( q_1^* \) to \( D_1 \) and \( q_2^* \) to \( D_2 \), and \( U_2 \) does not supply. The profits are thus

\[
\begin{align*}
U_1-D_1: & \, V^{FI}(c_1, c_2) - E, \text{ where } V^{FI}(c_1, c_2) = V^{PI}(c_1, c_2) \\
U_2-D_2: & \, V^{FI}(c_2, c_1) - E, \text{ where } V^{FI}(c_2, c_1) = V^{PI}(c_1, c_2).
\end{align*}
\]

Thus vertical integration by the high-cost supplier has no other effect than the efficiency loss. The reason is that \( U_2 \) did not supply \( D_1 \) and \( D_2 \) anyway. In particular, \( U_2 \) and \( D_2 \) do not have an incentive to integrate in the deterministic case if \( U_1 \) and \( D_1 \) have integrated. In contrast, with uncertain costs, bandwagoning may occur.

—Partial integration \( PI_2 \): Last, suppose that only \( U_2 \) and \( D_2 \) are integrated and that \( c_1 \leq c_2 \).

**Proposition 4**: Under \( PI_2 \) and \( c_1 \leq c_2 \) the allocation is the same as under \( NI \), except that \( U_2-D_2 \) incurs the efficiency loss \( E \). \( U_1 \) supplies \( q^* = q^*(c_1) \) to both \( D_1 \) and \( D_2 \), and \( U_2 \) does not supply. Industry output is \( 2q^* \) and profits are

\[
\begin{align*}
U_1: & \, U^{PI}(c_1, c_2) = U^{NI}(c_1, c_2) \\
D_1: & \, D^{PI}(c_2, c_1) = D^{NI}(c_1, c_2) \\
U_2-D_2: & \, V^{PI}(c_2, c_1) = E, \text{ where } V^{PI}(c_2, c_1) = V^{NI}(c_1, c_2).
\end{align*}
\]
As in proposition 3, vertical integration by the high-cost supplier has no other effect than the efficiency loss.\(^{20}\)

Next consider the ex ante stage. This is trivial when \(c_1\) and \(c_2\) are deterministic and investment costs are zero. Firms \(U_2\) and \(D_2\) have no incentive to integrate, whether or not \(U_1\) and \(D_1\) have. Thus the possible equilibrium industry structures are nonintegration and partial integration by \(U_1\) and \(D_1\). The latter will occur if and only if \(U_1-D_1\)'s profit is higher under partial integration than nonintegration, that is,

\[
V^{pl}(c_1, c_2) - [U^{ni}(c_1, c_2) + D^{ni}(c_1, c_2)] - E > 0.
\]

This completes the analysis of the case with deterministic marginal costs and zero investment costs. The next section considers uncertain marginal cost and positive investment cost. Because the section is more difficult than the others, the first-time reader may well wish to skip to the subsequent section, "Bargaining Effects."

Ex Post Monopolization: Uncertainty and Positive Investments

Costs \(c_1\) and \(c_2\) are now uncertain ex ante but are known ex post. In the certainty case with \(c_1 \leq c_2\), firm \(U_2\) had no incentive at all to remain

\(^{20}\) Some have questioned how our analysis would change if \(D_1\) and \(D_2\) competed à la Bertrand instead of à la Cournot in the downstream market. Note that this would involve a radical change in the timing of production and sales. Given our assumption that upstream firms must first ship the intermediate good to downstream firms, and that downstream firms then transform this good into final output, the downstream market game is played by firms with capacity constraints, and as noted previously, the outcome will inevitably be Cournot if \(c_1\) and \(c_2\) are high enough.

It is also worth giving the flavor of the argument as to why there may exist no pure strategy equilibrium in simple contracts under \(Pl_2\) (see note 19). Firm \(U_2\) can try to reduce industry output by offering \(q_{21} < q^*\) to \(D_1\) at the money-losing price \(t_{21} < c_2 q_{21}\), such that \(D_1\) makes more profit accepting \(U_2\)'s offer than \(U_1\)'s. While such a strategy would be too costly in terms of production cost for \(U_2\) if \(c_2\) is much larger than \(c_1\), it may become optimal for \(U_2\) if \(c_2\) is close to \(c_1\). Such a strategy is unlikely to succeed in practice. Basically, \(U_2\) bribes \(D_1\) to purchase a low output. But \(D_1\) would always go back to \(U_1\) to buy more output and bring itself to the reaction curve \(R_1\). If such recontracting is feasible, \(U_2\)'s counterstrategy does not succeed in bringing industry output below \(2q^*\). The possibility of \(D_1\)'s getting more from \(U_1\) is formalized in the equilibrium of our one-shot contracting game by \(U_1\)'s sleeping clause, allowing \(D_1\) to complement to \(q^*\) its purchases from \(U_2\).
in the industry, and so with $I > 0$ it would have exited. This feature disappears once $c_1$ and $c_2$ are stochastic. Because $c_2 < c_1$ with some probability, $U_2$ has an incentive to stay to take advantage of realizations in which it is the more efficient firm, as long as $I$ is small. We start by considering the case in which investment costs $I$ and $J$ are small enough that none of the four parties has an incentive to exit.

The Ex Ante Stage When Investment Costs Are Small

To analyze the case in which $c_1$ and $c_2$ are uncertain, we use the following corollary of propositions 1 through 4: $U_i-D_i$'s gain from integration is independent of whether $U_j$ and $D_j$ merge. This is not to say they are indifferent as to $U_j$’s and $D_j$’s integration decision; rather, integration by $U_j$ and $D_j$ implies the same decrease in the aggregate profit of $U_i$ and $D_i$ whether $U_i$ and $D_i$ are integrated or not.

For $c_i \leq c_j$, define the ex post gain from integration for $U_i$ and $D_i$ as

$$g(c_i, c_j) = V^{PL}(c_i, c_j) - [U^{NL}(c_i, c_j) + D^{NL}(c_i, c_j)]$$

$$= V^{FL}(c_i, c_j) - [U^{PL}(c_i, c_j) + D^{PL}(c_j, c_i)].$$

Note that $g(c, c) = 0$ for all $c$. For $c_i \geq c_j$ the ex post gain from integration is $g(c_i, c_j) = 0$. The ex ante or expected gain from integration for $U_i-D_i$ is thus

$$G(F_i, F_j) = \mathbb{E}g(c_i, c_j) = \mathbb{E}_{c_i \leq c_j} g(c_i, c_j).$$

The deterministic case suggests that the efficient firm gains more from integration than the inefficient one, which does not gain anything. The same holds in the uncertainty case. The natural definition of efficiency refers to first-order stochastic dominance.

**Definition:** $U_1$ is more efficient than $U_2$ if $F_1(c) \succeq F_2(c)$ for all $c$ (with at least some strict inequality).

**Proposition 5:** Suppose that $U_1$ is more efficient than $U_2$ and that either $[\underline{c}, \overline{c}]$ is sufficiently small where $[\underline{c}, \overline{c}]$ is the support of $F_1$ and $F_2$ (small uncertainty),
or \( c_i = c \) with probability \( \alpha_i \) and \( = +\infty \) with probability \((1 - \alpha_i)\) where \( \alpha_1 > \alpha_2 \) (large uncertainty).

Then \( U_1 \) has more incentive to integrate than \( U_2 \):
\[
G(F_1, F_2) > G(F_2, F_1).
\]

The proof of Proposition 5 is in appendix D.

Next, consider the loss, \( L(F_i, F_j) \), incurred by \( U_i \) and \( D_i \) when \( U_j \) and \( D_j \) merge. Propositions 1 through 4 imply that this loss is independent of whether \( U_i \) and \( D_i \) are integrated. Define, for \( c_i > c_j \),
\[
\ell(c_i, c_j) \equiv D^{NI}(c_i, c_j) - D^{PI}(c_j, c_i) \equiv V^{PI}(c_i, c_j) - V^{FI}(c_i, c_j);
\]
and, for \( c_i \leq c_j \), \( \ell(c_i, c_j) = 0 \).

Last define
\[
L(F_i, F_j) \equiv \mathcal{G} \ell(c_i, c_j) = \mathcal{G}_{\{c_i \geq c_j\}} \ell(c_i, c_j).
\]

**Proposition 6:** Suppose that \( U_1 \) is more efficient than \( U_2 \) and that one of the two assumptions of proposition 5 (small uncertainty or large uncertainty) holds. Then \( U_1 \) and \( D_1 \) lose less from the integration of \( U_2 \) and \( D_2 \) than \( U_2 \) and \( D_2 \) lose when \( U_1 \) and \( D_1 \) integrate:
\[
L(F_1, F_2) \leq L(F_2, F_1).
\]

Proposition 6 is proved in appendix D.

Under the assumptions of propositions 5 and 6, it is straightforward to solve the merger game. Let \( G_i \equiv G(F_i, F_j) \) and \( L_i \equiv L(F_i, F_j) \), where, by propositions 5 and 6, \( G_1 \geq G_2 \) and \( L_1 \leq L_2 \).

—Case 1: \( G_1 < E \), which implies \( G_2 < E \). In this case, \( U_1 \) and \( U_2 \) have a dominant strategy not to integrate. The industry structure is nonintegration.

—Case 2: \( G_1 - L_1 > E \). In this case it is a dominant strategy for \( U_1 \) to integrate. There are two subcases: if \( G_2 < E \), the outcome is \( PI_1 \); if \( G_2 > E \), the outcome is \( FI \). A further distinction can be made between eager bandwagon, which arises when \( U_2 \) and \( D_2 \) prefer a fully integrated industry to a nonintegrated industry \((G_2 - L_2 > E)\), and reluctant bandwagon, which arises when \( U_2 \) and \( D_2 \) merge but would have preferred the industry to remain nonintegrated \((G_2 - L_2 < E)\).

—Case 3: \( G_1 - L_1 < E < G_1 \). In this case, firm \( U_1 \) wants to integrate only if \( U_2 \) does not jump on the bandwagon. Thus if \( G_2 < E \), firm \( U_1 \)
integrates and the industry structure is $PI_1$, and if $G_2 > E$, firm $U_1$ refrains from integrating because this would trigger full integration. The industry structure is $NI$.

The stochastic-cost case is summarized in proposition 7.

**Proposition 7:** Suppose that $U_1$ is more efficient than $U_2$ and that small uncertainty or large uncertainty holds. Then if $G_1 < E$, or $G_1 - L_1 < E < G_1$ and $G_2 > E$, the industry structure is nonintegration. If $G_1 - L_1 > E$ and $G_2 < E$, or $G_1 - L_1 < E < G_1$ and $G_2 < E$, the industry structure is partial integration by $U_1$ and $D_1$. If $G_1 - L_1 > E$ and $G_2 > E$, the industry structure is full integration.

A welfare comparison of the different industry structures is simple in the case where $I$ and $J$ are sufficiently small that none of the four parties ever exits. The notion of welfare is the sum of consumer and producer surplus.

**Proposition 8:** In the absence of exit, any industry structure involving vertical integration ($PI_1$, $PI_2$, or $FI$) is socially dominated by the nonintegrated industry structure $NI$.

**Proof:** Vertical integration implies two welfare losses: the efficiency loss, which is $E$ under $PI_1$ and $PI_2$ and $2E$ under $FI$, and output contraction—that is, $q^*_1(c_1, c_2) + q^*_2(c_1, c_2) < 2q^*(c_i)$ if $c_i < c_j$ and either regime $PI_i$ or $FI$ holds. See propositions 2 through 4. Q.E.D.

We turn now to the case where $I$ and $J$ may be large. Because the possibility of exit must now be allowed for, we start by solving the ex post stage when exit has occurred.

*The Ex Post Stage after Ex Ante Monopolization*

Assume without loss of generality that $U_1$ and $D_1$ have integrated, and this causes $D_2$ or $U_2$ or both to exit, leading to ex ante monopolization. The three subcases are denoted by $M_{ud}$ (both $U_2$ and $D_2$ have exited), $M_d$ (only $D_2$ has exited), and $M_u$ (only $U_2$ has exited).

—Upstream and downstream monopolization ($M_{ud}$) or upstream monopolization ($M_u$). If $U_1$ and $D_1$, which have integrated, are monopolists in their respective industry segments, and $U_1$ has marginal
cost \( c_1 \), then \( U_1-D_1 \)'s profit is \( V^{M_d}(c_1) - E \), where \( V^{M_d}(c_1) = \pi^m(c_1) \).

The same holds if \( U_2 \) only has exited because \( U_1 \) supplies only its internal unit \( D_1 \); hence \( V^{M_d}(c_1) = V^{M_d}(c_1) \).

**—Downstream monopolization (\( M_d \)).** Suppose that only \( D_2 \) has exited. If \( c_1 \leq c_2 \), then \( D_1 \) procures internally and \( U_1-D_1 \)'s profit is \( V^{M_d}(c_1, c_2) - E \), where \( V^{M_d}(c_1, c_2) = \pi^m(c_1) \), and \( U_2 \)'s profit, \( U^{M_d}(c_2, c_1) \), is equal to zero.

If \( c_1 > c_2 \), then \( U_2 \) makes an offer to supply \( q^m(c_2) \) to \( D_1 \) at price \( t_{21} = P[q^m(c_2)]q^m(c_2) - \pi^m(c_1) \). Hence the profits for \( U_1-D_1 \) are \( V^{M_d}(c_1, c_2) - E \), where \( V^{M_d}(c_1, c_2) = \pi^m(c_1) \). For \( U_2 \) they are \( U^{M_d}(c_2, c_1) = \pi^m(c_2) - \pi^m(c_1) \).

We return now to the ex ante stage. We consider first the case where \( J \) is large but \( I \) is small, so that only downstream firms exit.

**I ‘‘Small,’’ J ‘‘Large’’ (Possibility of Ex Ante Downstream Monopolization)**

Assume that downstream firms’ investment is large in the sense that \( J > \mathbb{E}D^P(c_1, c_2) \), where \( \mathbb{E} \) is the expectation with respect to \( c_1 \) and \( c_2 \), while the upstream firms’ investment remains small. Throughout we assume that none of the firms exits in step 2 under nonintegration.

**ASSUMPTION 6: Viability under nonintegration.** For all \( i \) and \( j \), \( \mathbb{E}U^{NI}(c_i, c_j) \geq I \) and \( \mathbb{E}D^{NI}(c_i, c_j) \geq J \).

We first analyze when a \( U \) wants to rescue a failing \( D \) by merging with it; this may happen sometimes even though \( U \) and \( D \) would not want to merge if \( D \) were viable (we call this forced bandwagon).

**—When a \( U \) wants to rescue a failing \( D \).** When \( U_i \) and \( D_i \) integrate, only \( D_j \) suffers directly. Its loss is equal to \( L_j \). This may lead \( D_j \) to exit if its new expected profit falls below \( J \) and if \( U_j \) does not come to its rescue by merging with it. A merger gives \( D_j \) an incentive to invest because, given profit sharing, investment costs can be split between \( D_j \) and \( U_j \). Firm \( U_j \) cannot come to \( D_j \)'s rescue by subsidizing its investment cost because investment is not contractible. The only thing it can do is to merge at a reasonable price.

A crucial factor for knowing whether \( U_j \) and \( D_j \) merge when \( U_i \) and
$D_i$ have merged is whether $U_j$ is made better off by $D_j$'s exit. To simplify the notation a bit, let $\mathcal{U}_j^{M_i} = \mathcal{E}U^{M_i}(c_j, c_i)$ denote $U_j$'s expected profit when $D_j$ exits; $\mathcal{U}_j^{PI} = \mathcal{E}U^{PI}(c_j, c_i)$ be $U_j$'s expected profit under partial integration and no ex ante monopolization; $\mathcal{V}_j^{FI} = \mathcal{E}V^{FI}(c_j, c_i)$ be $U_j$-$D_j$'s expected profit under full integration; and $\mathcal{D}_j^{PI} = \mathcal{E}D^{PI}(c_i, c_j)$ be $D_j$'s expected profit under partial integration if it stays. These expected profits are computed assuming that $U_i$ and $D_i$ are integrated.

**PROPOSITION 9**: Following a $U_i$-$D_i$ merger, $U_j$ would prefer $D_j$ to exit ($\mathcal{U}_j^{M_i} > \mathcal{U}_j^{PI}$) in the case of large uncertainty. It would prefer $D_j$ to stay ($\mathcal{U}_j^{M_i} < \mathcal{U}_j^{PI}$) in the case of small uncertainty.

Proposition 9 (proved in appendix D) indicates when $U_j$ would like to keep an industrial base downstream. When it has a large cost advantage over $U_i$, which may arise in the case of large uncertainty, $U_j$ can obtain the monopoly profit if it deals with a single downstream firm; while it cannot commit not to supply both downstream firms if $D_j$ stays around. We call this the commitment effect. If $U_j$ has only a small cost advantage over $U_i$, Bertrand competition between the upstream firms implies that $U_j$'s profit is approximately $2q^*(c_j)(c_i - c_j)$ when both downstream firms are around, where $q^*(c_j)$ is the symmetric Cournot output for cost $c_j$; and $q^m(c_j)(c_i - c_j)$ when only $D_j$ is around, where $q^m(c_j)$ is the monopoly output at cost $c_j$. Because the Cournot industry output exceeds the monopoly output, $U_j$ is then better off facing two downstream units. We call this the demand effect.

--- **Forced bandwagon.** Next suppose that $U_i$ and $D_i$ have merged. We say that forced bandwagon by $U_j$ and $D_j$ occurs if the following three conditions hold: (a) $D_j$ is no longer viable by itself ($J > \mathcal{D}_j^{PI}$); (b) $U_j$ and $D_j$ are better off integrating than letting $D_j$ exit ($\mathcal{V}_j^{FI} - E - J > \mathcal{U}_j^{M_i}$); (c) $U_j$ and $D_j$ would not want to merge if $D_j$ were viable ($\mathcal{U}_j^{PI} + \mathcal{D}_j^{PI} - J > \mathcal{V}_j^{FI} - E - J$).

**PROPOSITION 10**: After $U_i$ and $D_i$ have merged: (i) a necessary condition for forced bandwagon is that $U_j$ would prefer $D_j$ not to exit ($\mathcal{U}_j^{PI} > \mathcal{U}_j^{M_i}$); and (ii) conversely, if $\mathcal{U}_j^{PI} > \mathcal{U}_j^{M_i}$, there exists $(E, J)$ such that forced bandwagon occurs.

**PROOF**: For (i), add (a), (b), and (c); for (ii), straightforward.

Q.E.D.
Propositions 9 and 10 together say that forced bandwagon cannot occur for large uncertainty, but may occur for small uncertainty because the nonintegrated upstream supplier is concerned about keeping an industrial base.

—The merger game. The merger game with large downstream investments involves many cases, including preemption and war-of-attrition games. See appendix E.

I ‘large,’ J ‘large’ (Possibility of Ex Ante Upstream and Downstream Monopolization)

We do not treat the case of general (large) investments upstream and downstream, but instead content ourselves with the following observation. When $U_i$ and $D_i$ merge, $U_j$ may suffer indirectly through the exit of $D_j$ (see proposition 9), and may exit itself. Given that $D_j$ exits, the exit of $U_j$ can only hurt $U_i-D_i$ because the integrated firm can always refuse to trade with $U_j$. It is therefore conceivable that $U_i$ and $D_i$ might refrain from integrating because this would trigger a chain of exits and reduce the industrial base upstream. In the variant of this section, however, this phenomenon does not arise because it is assumed that the upstream firms set prices. Hence, when $U_j$ is more efficient than $U_i$, it makes an offer to $U_i-D_i$ that makes the integrated firm indifferent between accepting the offer and using the internal technology. Thus $U_i-D_i$ does not benefit from $U_j$’s not exiting. But if the bargaining power were more evenly distributed, the phenomenon could occur. We will return to these ideas in the section ‘‘Bargaining Effects: Scarce Supplies.’’

Bargaining Effects

The previous sections focused on the idea that an upstream firm and a downstream firm might integrate to reduce their willingness to supply a rival downstream firm, thus enabling them to monopolize, at least partially, the downstream market. The next two sections analyze a different mechanism by which foreclosure can occur: via bargaining effects. We argue that an upstream firm and a downstream firm may merge to ensure that they trade with each other, that is that the upstream
firm channels scarce supplies to its downstream partner rather than to a downstream competitor and that the downstream firm satisfies its scarce needs by purchasing from its upstream partner rather than an upstream competitor. This can benefit the merging firms in two ways. First, to the extent that rival firms were obtaining some profit from trading with the merging partners, the merger will increase the merging firms' share of total profit. Second, the profits of rival firms may fall below the critical level at which they are covering their costs, and they may exit the market. The merging firms may then succeed in monopolizing the market ex ante.

Two variants capture these ideas. The first focuses on a downstream firm with scarce needs that favors its upstream partner. The second focuses on an upstream firm with scarce supplies that favors its downstream partner. The effects are treated separately because they have somewhat different implications and because the analysis is less burdensome that way. Obviously, in many real situations one would expect to find both effects.

**Bargaining Effects: Scarce Needs**

Assume, as before, two upstream firms and two downstream firms. In this variant, the downstream firms are not directly hurt by vertical integration and it can be assumed without loss of generality that their investment is equal to zero. Denote the investment cost of upstream firm $U_i$ by $I_i$ ($i = 1, 2$), where, without loss of generality, $I_1 \leq I_2$. To abstract from the ex post monopolization issues discussed earlier, we suppose that $U_1$ and $U_2$ have the same constant marginal cost $c$. Earlier assumptions predicted that nonintegration would be the outcome. However, we now drop the assumption that the upstream firms make independent and simultaneous take-it-or-leave-it offers to the downstream firms, supposing instead that contracts are achieved by bargaining. To be more specific, each nonintegrated upstream firm negotiates with each downstream firm to be its supplier. Moreover, the bargaining of an independent $U_i$ with $D_1$ is independent of the bargaining of $U_i$ with $D_2$.21 Finally, the competition of the upstream firms is not so fierce that

21. If $U_i$ and $D_i$ are integrated, bargaining between them over price is irrelevant, given our assumption that managers of $U_i$ and $D_i$ both get a fraction of total profit. In this case, $U_i-D_i$ will still want to compete with $U_j$ to supply $D_j$, assuming $U_j$ has not exited.
their profits are completely eliminated; instead we suppose that a constant fraction, \( \beta \), of the surplus from supplying a downstream firm accrues to each upstream firm, where \( 0 < \beta < 1/2 \) (so the fraction of surplus accruing to the downstream firm is \( 1 - 2\beta \)).\(^{22}\) We will also sometimes need to consider the case in which there is only one upstream firm in the market. We assume that this upstream firm captures a fraction, \( \beta' \), of the surplus from supplying a downstream firm, where \( \beta' > 2\beta \), so that a downstream firm does strictly worse bargaining with one upstream firm than with two.

**Remark:** The scarce needs variant can be reinterpreted as applying to a situation in which the upstream firms supply a piece of machinery or a technology that allows the downstream firms to produce at marginal cost \( c \). Each downstream firm has a unit demand for the machinery or the technology. In this reinterpretation the sense in which needs are scarce is particularly clear.

**Nonintegration**

Suppose for the moment that both upstream firms invest under nonintegration. Since \( U_1 \) and \( U_2 \) have the same marginal cost, the reaction curves \( R_1 \) and \( R_2 \), defined earlier, are the same: \( R_1(q) = R_2(q) = R(q) \). The equilibrium under nonintegration is described in the next proposition.

**Proposition 11:** Under nonintegration, \( D_1 \) and \( D_2 \) each buy \( q^* \) from the upstream firms, where \( q^* \) is the Cournot level corresponding to marginal cost \( c \): \( q^* = R(q^*) \). The surplus to be shared among each downstream firm and \( U_1 \) and \( U_2 \), given that the rival downstream firm chooses \( q^* \), is \( P(2q^*)q^* - cq^* = \pi^d \), and this is divided in the proportions \( (1 - 2\beta) \), \( \beta \), and \( \beta \) respectively. Total output is \( 2q^* \) and profits are

\[
\begin{align*}
U_i: & \quad U^{NI} = \beta \pi^d + \beta \pi^d = 2\beta \pi^d \\
D_i: & \quad D^{NI} = (1 - 2\beta) \pi^d.
\end{align*}
\]

22. Here, \( \beta \) can be understood as the expected share of the surplus that \( U_i \) obtains rather than the actual share. For example, one interpretation is that each upstream firm wins the competition to supply a particular downstream firm with probability \( 1/2 \); the winner receives a share, \( 2\beta \), of profit and the loser receives nothing.
The proof of this proposition is straightforward. Let $q_1$ and $q_2$ be the amounts that $D_1$ and $D_2$ are expected to purchase in equilibrium. Then $D_1$ in combination with either $U_1$ or $U_2$, or both, can, taking $q_2$ as given, achieve a total surplus of $\max_q [P(q + q_2)q - cq]$. The solution to this maximization problem is $q_1 = R(q_2)$. By a similar argument, $q_2 = R(q_1)$. It follows that $q_1 = q_2 = q^*$. The remainder of proposition 11 follows from the assumptions about bargaining and the division of surplus.

**Full Integration**

Consider next full integration, maintaining for the moment the assumption that $U_1$ and $U_2$ both invest. The only change caused by full integration is that $D_1$ will obtain all its supplies from its partner $U_1$, and $D_2$ will obtain all its supplies from $U_2$. There is no reason to buy externally because internal production is as cheap. This does not change equilibrium output levels because the best reaction for $U_i-D_i$ to an expected purchase of $q_j$ by $D_j$ is $R(q_j)$. Hence $q_1 = R(q_2)$ and $q_2 = R(q_1)$, that is, $q_1 = q_2 = q^*$. Firms $U_1$ and $D_1$ will together share the profit $\pi^d$, and similarly so will $U_2$ and $D_2$. From these profits must be subtracted the integration costs $E$. The outcome is summarized in proposition 12.

**PROPOSITION 12:** Under full integration, $D_i$ buys $q^*$ from upstream firm $U_i$ ($i = 1, 2$), where $q^* = R(q^*)$. Total output is $2q^*$ and profits are

- $U_1-D_1$: $V^{FI} = \pi^d - E$
- $U_2-D_2$: $V^{FI} = \pi^d - E$.

Note that the profits of $U_i-D_i$ are the same under full integration as under nonintegration, except for the integration cost.

**Partial Integration**

Suppose next that $U_i$ and $D_i$ integrate, $U_j$ and $D_j$ remain separate, and $U_i$ and $U_j$ both continue to invest. $U_i$ will now supply all of $D_i$'s needs, putting $D_i$ on its reaction curve $R(q_i)$; but, as in the case of nonintegration, $U_i$ and $U_j$ will compete for $D_j$'s custom. The latter
conclusion follows from the fact that, because $U_i$ and $U_j$ have the same marginal costs, $U_i$ cannot gain ex post from refusing to deal with $D_j$ or restricting its supplies to $D_j$. $U_j$ alone will agree to put $D_j$ on its reaction curve $R(q_i)$, which is the same outcome that occurs if $U_i$ and $U_j$ are both willing to supply $D_j$.

This argument shows that $q_i = R(q_i)$ and $q_j = R(q_i)$, that is, $q_i = q_j = q^*$. Although partial integration does not change output levels, it does affect the division of surplus. $U_j$ will lose the $\beta \pi^d$ it earned from supplying $D_i$ under nonintegration (that is, $U_i$ and $D_i$ will now divide $\pi^d$ between them); while the gains from trade that $D_j$ can realize in combination with $U_i$ or $U_j$ or both will be shared in the proportions $1 - 2\beta$, $\beta$, and $\beta$ respectively.

**Proposition 13:** Under partial integration, $D_i$ buys $q^*$ from $U_i$ and $D_j$ buys $q^*$ from $U_i$ or $U_j$ or both, where $q^* = R(q^*)$. Total output is $2q^*$ and profits are

\[
U_i-D_j: V_{\text{PL}}^i = (1 + \beta) \pi^d - E \\
U_j: U_{\text{PL}}^j = \beta \pi^d \\
D_j: D_{\text{PL}}^j = (1 - 2\beta) \pi^d.
\]

The combined profits of $U_i$ and $D_i$ are higher by $\beta \pi^d - E$ under partial integration than under nonintegration. On the other hand, the profits of $U_j$ and $D_j$ are lower by $\beta \pi^d$.

**Ex Ante Monopolization**

So far we have supposed that $U_1$ and $U_2$ invest under both integration and nonintegration. The final structure we consider is one in which the integration of $U_1$ and $D_1$ causes $U_2$ to exit (the mirror image case in which $U_2$, the firm with higher investment costs, merges with $D_2$ and $U_1$ exits will turn out to be irrelevant). This case leaves the single supplier, $U_1$, facing $D_1$ and $D_2$, one of which is its partner. We can apply proposition 1 to learn the outcome: $U_1$ will supply only $D_1$ and will monopolize the market; that is $U_1-D_1$ will choose the output level $q^m$ that maximizes $P(q)q - cq$.

Denote monopoly profit, $P(q^m)q^m - cq^m$ by $\pi^m$. 
**Proposition 14:** Under ex ante monopolization (integration by $U_1$ and $D_1$ and exit by $U_2$), $D_1$ buys $q^m$ from $U_1$, where $q^m$ maximizes $P(q)q - cq$, and $D_2$ buys nothing. Total output is $q^m$ and profits are

$$U_1-D_1: V_{M_u} = \pi^m - E$$

$$U_2: \text{zero}$$

$$D_2: \text{zero}.$$

We will assume in what follows that the profits of $U_1$ and $D_1$ are higher under ex ante upstream monopolization than under nonintegration. That is:

$$V_{M_u} = \pi^m - E > \pi^d.$$  

If this were not the case, integration would not be profitable under any conditions in the model of this section.23

**The Investment Decision**

Let us reconsider the assumption that upstream firms invest. Under nonintegration, $U_1$ and $U_2$ cover their costs and invest as long as

$$2\beta \pi^d > I_2.$$

We assume condition 4 in what follows.

Consider full integration. Here investment is less an issue. Full integration plus exit by $U_i$, say, could never be a correctly anticipated equilibrium outcome because, given that $D_i$ will not be supplied by $U_j$ and will make zero profits, $U_i$ and $D_i$ could do better by staying separate and saving their merger costs $E$.

Consider next partial integration, in particular the case in which $U_1$ and $D_1$ merge but $U_2$ and $D_2$ stay separate (the logic in the reverse case is similar). Under these conditions $U_2$ may or may not invest. It is easily seen, however, that $U_1$ invests. In particular, suppose the contrary: $U_1$ does not invest, but $U_2$ does. (If $U_2$ does not invest, $U_1-D_1$'s profits are automatically zero if $U_1$ does not invest; hence it is better for $U_1$ to invest.) Then ex post a single nonintegrated firm, $U_2$, will face two downstream firms, $D_1$ and $D_2$. Applying the same logic

23. In particular, $V_{M_u} \leq \pi^d \Rightarrow \beta \pi^d < E$, because $\pi^m > 2\pi^d$. That is, the net gain to $U_i$ and $D_j$ from integrating when $U_j$ and $D_j$ stay separate is negative.
as in proposition 1, we see that $U_2$ will supply $q^*$ to both $D_1$ and $D_2$. Moreover, given the assumption about one-on-one bargaining, $D_1$ and $D_2$ will obtain a share $(1 - \beta')$ of the surplus $\pi^d$, and $U_2$ will obtain the remainder. Thus $U_1-D_1$’s profits will be $(1 - \beta')\pi^d - E$. But because $I_1 \leq I_2$ and $\beta' > \beta$ and because of condition 4, $(1 - \beta')\pi^d < (1 + \beta)\pi^d - I_1$, which ensures that $U_1-D_1$ can do better by investing (see proposition 13). Thus it is never profitable for $U_1$ and $D_1$ to merge if $U_1$ does not invest.

The Merger Game

We treat the merger game as in the ex post monopolization variant. In particular, we suppose that the merger is irreversible and that if $U_i$ and $D_i$ merge, $U_j$ and $D_j$ can respond instantaneously by merging too. Under these assumptions full integration will never be an equilibrium outcome in the present variant. Neither $U_1$ and $D_1$ nor $U_2$ and $D_2$ will merge if the other pair follows suit because, by propositions 11 and 12, the final profit of each pair, $U_i-D_i$, will be less than the combined profits of $U_i$ and $D_i$ under nonintegration.

Partial integration without exit is also not possible. As in the ex post monopolization variant, the gain from the merger of $U_j$ and $D_j$ is the same whether $U_i$ and $D_i$ are integrated or not. This is given by $\beta \pi^d - E$. If this gain is positive, then $U_j$ and $D_j$ will follow suit if $U_i$ and $D_i$ merge. If it is negative, then $U_j$ and $D_j$ will not follow suit, and $U_i$ and $D_i$ will prefer nonintegration to partial integration.

Thus the only reason for $U_i$ and $D_i$ to merge is if the response of $U_j$ is to exit. In other words, the final outcome of the merger game will be either nonintegration or ex ante monopolization.

Proposition 15 shows which of these outcomes will occur. In the formal statement of the proposition we suppose that $U_1$ and $D_1$ merge if any merger occurs at all. It turns out that in case 2 of the proposition, there can be another equilibrium in which $U_2$ and $D_2$ merge and $U_1$ exits. This equilibrium is not compelling, however, because in the continuous-time model described earlier, $U_1$ and $D_1$ would preempt $U_2$ and $D_2$ by merging before date 0.

**Proposition 15:** Assume conditions 3 and 4. Suppose also that $U_1$ and $D_1$ decide first whether to merge, and if, and only if, they merge, $U_2$ and $D_2$ can respond by merging too. Then:
1. The merger game will result in nonintegration if
   (a) $\beta \pi^d > E$ and $\pi^d - E > I_2$; or
   (b) $\beta \pi^d < E$, and $\pi^d - E > I_2$; or
   (c) $\beta \pi^d < E$, $\beta \pi^d > I_2$.
2. The merger game will result in a merger of $U_1$ and $D_1$ and exit of $U_2$ if
   (a) $\beta \pi^d > E$, $\pi^d - E < I_2$; or
   (b) $\beta \pi^d < E$, $\pi^d - E < I_2$ and $\beta \pi^d < I_2$.

As long as "probability zero" cases of equality ($\beta \pi^d = E$ and so forth) are ruled out, these cases are exhaustive.

The proof of proposition 15 is straightforward. In case 1(a), $\beta \pi^d > E$ implies that $U_2$ and $D_2$ will find it profitable to bandwagon if $U_1$ and $D_1$ merge, unless $U_2$ exits. Because full integration is unprofitable for $U_1$ and $D_1$, they will merge only if $U_2$ exits (that is, only if $\pi^d - E < I_2$). In 1(b), $U_2$-$D_2$'s profits are positive under full integration ($\pi^d - E > I_2$), and hence $U_1$-$D_1$ cannot force exit by $U_2$. Therefore $U_1$ and $D_1$ prefer not to integrate. In 1(c), $U_1$-$D_1$ again cannot force exit by $U_2$ because $U_2$ can cover its investment costs by staying independent. Again $U_1$ and $D_1$ choose not to integrate.

Case 2 consists of the complementary region in parameter space to case 1; that is, it consists of those subcases in which the merger of $U_1$ and $D_1$ will cause $U_2$ to exit. Under these conditions, integration is profitable for $U_1$ and $D_1$ (by condition 3).

In case 2 the model may be consistent with another outcome: $U_2$ and $D_2$ merge and $U_1$ exits. In the continuous-time version of the model described earlier, however, this would lead to a preemption game that $U_1$ and $D_1$ would win by merging at date $-T$, where $T$ satisfies

\[-E + e^{-rT} \left(\frac{\pi^m - I_2}{r}\right) = 0.\]

Note that the discounted profit of $U_1$-$D_1$ at date 0 in this equilibrium is $(I_2 - I_1)/r$. For this reason the possibility that $U_2$ and $D_2$ merge and force exit of $U_1$ is ignored.

Remark: In this scarce needs variant, partial integration (without exit) and bandwagon (full integration) are not possible outcomes. However, there is another version of the scarce needs variant in which these outcomes can occur. Suppose that there are limits on how much $D_1$ and
$D_2$ can purchase from the upstream firms, perhaps because they have limited storage. If $D_1$ has more storage space than $D_2$, $U_1$ may merge with $D_1$ to cut $U_2$ out of the gains from trading with $D_1$. Moreover, this can be profitable even if $U_2$ and $D_2$ respond by merging to cut $U_1$ out of the gains from trading with $D_2$.

Rather than analyze a model of this type, we analyze shortly a symmetric version of it in which the upstream firms have scarce capacities. See the section called "Bargaining Effects: Scarce Supplies."

**Welfare**

The welfare effects of merger are straightforward in this variant. Merger followed by exit leads to lower output ($q^m$ vs. $2q^*$) and higher prices for consumers. So consumer surplus falls. Producer surplus, however, rises, and in some cases total surplus may also rise because of the saving in the exiting firm’s investment cost.24

**Bargaining Effects: Scarce Supplies**

The second bargaining effect is scarce supplies, the situation in which the upstream firms are capacity-constrained and integration occurs to ensure that an upstream firm channels its scarce supplies to its downstream partner. Suppose that the two upstream firms, $U_1$ and $U_2$, have exogenously given capacities $\bar{q}_1$ and $\bar{q}_2$, respectively. Assume that $U_1$ is bigger than $U_2$ and thus $\bar{q}_1 > \bar{q}_2$. To simplify, suppose that $U_i$’s marginal cost of production is zero up to its capacity constraint $\bar{q}_i$ ($i = 1, 2$) and that

$$Q = \bar{q}_1 + \bar{q}_2 \leq q^m = \arg \max P(q)q.$$  

Condition 5 ensures that there is no motive to monopolize the market ex post by restricting output. Given the condition, even if there were

24. For example, let $p = a - bQ$, $\beta = \frac{1}{2}$. Then $q^m = (a - c)/2b$ and $q^* = (a - c)/3b$. Total surplus if $U_1$ and $D_1$ merge and $U_2$ exits is $W_m = \frac{3}{2} [(a - c)^2/b] - I_1 - I_2$. Total surplus under duopoly is $W_d = \frac{4}{9} [(a - c)^2/b] - I_1 - I_2$. If $E$ is small and $\pi^d - E < I_2$, it is easy to check that $U_1$ and $D_1$ will merge and $U_2$ will exit; and $W_m > W_d$. These conditions are also consistent with $2\beta \pi^d > I_2$, that is, with both firms investing under nonintegration.
only one downstream firm, it would wish to purchase and sell on the downstream market all the output that \( U_1 \) and \( U_2 \) have available.

Condition 5 is a simplifying assumption that will fail to be satisfied in many markets.\(^2\) In the absence of the condition, aspects of both previous variants come into play (\( \bar{q}_i = \infty \) and \( \bar{q}_j = 0 \) arises in the large-uncertainty case of the ex post monopolization variant and \( \bar{q}_i = \bar{q}_j = \infty \) in the scarce needs variant). A new possibility must also be dealt with: a downstream firm may try to purchase more supplies than it needs and destroy some of them to keep them out of the hands of a rival (in principle, each firm would like to destroy \( \bar{Q} - q' \) if it could buy all the supplies). If condition 5 holds, such a strategy is never optimal. We should also stress that we are confident that our results will continue to be relevant when condition 5 does not hold.

Although \( D_1 \) and \( D_2 \) compete for supplies, they do not really compete on the product market. As long as no upstream firm exits, each unit of the intermediate good has a fixed value, \( P(\bar{Q}) \), for the downstream firms. Thus, if upstream investment costs are small enough and ex ante monopolization is not an issue, the scarce supplies model applies to industries in which the downstream firms are in separate product markets.

Because only the nonintegrated downstream firms are hurt by integration in this variant, it is natural to assume that only \( D_i \) has to invest to operate (but see the remark after proposition 20, which discusses upstream investments). We denote \( D_i \)'s investment cost by \( J \) (assumed to be independent of \( i \)).

Bargaining is modeled in a way similar to that for the case of scarce needs. The roles of the upstream and downstream firms are reversed. The downstream firms are assumed to negotiate with each independent upstream firm to purchase supplies, and the bargaining of \( D_i \) with \( U_1 \) is independent of the bargaining of \( D_i \) with \( U_2 \). We suppose that a fraction, \( \beta \), of the surplus from \( U_i \)'s supplying \( D_1 \) or \( D_2 \) accrues to each of \( D_1 \) and \( D_2 \), and the remaining fraction \( (1 - 2\beta) \) accrues to \( U_i \). We will also sometimes want to consider the case where a single downstream firm bargains with \( U_i \). Under these conditions, again by analogy to the discussion of scarce needs, the downstream firm receives a fraction, \( \beta' \), of the surplus, and \( U_i \) receives \( 1 - \beta' \), where \( \beta' > 2\beta \).

\(^2\) We expect the condition to hold if the cost of building capacity is large.
Nonintegration

Suppose for the moment that both downstream firms invest under nonintegration. The proposition that characterizes equilibrium in this case is immediate.

**Proposition 16:** Under nonintegration, the downstream firms buy the total available capacity, $\bar{Q}$, from the upstream firms. The surplus to be shared between each upstream firm and $D_1$ and $D_2$ is $P(\bar{Q})\bar{q}_i$. This is divided in the proportions $(1 - 2\beta)$, $\beta$, and $\beta$, respectively. Profits are

$$U_i: U_i^{NI} = (1 - 2\beta) P(\bar{Q})\bar{q}_i$$

$$D_i: D_i^{NI} = \beta P(\bar{Q})(\bar{q}_1 + \bar{q}_2) = \beta P(\bar{Q})\bar{Q}.$$ 

**Full Integration and Partial Integration**

Next consider full integration and partial integration, maintaining for the moment the assumption that $D_1$ and $D_2$ invest. If $U_i$ and $D_i$ and $U_j$ and $D_j$ both merge, $U_i$ will sell all its supplies to $D_i$ and $U_j$ all its supplies to $D_j$. If $U_i$ and $D_i$ merge and $U_j$ and $D_j$ do not, $U_i$ will sell all its supplies to $D_i$, and $D_i$ and $D_j$ will compete for $U_j$'s supplies.

The outcomes in these cases are summarized in propositions 17 and 18.

**Proposition 17:** Under full integration, $D_i$ buys $\bar{q}_i$ from $U_i$ ($i = 1, 2$) and profits are

$$U_i-D_j: V^{fi}_{i} = P(\bar{Q})\bar{q}_i - E(i = 1, 2).$$

**Proposition 18:** Under partial integration ($U_i$ and $D_i$ merge, $U_j$ and $D_j$ do not), $D_i$ buys $\bar{q}_i$ from $U_i$, and $D_i$ and $D_j$ compete to buy $U_j$'s supplies, $\bar{q}_j$, sharing the surplus from this transaction in the proportions $\beta$, $\beta$, and $1 - 2\beta$, respectively. Profits are

$$U_i-D_i: V^{pl}_{i} = P(\bar{Q})(\bar{q}_i + \beta\bar{q}_j) - E$$

$$U_j: U_j^{pl} = (1 - 2\beta) P(\bar{Q})\bar{q}_j$$

$$D_j: D_j^{pl} = \beta P(\bar{Q})\bar{q}_j.$$
Propositions 16 through 18 show that the gain to \( U_1 \) and \( D_1 \) from integrating while \( U_2 \) and \( D_2 \) do not is \( \beta P(\overline{Q})\overline{q}_1 - E \), which is the share of surplus that \( D_2 \) used to get from buying \( U_1 \)'s supplies, but which is now divided between \( U_1 \) and \( D_1 \). The gain to \( U_2 \) and \( D_2 \) of jumping on the bandwagon is \( \beta P(\overline{Q})\overline{q}_2 - E \). In other words, as in the previous two variants, the benefits to \( U_i \) and \( D_i \) of integrating are independent of whether \( U_j \) and \( D_j \) integrate (this ignores the possibility that integration by \( U_i \) and \( D_i \) causes \( D_j \) to exit). In contrast to the scarce needs case, however, \( U_1 \) and \( D_1 \) may gain from integrating even if \( U_2 \) and \( D_2 \) follow suit because

\[
V_{1}^{FI} - (U_{i}^{NI} + D_{i}^{NI}) = \beta P(\overline{Q})(\overline{q}_i - \overline{q}_j) - E,
\]

which may be positive if \( \overline{q}_1 \) is sufficiently larger than \( \overline{q}_2 \) (however, the same formula shows that \( U_2 \) and \( D_2 \) cannot gain from integrating if \( U_1 \) and \( D_1 \) follow suit, given \( \overline{q}_2 < \overline{q}_1 \)).

Propositions 16 through 18 also tell us that a merger by \( U_1 \) and \( D_1 \) reduces \( D_2 \)'s profits, but does not have a direct effect on \( U_2 \)'s profits (compare \( U_2^{FI} \) and \( U_2^{NI} \)). The reduction in \( D_2 \)'s profit may cause \( D_2 \) to exit, a case we consider next.

**Ex Ante Monopolization (Exit by \( D_2 \))**

With \( D_2 \) exiting, \( D_1 \) receives \( U_1 \)'s supplies automatically (since they are merged) and negotiates to buy \( U_2 \)'s supplies too. An important difference between this case and previous ones is that if \( D_1 \) declines to buy \( U_2 \)'s supplies, they disappear from the market. Hence the gains that \( D_1 \) can achieve from trading with \( U_2 \) are \( P(\overline{Q})\overline{q}_1 - P(\overline{q}_1)^{\overline{q}_1} \), rather than \( P(\overline{Q})(\overline{Q} - \overline{q}_1) = P(\overline{Q})\overline{q}_2 \). Given one-on-one bargaining, a fraction, \( \beta' \), of these gains goes to \( D_1 \) and a fraction \( (1 - \beta') \) to \( U_2 \).

**Proposition 19:** Under the integration of \( U_1 \) and \( D_1 \) and exit by \( D_2 \), \( D_1 \) buys \( \overline{q}_1 \) from \( U_1 \) and \( \overline{q}_2 \) from \( U_2 \). Profits are

\[
U_1-D_1: V_{1}^{M_d} = P(\overline{q}_1)^{\overline{q}_1} + \beta'[P(\overline{Q})\overline{Q} - P(\overline{q}_1)^{\overline{q}_1}] - E
\]

\[
U_2: U_{2}^{M_d} = (1 - \beta') \{P(\overline{Q})\overline{Q} - P(\overline{q}_1)^{\overline{q}_1}\}
\]

\( D_2 \): Zero.

As in the case of scarce needs we suppose that \( U_1-D_1 \)'s profits are higher under ex ante monopolization than under nonintegration. That is:
(6) \[ V_1^{M_i} = P(q_1)q_1 + \beta'[P(Q)Q - P(q_1)q_1] - E \]
\[ > (1 - 2\beta)P(Q)q_1 + \beta P(Q)Q. \]

The right side of the equation is decreasing in \( \beta \) because \( Q < 2\overline{q}_1 \) and so reaches a maximum \( P(Q)\overline{q}_1 \) when \( \beta = 0 \). Hence the condition certainly holds if \( E \) is small enough. If the condition fails to hold, neither \( U_1 \) and \( D_1 \) nor \( U_2 \) and \( D_2 \) will ever have an incentive to integrate in the present model.

The Investment Decision

Let us reconsider the assumption that downstream firms invest. Under nonintegration, \( D_1 \) and \( D_2 \) cover their costs and invest as long as

(7) \[ \beta P(Q)Q > J. \]

We assume condition 7 in what follows.

Under full integration, it is not difficult to show that it will never pay \( D_i \) to exit for some \( i \). (Obviously, it would not pay \( D_1 \) and \( D_2 \) both to exit since then there would be no market.) In particular, \( U_i \) and \( D_i \) would do better not to merge at all if merger leads to \( D_i \)’s exit. The result of \( D_i \)’s exit would be that \( U_i \) would sell \( \overline{q}_i \) to \( D_j \), receiving a fraction \( (1 - \beta') \) of the surplus. \( U_i-D_i \)'s total profits would be \( (1 - \beta') [P(Q)Q - P(q_i)q_i] - E \), as opposed to \( P(Q)\overline{q}_i - E - J \) if \( D_i \) invests. Because \( P(q_j) \leq P(Q) \), \( D_i \)'s exit increases \( U_i-D_i \)'s profit only if \( J > \beta'P(Q)\overline{q}_i \). But in the latter case, \( D_i \) would exit if \( U_i \) and \( D_i \) were not integrated, given that \( U_j-D_j \) are integrated; and thus \( U_i \) would enjoy profit \( (1 - \beta') [P(Q)Q - P(q_i)q_i] > (1 - \beta') [P(Q)Q - P(q_j)q_j] - E \) by not merging with \( D_i \). Thus \( U_i \) would be better off refusing to merge with \( D_i \).

Consider finally partial integration, in particular where \( U_1 \) and \( D_1 \) merge but \( U_2 \) and \( D_2 \) stay separate (the logic in the reverse case is the same). Under these conditions \( D_2 \) may or may not invest. It is easily seen, however, that \( D_1 \) invests (if the \( U_1-D_1 \) merger is worthwhile at all). In particular, note that, by the same argument as in the full integration case, if \( D_1 \) exits, \( U_1-D_1 \)'s profit equals \( (1 - \beta') [P(Q)Q - P(q_2)q_2] - E \). But this is smaller than \( U_1 \)'s profit in the worst possible scenario if \( U_1 \) and \( D_1 \) do not integrate, \( (1 - \beta') [P(Q)Q - P(q_2)q_2] \), which occurs if \( U_2 \) and \( D_2 \) integrate and \( D_1 \) exits.
The Merger Game

Again suppose that a merger is irreversible and that if $U_i$ and $D_i$ merge, $U_j$ and $D_j$ can respond instantaneously by merging too. As in proposition 15, we suppose first that $U_i$ and $D_i$ merge if any merger occurs at all; we then check that $U_2-D_2$ will not preempt $U_1-D_1$. It is clear that the worst outcome for $U_1-D_1$ is if $U_2$ and $D_2$ decide to merge. The reason is that in this case $U_2$’s supplies are denied to $D_1$ but at the same time they are sold on the market and so depress output price. Hence if $U_1-D_1$’s profits rise because of the merger even in this case, $U_1$ and $D_1$ will certainly merge: doing so is a dominant strategy. From propositions 16 and 17, we conclude that if

$$
\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) > E,
$$

$U_1$ and $D_1$ certainly merge. On the other hand, if

$$
\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) < E,
$$

the $U_1-D_1$ merger will depend on the response of $U_2$ and $D_2$. Proposition 20, which is proved in appendix F, provides a full characterization of the different cases. Let $X = P(\bar{Q})\bar{q}_2 - J - E$ and $Y = (1 - \beta')[P(\bar{Q})\bar{q}_1 - P(\bar{q}_1)\bar{q}_1]$.

**Proposition 20:** Suppose $U_1$ and $D_1$ decide first whether to merge, and if and only if they merge, $U_2$ and $D_2$ can respond by merging too. Then:

1. If $\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) > E$ and $\beta P(\bar{Q})\bar{q}_2 > E$, then $U_1$ and $D_1$ will merge and
   (a) $U_2$ and $D_2$ will also merge if $X > Y$ (reluctant bandwagon).
   (b) $D_2$ will exit if $X < Y$.

2. If $\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) > E$ and $\beta P(\bar{Q})\bar{q}_2 < E$, then $U_1$ and $D_1$ will merge and
   (a) $U_2$ and $D_2$ will stay independent, with $D_2$ investing if $\beta P(\bar{Q})\bar{q}_2 > J$.
   (b) $D_2$ will exit if $\beta P(\bar{Q})\bar{q}_2 < J$ and $X < Y$.
   (c) $U_2$ and $D_2$ will merge if $\beta P(\bar{Q})\bar{q}_2 < J$ and $X > Y$ (forced bandwagon).

3. If $\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) < E$,
   (a) $U_1$ and $D_1$ will merge and $D_2$ will exit if $\beta P(\bar{Q})\bar{q}_2 < J$ and $X < Y$. 

(b) $U_1$ and $D_1$ will merge, $U_2$ and $D_2$ will stay separate, and $D_2$ will not exit if $\beta P(\overline{Q})\overline{q}_2 > J$, $\beta P(\overline{Q})\overline{q}_2 < E$, and $\beta P(\overline{Q})\overline{q}_1 > E$.

(c) No merger will occur if $\beta P(\overline{Q})\overline{q}_2 < J$ and $X > Y$, or $\beta P(\overline{Q})\overline{q}_2 > J$ and $\beta P(\overline{Q})\overline{q}_2 > E$, or $\beta P(\overline{Q})\overline{q}_2 > J$, $\beta P(\overline{Q})\overline{q}_2 < E$, and $\beta P(\overline{Q})\overline{q}_1 < E$.

Note that a $U_1$-$D_1$ merger will certainly occur if $\overline{q}_1$ is very large relative to $\overline{q}_2$, that is, if $q_1 = \overline{Q}$, $\overline{q}_2 = 0$. This is because condition 6 implies that $\beta P(\overline{Q}) = \beta P(\overline{Q}) (\overline{q}_1 - \overline{q}_2) > E$. However, a $U_1$-$D_1$ merger can also occur even if $\overline{q}_1$ and $\overline{q}_2$ are quite close if the shift in surplus away from $D_2$ is just enough to cause $D_2$’s profits to fall below $J$ and lead to $D_2$’s exit—for example, consider 3(a) in proposition 20 and suppose $\beta P(\overline{Q})\overline{q}_2 = J$, $P(\overline{Q}) = P(\overline{q}_1)$, and $\beta'$ is very small.

Eager bandwagon is never an outcome in this model. $U_2$ and $D_2$ are never better off under full integration than under nonintegration; this follows because condition 8 cannot hold when $\overline{q}_1$ and $\overline{q}_2$ are interchanged. However, reluctant bandwagon occurs in 1(a) and the forced bandwagon in 2(c) of proposition 20. (The characterization of the different types of bandwagon is not contained in the proof of proposition 20 but is left to the reader.)

So far we have assumed that $U_1$ and $D_1$ move to merge first. Might $U_2$ and $D_2$ want to preempt a $U_1$-$D_1$ merger? Clearly there is no advantage to preemption if $U_1$ and $D_1$ decide to merge anyway; $U_2$ and $D_2$ would do better to let $U_1$ and $D_1$ merge first and then select a best response. This means that preemption is useless in cases 1 and 2 of proposition 20 because a $U_1$-$D_1$ merger is a dominant strategy. In case 3(c) preemption is unnecessary because no merger occurs anyway. This leaves 3(a) and 3(b). Case 3(b) implies that $\beta P(\overline{Q})\overline{q}_1 > J$, that is, $D_1$ does not exit if $U_2$ and $D_2$ merge; moreover, $\beta P(\overline{Q})\overline{q}_1 > E$, so $U_1$ and $D_1$ will jump on the bandwagon. Hence preemption does not prevent merger here. This leaves case 3(a). It is easy to check that in the continuous-time preemption game described in the discussion of the model, $U_1$ and $D_1$ have more incentive to integrate, and their merger preempts $U_2$-$D_2$, except possibly in the following subcase: if $\beta P(\overline{Q})\overline{q}_1 < J$ ($D_1$ exits if $U_2$ and $D_2$ merge and $U_1$ does not rescue $D_1$), and $P(\overline{Q})\overline{q}_1 - J - E < (1 - \beta') [P(\overline{Q})\overline{q} - P(\overline{q}_1)\overline{q}_1]$ ($U_1$ does not rescue $D_1$), the incentives for $U_1$-$D_1$ to preempt $U_2$-$D_2$ and for $U_2$-$D_2$ to preempt $U_1$-$D_1$ are equal. Whoever preempts the other, the nonintegrated down-
stream firm exits, and preemption is a zero-sum game—what one gains, the other loses. Preemption then occurs at the date at which each is indifferent between preemption or not preemption.\textsuperscript{26}

Finally, in contrast to the earlier ex post monopolization scenario, there are no ‘‘public good’’ aspects to mergers here: the nonmerging downstream firm suffers from lack of supplies, and the nonmerging upstream firm may suffer from the exit of its downstream partner. Neither pair, \(U_1\) and \(D_1\) nor \(U_2\) and \(D_2\), ever wants the other pair to move first, and there cannot be a war of attrition.

\textbf{Remark.} To keep the variant relatively simple, we have ignored upstream investments. An implication of this is that vertical mergers have no effect on consumers: in all the subcases of proposition 20, \(Q\) units are supplied to consumers and price is \(P(Q)\). Allowing upstream investments would not alter the first-round effects of a \(U_1-D_1\) merger because it has no effect on \(U_2\)’s profits. However, if \(D_2\) exits as a result of the merger, this \textit{will} reduce \(U_2\)’s profits and might cause \(U_2\) to exit. In other words, a sequence of exits is a possible outcome when upstream and downstream firms both invest. Under these conditions, supplies will disappear from the market and consumer prices will rise.

There is another new possibility that arises when upstream firms invest. Whereas \(U_1-D_1\) always benefits from \(D_2\)’s exit (this increases \(D_1\)’s monopsony power), \(U_1-D_1\) may suffer from \(U_2\)’s exit because scarce supplies disappear from the market. Hence in some cases \(U_1\) and \(D_1\) may refrain from merging in order to keep \(U_2\) alive.\textsuperscript{27}

\textbf{Welfare}

The welfare effects of a merger are straightforward in the scarce supplies variant. Since, in the absence of upstream investments, total

\textsuperscript{26} U_1-D_1 and U_2-D_2 then have equal probabilities of preempting: see Fudenberg and Tirole (1985) for the formalization of the continuous-time preemption strategies. The date \((-T)\) at which preemption occurs is given by \(E = e^{-\gamma}(P(q_1)\bar{q}_1 + \beta' P(\bar{Q}) - P(q_1)\bar{q}_1) = (1 - \beta') [P(\bar{Q})Q - P(q_1)\bar{q}_1],\) where \(E\) is now taken to be a stock rather than a flow.

\textsuperscript{27} One case in which \(U_1-D_1\) will barely be hurt by \(U_2\)’s exit is when \(P(q_1)\bar{q}_1 \approx P(\bar{Q})\bar{Q}\). This is because even if \(D_2\) and \(U_2\) exit, \(U_1-D_1\) achieves \(P(q_1)\bar{q}_1 - E\), and this is almost as much as is received if only \(D_1\) exits (\(V^{\text{U-M}}\)). Hence for this case the presence of upstream investments will not change the analysis at all. Moreover, if \(\beta P(\bar{Q})\bar{q}_2 < J\) and \((1 - \beta') [P(\bar{Q})Q - P(q_1)\bar{q}_1] < I,\) that is if \(D_2\) and \(U_2\) both exit, there will be a clear effect on consumers from the \(U_1-D_1\) merger: output will fall from \(\bar{Q}\) to \(\bar{q}_1\), and price will rise from \(P(\bar{Q})\) to \(P(q_1)\).
output is always $Q$, consumers neither gain nor lose from mergers. Firms lose in the aggregate to the extent that merger costs are incurred, but gain to the extent that investment costs $J$ are saved (for example, if $U_1$ and $D_1$ merge and $D_2$ exits, the net gain is $J - E$). Since, under partial or full integration, merger costs are incurred without investment costs being saved, these cases are always dominated by nonintegration.

Once upstream investments are allowed, consumers will generally be affected by mergers. In particular, under the maintained hypothesis that all firms invest under nonintegration, a $U_1-D_1$ merger that leads to the exit of both $D_2$ and $U_2$ will cause a fall in total supply from $Q$ to $q_1$, and a corresponding price rise from $P(Q)$ to $P(q_1)$.

**Extensions**

We mention two brief possible extensions of the model. First, our analysis is couched in terms of integration between a supplier and a buyer. However, the ex post monopolization variant seems likely to extend to integration between two manufacturers of complementary products. Suppose manufacturer $A_1$ merges with $B_1$. By doing this, $A_1$ makes it credible that it will give information about developments of its products only to $B_1$, thus allowing $B_1$ an early start in the design of compatible complements.

Suppose first that $A_1$ is a monopolist in the $X$ market (this situation is analogous to the essential facility case). Two firms, $B_1$ and $B_2$, produce goods $Y_1$ and $Y_2$ that are complements to $X$. An unintegrated $A_1$ has an incentive to provide both $Y$ manufacturers with information about its product developments in order to create low costs and competition in the $Y$ market and consequently to be able to charge a high price for good $X$. But total industry profit (from goods $X$ and $Y$) can often be raised by raising prices in the $Y$ market. For instance, if $Y_1$ and $Y_2$ are good substitutes, the prices in the $Y$ market under Bertrand competition are close to marginal cost. However, if consumers are heterogeneous and have different demands in the $Y$ market, optimal second-degree price discrimination requires prices well above marginal cost. Another reason why higher prices in the $Y$ market might increase

---

28. See, for example, Tirole (1988, pp. 145–47).
industry profits is that some consumers may want to consume the $Y$ good only. In either case, integration of $A_1$ and $B_1$ credibly commits $A_1$ not to give early information to $B_2$ if $Y_1$ and $Y_2$ are good substitutes. This enables $B_1$ to raise its price.

Second, assume that $A_1$, which produces $X_1$, faces competition by $A_2$, which produces $X_2$. The rival $A_2$ will be indirectly hurt by $A_1$-$B_1$’s integration even though it may not need any information from $B_1$. On the one hand, the increase in the price of $Y_1$ reduces the number of consumers who want to mix and match $X_2$ and $Y_1$. On the other hand, if $B_2$ exits, there are no more consumers who want to mix and match $X_2$ and $Y_2$. Several outcomes may result. $A_2$ may exit; or it may be forced to bandwagon by coming to $B_2$’s rescue.

As very tentative illustrations (tentative because we have not studied the industries in detail), consider IBM’s limiting early announcements of its developments in computer technology to its disk drive subsidiary or airlines’ that offer complementary flights and merge to gain market power by facilitating exclusive coordination of schedules at hubs.

We have also assumed that the upstream firms are subject either to constant returns to scale (first two variants) or to decreasing returns to scale (third variant). An interesting extension of the model would allow for upstream increasing returns to scale over some range, as in the case of a U-shaped cost curve. A (possibly hypothetical) illustration is the following: by buying supercomputers exclusively from Japanese manufacturers (as a result of vertical integration, for example) Japanese owners of supercomputers reduce the size of the market for U.S. supercomputer manufacturers, whose unit production costs therefore rise. As a consequence, U.S. consumers of supercomputers forgo some use of them and hence are at a disadvantage relative to their Japanese competitors in the product market. This illustration is similar to our ex post monopolization variant, except that vertical integration not only enables the most efficient supplier, which is ex post the Japanese manufacturers of supercomputers, to commit to restrict supplies to U.S. consumers of supercomputers, but also creates the upstream cost differential that was assumed exogenous in the discussion of ex post monopolization. The illustration also possesses some features of our scarce needs variant.29

29. In that variant a merger between an upstream and a downstream firm could disadvantage the rival downstream firm by causing exit of the rival upstream firm. This is an extreme example of an increase in the upstream firm’s unit production costs.
Applications

This section applies the model to three industries. The discussion is only meant to suggest how one might analyze these industries using the framework just presented. Of course, the evidence on vertical integration in these industries was not collected with this kind of model in mind.

Case 1: The Cement and Ready-Mixed Concrete Industries

The cement industry consists of kilns and mills that convert limestone, clay, and gypsum into cement. The ready-mixed concrete industry combines cement, sand, aggregates, and water to make concrete. In the early 1960s a great deal of vertical integration occurred between the cement and the ready-mixed industries. In particular, a large number of cement companies integrated forward by acquiring ready-mixed concrete companies. This heightened merger activity attracted the attention of the Federal Trade Commission, and it conducted an inquiry resulting in the Economic Report on Mergers and Vertical Integration in the Cement Industry in 1966.

Characteristics of the Cement and Concrete Industries. Cement is a very homogeneous commodity; it is manufactured to strict specifications, there are no problems of customer-specific investment, and any ready-mixed concrete manufacturer can easily turn to an alternative supplier.

Because of large minimum efficient scale, concentration in the cement industry was very high in the 1960s. Since cement is bulky and costly to transport, 90 percent of it was shipped 160 miles or less. And even at the state level, which may be larger than actual market areas, in only 6 percent of the states did the four largest suppliers account for less than 50 percent of cement shipments.

Concentration in the ready-mixed concrete industry was apparently lower; however, the industry consisted of a few large firms handling large contracting jobs such as highways and bridges and many small firms handling smaller jobs. As a result, in 17 of 22 metropolitan areas for which the FTC had data, the 4 leading ready-mixed companies accounted for 50 percent or more of ready-mixed sales. In eight of

these markets the four largest companies accounted for 75 percent or more.31

The period immediately after World War II saw a steady growth in demand for cement with no corresponding increase in capacity. As a result, by 1955 cement mills were operating at 94 percent of capacity.32 In response, existing cement mills were expanded and new mills constructed so that by 1960 the capacity utilization rate was down to 74 percent.

The merger wave seems to have been triggered by significant excess capacity among cement mills. From 1955 to 1965 the cement industry expanded capacity by 60 percent—twice as fast as actual shipments of cement grew during the decade.33 This burst in cement mill construction and expansion was a response to high-capacity utilization levels in the early 1950s, which resulted in spot shortages of cement. Demand continued to grow throughout the 1960s, but because so much new capacity was brought on line, cement manufacturers saw their excess capacity cut into industry profits. Eighty percent of the vertical acquisitions occurred when market conditions were weak, and 37 of 55 took place in markets with above-average excess capacity.34 The overcapacity was also aided by technological change that made newer cement mills cheaper to operate and made it feasible to build larger plants. By modernizing to cut costs, cement makers contributed to the industrywide overcapacity. Neither demand conditions nor innovations in the concrete market seem to have played an important role in triggering mergers.

Pattern of Integration. The 1960s witnessed a wave of acquisitions of concrete manufacturers by cement producers. The acquired ready-mixed companies made between 19 percent and 45 percent of total sales in their respective market areas.35

It is generally agreed that each acquiring cement producer hoped to assure itself of guaranteed outlets.36 Efficiency reasons do not seem to have been an important factor.37

Bandwagoning occurred in many markets. All the executives’ comments point to the fact that many companies had been driven to purchase their customers because their competitors were doing likewise. For example, in its Annual Report of 1963, the Alpha Portland Cement Company stated, “Vertical integration within our industry has been on the increase in recent years. Alpha is presently not inclined to integrate vertically. However, if our position in the industry is put in jeopardy as a result of such corporate arrangements, there will be no alternative but to make similar moves.”

Wilk (1968) also cites evidence that many cement firms dropped out of a market after a large customer had been bought out by competing cement manufacturers.

**Link with Analysis.** The pattern of integration in the industry suggests that the relevant variant is the scarce needs one (see in particular the extension of the scarce needs model in which downstream firms have limited capacity). Upstream firms were eager to assure themselves of a downstream outlet. The bottleneck seems to have been the downstream industry.

Also consistent with the scarce needs variant are that the complaining firms were cement producers and that the mergers affected the largest ready-mixed concrete firms.

One prediction of the scarce needs model is not borne out by the facts. Although the ready-mixed companies that had been acquired increased from 37 percent to 69 percent the fraction of their supply obtained from the acquiring cement companies after the mergers, as the theory would predict, they still purchased from other cement suppliers. The scarce needs variant has all supplies produced by the internal manufacturer. This particular prediction, however, relies on constant returns to scale upstream; and although there was excess capacity in the cement industry, there may have been capacity constraints

39. Although the scarce supplies variant is clearly ruled out by the existence of excess capacity in the cement industry, the evidence against ex post monopolization is more circumstantial and consists mainly of the fact that upstream firms were the ones that complained. To get more factual evidence against ex post monopolization, one would have to show that there were only small differences in marginal costs among the upstream firms, or at least that the upstream firms that first merged were not the most efficient ones.
for some individual cement producers. The theory of scarce needs could be modified by increasing the number of upstream firms and allowing for individual but not industry capacity constraints to account for the possibility of outside supplies.

Based on the executives’ interviews and annual reports, the relevant bandwagoning behavior seems to have been reluctant.41

Why did integration take place in the 1960s and not earlier? A primary determinant of the merger activity was the excess capacity in the cement industry that appeared then. Before this wave of forward integration, there were some instances of backward integration into cement manufacture by concrete makers. Typically, a large concrete maker would build a modern cement mill from scratch and use most of the cement to meet its own needs. These backward moves were initiated during the late 1950s, when cement was very profitable because of the limited capacity in the industry. Concrete makers’ profits were squeezed by the high price of cement and the highly competitive nature of the concrete business, which held prices down. That is, the relevant model for the late 1950s may have been the scarce supplies variant. However, the gains from foreclosure seem to have been smaller than in the 1960s.

Finally, it would be interesting to know whether the Federal Trade Commission and the various commentators, in dismissing efficiency reasons for mergers, recognized the possibility of holdup problems in the cement industry. It is possible that at a time of excess capacity, a number of cement producers were no longer viable; they would have exited if they could not have combined with a concrete firm. This would provide an efficiency motive for mergers, which might offset the foreclosure effects emphasized here. More information is required to tell whether this efficiency effect could have been large. As noted in the introduction, however, the fact that the mergers involved large cement and concrete firms provides some support for foreclosure as the relevant effect.

**Case 2: Computer Reservation Systems**

Computer reservation systems (CRS) book airline seats electronically. The CRS industry was vertically integrated with airlines from its inception, and the two largest systems are Sabre, owned by American

Airlines, and Apollo, controlled by United Airlines. TWA, Texas Air, and Delta have competing CRS. Although the systems typically listed flights of airlines other than the ones that controlled them, by 1984 there were widespread complaints that they were biased in favor of the host airlines, neutral vis-à-vis the airlines that did not compete with the hosts, and biased against airlines that did compete with the hosts. The bias was partly monetary. In 1981–82 American charged Eastern Airlines $0.24 for each booking on Sabre, Delta Airlines $1.32, and New York Air $2.00. Eastern was a large carrier that did not compete fiercely with American. It was charged a low rate to give Sabre wider coverage, making the CRS more attractive to travel agents. Delta competed with American at its Dallas hub, and there is evidence that American wanted to drive Delta out of Dallas–Fort Worth. New York Air was a price cutter. Another important element of discrimination concerned the order the flights were displayed on the travel agent’s screen. This order is crucial because agents have little time or willingness to screen through several displays. Being listed near the top provides a major competitive advantage for an airline.

In 1984 eleven airlines that were not integrated into the systems filed an antitrust suit against American and United, charging them with monopolization of CRS. In November 1984 the Civil Aeronautics Board established regulations to guarantee more equal access.

ANALYSIS. One way of looking at the industry is to regard the CRS as an upstream firm with, possibly, scarce supplies. The system supplies an input (flight booking) to downstream firms, the airlines, which set prices for flights. For simplicity, we will use the paradigm of an upstream monopolist (an essential facility) serving several downstream competitors. Clearly there is competition among computer reservation systems, but this competition is imperfect. Furthermore, a travel agent usually consults a single CRS when serving a customer.

What are the efficiency gains of vertical integration? They do not seem substantial, but they may exist, and further research is needed to see whether this is the case. The integrated CRS and airline can derive

43. It is sometimes argued that computer interconnections between the CRS and the airlines can be improved through vertical integration; it is unclear, however, why the same coordination could not be achieved under nonintegration via a contract.
three other types of benefits. First, the host airline may favor its own flights by biasing display in their favor; this gives rise to an ex post monopolization effect. Second, the host airline may acquire real-time access to all prices and seat availability and thus get an edge over its competitors. The implications of this effect are less clear than those of the first, but they relate to an ex post competitive advantage as well. Third, the integrated CRS will give priority to the host airline and thus does not leave bargaining rents to other airlines.

How do the first and third gains fit in the model? To take an extreme example, suppose there is a single CRS and two airlines. Assume first that there are two priority lines on the screen allowing the CRS to display two flights (other lines require another display for the travel agent and do not sell in this extreme case). Assume also that priority is not contractible. A customer’s preferred departing time to go from city A to city B is noon, and the two airlines each have such a flight. A nonintegrated CRS will list the two flights. (The CRS is actually indifferent between doing this and listing two flights of the same airline because it does not receive compensation for priority, but it is reasonable to assume that it displays the noon flights of the two airlines if it receives some small benefit from pleasing travel agents or helping both airlines stay alive.) Knowing this, the two airlines will compete fiercely in the price of their noon flight. But if the first airline and the CRS merge, the CRS will show this airline’s noon and 2:00 p.m. flights and will relegate the other airline’s noon flight to a lower, nonselling ranking. Facing less competition, the first airline can raise its price on the noon flight, and the customers as well as the rival airline are hurt.

This is an example of ex post monopolization. Take now another extreme case in which there is a single priority line on the screen (all other lines are not conspicuous enough to sell), and priority can be contracted between an airline and the CRS. The situation in which the unintegrated CRS is unable to commit to give priority to a single airline disappears. This then is the scarce supplies variant. An unintegrated CRS leaves some bargaining gains to each airline when selling the scarce supply; one airline’s gain can be recaptured if the CRS vertically integrates with the other airline.

The assumptions underlying ex post monopolization and scarce supplies here seem inconsistent. However, reality is a mixture of the two situations. First, priority was partly contractible before 1984. The or-
dering of display was computed through a complex system of penalties, one for not being the host airline, another depending on the difference between the actual flight departure time and a customer’s desired departure time, a third for stops and connecting flights, and so forth. Airlines could reduce the level of nonhost penalty by becoming cohosts. However they could not fully contract on priority because the CRS could often make minor adjustments to its algorithm to decide which connections were listed, change the algorithm when introducing new flights, issue boarding passes only for the host airline, shave schedule times, break ties in favor of airlines who have certain flight numbers, and so forth. Thus priority had both contractible and noncontractible elements. Second, whether the supply of screen space for relevant flights is scarce depends on the route, the time of day, the season, and so forth. Thus one would expect space on the screen sometimes to be scarce, as in the one-line example, and sometimes not, as in the two-line example.\footnote{The contracting difficulties may also offer clues as to why the vertically integrated outcome could not have been achieved through an exclusive-dealing contract between the CRS and the airline. After all, discriminatory rates and penalties resemble partial exclusive dealing. One issue with exclusive dealing is that ideally an independent CRS would have liked to give a low penalty level to an airline \textit{together} with the commitment to impose high penalty levels to rival airlines. Such an exclusionary practice, like other forms of exclusive-dealing contracts, would probably have been frowned on by the courts. Another issue is that display bias is only partially contractible, so that some of the private gains to exclusionary behavior are best realized through vertical integration. And indeed, only one short-lived attempt to compete was made by a CRS not owned by an airline, which suggests that integrated CRS yielded more profits.}

\textit{Case 3: Terminal Railroad Case}

Terminal Railroad is the quintessential example of an essential facility:

The Terminal Company controlled a bridge across the Mississippi River, and the approaches and terminal at St. Louis, a very significant junction point for competing railroads. That company had every incentive to serve equally all railroads entering or leaving St. Louis, charging whatever the market or regulatory agencies would bear. However, once the Terminal Company was acquired by several of those railroads, the new owners might use their control over it to exclude or prejudice their rivals. Rather than order dissolution of the combination, with restoration of the Ter-
minal Company’s independence, the Supreme Court required the members to admit their railroad competitors to their consortium. Although the Court did not use the word, we might describe the Terminal Company’s bridge, tracks, and terminals as “essential facilities” that had to be shared with competitors.  

One can view the Terminal Company as an upstream monopolist and the competing railroads as downstream rivals. Note that strategic vertical integration by an upstream essential facility cannot be driven by scarce needs downstream. Because there is a single supplier, integration of a $U$ and a $D$ appropriates no bargaining surplus from other suppliers. Thus, absent efficiency gains, forward integration by an upstream monopolist may be driven either by the ex post monopolization effect or by the scarce supplies effect.

Scarce supplies seemed to play no role in this case. According to Areeda and Hovenkamp (1987, ¶736.1b), the Terminal Company’s “minimum efficient scale could accommodate all the traffic.” Although there is little evidence, efficiency considerations also seemed secondary. Furthermore, if there had been efficiency gains from vertical integration, one would have to explain why these gains would not also have applied to the excluded railroads, in which case joint ownership of the Terminal Company by all the railroads would have been optimal. Thus a first look at the Terminal Railroad case suggests that the motive for integration was to monopolize the rail market around St. Louis.

**Review of the Literature**

This section compares our analysis with those in the literature on vertical integration and foreclosure, in particular the contributions of Ordover, Saloner, and Salop (1990); Salinger (1988); and Bolton and Whinston (1989).

The model presented by Ordover, Saloner, and Salop is, in effect, a special case of our first variant in which $c_1 = c_2$. In contrast to our analysis, they find that vertical integration can be profitable under these

46. See the discussion in Hart and Moore (1988, section 4.4).
conditions. The authors argue that, under nonintegration, price competition in the intermediate and output markets leads to the standard Bertrand product-market outcome. If upstream firm $U_1$ and downstream firm $D_1$ merge, $U_2$ can raise its input price to $D_2$ because $U_1$ will no longer be as anxious to supply the rival downstream firm $D_2$ as before. This gives $D_2$ a disadvantage as a competitor in the product market and allows $U_1-D_1$ to increase market share and make positive profit.\(^{47}\) In other words, vertical integration forecloses product-market competition by ‘raising rivals’ costs.’\(^{48}\)

The authors’ analysis makes implicit assumptions about commitment and contracting possibilities that are questionable. They assume that when $U_1$ and $D_1$ merge they can commit not to supply rival $D_2$ at a price below $\bar{p}$, where $\bar{p}$ is a choice variable for $U_1$ and $D_1$. Then $U_2$ and $D_2$ decide whether to merge. The authors show that $U_1$ and $D_1$ commit to a price $\bar{p}$ above marginal cost $c$. In equilibrium, $U_2$ slightly undercut $\bar{p}$ to $\bar{p} - \epsilon$ and supplies $D_2$. Thus $U_1-D_1$ has succeeded in raising $D_2$’s marginal cost. However, $\bar{p}$ cannot be too large because the shrinking of $D_2$’s market share would induce $U_2$ and $D_2$ to merge as well.

There are two problems with this reasoning. First, if two-part tariffs are allowed, as in our analysis, $U_2$ and $D_2$ always have an incentive to transfer the intermediate good at marginal cost and bargain over a fixed fee. Thus in the presence of two-part tariffs, $U_1-D_1$ cannot affect $D_2$’s marginal cost and hence market competition. Second, the commitment of $U_1-D_1$ is unlikely to be believable. Why would $U_1-D_1$ not undercut $U_2$ by $\epsilon$ in turn? The effect on $D_2$’s reaction curve is negligible (of the order of $\epsilon$), while $U_1-D_1$’s increased profit from supplying $D_2$ is significant (it is approximately $(\bar{p} - c)q$, where $q$ is the quantity $U_2$ sells to $D_2$). Thus $U_1-D_1$ can gain from such a deviation ex post, and any commitment ex ante not to make such a deviation lacks credibility. This is in spite of the fact that competitive undercutting of this type leads inexorably to the Bertrand outcome and thus eliminates all the benefits from the integration of $U_1$ and $D_1$.

We are not suggesting that it is never feasible for an upstream firm to commit to charge high prices to a downstream firm. One way this

\(^{47}\) In fact, because competition between $U_1$ and $U_2$ becomes less fierce, the nonintegrated upstream firm $U_2$ also benefits from the merger (makes a profit) in equilibrium.\(^{48}\) Salop and Scheffman (1983).
could be achieved is via a form of exclusive-dealing contract (see appendix C); another is through reputation. What is unclear from Ordover, Saloner, and Salop, however, is the mechanism for enforcing commitments and why $U_1$ and $D_1$ need to merge to take advantage of this mechanism. That is, if exclusive-dealing contracts are feasible, why cannot $U_1$ write such a contract with $D_1$ to restrict supplies to $D_2$ while remaining independent?\textsuperscript{49}

The authors also obtain different conclusions from ours. Our model explains why firms sometimes respond to a merger by themselves merging, how it can be profitable for an integrated upstream firm to sell to a rival downstream firm, and why an upstream firm and downstream firm may merge to drive a rival out of the market. In contrast, bandwagoning never occurs in their model (at most one pair of firms is integrated). Integrated and nonintegrated firms never trade with each other and, because a nonintegrated upstream firm benefits from its rival’s integration, an upstream firm might refrain from integration in order to monopolize the market ex ante (in the presence of investment costs). Finally our model yields predictions on which firms are more likely to integrate (those with lower marginal costs, lower investment costs, or higher capacities), whereas Ordover, Saloner, and Salop are silent on this because they consider identical firms.\textsuperscript{50}

\textsuperscript{49} Several papers have in fact studied the use of exclusive-dealing contracts to foreclose markets. See Comanor and Frech (1985), Mathewson and Winter (1986), and Schwartz (1987). These papers, however, put restrictions on the types of nonexclusive-dealing contracts that can be offered. Also see Krattenmaker and Salop (1986) for a very good discussion of the law and economics of exclusive dealing.

\textsuperscript{50} Salinger’s (1988) model is similar to that of Ordover, Saloner, and Salop in several respects. He makes the same technological assumptions they do but assumes that a large number of upstream and downstream firms interact in an anonymous market. The downstream firms take the price of the intermediate good as given in their input decisions, but act as Cournot oligopolists in the consumer-good market. The upstream firms in turn act as Cournot oligopolists in the intermediate-good market. Salinger argues that, if $U$ and $D$ merge, $U$ no longer supplies the intermediate good to the anonymous market, preferring instead to channel it to $D$. Similarly, $D$ no longer purchases input in the anonymous market, preferring instead to be supplied by $U$. A strategy that Salinger’s upstream Cournot assumption does not permit is for an integrated supplier to undercut its nonintegrated rivals slightly, so that nonintegrated purchasers buy the same total amount as before but now buy from the integrated supplier. Yet a price-cutting strategy seems natural, particularly in the context of many trading relationships between upstream and downstream firms that are personalized rather than anonymous, and where price setting, possibly in conjunction with quantity setting, seems more plausible than pure quantity setting.
A recent paper by Bolton and Whinston (1989), written independently of ours, studies the motives for vertical integration from the perspective of incomplete contracting, but mainly in a situation in which downstream firms operate in different product markets. The authors' basic model consists of two downstream firms, $D_1$ and $D_2$, and one upstream firm, $U$. The downstream firms make variable investments specific to the upstream firm, but the upstream firm does not invest. Each downstream firm requires one unit of intermediate good from the upstream firm ex post; the upstream firm can satisfy both downstream firms in some states of the world, but in others it has only one unit of intermediate good available. Long-term contracts cannot be written, and ex post bargaining is modeled as an extensive-form game in which the ability of the upstream firm to sell to $D_j$ plays the role of an outside option in the bargaining between it and $D_i$. In contrast to our model, investment costs are not shared under integration and the returns to investment are completely appropriated by a firm's owner.

When the upstream firm has only one unit of intermediate good available, the authors' model is close to our scarce supplies variant. The motive for integration is different, however. If $D_1$ buys $U$, this has no direct effect on $D_2$'s investment decision because, assuming the outside option binds, if $D_2$ values the intermediate good more than $D_1$ does, $D_2$ will continue to buy it at a price equal to that $D_1$ is willing to pay. However, there is an indirect effect because $D_1$ now appropriates all the returns from $U$'s bargaining with $D_2$ and so has an incentive to invest more to increase these returns. This in turn causes $D_2$ to invest less.\(^{51}\)

Given that the motive for integration is different in their model, it is not surprising that Bolton and Whinston also reach different conclusions. They find that when outside options are binding in the bargaining process, nonintegration is socially optimal. The reason is that because each downstream firm pays an input price determined by the other downstream firm's willingness to pay, it receives at the margin the full increase in the marginal product of its investment. (In contrast, in the discussion of scarce supplies we find that either nonintegration or vertical integration and exit can be socially optimal.) However, when

---

51. Bolton and Whinston also consider a form of bandwagoning, whereby a merger of $U$ and $D_1$ causes $D_2$ to build upstream capacity so as to supply its internal needs.
outside options are binding, nonintegration is not privately optimal in their model: by integrating, $U$ and one of the downstream firms can make themselves better off at the expense of the other downstream firm. In fact, Bolton and Whinston find that the only privately optimal arrangements involve vertical integration by $U$ and one of the $D$’s, or complete integration of $U$, $D_1$, and $D_2$. In contrast, we do not allow complete integration, and find that when $q_2 = 0$ (that is, when there is only one upstream firm), either nonintegration or integration between $U$ and $D_i$, with or without exit of $D_j$, can be privately optimal.

A final difference between the two models is that in Bolton and Whinston consumer surplus is independent of ownership structure (for example, if downstream firms make take-it-or-leave-it offers to consumers, consumer surplus equals zero). In our scarce supplies variant, exit by a downstream firm can lead to exit by an upstream firm, and thereby to a decrease in total supplies and a decrease in consumer surplus.

Appendix A: Proofs of Propositions 1–4

Proof of Proposition 1. The strategies are: $U_1$ offers to sell $q^*$ units at price $t^*$ to each $D_j$ (formally: $t_{1j}(q_{1j}) = t^*$ if $q_{1j} = q^*$, $= + \infty$ if $q_{1j} \neq q^*$). $U_2$ offers to supply each $D_j$ at marginal cost, that is, $t_{2j}(q_{2j}) = c_2q_{2j}$ for $j = 1, 2$. Each downstream firm accepts $(t^*, q^*)$ in equilibrium. If one of the upstream firms offers another contract to $D$, this $D$ continues to anticipate output $q^*$ by its rival and maximizes its profit: that is, it maximizes $r(q_{1j} + q_{2j}, q^*) - t_{1j}(q_{1j}, q_{1j}) - t_{2j}(q_{2j}, q_{2j})$ subject to $q_{1j} + q_{2j} \leq \max(q_{1j}, q_{2j})$. A downstream firm’s behavior is obviously optimal given the offers it faces and given that it expects its rival to purchase $q^*$.

Can $U_2$ deviate and make a positive profit? For instance, can it sell $q_{22}$ at price $t_{22} > c_2q_{22}$ to $D_2$? Note that $D_2$ can guarantee itself $D^{NL}(c_1, c_2) = r[R_2(q^*), q^*] - c_2R_2(q^*)$ by refusing $U_2$’s offer and purchasing $q^*$ at price $t^*$. Because $R_2(q^*)$ is the best response to $q^*$ for marginal cost $c_2$, firm $D_2$ would get strictly less than $D^{NL}(c_1, c_2)$ by buying $q_{22}$ at price $t_{22} > c_2q_{22}$ and rejecting offer $(q^*, t^*)$ from $U_1$. Similarly, because $R_2(q^*) \leq q^*$ (as $c_2 \geq c_1$), $q_{22} = 0$ maximizes $r(q^* + q_{22}, \ldots)$.
Last, can $U_1$ increase its profits? No, because it is already maximizing $t_{ij} - c_1 q_{ij}$, subject to the constraint $r(q_{ij}, q^*) - t_{ij} \geq r[R_2(q^*), q^*] - c_2 R_2(q^*)$ over pairs $(q_{ij}, t_{ij})$. Thus it extracts the maximum feasible surplus from each $D_j$, given that the latter can buy at marginal cost $c_2$ and expects its rival to buy $q^*$.

**Proof of Proposition 2.** The strategies are: $U_1$ offers $q^*_2$ at price $t_2^*$ to $D_2$. $U_2$ offers to supply at marginal cost: $t_2(j, q_2j) = c_2q_2j$ for all $j$ and $q_2j$. In equilibrium, $D_2$ buys $q^*_2$ from $U_1$ and 0 from $U_2$. Again, it is clear that $D_2$ acts optimally given the contract offers and the anticipation that $D_1$ procures $q^*_1$ internally.

Can $U_2$ make a strictly positive profit? Suppose that $U_2$ makes a different offer and $D_2$ buys $q^*_2$ at price $t_{22} > c_2q_{22}$ from $U_2$. Then $D_2$’s profit is $\max[r(q^*_2 + q_{22}, q^*_1) - c_2 q^*_2 - t_{22}, r(q_{22}, q^*_1) - t_{22}]$. Because $q^*_2 = R_2(q^*_1)$ and $t_{22} > c_2q_{22}$, this profit is strictly lower than $\Pi(c_1, c_2)$, and $D_2$ is better off rejecting $U_2$’s contract after all.

Can $U_1-D_1$ make more than $\Pi(c_1, c_2)$? Suppose that $U_1$ offers a different contract to $D_2$. Let $(Q_1, Q_2)$ denote the resulting outputs for $D_1$ and $D_2$, which we for the moment assume deterministic. First, note that $Q_1 = R_1(Q_2)$, because $U_1-D_1$ can procure internally at marginal cost $c_1$ and externally at marginal cost $c_2$. Furthermore, $Q_2 \geq R_2(Q_1)$ because $D_2$ can buy any amount from $U_2$ at marginal cost $c_2$. We thus have $Q_1 \leq q^*_1$, $Q_2 \geq q^*_2$ and $Q_1 + Q_2 \geq q^*_1 + q^*_2$ from $|dR_1/dq_2| < 1$ (see figure 1). Thus industry profits are lower than in our presumed equilibrium. Yet $D_2$ can guarantee itself $\Pi(c_1, c_2)$ because by turning down $U_1$’s offer it obtains

$$\max_{q_{22}}[r(q_{22}, Q_1) - c_2 q_{22}] \geq \max_{q_{22}}[r(q_{22}, q^*_1) - c_2 q_{22}] = \Pi(c_1, c_2).$$

Hence industry profits have fallen, while $U_2$ and $D_2$ are at least as well off. Hence $U_1-D_1$ cannot increase its profit. This reasoning extends straightforwardly to random outcomes $(\hat{Q}_1, \hat{Q}_2)$. First note that $\hat{Q}_1$ is necessarily deterministic (equal to some $Q_1$) as it maximizes the strictly concave function

$$\mathbb{E}[r(Q_1, \hat{Q}_2) - c_1 Q_1],$$
where \( \mathbb{E} \) denotes the expectation operator. Furthermore, any realization \( Q_2 \) of \( \hat{Q}_2 \) exceeds \( R_2(Q_1) \). Let \( Q_2 \) be the infimum in the support of \( \hat{Q}_2 \). Then \( Q_2 \geq R_2(Q_1) \) and \( Q_1 \leq R_1(Q_2) \) (recall that reaction curves are downward sloping). This implies that \( Q_2 \geq q_2^* \) and \( Q_1 \leq q_1^* \) (see figure 1). Hence, \( D_2 \) can guarantee itself \( D^{pl}(c_1, c_2) \). Let \( Q_2^* = \mathbb{E}[\hat{Q}_2] \geq q_2^* \) denote the expectation of \( \hat{Q}_2 \). Our assumption that a firm’s marginal revenue is convex in the other firm’s output and the fact that marginal revenue is decreasing in \( Q_1 \) imply that \( Q_1 \geq R_1(Q_2) \). This inequality, together with \( Q_2^* \geq R_2(Q_1) \), implies that \( Q_1 + Q_2^* \geq q_1^* + q_2^* \) (see figure 1). Last, because the industry profit function is concave in total output, the upper bound on industry profit, which presumes production efficiency, satisfies \( \mathbb{E}[P(Q_1 + \hat{Q}_2) (Q_1 + \hat{Q}_2) - c_1(Q_1 + \hat{Q}_2)] \leq P(Q_1 + Q_2^*) (Q_1 + Q_2^*) - c_1(Q_1 + Q_2^*) \leq P(q_1^* + q_2^*) (q_1^* + q_2^*) - c_1(q_1^* + q_2^*) \). Hence, industry profit is smaller, and so is the profit of \( U_1-D_1 \). Q.E.D.

**Proof of Proposition 3.** In equilibrium \( U_1 \) produces internally \( q_1^* \) and offers to supply \( q_2^* \) to \( D_2 \) at price \( t_2^* = c_2q_2^* \). \( U_2 \) does not supply. The proof is essentially that of proposition 2. The only possible point of departure comes from the fact that \( U_2 \) “supplies” \( D_2 \) instead of externally. But this makes no difference for the proof that \( U_1-D_1 \) cannot raise its profit because \( D_2 \) can already buy at marginal cost \( c_2 \) from \( U_2 \) under \( P_{ij} \). We only have to check that \( U_2-D_2 \) cannot raise its profit by making an alternative offer to \( D_1 \). Suppose it does so. Because \( D_1 \) and \( D_2 \) can purchase internally at marginal cost, we have \( Q_1 \geq R_1(Q_2) \) and \( Q_2 \geq R_2(Q_1) \) (the case of random \( \hat{Q}_1 \) and \( \hat{Q}_2 \) is solved as in proposition 2). Thus industry profit can only be lower than the one obtained in proposition 3. It thus suffices to check that even if \( U_2 \) changes its contract offer to \( D_1 \), which was to supply at marginal cost \( c_2 \), still \( U_1-D_1 \) can guarantee itself \( V^{FI}(c_1, c_2) \) (gross of the efficiency loss). To see this, note that if \( Q_2 > R_2(Q_1) \), it is profitable for \( U_2 \) to supply \( D_2 \) any positive amount internally, and so \( Q_2 = q_2^* \); but then \( U_1-D_1 \) can get \( V^{FI}(c_1, c_2) \) by not buying from \( U_2 \) and producing \( q_1^* \) internally. On the other hand, if \( Q_2 = R_2(Q_1) \), \( Q_2 \leq q_2^* \) because \( Q_1 \geq R_1(Q_2) \) (see figure 1) and again \( U_1-D_1 \) can get \( V^{FI}(c_1, c_2) \). Hence, \( U_2-D_2 \) cannot gain by offering a different contract to \( D_1 \). Q.E.D.

**Proof of Proposition 4.** Consider the following strategies: \( U_2 \) offers to sell at marginal cost \( c_2 \) to \( D_1 \) up to \( q^* \). Thus \( t_{21}(q_{21}) = c_2q_{21} \) for
\( q_{21} \leq q^* \), \( = + \infty \) for \( q_{21} > q^* \). \( U_1 \) offers to sell \( q^* \) at price \( t^* \) to \( D_2 \) (where \( q^* \) and \( t^* \) are as in proposition 1). \( U_1 \) offers to sell either \( q^* \) at price \( t^* \) or \( q_{11} \) at price \( t_{11}(q_{11}, q^*) = r(q^*, q^*) - r(q^* - q_{11}, q^*) \) to \( D_1 \) if \( D_1 \) can exhibit total output \( \hat{Q}_1 \geq q^* \). In equilibrium \( D_1 \) buys \( q^* \) at price \( t^* \) from \( U_1 \).

Note that \( U_1 \) simply offers to make up the difference to \( q^* \) if \( D_1 \) does not buy \( q^* \) from \( U_2 \). First, we show that \( D_1 \) cannot increase its profit. From the definition of \( t_{11} \), if \( D_1 \) buys \( q_{21} \leq q^* \) from \( U_2 \), then \( D_1 \) has the same profit whether it buys the complement to \( q^* \) from \( U_1 \) or not. Its profit is thus \( r(q_{21}, q^*) - c_2q_{21} \leq r[R_2(q^*), q^*] - c_2q^* = D^{NI}(c_1, c_2) \) (by definition of \( R_2 \)). Second, the proof that \( U_1 \) cannot make more than \( U^{NI}(c_1, c_2) \) is the same as that in proposition 1: \( U_2 \) and \( D_2 \) are now integrated, but \( U_2 \) continues to supply \( D_2 \) at marginal cost \( c_2 \).

The third and most difficult part of the proof consists of showing that \( U_2-D_2 \) cannot make more than \( D^{NI}(c_1, c_2) \). Suppose that \( U_2 \) makes a different contract offer to \( D_1 \). Suppose first that there exists no \((q_{21}, t_{21})\) in the new contract, such that \( r(q_{21}, q^*) - c_2q_{21} \leq r[R_2(q^*), q^*] - c_2q^* = D^{NI}(c_1, c_2) \). Then specify that \( D_1 \) turns down \( U_2 \)'s contract offer and buys \( q^* \) from \( U_1 \), and that \( D_2 \) also buys \( q^* \) from \( U_1 \) and does not produce internally. This is clearly a continuation equilibrium, and it gives the same profit to \( U_2-D_2 \) as before. Thus assume that there exists \((q_{21}, t_{21})\) such that \( r(q_{21}, q^*) - c_2q_{21} > D^{NI}(c_1, c_2) \). The definition of \( D^{NI}(c_1, c_2) \) implies that \( U_2-D_2 \) does not make money on the trade because \( t_{21} \leq c_2q_{21} \). Suppose \( q_{21} \leq q^* \) and then specify that \( D_1 \) buys \( q_{21} \) at price \( t_{21} \) from \( U_2 \) and buys \( q^* - q_{21} \) at price \( t_{11}(q^* - q_{21}, q^*) \) from \( U_1 \), and that \( D_2 \) buys \( q^* \) at price \( t^* \) from \( U_1 \) and does not produce internally. Again, this continuation equilibrium yields at most \( D^{NI}(c_1, c_2) \) to \( U_2-D_2 \). Or suppose \( q_{21} > q^* \), and assume that in equilibrium \( D_1 \) buys \( q_{21} \) from \( U_2 \) (the case of a random strategy for \( D_1 \) is treated as in proposition 2). Then \( D_1 \)'s total output \( Q_1 \geq q_{21} \) and the profit of \( U_2-D_2 \) is at most \( \max_{q_2} [r(Q_2, q_{21}) - c_2Q_2] \leq r[R_2(q^*), q^*] - c_2R_2(q^*) = D^{NI}(c_1, c_2) \).

Buying \( q^* \) from \( U_1 \) is not a best response to \( Q_1 \) as it yields \( r(q^*, Q_1) - t^* = r(q^*, Q_1) - r(q^*, q^*) + D^{NI}(c_1, c_2) < D^{NI}(c_1, c_2) \). We thus conclude that \( U_2-D_2 \) cannot increase its profit beyond \( D^{NI}(c_1, c_2) \).

Q.E.D.
Appendix B: Uniqueness in the Ex Post Monopolization Variant

We look at (perfect Bayesian) equilibria in the following class:
Restriction 1. The equilibrium is in pure strategies.
Restriction 2. Market-by-market bargaining: when a downstream firm, $D_k$, receives an out-of-equilibrium offer from an unintegrated upstream firm, $U_i$, it does not change its beliefs about $U_i$’s offer to $D_\ell (\ell \neq k)$.
Restriction 3. No-money-losing offers: an unintegrated firm does not make an offer at a price below marginal cost—that is, one that would lose money if accepted; $t_{ij}(q_{ij}, \hat{Q}_{ij}) \geq c_iq_{ij}$ for all $i, j, q_{ij}$, and $\hat{Q}_{ij}$.

Let us comment on restrictions 2 and 3. Restriction 2, although not implied by perfect Bayesian equilibrium, is a natural one. An unintegrated $U$ makes secret and independent offers to two downstream firms and tries to extract the best deal from each of them. Because there is no information leakage from one customer to the other, the unintegrated $U$ has no incentive to change the offer to $D_\ell$ when it changes its offer to $D_k$ (and indeed equilibrium behavior requires that it does not do so if its offer to $D_\ell$ is uniquely optimal). No such restriction can be imposed for an integrated $U$. When it changes its offer to its subsidiary’s rival, it also wants to change its supply to its subsidiary, with whom it shares profit.

Given restriction 2, restriction 3 is in the spirit of trembling-hand perfection of not allowing a player to play a weakly dominated strategy.\textsuperscript{52} An offer that contains a money-losing pair is worse for $U$ than the same offer without it if there is a small probability that the downstream firm chooses this money-losing pair.\textsuperscript{53}

The equilibria described in the discussion of the ex post monopolization variant satisfy restrictions 1 through 3.

\textbf{PROPOSITION A:} Under NI, FI, PI\textsubscript{1}, M\textsubscript{u}, M\textsubscript{d}, and M\textsubscript{ud} there exists a single perfect Bayesian equilibrium satisfying restrictions 1 through 3.
Under PI\textsubscript{2} the equilibrium described in proposition 4 is undominated in the set of perfect Bayesian equilibria satisfying 1 through 3. Fur-

\textsuperscript{52} Selten (1975).
\textsuperscript{53} One might think that including the money-losing pair could act as a “sunspot” and induce the downstream firm to choose from among the non-money-losing pairs the one that $U$ prefers. However, this selection can also be made directly by $U$ by offering a single best pair to the downstream firm.
thermore, any other equilibrium satisfying 1 through 3, if one exists, has $U_2$ supplying at a loss to $D_1$, and $D_1$ producing more than $q^*$, and the integrated firm $U_2$-$D_2$ making less profit than in the equilibrium of proposition 4.

We have been unable to prove or disprove uniqueness in the class considered under $PI_2$. But if other equilibria exist, they are somewhat pathological: $U_2$ supplies at a loss its subsidiary’s rival. Such behavior might be plausible if $D_1$ bought from $U_2$ a quantity less than $q^*$ and bought nothing from $U_1$. However, $D_1$ ends up buying more than $q^*$, the amount it buys from $U_1$ in the equilibrium of proposition 4.

**Proof of Proposition A.** Let $q_1$ and $q_2$ denote the final outputs of $D_1$ and $D_2$, and let $q_{ij}$ be $U_i$’s supply to $D_j(q_j = q_{1j} + q_{2j})$.

**Nonintegration.** Under market-by-market bargaining (restriction 2), $U_1$ and $U_2$ are competing à la Bertrand for each $D_j$ separately. For instance, $D_1$’s beliefs about $q_2$ are fixed in a given equilibrium and do not depend on $U_1$’s and $U_2$’s offers to $D_1$. $U_1$’s best offer is then trivially the best reaction $R_1(q_2)$ to $q_2$, at the highest price such that $D_1$ does not want to buy from $U_2$. And symmetrically for $D_2$. Hence the equilibrium outputs are $q_1 = q_2 = q^*$ and the transfers to $U_1$ equal $t^*$, where $q^*$ and $t^*$ are given in proposition 1.

**Full integration.** Because integrated downstream firms can procure internally at marginal cost, $q_i \geq R_i(q_j)$, aggregate profit, gross of integration cost, $\pi_1 + \pi_2$, satisfies $\pi_1 + \pi_2 \leq r(q_1^*, q_2^*) + r(q_2^*, q_1^*) - c_1(q_1^* + q_2^*)$, with equality only if $q_1 = q_1^*$ and $q_2 = q_2^*$, that is only if $q_1 = R_1(q_2^*)$ and $q_2 = R_2(q_1^*)$. It thus suffices to show that $U_1$-$D_1$ can guarantee itself $\pi_1^* = r(q_1^*, q_2^*) - c_1q_1^* + (c_2 - c_1)q_2^*$, and that $U_2$-$D_2$ can guarantee itself $\pi_2^* = r(q_2^*, q_1^*) - c_2q_2^*$. If this is so, the equilibrium outputs and profits are as in proposition 4.

Suppose that firm $U_1$ offers to supply $D_2$ up to $q_2^*$ at price $t''_{12}(q_{12})$, where

$$t''_{12}(q_2^*) = c_2q_2^*, \quad t''_{12}(q_{12}) < c_2q_{12} \quad \text{and} \quad \lim_{n \to \infty} t''_{12}(q_{12}) = c_2q_{12}$$

for $0 < q_{12} < q_2^*$. That is, $U_1$ offers to undercut $U_2$ slightly up to $q_2^*$. Figure 2 exhibits $D_2$’s reaction curve, $R_2^*(q_1)$, coming from the maximization of $r(q_2, q_1) - c_2q_{22} - t''_{12}(q_{12})$, subject to $q_{12} + q_{22} = q_2$. $R_2^*$ coincides with $R_2$ for $q_1 \leq q_1^*$ and, for $n$ sufficiently large, is close
to $R_2$ for $q_1 \geq q_1^*$. Note that $r''_{12}(\cdot)$ can be chosen so that $R_2^n$ is continuous, which we will assume.

Because $U_2$ may make an offer to $D_1$, the reaction curve $\tilde{R}_1(q_2)$ of $D_1$ is obtained by solving

$$\max[r(q_1, q_2) - c_1 q_{11} - t_{21}(q_{21}, \hat{Q}_{21})],$$

subject to

$$\begin{cases} q_{11} + q_{21} = q_1 \\ q_1 \geq \hat{Q}_{21} \end{cases}$$

where we adopt the convention that $t_{21}(0, \cdot) = 0$. ($\hat{Q}_{21}$ denotes the quantity exhibited to $U_2$ by $D_1$.) By the standard revealed preference argument, $\tilde{R}_1(q_2)$, which need not be single-valued, is monotonic (non-increasing). Furthermore $\tilde{R}_1(q_2) \geq R_1(q_2)$ because $D_1$ can always re-
frain from buying from $U_2$. The crucial feature of $\bar{R}_1$ is that it admits only horizontal jumps. Therefore $\bar{R}_1$ and $R_2^*$ intersect for some $q_2 \leq q_2^*$ (there may exist several such intersections, but they all share this property); see figure 2. This implies that by buying $q_{21} = 0$, the merged $U_1-D_1$ can guarantee itself at least $\pi_1^*$ by offering the above contract to $D_2$. The reasoning for why $U_2-D_2$ can guarantee itself $\pi_2^*$ is symmetrical. It suffices that $U_2$ offer no contract to $D_1$.

**Partial Integration $PI_1$.** First note that the no-money-losing-offer assumption implies that in equilibrium $q_1 = R_1(q_2)$. Because the un-integrated $U_2$ does not supply under marginal cost $c_2$, and $D_1$ can procure internally at marginal cost $c_1 < c_2$, firm $D_1$ only purchases internally and has reaction curve $R_1$. Next we claim that $q_2 \geq R_2(q_1)$, for if $q_2 < R_2(q_1)$, then $U_2$ could increase its profit by offering to put $D_2$ on its reaction curve. More precisely, if $q_{12}, q_{22}, \hat{q}_{12},$ and $\hat{q}_{22}$ maximize $r(q_2, q_1) - t_{12}(q_{12}, \hat{q}_{12}) - t_{22}(q_{22}, \hat{q}_{22})$, such that $q_{12} + q_{22} = q_2$ and $q_2 \geq \hat{q}_{12}, \hat{q}_{22}$, then the contract "$R_2(q_1) - q_{12}$ at price $t_{22}(q_{22}, \hat{q}_{22}) + r[R_2(q_1), q_1] - r(q_2, q_1) - \epsilon$" offered by $U_2$, where $\epsilon$ is positive and small, is strictly preferred by $D_2$ to rejecting the contract and buying from $U_1$ only, and yields a strictly higher profit to $U_2$, as is easily checked.

Because $q_1 = R_1(q_2)$ and $q_2 \geq R_2(q_1)$, $q_1 \leq q_1^*$ and $q_2 \geq q_2^*$ and $\pi_1 \leq \pi_1^*$ with equalities if, and only if, $\pi_1 = \pi_1^*$. To show that $U_1-D_1$ can guarantee itself $\pi_1^*$, note that if it offers the schedule $t_{12}^*(q_{12}) = (c_2 - \epsilon)q_{12}$ for all $q_{12}$ to $D_2$, where $\epsilon$ is small, $D_2$ will never buy from $U_2$, which makes no money-losing offer, and has reaction curve $R_2^*(\cdot)$ converging uniformly to $R_2(\cdot)$ when $\epsilon$ tends to 0. Thus as $\epsilon$ tends to 0, the Nash equilibrium when $t_{12}^*(\cdot)$ is offered to $D_2$ by $U_1$ converges to $(q_1^*, q_2^*)$. We thus conclude that the unique equilibrium satisfying our restrictions is the one exhibited in proposition 2.

**Partial Integration $PI_2$.** Note first that market-by-market bargaining for $U_1$ implies that $U_1$ sells $q_2 = R_1(q_1)$ at price $c_2q_2$ to $D_2$. Second, we claim that $q_1 \geq R_1(q_2)$. Otherwise, $U_1$ would put $D_1$ on its reaction curve $R_1(q_2)$—again, we invoke market-by-market bargaining. Furthermore, $q_1 = R_1(q_2)$ if $q_{21} = 0$. We thus conclude that either $q_1 = q_2 = q^*$ and $U_1$ supplies $q^*$ at price $c_2q^*$ to both $D_1$ and $D_2$, or $q_1 > q^* > q_2$ and $q_{21} > 0$. Uniqueness under $M_u$, $M_d$, and $M_{ud}$ is straightforward.

Q.E.D.
Appendix C: Exclusive Dealing

We analyze exclusive dealing in the context of the model of ex post monopolization. We solve the ex ante stage with deterministic costs under exclusive dealing (ED) to point to the essential difference between ED and vertical integration as means of foreclosing markets. In our model, an exclusive-dealing contract between $U_1$ and $D_1$ allows $U_1$ to commit not to supply $D_2$.

In a nutshell, ED has two drawbacks and one advantage relative to vertical integration. It dominates vertical integration in that it allows firms to remain independent and avoids the incentive loss $E$. The first drawback, which we will not study but could be represented by a constant loss, $K$, given our constant-returns-to-scale assumption, is associated with the loss of gains from trade between the upstream firm $U_1$ and third parties (firms outside the industry). Such a loss occurs either if shipments by $U_1$ cannot be monitored by $D_1$ or if arbitrage between third parties and $D_2$ cannot be prevented. Then the only credible way for $U_1$ to cease trading with $D_2$ is if $U_1$ promises not to trade with anybody but $D_1$. Second, and more important from the point of view of our model, ED implies production inefficiency. Precisely when $U_1-D_1$ gain by foreclosing the market ($c_1 < c_2$), ED forces $D_2$ to buy from the high-cost supplier. Hence under ED, $U_1-D_1$ loses the profit $(c_2-c_1)q_2^*$ obtained by selling $q_2^*$ to $D_2$. Thus, ignoring the cost $K$ of not trading with third parties (for example, if no third party exists), the total profit of $U_1$ plus $D_1$ when $c_1 \leq c_2$ is

$$V_{ED}(c_1, c_2) = r(q_1^*, q_2^*) - c_1 q_1^*$$

under ED, and

$$V_{PI}(c_1, c_2) - E = V_{FI}(c_1, c_2) - E$$

$$= r(q_1^*, q_2^*) - c_1 q_1^* + (c_2 - c_1)q_2^* - E$$

under vertical integration.

Now, suppose that costs are deterministic and that $c_1 \leq c_2$. Propositions 3 and 4 imply that $U_2$ and $D_2$ have no incentive to integrate whether $U_1$ and $D_1$ are integrated or not. It is easy to see that $U_2$ and $D_2$ have no incentive to sign an ED contract either. Assuming no exit occurs, the only possible industry structures are NI, PI, and ED (ED
contract between $U_1$ and $D_1$). The optimal choice for $U_1$-$D_1$ between these three structures is given by proposition B.

**PROPOSITION B:** Consider the deterministic case in which there are no investments and thus exit does not occur. Assuming $c_1 < c_2$, either of the three possible industry structures—$NI$, $PI_1$, and $ED_1$—may be optimal for $U_1$-$D_1$, and thus arise. In particular:

(i) If $c_2$ is close to $c_1$, $NI$ is preferred to $ED_1$ by $U_1$-$D_1$ if the demand function is linear.

(ii) If $c_2$ is much larger than $c_1$, $ED_1$ is preferred to $NI$.

(iii) If $E$ is small, $PI_1$ dominates both $NI$ and $ED_1$.

(iv) If $E$ is large, $PI_1$ is dominated by both $NI$ and $ED_1$.

**PROOF:** (i) Note that $V_{ED}(c_1, c_1) = U_{NI}(c_1, c_1) + D_{NI}(c_1, c_1)$, so $U_1$-$D_1$ is indifferent between $NI$ and $ED_1$ in the symmetric case. Raising $c_2$ above $c_1$, we obtain from the envelope theorem:

$$\frac{\partial V_{ED}}{\partial c_2} = \frac{\partial r}{\partial q_2} \frac{\partial q_2^*}{\partial c_2} = P'(q_1^* + q_2^*) q_1^* \frac{\partial q_2^*}{\partial c_2},$$

while

$$\frac{\partial (U_{NI} + D_{NI})}{\partial c_2} = R_2[q^*(c_1)].$$

Because at $c_2 = c_1$, $q^*(c_1) = R_1(q_2^*) = R_2[q^*(c_1)]$, for linear demand one has

$$\left. \frac{\partial V_{ED}}{\partial c_2} \right|_{c_2 = c_1} < \left. \frac{\partial (U_{NI} + D_{NI})}{\partial c_2} \right|_{c_2 = c_1}.$$

(ii) Fixing $c_1$, define $\bar{c}_2$ as the lowest value of $c_2$, such that $q_2^*(c_2, c_1) = 0$. For $c_2 \geq \bar{c}_2$, $ED_1$ allows $U_1$-$D_1$ to obtain the monopoly profit $\pi^m(c_2)$, while $U_{NI}(c_1, c_2) + D_{NI}(c_1, c_2)$ is bounded away from this monopoly profit (see proposition 1).

(iii) It suffices to show that $U_1$-$D_1$ strictly prefers vertical integration for $E = 0$ (by continuity, this will also hold for $E$ = small). That $PI_1$ strictly dominates $NI$ for $U_1$-$D_1$ when $E = 0$ results from proposition 2. And $V_{PI}(c_1, c_2) = V_{ED}(c_1, c_2) + (c_2 - c_1)q_2^*$ implies that $PI_1$ dominates $ED_1$ when $E = 0$ (for $c_2 < \bar{c}_2$; for $c_2 \geq \bar{c}_2$, $ED_1$ and $PI_1$ are equivalent if $E = 0$).
(iv) Trivial \((E\text{ is incurred only under vertical integration})\).

\[ \text{Q.E.D.} \]

Appendix D: Proofs of Propositions 5, 6, and 9

Proof of Proposition 5

Proposition 5 is trivial in case ii. From propositions 1 through 4, the gain from integration occurs when \(c_i = c\) and \(c_j = +\infty\), which has a higher probability for \(i = 1\) than for \(i = 2\).

To prove the proposition for small uncertainty (case i) we first show that \(g(c_i, c_j)\) is decreasing in \(c_i\) and increasing in \(c_j\). Using the definition of \(g(c_i, c_j)\) and the envelope theorem, we have (for \(c_i < c_j\)):

\[
\frac{\partial g(c_i, c_j)}{\partial c_j} = P'(q_i^* + q_j^*)q_i^* \frac{\partial q_i^*}{\partial c_j} + q_j^* + (c_j - c_i) \frac{\partial q_j^*}{\partial c_j} - R_j[q^*(c_i)].
\]

In particular, because \(q_j^*(c, c) = R_j[q^*(c)]\),

\[
\left. \frac{\partial g(c_i, c_j)}{\partial c_j} \right|_{c_j = c_i} = P'[2q^*(c_i)]q^*(c_i) \frac{\partial q_j^*}{\partial c_j} > 0.
\]

Hence, \(g(c_i, c_j)\) is increasing in \(c_j\) for small uncertainty. Next, we have

\[
\frac{\partial g(c_i, c_j)}{\partial c_i} = -q_i^* - q_j^* + [P'(q_i^* + q_j^*)q_i^* + (c_j - c_i)] \frac{\partial q_i^*}{\partial c_i}
\]

\[
- \left( 2P'[2q^*(c_i)]q^*(c_i) \frac{dq^*}{dc_i} - 2q^*(c_i) - P' \{q^*(c_i) \right)
\]

\[+ R_2[q^*(c_i)]R_2(q^*) \frac{dq^*}{dc_i} \right).
\]

In particular,

\[
\left. \frac{\partial g(c_i, c_j)}{\partial c_i} \right|_{c_i = c_j} = P'[2q^*(c_i)]q^*(c_i) \left[ \frac{\partial q_j^*}{\partial c_i} - \frac{dq^*(c_i)}{dc_i} \right] < 0,
\]

because \(\partial q_i^*/\partial c_i - dq^*(c_i)/dc_i > 0\).
Last, if \( \frac{\partial g}{\partial c_i} < 0 \) and \( \frac{\partial g}{\partial c_j} > 0 \), then \( G(F_i, F_j) > G(F_j, F_i) \) if \( F_i \) first-order stochastically dominates \( F_j \). Because \( \frac{\partial g}{\partial c_j} > 0 \), then \( G(F_i, F_j) > G(F_i, F_i) \). And because \( \frac{\partial g}{\partial c_i} < 0 \), then \( G(F_i, F_i) > G(F_j, F_i) \). Q.E.D.

**Proof of Proposition 6**

The inequality in proposition 6 is an equality in the large uncertainty case. \( U_i \) and \( D_i \) might suffer from integration by \( U_j \) and \( D_j \) only if \( c_i = +\infty \) and \( c_j = c \) (see propositions 1 through 4). But in this case \( D_{Ni}(+\infty, c) = 0 \) anyway.

Consider next small uncertainty: as in proposition 5, our strategy is to show that \( \ell(c_i, c_j) \) is decreasing in \( c_j \) and increasing in \( c_i \) for \( c_i > c_j \). We have

\[
\frac{\partial \ell(c_i, c_j)}{\partial c_j} = P'[R_i[q^*(c_j)] + q^*(c_j)]R_i[q^*(c_j)] \frac{dq^*(c_j)}{dc_j} - P'(q_i^* + q_j^*)q_i^* \frac{dq_j^*}{dc_j}.
\]

At \( c_j = c_i \),

\[
\frac{\partial \ell(c_i, c_j)}{\partial c_j} \bigg|_{c_i=c_j} = P'[2q^*(c_j)]q^*(c_j) \left( \frac{dq^*(c_j)}{dc_j} - \frac{dq_j^*}{dc_j} \right) < 0,
\]

because \( dq^*/dc_j - dq_j^*/dc_j > 0 \), as is easily seen on a diagram. Hence, in the small uncertainty case, \( \ell(c_i, c_j) \) is decreasing in \( c_j \). Next,

\[
\frac{\partial \ell(c_i, c_j)}{\partial c_i} \bigg|_{c_i=c_j} = - R_i[q^*(c_j)] + q_i^* - P'(q_i^* + q_j^*)q_i^* \frac{dq_j^*}{dc_i}.
\]

In particular,

\[
\frac{\partial \ell(c_i, c_j)}{\partial c_i} \bigg|_{c_i=c_j} = - P'[2q^*(c_j)]q^*(c_j) \frac{dq_j^*}{dc_i} > 0.
\]

Hence, \( \ell(c_i, c_j) \) is increasing in \( c_i \) in the case of small uncertainty.

Last, if \( \frac{\partial \ell}{\partial c_j} < 0 \), then \( L(F_1, F_2) < L(F_1, F_1) \). And because \( \frac{\partial \ell}{\partial c_i} > 0 \), then \( L(F_1, F_1) < L(F_2, F_1) \). Q.E.D.
Proof of Proposition 9

In the case of large uncertainty,

$$\mathcal{U}_j^{Md} = \alpha_j(1 - \alpha_i)\pi^m(c) > \mathcal{U}_j^{PI} = \alpha_j(1 - \alpha_i)[2\pi^d(c)],$$

where $\pi^d(c) \equiv r[q^*(c), q^*(c)] - cq^*(c)$ and $\pi^m(c) \equiv \max [r(q, 0) - cq]$.

Suppose that $c_j \leq c_i$. Then

$$U_j^{PI}(c_j, c_i) - U_j^{Md}(c_j, c_i) = 2\{r[q^*(c_j), q^*(c_j)] - c_jq^*(c_j)\} \tag{1}$$

$$- 2\{r[R_iq^*(c_j)], q^*(c_j)\} - c_jR_i[q^*(c_j)]$$

$$- [\pi^m(c_j) - \pi^m(c_i)].$$

Keeping $c_j$ constant, let us take the derivative of this expression with respect to $c_i$ at $c_i = c_j$:

$$\frac{\partial [U_j^{PI}(c_j, c_i) - U_j^{Md}(c_j, c_i)]}{\partial c_i} \bigg|_{c_i = c_j} = 2q^*(c_j) - q^m(c_j) > 0,$$

where we use the fact that in a symmetric Cournot equilibrium, total output exceeds the monopoly output. But $U_j^{PI}(c_j, c_i) = U_j^{Md}(c_j, c_i) = 0$ for $c_j \geq c_i$. Hence, $U_j^{PI}(c_j, c_i) > U_j^{Md}(c_j, c_i)$ for $c_j < c_i$ and $(c_i - c_j)$ small, which proves the result in the case of small uncertainty.

Q.E.D.

Appendix E: War of Attrition and Preemption in the Ex Post Monopolization Variant Merger Game

We assume large downstream investments ($\mathbf{\mathcal{D}}^p_j < J$), and small upstream investments so as to focus on downstream monopolization, and show that two polar cases of merger dynamics, war of attrition and preemption, may arise. If uncertainty is large, a low-cost upstream firm is a monopolist when its rival’s cost is high. The low-cost firm’s problem is then to commit not to supply both downstream firms. One possibility for commitment is that the low-cost supplier is integrated. Another is that one of the downstream firms has exited. The upstream firm then benefits from downstream monopolization and does not want to rescue
a failing downstream firm (proposition 9). In this respect ex ante monopolization by vertical integration resembles a public good. Both upstream firms benefit from it, and each firm prefers the other to trigger downstream exit and incur the integration cost. This suggests the possibility of a war of attrition between the upstream firms. There is a second consideration, however. When both upstream firms’ costs are low, the remaining buyer after ex ante monopolization enjoys a monopoly profit on the product market. Obviously, each downstream firm would like to be the one that enjoys this monopoly profit, which suggests that the merger game might resemble a game of preemption. We show by means of symmetric examples that there is indeed a conflict between these two effects. In the relevant range for the integration cost, firms will wage a war of attrition if the integration cost is high, and will try to preempt each other if the integration cost is low, resulting in late and early vertical integration respectively.

Consider, in the ex post monopolization variant, a slight modification of the symmetric, large uncertainty case. Let \( c_i = c \) with probability \( \alpha \), and \( c' \) with probability \( (1 - \alpha) \). Before, we assumed that \( c' = +\infty \). Let us assume that \( c' \) is slightly smaller than \( \bar{c} \), where \( \bar{c} \) is the smallest marginal cost, such that the Cournot output of a firm with cost \( \bar{c} \) facing a firm with cost \( c \) is equal to zero. The purpose of having \( c' \) lower than \( \bar{c} \) is to allow downstream firms to suffer from integration. Let \( q^*(c) \) denote the Cournot output when both firms have cost \( c \). Let

\[
\pi^d(c) \equiv r[q^*(c), q^*(c)] - cq^*(c)
\]

denote the Cournot profit. And let

\[
\pi^m(c) \equiv \max_q [r(q, 0) - cq]
\]

denote the monopoly profit. In this symmetric example, we drop the subscripts under the expected profit functions. The reader will easily check that the expected profits are:

\[
\begin{align*}
\mathbb{E}^N_I & = \alpha(1 - \alpha)2[\pi^d(c) - D^N(c, c')] \\
\mathbb{D}^N_I & = \alpha^2\pi^d(c) + (1 - \alpha)^2\pi^d(c') + 2\alpha(1 - \alpha)D^N(c, c') \\
M_d & = \mathcal{U}^M_d = \alpha\pi^m(c) + (1 - \alpha)\pi^m(c') \\
\mathcal{U}^M_d & = \alpha(1 - \alpha)[\pi^m(c) - \pi^m(c')] < \mathcal{U}^M_d.
\end{align*}
\]
(Partial integration is not feasible if, as we will assume, \( J \) is sufficiently big. Also, full integration will not occur if \( c' \) is close to \( \bar{c} \), from propositions 9 and 10.) Because \( D^{PI}(c, c') < D^{NI}(c, c') \) (from proposition 2), for any \( \alpha \), there exists \( J \) such that \( \mathbb{D}^{NI} > J > \mathbb{D}^{PI} = \alpha^2 \pi^d(c) + (1 - \alpha)^2 \pi^d(c') + \alpha(1 - \alpha)D^{NI}(c, c') + \alpha(1 - \alpha)D^{PI}(c, c') \). Furthermore, a merger implies exit of the unmerged downstream firm. Knowing that \( \mathcal{V}^{M_d} > \mathcal{U}^{NI} + \mathbb{D}^{NI} \), let us choose \( E \) such that

\[
\mathcal{V}^{M_d} - E > \mathcal{U}^{NI} + \mathbb{D}^{NI},
\]

so that nonintegration cannot be an equilibrium of the merger game. We must further distinguish two cases.

**Case 1:** \( \mathcal{V}^{M_d} - E - J < \mathcal{U}^{M_d} \). In this case, every firm likes ex ante monopolization, but each would like the other to merge because the integration cost is high. Ex ante monopolization is a “public good.” Although our reduced form for the merger game yields two pure-strategy equilibria (\( U_1 \) and \( D_1 \) merge and \( U_2 \) and \( D_2 \) merge), in this case the reduced-form representation of the game is inadequate. In real time we would expect a war of attrition. To be more precise, suppose that all payoffs are flow payoffs (as discussed in the description of the merger game), and let \( e \) denote the flow equivalent of the integration cost at rate \( r \): \( e = rE \). Case 1 can then be described by \( \mathcal{V}^{M_d} - e - J < \mathcal{U}^{M_d} \).

In the symmetric equilibrium of the war of attrition, each \( U_i-D_i \) randomizes between integrating and not integrating at each instant, conditionally on no one having merged yet. That is, if the game takes place on \([0, +\infty)\), the probability of integration by \( U_i \) and \( D_i \) between \( t \) and \( t + dt \) conditional on no merger having yet occurred is \( xdtdt \), where \( x \) is given by

\[
x \left[ \frac{\mathcal{U}^{M_d} - (\mathcal{V}^{M_d} - e - J)}{r} \right] = (\mathcal{V}^{M_d} - e) - (\mathcal{U}^{NI} + D^{NI}).
\]

The left-hand side represents the benefits of not integrating times the per-unit-of-time probability that the rival integrates, and the right-hand side denotes the gain from monopolizing the industry. The war of attrition is shorter (\( x \) is larger) when the integration cost is larger.

**Case 2:** \( \mathcal{V}^{M_d} - E - J > \mathcal{U}^{M_d} \). In this case, each firm prefers to be the one that triggers ex ante monopolization. Again, the reduced-form representation of the merger game is not adequate. In real time the
game would resemble a preemption game, and rent dissipation would occur. To see this, suppose that the game is played in continuous time, with the payoffs standing for flow payoffs; thus case 2 corresponds to \((\psi^M_d - J - \psi^M_d)/r > E\). Assume that the market opens at date 0, but mergers can occur before date 0. We claim that some \(U_i \cdot D_i\) merges at date \(-T\) (triggering \(D_j\)'s exit), where \(T\) is such that \(U_j\) and \(D_j\) are indifferent between preemptsing \(U_i \cdot D_i\) by merging at \(-(T + \epsilon)\) and letting \(U_i \cdot D_i\) preempt:

\[-E + e^{-rT} \left(\frac{\psi^M_d - J - \psi^M_d}{r}\right) = 0.54\]

In equilibrium, the firms’ profits from ex ante monopolization are dissipated through wasteful early integration.

**Appendix F: Proof of Proposition 20**

We argued in the text that, if \(\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) > E\), it is a dominant strategy for \(U_1\) and \(D_1\) to merge. What is the response of \(U_2\) and \(D_2\)? If \(\beta P(\bar{Q})\bar{q}_2 > E\), then propositions 16 and 18 tell us that, given that \(D_2\) invests, \(U_2\) and \(D_2\) will prefer to merge. Hence either \(U_2\) and \(D_2\) will integrate or \(D_2\) will exit, depending on whether \(P(\bar{Q})\bar{q}_2 - J - E \geq (1 - \beta') [P(\bar{Q})Q - P(\bar{q}_1)\bar{q}_1]\). The left-hand side of this inequality represents the profits of \(U_2\) and \(D_2\) if they merge, while the right-hand side represents \(U_2\)'s profits if \(D_2\) exits; it is easy to see that if the left side is less than the right, \(D_2\) will choose to exit.

On the other hand, if \(\beta P(\bar{Q})\bar{q}_2 < E\), and if partial integration is viable, that is, \(\beta P(\bar{Q})\bar{q}_2 > J\), then \(U_2\) and \(D_2\) will not merge. However, if \(\beta P(\bar{Q})\bar{q}_2 < J\), then \(U_2\) can either let \(D_2\) exit and make profit \((1 - \beta') [P(\bar{Q})Q - P(\bar{q}_1)\bar{q}_1]\) or rescue \(D_2\) by a merger and make profit \(P(\bar{Q})\bar{q}_2 - J - E\). \(U_2\) will choose whichever strategy is more profitable.

Consider next the case \(\beta P(\bar{Q})(\bar{q}_1 - \bar{q}_2) < E\). Now the decision of \(U_1\) and \(D_1\) to merge will depend on the response of \(U_2\) and \(D_2\). A

54. See Fudenberg and Tirole (1985) for a similar treatment in the context of the adoption of a new technology and for a full description of the equilibrium strategies.
comparison of $V_1^{FI}$ and $U_1^{NI} + D_1^{NI}$ shows that, given condition 9, $U_1$ and $D_1$ will only merge if $U_2$ and $D_2$ remain separate, with $D_2$ possibly exiting. In fact we know from condition 6 that $D_2$’s exiting is a sufficient condition for $U_1$ and $D_1$ to merge. On the other hand, if $D_2$ remains independent and continues to invest, $U_1$ and $D_1$ will merge if, and only if, $\beta P(Q)\bar{q}_1 > E$, because this guarantees that $U_1$-$D_1$’s profits are higher under partial integration than under nonintegration. This yields case C.

Q.E.D.
Comments and Discussion

Comment by Dennis W. Carlton: Hart and Tirole have written a very clever paper to show how vertical integration can be privately desirable yet socially undesirable. Even though the paper is difficult, it is certainly valuable because vertical foreclosure is an important issue in antitrust enforcement. The authors use a very general model and make a lot of assumptions (or avoid making assumptions) and that means it is difficult for them to get their results simply. In some sense, they may make more work for themselves than they have to.

I discuss three main topics. First, I discuss how to formulate the game that the firms play and suggest ways that vertical integration can influence that formulation. Some of these ways are not analyzed directly by Hart and Tirole, but they could be especially important and have direct effects for the problem analyzed.

Second, I discuss some standard motives for vertical integration that are specifically abstracted from in the paper. These motives have a tremendous effect on how one interprets the results of the paper—in particular, the authors’ suggestions about when the anticompetitive effects of vertical integration would be most severe.

Finally, I conclude with comments on the authors’ empirical examples of vertical integration.

Model Formulation

In the model there is duopoly at the upstream and downstream stages. Therefore to make any progress, one must specify the game played at
each stage in both the integrated and nonintegrated cases. I am always uncomfortable when the strategy space is specified because it is often not obvious. Is it Bertrand in prices? Is it Cournot in quantities? If the point of a paper is to show that something is theoretically possible—for example, that socially undesirable vertical foreclosure could occur—then the paper is interesting only as long as the strategy space is not too outlandish. If the point of a paper is to show that foreclosure is not only theoretically possible but actually occurs and if the paper will be used for policy recommendations, then it matters very much what the strategy space is. I am especially wary when I know that the results may change significantly if there are changes in the strategy space.

To illustrate my point regarding sensitivity of results, I will use a paper by Ordover, Salop, and Saloner to which Hart and Tirole refer. In that paper foreclosure occurs if there is Bertrand competition but not if there is Cournot competition. Hart and Tirole criticize the authors for assuming that firms can make certain binding commitments, and Hart and Tirole obtain their results in a more general strategy space than Ordover, Salop, and Saloner. But commitments may possibly be made in more complicated ways than the models of Hart and Tirole allow. Their results would change dramatically if such commitments were allowed and, as in the paper by Ordover, Salop, and Saloner, would depend on whether the game is Bertrand or Cournot.

The theoretical results on vertical foreclosure depend critically on assumptions about strategies and commitments that I find hard to validate empirically. I am left in the undesirable situation of understanding the theoretical possibility of socially undesirable vertical foreclosure, but not being able to identify it when I see it.

Hart and Tirole emphasize that their approach, in contrast to that of others, relies on commitments that are credible. This is a virtue of their paper. However, vertical integration can enable a firm to make many more commitments than the ones Hart and Tirole analyze. The ability to make such commitments will influence outcomes and thereby influence the incentive to engage in vertical integration.

Let me explain how vertical integration can affect the credibility of a commitment. Vertical integration can eliminate opportunism and thereby allow greater specialization of assets to occur. When specialization occurs, products can be more idiosyncratic and can be more differen-
tiated. If products become more differentiated, the force of Bertrand competition can be lessened. Therefore, vertical integration can be a way for firms to commit not to produce identical products, which would be beneficial to them because it would lessen competition.

A second commitment from vertical integration is also related to specialization. If the upstream product produced by the vertically integrated firm becomes more specialized, it may not be as useful to an unintegrated firm. In fact, if the integrated firm chose to, it could make the input useless to the unintegrated firm. Therefore, vertical integration is a way in which the integrated firm can create a credible commitment that it will not supply an unintegrated firm.

Finally, there are situations in which a rival will not rely on a competitor to supply its product. Customers interested in obtaining a second source often want to make sure that it has a supply completely separate from that of the first source. This is another way in which vertical integration could result in a commitment not to supply.

When Anticompetitive Foreclosure Is Likely to Occur

There are two standard reasons, not discussed in the paper, for vertical integration. One has to do with variable proportions and the other with double markup. Hart and Tirole eliminate these reasons by assuming the possibility of two-part tariffs. It is proper for them to ignore the standard reasons because they focus on the incentive to integrate vertically that arises solely from strategic considerations related to foreclosure. In terms of logic, what they are doing is perfectly reasonable.

But if the relevance of their results for policymaking is to be considered, these standard reasons must be taken into account because two-part tariffs may not be in use, and price may exceed marginal cost. Any time an input supplier is charging a downstream firm a price different from marginal cost, there are incentives for vertical integration to eliminate the double markup or, if there are variable proportions, to induce efficient input ratios of capital and labor. Hart and Tirole suggest that policymakers should be especially alert to anticompetitive foreclosure when vertical integration occurs and one of the firms is especially efficient. But this is precisely the situation in which efficiency gains
from vertical integration are greatest because price exceeds marginal cost and there are variable proportions or a double markup.

Therefore, I would take exception to their recommendation that the burden of justifying a vertical merger should be borne by the merging firms if one of the merging firms is especially efficient. First, as I have explained, it is not clear that the anticompetitive harm, compared with the efficiency gains, is greatest when one of the integrating firms is very efficient. Second, shifting the burden of justifying a merger onto the merging parties may be counterproductive. Suppose a vertical merger will create efficiencies. I have little faith that economists can always convince a government enforcement agency ex ante of efficiencies. Enforcement agencies are thus appropriately skeptical when they see such demonstrations. Shifting the burden, as Hart and Tirole suggest, could result in fewer efficient vertical mergers.

**Empirical Examples**

Most theoretical models stress asymmetries of cost and information among firms. Yet these asymmetries appear to play no role in the empirical discussion of this paper. Moreover, all the theoretical models assume that two-part tariffs are used. Were two-part tariffs used in any of the empirical cases studied here?

**Comment by Oliver E. Williamson:** This paper works out of an incomplete contracting setup, broadly in the spirit of Grossman and Hart (1986). Because the modeling of incomplete contracting is a formidable task, simplification is greatly needed. Simplification is accomplished by focusing on competition and exchange between two successive stages of production, both of which are organized as duopolies.

That the analysis of even a successive duopoly is complicated is borne out by the length of the paper. Indeed, keeping the three variants of the authors’ model straight puts a real burden on the reader. That, however, is in the nature of the problem. They have done all that can be reasonably expected to relieve these burdens by their meticulous procedure. Their comparison and contrast of their treatment with the recent literature reveals, I think, the advantage of addressing the issues on their terms.
Nevertheless, I have two reservations. First, and most important, the public policy tone of the paper seems wrong. Second, the way the authors characterize the benefits and costs of integration (effects of market power aside) are restrictive.

Antitrust analysis and enforcement have come a long way from the 1960s. Then, possible economies associated with new forms of organization (vertical, horizontal, and conglomerate mergers and nonstandard forms of contracting) were held in low regard, and monopoly power was ascribed to market shares, even small shares, in what were often contrived definitions of relevant markets. Although some would contend that the pendulum has swung too far, the excesses of the 1990s are to be preferred to those of the 1960s. One of the reasons for the improvement is that antitrust enforcement in the 1990s is much more informed by the relevant economic theories.

Hart and Tirole inform and refine our understanding of the trade-offs posed by vertical integration. I would urge, however, that applications of their results be restricted to circumstances that closely approximate those of the model. Their use of the model to interpret the reorganization of the cement industry in the 1960s suggests wider scope for the model and a more elastic approach to public policy than I believe is appropriate.

Their model examines consequences for market power and efficiency arising from vertical mergers between successive duopolists. Although the qualitative effects the authors display arguably apply outside these very special circumstances, concerns about monopoly power are nonetheless attenuated as the number of firms increases or as entry becomes easier. If nontrivial market power indicia need to be exceeded before antitrust enforcement resources are properly mobilized, which is surely judicious, then the first question is whether the cement industry crossed the threshold. That the conditions of concentration and entry at both the cement and ready-mixed stages in relevant geographic markets exceeded the threshold is not demonstrated and is, I think, doubtful.

Hart and Tirole propose that the scarce needs variant of their model is the one applicable to cement. They aver in this connection that the “bottleneck seems to have been the downstream industry.” A striking feature of the cement industry in the 1960s, however, is that there was significant excess capacity.¹ Possibly that excess was less in the ready-

¹ Allen (1971).
mixed stage than in the manufacture of cement. There is no indication whatsoever, however, that the bottleneck was constraining. Furthermore, temporary bottlenecks that are easily relieved by low-cost entry—possibly financed by efficient upstream cement suppliers who are the intended victims of ready-mixed firm foreclosures—are scarcely constraints at all.

But so what? The hazard is that the elastic use of the model by the authors encourages the even more expansive use of it by others, especially those directly involved in enforcing antitrust laws. If such uses are not what Hart and Tirole intend, then they should restrict their applications to circumstances that more closely approximate the conditions of the model, as is arguably the case for the other two examples that they discuss.

My second concern is that the paper by Grossman and Hart (1986) out of which Hart and Tirole work characterizes the efficiency gains and losses from vertical integration in a very special way. Specifically, Grossman and Hart (and Hart and Tirole) assume that managers of firms and of divisions are compensated very similarly—namely, that they appropriate the net receipts of the stage of production to which they are assigned. Thus these managers face high-powered incentives. The source of the efficiency gains and losses of vertical integration under this setup turn entirely on the different ex ante investment distortions that alternative forms of ownership induce.

That is an important result. As I have argued elsewhere, however, the managers of internal divisions do not face the same high-powered incentives as the owners of independent firms: these internal incentives are more easily corrupted, internal organization has access to control instruments that are superior to the market, the deliberate attenuation of incentives promotes easier and better ex post bilateral adaptation.² An important contributing factor to these differences between market and hierarchy is that each mode faces a different system of contract law. The courts treat disputes over prices, quality, delays, and so forth that arise between firms differently than they do identical disputes that arise within firms. They will routinely hear the former, but they refuse to give standing to the latter. In effect, the rule of law that applies to internal disputes (of an instrumental kind) is that of forbearance.³ That

is why fiat is an important instrument for dealing with internal disputes and distinguishes internal from market organization, earlier claims to the contrary notwithstanding.  

The upshot is that the efficiency gains and losses of vertical integration are different from those that Hart and Tirole address. Specifically, the main efficiency trade-off with which vertical integration needs to be concerned is between incentive intensity (where market procurement enjoys the advantage) and bilateral adaptability (where the advantage accrues to internal organization as a condition of bilateral dependency builds up). Possibly these efficiency features will play out very similarly to those of concern to Hart and Tirole when examined in a combined market power–efficiency framework. That, however, is conjectural.

In any event, it has long since been conceded that the die-hard branch of the Chicago School erred on the pure logic of vertical integration and vertical market restrictions: there really can be anticompetitive effects. Although new demonstrations of the die-hard error may add to our understanding, the instinct that antitrust should proceed in the vertical area with great caution is not upset. In the degree to which an elastic application of new models encourages an expansive antitrust enforcement program, errors of excess are certain to result. That is easily avoidable by applying the lessons of the new models in carefully delimited ways.

**General Discussion:** Many of the participants commented on the assumptions underlying the authors’ models. Daniel Spulber emphasized the important contribution made by the Hart-Tirole model. He noted that the model assumes barriers to entry and exit and said that in the absence of such barriers, there is the possibility of entry of efficient competitors who will supply the downstream firm that is left out after


5. The term die-hard Chicagoan is that of Richard Posner, who defines such a person as one “who has not accepted any of the suggested refinements of or modifications in Director’s original ideas” (Posner, 1979, p. 932). Posner is somewhat reluctant to grant that vertical integration could disadvantage rivals but concedes in a footnote that capital costs would be adversely affected (p. 936) and subsequently elevates this admission to the text (p. 945). For a more expansive discussion of this and other possible costs of vertical integration, see Williamson (1974, pp. 1456–63).
a vertical merger occurs. The model also assumes that the strategies of downstream firms are exogenous. Spulber commented that after a vertical merger, the manner in which downstream firms compete might change, which would affect any evaluation of vertical mergers. Spulber also noted that the model used constant returns to scale, and he speculated that if economies of scale existed, monopoly gains from vertical mergers might be even greater.

Michael Salinger praised the model because it allowed two-part tariffs and relaxed assumptions about contracts. He noted, however, that the perfect two-part tariffs that result from the model could not exist under any reasonable set of assumptions. According to Salinger, allowing for perfect two-part tariffs eliminates the success of markup as an issue in vertical integration.

Steven Salop believed that the model was much better than the traditional one of the Chicago School. The Chicago School critique of vertical integration—that no monopoly power could be gained through a vertical merger—was based, he said, on an overly simple model (with fixed proportions, constant returns, and so forth). Salop also noted that Hart and Tirole built a model in which the players are psychologically, legally, and informationally unable to make commitments. Though he believed that contract law made it possible to make and enforce commitments, he was pleased that the authors were working out the implications.

Michael Whinston said that once it is acknowledged that vertical structures involve multilateral supply relations and that there are incomplete contracts, there will be ex post externalities. There will be a difference between what is privately optimal and what is socially optimal, which means that the parties left out of a vertical merger—the other competitors and the consumers—might be hurt by it.
References


Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. 1978. "Vertical


