Extending Market Power through Vertical Integration*

by

Catherine de Fontenay and Joshua S. Gans**

First Draft: 15th October, 1998
This Version: 23rd February, 1999

This paper derives a model of vertical integration when it is difficult to write binding long-term supply price contracts. Thus, a vertically separated monopolist is vulnerable to hold-up. Without integration, we demonstrate that a bottleneck monopolist has an incentive to encourage more firms in a related segment than would arise in a pure monopoly. Having more firms mitigates the hold-up power of any one. This, however, distorts the cost structure of the industry toward greater industry output and, hence, lowers final good prices. Vertical integration mitigates the hold-up problem faced by the monopolist. It allows it to generate and appropriate a greater share of monopoly profits. Horizontal competition mitigates the anti-competitive effects of integration. Journal of Economic Literature Classification Number: L42

Keywords. vertical integration, monopolisation, exclusion, hold-up

* The authors would like to thank Tim Bresnahan, Stephen King, Preston McAfee, Scott Stern, and Jeff Zwiebel. Responsibility for all errors remains our own. The latest version of this paper is available at http://www.mbs.unimelb.edu.au/home/jgans.

** University of New South Wales and University of Melbourne, respectively. Contact e-mail addresses: C.de.Fontenay@unsw.edu.au and J.Gans@mbs.unimelb.edu.au. Contact addresses: Catherine de Fontenay, School of Economics, UNSW, Sydney, 2052, NSW, Australia; Fax: +61-2-9313-6337; and Joshua Gans, Melbourne Business School, 200 Leicester Street, Carlton, Victoria, 3053, Australia; Fax: +61-3-8660-2047.
I. Introduction

This paper examines the strategic motives for a monopolist in one segment of a vertical production chain to integrate into other segments. Strategic integration — as opposed to integrating for reasons of technical efficiency, or to induce investment by a partner — refers to a monopolist’s efforts to extend, or “leverage,” monopoly power into another market. A canonical example has an upstream monopolist, who controls an essential input, integrating and then excluding downstream firms from access to that input. Gas, electricity, local-loop telecommunications, rail and other networks are examples of inputs that cannot easily be duplicated, and are in some instances essential. Competition authorities have long been concerned that monopoly owners of such essential facilities are able to monopolize other markets downstream or upstream by vertically integrating.

Despite these regulatory concerns, their theoretical underpinnings have, until recent times, been weak. The Chicago School, in particular Bork (1978) and Posner (1976), argued that bottleneck monopolists can engender and appropriate monopoly profits from a downstream (or upstream) segment without resorting to integration. This means that vertical integration does not involve any more or less monopolisation than vertical separation. Integration, therefore, occurs only for reasons of technological efficiency and should not be opposed.

Recently, economists have begun to focus on contractual incompleteness as a reason why a vertically separated monopolist is unable (in the absence of integration) to leverage its monopoly power downstream. Rey and Tirole (1996) provide the
representative model of that literature.\(^1\) In their model, an upstream monopolist cannot sign contracts with independent downstream firms that implement monopoly pricing for the final good. This is because it cannot commit to abstain from secretly discounting to any one downstream firm, in a form of postcontractual opportunism. Hence, a given downstream firm realises that, having signed a contract consistent with a downstream monopoly, the upstream monopolist has an incentive to sign contracts with other downstream firms that maximise their joint profits. This would involve output higher than the monopoly level and reduced profits for downstream firms. The only equilibrium involves each downstream firm signing supply contracts consistent with downstream (oligopolistic) competition.\(^2\) In this environment, by giving it a direct interest in downstream profitability, vertical integration helps commit the monopolist not to engage in secret discounting. Integration, therefore, allows the monopolist to extend its power downstream.

The contractual incompleteness that drives this result has two dimensions. First, the monopolist is unable to write supply contracts contingent on downstream measures of profitability (this leads to the potential for secret discounting). Second, the monopolist cannot sign an exclusive contract with one or more downstream firms. Otherwise, it could simply exclude all but one firm, as a means of committing not to offer quantity discounts to other downstream firms. Thus, direct ownership of some (or all) downstream firms is the only way in which a monopolist can resolve its commitment problem and not favour some firms with secret discounts. Of course, integration only imperfectly resolves this commitment problem because the monopolist cannot commit

\(^1\) Other examples include Hart and Tirole (1990), O’Brien and Shaffer (1992) and McAfee and Schwartz (1994).
not to favour its own downstream units when independent units exist (Hart and Tirole, 1990).

We examine the incentives for vertical integration in a contracting environment that focuses on longer-term contracting possibilities. Information problems, such as secret discounts, are assumed away, and all supply contracts are assumed observable. The critical feature is that a monopolist cannot write enforceable long-term price/quantity contracts. Thus, having observed other contractual arrangements, any party can choose to renegotiate contractual terms ex post. The number of downstream firms a monopolist deals with is also non-contractible. So while the monopolist cannot write contracts that exclude downstream firms it can choose to exclude any firm. Because the production technology is assumed concave at the firm level, the monopolist is not indifferent as to the number of firms it supplies, and will never supply an infinite number of firms. The model takes the set of firms with productive capabilities to be fixed during the production phase, as do most models, but assumes that the size of the set was determined by the monopolist’s predicted demand for suppliers. As such, it is quite feasible for the monopolist to supply firms in a manner consistent with maximising industry profits. However, given that, after the number of firms with access to the monopolist’s input has been fixed prices are determined in ex post negotiations, the monopolist chooses its number of suppliers with a view to the outcome of those negotiations. Limiting the number of downstream firms confers bargaining power on those firms in ex post price negotiations. Therefore, the monopolist is subject to ex-post hold-up, the traditional form of postcontractual opportunism (Williamson, 1975). This alters its exclusionary choices.

---

2 This equilibrium involves the selection device of passive beliefs. For a discussion see McAfee and Schwartz (1994).
If the monopolist were not subject to hold-up (e.g., if it could make take-it-or-leave-it price offers giving it all the ex post bargaining power), it would choose the number of firms downstream to minimize the production costs associated with producing the monopoly output.\footnote{Rey and Tirole (1996) assume that the upstream monopolist can make such take-it-or-leave-it offers. Hence, they assume away hold-up problems that arise from exclusion with imperfect price commitment.} In actuality, the monopolist might be better off distorting downstream industry structure (away from its monopoly level) and supplying more firms. Expanding the number of downstream firms reduces the hold-up power of any one, giving the monopolist greater revenues. However, this expansion involves higher fixed costs downstream and, under commonly used production technologies, lower downstream costs at the margin. Ex post profit maximisation, therefore, involves greater production and lower industry profits than would be case under monopoly.

Vertical integration has a powerful role to play in supply interactions. Because other inputs are assumed to be non-specific in this model, asset ownership is the only source of hold-up. Integration consequently commits the monopolist not to renegotiate terms with its own downstream units, and assures it of some production even if independent firms refuse to deal. Hence, integration increases the monopolist’s outside option in ex post price negotiations leading to higher supply prices. More importantly, in so doing, integration reduces the incentive of the monopolist to grant access to a larger number of downstream firms. So the effect of vertical integration is a reduction in industry output and higher final good prices. If there were no efficiency costs to integration, complete integration would become desirable. To focus on strategic considerations, we examine the incentive to integrate when integration is undesirable from an efficiency point of view. When integration is only mildly technically inefficient,
the outcome is partial integration, as the monopolist integrates some of the downstream
firms to depress the bargaining power of remaining firms.

The outline for the paper is as follows. In section I, we set up the model structure
focussing on the case of an upstream monopolist selling to a number of downstream
firms. To determine prices ex post, we model the multi-agent bargaining game using the
framework of Stole and Zwiebel (1996a, 1996b). Their model was concerned with the
situation of a monopsony employer bargaining with a fixed number of workers. The pool
of workers with firm-specific skills was fixed during production, conferring bargaining
power on workers in renegotiations. Stole and Zwiebel solve for the bargaining
outcomes that are robust to bilateral renegotiation. In section II, we consider a case
closest to theirs in that downstream firms are capacity constrained to unit production.
Hence, if a breakdown in negotiations led a downstream firm to exit production, the
upstream monopolist could not substitute for the lost production using remaining
downstream firms. There we demonstrate that, without integration, industry output
exceeds its monopoly level and that integration results in a reversion towards monopoly
output. Section III then turns to the general cost case. In contrast to Stole and Zwiebel’s
framework, some substitution is possible ex post. Nonetheless the monopolist still has an
incentive to distort the number of firms downstream. This leads to greater capitalisation
downstream and hence, higher output ex post than a monopoly situation. Integration
mitigates this effect and is, therefore, welfare reducing. Section IV demonstrates that, in
contrast to other results in the literature, a downstream bottleneck has precisely the same
effect as an upstream one. To examine the effect of horizontal competition, in Section V,
we consider briefly the case where the downstream firms attached to one upstream
supplier (as part of a network) compete with the downstream firms of a rival upstream
supplier. Industry output is higher (under non-integration and integration), integration incentives are diminished, and the integration decisions of upstream firms are strategic substitutes (i.e., integration helps rival upstream suppliers). A final section concludes and offers directions for future research.

II. Basic Set-Up

We consider the case of an upstream monopolist who sells a necessary input to downstream producers. For simplicity, we assume that the upstream monopolist faces no production costs. Firm $i$ requires $q_i$ units of the upstream input and $c(q_i)$ units of another input to produce $q_i$ units of output. The price of the second input is normalised to 1. If downstream firms produce quantities, $\sum_{i=1}^{N} q_i$, this results in a market price for the final good determined by $P = D(\sum_{i=1}^{N} q_i)$.

Let $\pi(q_i, q_{-i})$ be the profit of firm $i$ gross of any upstream input supply price. Industry profit is $\Pi(Q, N) = \sum_{i=1}^{N} \pi(q_i, q_{-i})$, where $Q = \sum_{i=1}^{N} q_i$ and $N$ is the number of downstream firms. Industry profit is maximised at $P^m$, $Q^m$, and $N^m$. Notice that if $c(q_i)$ takes the familiar linear form of $F + \theta q_i$ for constants $F$, $\theta > 0$, then only one downstream producer is used. With convexity in variable costs, production is optimally spread among more firms. If $c(q_i)$ is such that $c(1) = \theta > 0$ while $c(q_i) = \infty$ for $q_i > 1$, we have a capacity constrained technology where industry output equals $N$.

As noted in the introduction, the literature on vertical relations usually assumes that the monopolist can make take-it-or-leave it offers of input supply contracts to downstream firms. If such contracts are publicly observable, the monopolist optimally
selects $N''$ firms and makes an offer to each that ensures that each downstream firm produces only $Q'''/N''$, appropriating the total profit of each downstream firm. This results in the upstream monopoly being extended over the entire market. In this environment, there is no incentive for the monopolist to vertically integrate, or engage in any other form of vertical restriction, in order to appropriate a monopoly profit.

This paper enriches the bargaining game between the monopolist and each downstream firm. In particular, we do not simply assume that the monopolist can make take-it-or-leave-it offers to designated firms. Instead, it must negotiate with each firm individually and faces prohibitively large costs in expanding the number of firms it negotiates with ex post. Our conception here is of an environment in which downstream firms enter into production far less frequently than contracts are renegotiated (as a simplification, they are assumed only to enter once). Therefore, given an initial selection of downstream firms, contracts can be renegotiated at any time.

Specifically, the bargaining game has the following stages:

(i) The monopolist designates $N$ potential firms that it will supply.\(^4\)

(ii) The monopolist engages each firm in one-on-one negotiations over the input supply contract. This bargaining takes the non-binding form as specified by Stole and Zwiebel (1996a, 1996b).

(iii) Production and downstream competition begin.

The rationale behind stage 1 is as follows: while the monopsonist has some choice over the pool of potential firms at the beginning of the game, it cannot contract these choices at stage 1 nor expand them at stage 2 — although, as we will demonstrate, it can reduce

---

\(^4\) This is a truthful (since full-information) statement about the number of firms it wishes to supply in equilibrium. $N$ firms will develop (costless) productive capacity. The model can be extended to the case where productive capacity is costly. That cost introduces a complication in that downstream firms might be at risk of hold-up. Potential hold-up serves to restrict the size of $N$ to be less than the monopolist’s ex-post desired $N$. However, for relatively small capacity costs, the qualitative results do not change.
this pool. Essentially, it cannot easily replace the initial set of firms. This gives those firms some hold-up power.

Bargaining over contract terms takes place in a similar manner to the wage bargaining modelled by Stole and Zwiebel (1996a, 1996b), hereafter SZ. The monopolist and the firm bargain *bilaterally* over the contracts in sequential transactions. Long-term price and supply agreements are not possible. So we look for solutions that are *stable* in the sense that there is no desire on the part of either the monopolist or the individual firm to exercise their respective abilities to renegotiate supply contract terms. While it is possible for different firms to receive different contracts, given the symmetry between them, this does not occur in equilibrium.

### III. Capacity Constrained Case

The specific implications of the Stole and Zwiebel framework are best introduced with a simple model that assumes downstream firms are constrained to unit production. This is very close to the SZ context of a firm bargaining with many workers, who can produce a unit of labor or exit negotiations. Negotiations are sequential and bilateral; as a consequence, the solution is not necessarily within the bargaining core. They assume that the cost of bargaining is that negotiations might exogenously break down with some infinitesimal probability. Binmore, Rubinstein and Wolinsky (1986) demonstrated that two agents, bargaining under an infinitesimal risk of breakdown, agree on the Nash outcome; namely, that each agent receives the same benefit from agreement (where benefit is measured as the gain over what the agent would earn in the event of a
breakdown in bargaining). Negotiations can be reopened at any time, and are reopened if any agent is not receiving their Nash outcome.\footnote{In point of fact, SZ restrict the opportunities for renegotiation in the bargaining process to show that their results require very limited renegotiation. The results hold when more opportunities for renegotiation exist.}

We illustrate the model for the special case where each downstream firm is capacity constrained to produce only one unit of output per period for a cost $\theta$.\footnote{Recall that there are no additional fixed costs here. So $\theta$ may be interpreted as both a marginal cost for the industry and a fixed cost for the downstream firm.} This cost represents the outside option of an individual downstream firm. For the monopolist, when negotiations with an individual firm break down, in addition to not being able to supply that firm, it must also potentially renegotiate pricing arrangements with other firms. So a breakdown involves an inframarginal effect, altering the monopolist’s outside option. To see this, consider the monopolist negotiating with a single firm only over the supply price for a unit of the input, $\tilde{p}(1)$. We will assume that firms produce a homogenous product. If negotiations break down, the monopolist receives 0, while the firm avoids costs $\theta$. If supply takes place, the downstream firm earns revenues of $\pi(1) = P(1) - \theta$. Splitting the surplus (i.e., equating upstream profits of $\tilde{p}(1)$ to downstream profits of $\pi(1) - \tilde{p}(1)$) means that $\tilde{p}(1) = \frac{1}{2} \pi(1)$.

In contrast, consider the case when the monopolist supplies two firms. Each firm bargains \textit{bilaterally} with the monopolist. Nash bargaining implies that each firm splits with the monopolist the surplus created by the relationship -- but what is that surplus? Part of the benefit to the monopolist of supplying a second firm is that it reduces the bargaining power of the first firm. And if negotiations break down with one of the two firms, the monopolist must renegotiate with the remaining firm, who then receives $\pi(1)$ -
\( p(1) \), as above. Thus we obtain the price \( p(2) \) paid by each of the two firms by equating the benefit to the monopolist and to a firm:

\[
\text{Benefits to Monopolist} = \text{Benefits to a Firm} = 2p(2) - p(1) = \pi(2) - p(2)
\]

For the case of \( N \) downstream firms, the same recursive structure applies: one benefit of supplying the \( N^{th} \) additional firm is the amount by which it increases the payments from the remaining \((N - 1)\) firms. Let \( p(N) \) be the negotiated price when the monopolist deals with \( N \) firms.

\[
\begin{align*}
\text{Benefits to Monopolist} &= \text{Benefits to a Firm} = Np(N) - (N - 1)p(N - 1) = \pi(N) - p(N) \\
\Rightarrow \quad p(N) &= \frac{N-1}{N+1} p(N-1) + \frac{1}{N+1} \pi(N) \\
&= \frac{N-1}{N+1} p(N-2) + \frac{1}{N(N+1)} b(N-1)\pi(N-1) + N\pi(N) g \\
&= \ldots \\
&= \frac{1}{N(N+1)} \sum_{i=0}^{N} \pi(i) \\
&= \frac{1}{N(N+1)} \sum_{i=0}^{N} \Pi(i)
\end{align*}
\]

Therefore, the revenue accruing to the monopolist becomes:

\[
Np(N) = \frac{1}{N+1} \sum_{i=0}^{N} \Pi(i)
\]

Stole and Zwiebel demonstrate the robustness of results based on their bargaining mechanisms by considering the outcomes in alternative environments. For example, they allow for asymmetric ex post bargaining power, and heterogeneous agents. They also demonstrate that this bargaining mechanism yields payoffs for agents that are their Shapley values in the corresponding cooperative game. Shapley values have long held intuitive appeal in normative work on bargaining.

It is important to note that the qualitative result is not driven by any sort of “small numbers” property. Very similar results are derived for the case in which there are \( Nh \).
firms, each producing a quantity $h$ of the final good for a cost $\theta h$. In the extreme, one can derive the results for infinitesimally small firms, that is, firms that lies on a continuum $[0,N]$ and produce a total of $N$ units. In this case, the formula is derived from the calculus of variations:

$$N\tilde{p}(N) = \frac{1}{N} \sum_{i=0}^{N} \Pi(i) di .$$

In the current capacity-constrained example, the monopolist’s profit is

$$\frac{1}{N} \sum_{i=0}^{N} P(i) - \theta \zeta(i) ,$$

the average of industry profit (as it is in the discrete case).

This result is possible because bargaining power rests on inframarginal effects (the change in the supply price) rather than the marginal effect of any individual firm. Hence, downstream firms have some hold-up power even in the continuum case. We will assume that the pool of downstream firms is a continuum for the remainder of this section, for clarity of presentation.

**Monopolist’s Output Choice With No Integration**

Turning now to the level of production chosen by the monopolist, we can show that it produces up to the point at which downstream firms do not appropriate any industry rents. Let $\tilde{N}$ be the number of firms (alternatively, output level) that maximises the monopolist’s payoff, denoted $\nu(N) = N\tilde{p}(N)$. This yields the first order condition:

---

7 To see this note that if we equate the Benefits to the Monopolist with the Benefits to the $Nh$th firm, then

$$N\tilde{p}(N) - (N - h)\tilde{p}(N - h) = (\pi(N) - \tilde{p}(N))h$$

This implies that,

$$\pi(N) - \tilde{p}(N) = \lim_{h \to 0} \frac{N\pi(N) - (N-h)\tilde{p}(N-h)}{h} = \frac{\partial}{\partial N} (N\tilde{p}(N))$$

and the differential equation is solved by the stated formula. A formal proof of this is available from the authors.

8 All of the qualitative results below hold for the discrete case.
\[ \frac{1}{N} \Pi(N) - \frac{1}{N} \int_{0}^{\tilde{N}} \Pi(i) di = \frac{1}{N} \Pi(N) - \bar{p}(N) = 0 \]

or \( \bar{p}(\tilde{N}) = \pi(\tilde{N}) \). The profit earned by the monopolist is \( \nu(\tilde{N}) = (P(\tilde{N}) - \theta)\tilde{N} \); that is, it appropriates all downstream profits. The monopolist, anticipating SZ bargaining, differs from the take-it-or-leave-it monopolist by maximising an “average” industry profit rather than profit per se. So the monopolist produces beyond the point at which marginal profit, \( \Pi'(N) \), equals zero.

**Proposition 1.** Let \( \tilde{N} \) be the monopolist’s optimal choice of \( N \). Then, \( \tilde{N} \geq N^m \).

**Proof:** If \( \tilde{N} = N^m \), we would have \( \nu(N^m) = \Pi(N^m) \). However the concavity of industry profits means that \( \Pi(N) \) is strictly increasing until \( N^m \). Therefore, \( \Pi(N) > \frac{1}{N} \int_{0}^{\tilde{N}} \Pi(i) di = \nu(N) \), for all \( N \leq N^m \), a contradiction. It must be the case that \( \tilde{N} \) exceeds \( N^m \).

In this environment, the monopolist is only partly able to extend its monopoly power downstream. At the monopoly output level, \( N^m = Q^m \), the monopolist can raise its own profit by adding an additional firm. Production, therefore, takes place at a level above monopoly levels for the industry. This is because the monopolist expands the number of downstream firms in order to depress the bargaining power of downstream firms as a whole.

**Example 1.** Suppose that \( P(Q) = A - Q \) and \( A > \theta \). Then \( Q^m = \frac{1}{2}(A - \theta) \) while \( \tilde{Q} = \frac{3}{4}(A - \theta) \) compared with a social optimum of \( A - \theta \).

At this point, it is worthwhile comparing this result to previous results concerning contractual incompleteness and vertical relations. In Hart and Tirole (1990), O’Brien and Shaffer (1992), McAfee and Schwartz (1994), and Rey and Tirole (1998), the upstream monopolist is forced to increase output (or lower marginal supply price) so as to convince downstream firms, concerned about post-contractual opportunism, to accept offers. In
that environment, an upstream monopolist has an incentive to restrict the number of
downstream firms. However, it cannot commit to so doing, i.e., it cannot sign
exclusionary, or effectively exclusionary, contracts. Our model internalises all such
concerns as offers are public and the monopolist can choose to restrict the number of
downstream firms. In principle, it could commit ex ante to exclude any firm.
Nonetheless, it has an incentive to expand the number of such firms so as to mitigate the
hold-up power of any one. This is because the monopolist and any individual
downstream firm cannot commit not to renegotiate contractual terms ex post. In effect,
the monopolist is subject to ex post opportunism that it mitigates by making its input
available to more firms. Thus, monopoly power is not perfectly leveraged downstream
because of a lack of commitment on price as opposed to quantity.

*The Incentive to Integrate*

Downstream firms have the most bargaining power when they are few in number,
so that the monopolist is on the steeply rising portion of its profit function. Supplying
many firms moves the monopolist to a downward-sloping portion of its profit function;
thus the marginal productivity of each firm is negative, and the bargaining power of the
first few much attenuated. Heuristically, the monopolist is reducing the probability that
negotiations will break down with many producers, leaving it with a few firms (see Stole
and Zwiebel (1996b) for further intuition).

Overproduction is a particularly costly means of reducing the bargaining power of
downstream firms. Integration allows the monopolist to avoid facing a few firms on a
steep portion of its profit function. The monopolist can acquire a few firms, ensuring
demand from those firms; then independents can never threaten it with the loss of that
demand.
To demonstrate this, we assume away efficiency reasons for integration, by supposing it to be mildly technically inefficient. Then the monopolist will only integrate for strategic reasons. A vertically integrated firm is assumed to face an additional cost, $\Delta > 0$, per unit of production. While we treat this cost as exogenous here, it would be possible to interpret it as indicative of the notion that firms’ incentives to undertake effort — or non-contractible investment more generally — are diminished by integration (as in Williamson, 1985, or Grossman and Hart, 1986). With these assumptions, the stages of the bargaining/production game are as follows:

(i) The monopolist designates $N$ potential downstream firms that it will deal with and integrates $I$ of them.

(ii) The monopsonist and each independent firm engage in one-on-one negotiations over input supply terms. This bargaining is of the SZ type.

(iii) Production and downstream competition begin.

Exposition is simplified by our earlier assumption that there are no costly productive assets; hence there is no ex-ante asset market. It is also simplified because we need only consider the integration decision in the first stage. The reasoning behind this assumption is that downstream firms have ex-post bargaining power, and therefore integration will be costly ex-post if bargaining power extends to bargaining over assets, as seems likely. Therefore, the monopolist will always prefer costless ex-ante integration — that is, integration whose only cost is the loss of efficiency, $\Delta$.

---

9 Integration has such an effect on a firm when it loses “residual rights of control” over its production; in other words, it is no longer the residual claimant for profits. Implicitly we are assuming that integration does not confer the residual rights of control on any of the downstream firms, since it could not do so for more than one.
If \( I \) firms are integrated, we can derive the price paid by the \((N - I)\) independent firms, 
\[
\tilde{p}(N, I) \quad \text{for } N \leq I \leq N - 1.
\]

\[
\tilde{p}(N, I) = \frac{1}{(N - I)^2} \int_{I}^{N} \pi(i)di - \frac{I}{N - I} \pi(N).
\]

The monopolist’s profits \( v \) are now its revenues as a producer of \( I \) units, and the revenue it earns from \((N - I)\) downstream firms:

\[
v(N, I) = (N - I) \tilde{p}(N, I) + I \pi(N) - \Delta.
\]

\[
v(N, I) = \frac{1}{N - I} \int_{I}^{N} \Delta(i)di - \Delta
\]

Note that, after integration, the monopolist’s profits are the average from \( I \) to \( N \), rather than from 0 to \( N \).

Intuitively, one would expect that the profitability of integration is higher, the lower are the costs of integration, \( \Delta \). The following proposition confirms this.

**Proposition 2.** Let \( \tilde{N}(I) \) be the number of downstream firms the monopolist will trade with and \( \tilde{p}(\tilde{N}(I), I) \) the resulting prices. The monopolist’s integration choice is as follows:

(i) \( \text{for } \Delta = 0 \), \( \tilde{T} = \tilde{N} \quad (= N^m) \)

(ii) \( \text{for } 0 < \Delta < \tilde{p}(\tilde{N}(0), 0), \ 0 < \tilde{T} < \tilde{N} \)

(iii) \( \text{for } \Delta \geq \tilde{p}(\tilde{N}(0), 0), \ \tilde{T} = 0 \quad (\text{and } \tilde{N} = \tilde{N}(I = 0)) \).

**PROOF:** Note that the first order condition for the choice for \( N \), given any \( I \leq N \), implies that:

\[\pi(N) - \tilde{p}(N, I) = \lim_{h \to 0} (\frac{N - I - h}{N - I} \tilde{p}(N, I) + \frac{I}{N - I} \pi(N) - \Delta h) - (\pi(N) - \tilde{p}(N - h, I) + \frac{I}{N - h} \pi(N - h) - \Delta).
\]

In one-on-one bargaining, this is equated to the benefit received by the \( Nh \)th firm: \( (\pi(N) - \tilde{p}(N, I))h \). This implies that:

\[
\pi(N) - \tilde{p}(N, I) = \lim_{h \to 0} \frac{(N - I) \tilde{p}(N, I) + I \pi(N) - \Delta h}{h} = \frac{N}{\Delta} (N \tilde{p}(N, I) + I \pi(N)).
\]

Solving this differential equation and re-arranging gives the price in the text.

\(^{10}\) Again, from the calculus of variations, the benefits to the monopolist from the \( Nh \)th firm is the change in price received from independent suppliers plus the change in final price received by the \( I \) integrated units:

\[
((N - I) \tilde{p}(N, I) + I \pi(N) - \Delta) - ((N - I - h) \tilde{p}(N - h, I) + I \pi(N - h) - \Delta).
\]

In one-on-one bargaining, this is equated to the benefit received by the \( Nh \)th firm: \( (\pi(N) - \tilde{p}(N, I))h \). This implies that:

\[
\pi(N) - \tilde{p}(N, I) = \lim_{h \to 0} \frac{(N - I) \tilde{p}(N, I) + I \pi(N) - \Delta h}{h} = \frac{N}{\Delta} (N \tilde{p}(N, I) + I \pi(N)).
\]
\[ \Pi(\tilde{N}) = \frac{1}{N-1} \sum_{i} \Pi(i) di \iff \bar{p}(\tilde{N}) = \pi(\tilde{N}) . \]

Using this fact, we can see that:

\[
\frac{dv(N, I)}{dI} = \frac{\partial v(N, I)}{\partial I} = -\frac{\Pi(I)}{N-1} + \frac{1}{(N-1)^2} \sum_{i} \Pi(i) di - \Delta = 0
\]

\[ \Rightarrow \quad \Pi(\tilde{N}) - \Pi(\tilde{I}) = \Delta(\tilde{N} - \tilde{I}) \]

Note that because \( \Pi \) is concave, \( v \) is concave:

\[
\frac{d^2v(N, I)}{dI^2} = \frac{1}{(N-1)^2} \left[ -(N-1)\Pi'(I) + \Pi(N) - \Pi(I) \right] < 0 .
\]

Therefore the first-order condition implies that \( \tilde{I} = \tilde{N} \) when \( \Delta = 0 \); that \( 0 < \tilde{I} < \tilde{N} \) when \( \Delta \) is small; and that \( \tilde{I} = 0 \) when \( \Delta \geq \bar{p}(\tilde{N}(0), 0) \) (because in that case \( dv/dI < 0 \) for all \( I \)).

The proposition shows that the monopolist chooses partial integration when the efficiency costs of integration (\( \Delta \)) are small, and full integration when there are no such costs. When these costs are zero, complete integration is worthwhile as the monopolist does not face a hold-up problem. It does not need to expand production to reduce the bargaining power of independent firms. However, when \( \Delta \) exceeds average profits without integration, it is never worthwhile to integrate. This is because integration recovers the hold-up value of an individual firm. At the optimal \( N \), this is simply average profits. These results can be illustrated using our running example.

**Example 2.** Returning to our previous linear demand example, it is easy to show that \( \tilde{N} = \min \left[ \frac{4-\Delta + 2\Delta}{2}, \frac{3(A-\Delta)}{4} \right] \) and \( \tilde{I} = \max \left[ \frac{4-\Delta - 4\Delta}{2}, 0 \right] \). Note that as \( \Delta \) goes to zero, all downstream firms are integrated, while for \( \Delta \) too high (\( \geq \frac{4-\Delta}{4} \)), no integration occurs.

The following proposition characterizes the equilibrium in the interesting case; that of small efficiency costs \( \Delta \). Integration raises the equilibrium price at any \( N \), by reducing the downstream firms’ bargaining power. As a result, the monopolist has less
incentive to distort downstream supply towards excess supply, and the equilibrium output \(N\) is lower. Even under partial integration, the monopolist extracts the entire surplus at the equilibrium number of firms, and that surplus is now higher.

**Proposition 3.** For an intermediate value of \(\Delta\) (if \(0 < \Delta < \tilde{\Delta}(\tilde{N}(0),0)\)):

(i) \(\tilde{p}(N,I) > \tilde{p}(N,0), \forall N, \forall I < N\);

(ii) \(N^m < \tilde{N}(\tilde{I}) < \tilde{N}(0)\);

(iii) \(\tilde{p}(\tilde{N}(\tilde{I}), \tilde{I}) = \pi(\tilde{N}(\tilde{I})) > \pi(\tilde{N}(0)) = \tilde{p}(\tilde{N}(0),0)\).

**PROOF:** (i) Integration has the following effect on the negotiated price:

\[
\frac{\partial \tilde{p}(N,I)}{\partial I} = \frac{2}{(N-I)^2} \int \frac{1}{(N-I)} \int_{I}^{N} \Pi(i)di - \frac{\Pi(N) - \Pi(I)}{2} \tilde{p}
\]

Given the concavity of \(\Pi\), the average of its value from \(I\) to \(N\) (the first term in the brackets) is always greater than the average of its value at \(I\) and \(N\) (the second term in brackets).\(^{11}\)

(ii) For intermediate levels of \(\Delta\), the number of firms chosen exceeds the take-it-or-leave-it number, \(N^m\) (Proposition 2). Note that around the chosen level of \(N\), the mixed partial derivative of \(\nu\) with respect to \(I\) and \(N\) is negative. That is,

\[
\frac{\partial^2 \nu}{\partial I \partial N} = -\frac{2}{(N-I)^2} \int \frac{1}{(N-I)} \int_{I}^{N} \Pi(i)di - \frac{\Pi(N) + \Pi(I)}{2} \tilde{p}(N,I) \frac{\partial \tilde{p}(N,I)}{\partial I},
\]

which is negative for all \(I < \tilde{I}\). Hence, starting from \(I = 0\), as \(I\) increases, the marginal return to \(N\) falls. Hence, \(\tilde{N}(I) \leq \tilde{N}(0), \forall I < \tilde{I}\).

---

\(^{11}\) If \(\Pi\) is concave, any two points \(I\) and \(N\) must satisfy

\[\Pi(\lambda N + (1-\lambda)I) \geq \lambda \Pi(N) + (1-\lambda)\Pi(I), \forall \lambda \in [0,1].\]

Summing over all possible values of \(\lambda\) from 0 to 1 yields the following inequality:

\[
\int_{0}^{1} \Pi(\lambda N + (1-\lambda)I)d\lambda \geq \int_{0}^{1} \lambda \Pi(N) + (1-\lambda)\Pi(I)d\lambda
\]

Then a change of variables to \(I = \lambda N + (1-\lambda)I\) yields:

\[
\int_{0}^{1} \int_{I}^{N} \Pi(i)di \geq \frac{1}{N-I} \left( \frac{1}{N-I} \int \Pi(N) - \Pi(I) \right) \tilde{p}(N,I) \frac{\partial \tilde{p}(N,I)}{\partial I}.
\]
(iii) Recall from the proof of Proposition 2 that at the optimal $\tilde{N}$, $\tilde{p}(\tilde{N}) = \pi(\tilde{N})$. Therefore if $\tilde{N}(I) \leq \tilde{N}(0)$, $\tilde{p}(\tilde{N}(I), \tilde{I}) = \pi(\tilde{N}(I)) > \pi(\tilde{N}(0)) = \tilde{p}(\tilde{N}(0),0)$.

IV. General Cost Case

The capacity-constrained case highlights the monopolist’s desire to increase the number of downstream firms so as to mitigate the hold-up power of any one. This effect is less extreme for more general, convex cost technologies downstream. The exit of any one of $N$ firms will not reduce total output by $1/N$, because the monopolist adjusts by increasing the input supply to other firms. But other downstream firms are imperfect substitutes for the exiting firm, and therefore firms retain some hold-up power. As we demonstrate in this section, the monopolist still has a (somewhat mitigated) incentive to supply an excess number of firms, and to partially integrate.

Output Choice under Bargaining

Suppose that downstream firms all have symmetric quasi-convex costs, $c(q_i)$, that include some fixed costs. Integrated firms, however, have an additional cost of $\Delta(q_i)$. We show that without integration, the monopolist biases supply towards more downstream firms, higher output, and hence, lower final goods prices compared with the take-it-or-leave-it case.

Note that it is possible for the monopolist to alter the allocation of output among firms in the event negotiations with any one break down. It can substitute some of the ‘lost’ demand using the remaining firms. However, this substitution is imperfect and
total output does fall. We show that in each round of one-on-one bargaining, the monopolist and downstream firms agree to quantities that maximize industry profit.

**Proposition 4.** Suppose there are \( N \) downstream firms, of which \( I \) are integrated. Let \( q_i \) be the output of a downstream firm that is integrated, and \( q_j \) the output of a non-integrated firm. Under bilateral SZ bargaining, the quantities of input \( \{q_i,q_j\} \) supplied to downstream firms are those which maximize industry profits for a given \( N \) and \( I \). We denote these profits as \( \Pi(N,I) \):

\[
\Pi(N,I) = \max_{q_i,q_j} I \left[ P(q_i + (N-I)q_j) - c(q_i) - \Delta(q_i) \right] + (N-I) \left[ P(q_i + (N-I)q_j) - c(q_j) \right]
\]

**Proof:** Let \( \tilde{p} \) be the bargained price when the monopolist supplies \( \{q_i,q_j\} \), and let \( \hat{p} \) be the bargained price if negotiations break down with one of the \( (N-I) \) independent firms (subscripts are omitted, for clarity). The monopolist maximizes:

\[
\max_{q_i,q_j} \nu = (N-I)\tilde{p}q_i + I\left[ Pq_j - c(q_j) - \Delta(q_j) \right]
\]

subject to the bargaining constraint, namely that the benefit to the monopolist of an additional supplier be equal to the benefit to the supplier:

\[
\nu = \left( (N-I)\tilde{p}q_i + I\left[ \hat{p}q_j - c(q_j) - \Delta(q_j) \right] \right) = \left[ Pq_j - c(q_j) \right] - \tilde{p}.
\]

Reasoning by backward induction, we should take the value of \( \Phi \) as given in the above problem. If, given \( N \) downstream firms, negotiations were to break down irrevocably with one firm, the value of \( \Phi \) would be determined by the monopolist maximizing its profits subject to bargaining with the remaining downstream firms. Monopolist and firms alike expect \( \Phi \) to be so determined, and therefore they bargain taking \( \Phi \) as having that value. The bargaining equation defines \( \tilde{p} \):

\[
\Rightarrow \tilde{p}q_i = \frac{1}{N-I+1} \left[ Pq_i - c(q_i) - I \left[ \hat{p}q_j - c(q_j) - \Delta(q_j) \right] - \Phi \right]
\]

where \( P = Pa_j + (N-I)q_j \).

Therefore, the maximization problem becomes:
\[= \text{Max}_{q_i,q_j} \left[ \frac{1}{N-I+1} \left( Pq_i - c(q_i) - \Delta(q_i) \right) + (N-I) \left( Pq_j - c(q_j) \right) + (N-I) \Phi \right]. \]

The optimal \( \{q_i,q_j\} \) values which are clearly those which maximize \( \Pi(N,I) \).

\textit{Distortions Induced by Bargaining}

Although the choice of output is optimal for any given \( N \), the monopolist has an incentive to distort its ex ante choice of \( N \) so as to maximise the rents it receives, and to integrate for the purpose of reducing distortions. This is demonstrated in the following proposition.

\textit{Proposition 5.} Suppose that \( c(q_i) \) is convex for any strictly positive \( q_i \) and \( \lim_{q_i \to 0} c(q_i) > 0 \).

(i) In the absence of integration, the number of downstream firms the monopolist chooses to supply, \( N_i \), exceeds \( N^m \). Moreover, industry output, \( Q \geq Q^m \). These inequalities are strict for \( c(.) \) strictly convex.

(ii) Propositions 1, 2 and 3 hold for the general cost case as well.

PROOF: Given the assumptions of symmetry and assuming that the set of downstream firms is a continuum on \([0,N]\), the industry profit function becomes:

\[ \Pi(N) = \text{Max}_Q \left[ P(Q)Q - Nc(Q / N) \right]. \]

Observe that, by the envelope theorem, \( \Pi(N) \) is concave, given the convexity of the cost function:

\[ \frac{\partial \Pi(N)}{\partial N} = \frac{\partial \Pi(N)}{\partial N} = -c(Q / N) + c'(Q / N)Q / N. \]

The structure of the industry profit function implies that at \( N \to \infty, Q/N \to 0 \). Therefore the optimal \( N \) will be interior (i.e., will not be infinitely large) so long as there are some fixed costs, i.e., so long as \( \lim_{q_i \to 0} c(q_i) > 0 \).

Recall that \( N^m \) is the number of firms that maximises \( \Pi(N) \), and \( Q^m \) and \( P^m \) are the resulting industry quantity and price. Note that as \( N \) increases beyond \( N^m \) this increases industry output, \( Q \). That is, the marginal profit from \( Q \) is:

\[ P'(Q)Q + P(Q) - c'(Q / N), \]
where we have assumed, without loss in generality, that \( c(.) \) is twice continuously differentiable. The derivative of marginal profit with respect to \( N \) is 
\[ c''(Q/N)Q/N^2 \geq 0. \]
Hence, as \( N \) increases, the output that maximises industry profit also increases. Thus, so long as \( \bar{N} \geq N^m \), \( \bar{Q} \geq Q^m \).

So long as \( \Pi \) is concave, Propositions 1, 2 and 3 carry over immediately to the general-cost case. Note simply that the definition of \( \bar{p} \) and \( v \) in the capacity-constrained case (with or without integration) were written entirely in terms of maximized industry profits, \( \pi \) and \( \Pi \). Proofs of propositions were also in terms of industry profits. The same proofs apply to the general cost case, since we have shown that bargained outcomes bear the same relations to industry profits in the capacity-constrained and general-cost cases (Proposition 4).

As in section I, with general convex technologies, \( \bar{Q} \geq Q^m \). While the monopolist is able to choose a contract that implements \( Q^m \) downstream and hence, generate monopoly profits in the industry, its incentives are biased towards increasing the number of firms downstream so as to avoid being held up by any one of them. While the substitutability between the technologies mitigates some of this hold-up power, it does so imperfectly. As such, the monopolist chooses a downstream industry structure that increases the number of firms and hence, given the convex (variable) costs, it reduces the marginal cost of any given firm. This means that not only does the monopolist spread output over more firms, it has an incentive to increase volume, thereby lowering price and overall industry profit. The result is a more competitive industry structure.

The monopolist has an incentive to integrate, and reduce the expansion in output. Interestingly, however, integration does not distort output choices, as opposed to the findings of the secret-discount model (see Hart and Tirole, 1990). The bargaining power of downstream firms is determined by the monopolist’s profits if they exit production; therefore it generates no incentive to distort output choices once agreement has been
reached. This is a direct (and testable) consequence of postulating Nash-based, and hence efficient, bargaining.

V. Monopsony

An issue addressed by Rey and Tirole (1996) was whether upstream or downstream bottlenecks were more anti-competitive under non-integration. This question is important for industry policy. For example, should consumers deal directly with mobile phone carriers who receive access to a monopoly local loop? Or should gas pipeline owners be able to sell directly to gas users? Rey and Tirole (1996) argue that downstream bottlenecks are more problematic. They reason that while upstream monopoly power is limited by post contractual opportunism, a downstream monopolist is “naturally inclined to ‘internalize’ any negative externality” (p.21) between upstream suppliers.

In our framework, where the problem is renegotiation and hold-up rather than an inability to exclude, there is no difference between the outcomes that result when a monopolist has direct access to final customers and is a monopsonist to upstream suppliers versus the case considered in the previous section. To see this, consider again the capacity constraints case. Now we assume that the production cost $\theta$ is incurred at the upstream stage.

For instance, consider bargaining between the monopsonist and one supplier. If negotiations break down, the monopsonist receives 0 while the supplier avoids costs $\theta$. If supply takes place, the supplier earns $\bar{p}(1) - \theta$ while the monopsonist earns $P(1) - \bar{p}(1)$. 
Splitting the surplus means that \( \bar{p}(1) = \frac{1}{2} I P(1) + \theta \). Therefore the monopsonist’s profits are half of industry profits: \( v(1) = \frac{1}{2} I P(1) - \theta = \frac{1}{2} \pi(1) \).

When there are two suppliers, the monopsonist receives this profit in the event negotiations with one breaks down. If negotiations are successful it pays \( \bar{p}(2) \) to each supplier and earns total profits of \( P(2) - 2 \bar{p}(2) \). Splitting the surplus allows us to solve for \( \bar{p}(2) \):

\[
\begin{align*}
P(2) - 2 \bar{p}(2) - \frac{1}{2} I P(1) - \theta &= \nabla(2) - \theta \\
\Rightarrow \bar{p}(2) &= \frac{1}{2} B(2) - \frac{1}{2} P(1) \hat{g} \frac{1}{2} \theta
\end{align*}
\]

The monopsonist, therefore, earns \( v(2) = \frac{1}{2} I P(2) + P(1) - \theta = \frac{1}{2} \Pi(2) + \Pi(1) \). Working recursively, we have for \( N \) suppliers:

\[
v(N) = \frac{1}{N + 1} \sum_{i=1}^{N} \Pi(i)
\]

the same payoff that the monopolist would receive if it did not have direct access to final customers.

In contrast to the conclusions of Rey and Tirole (1996), the incidence of price from final consumers does not alter the strategic incentives of a monopolist to extend their monopoly power vertically or integrate into other segments. This is because, in either case, the monopolist has precisely the same ability to internalise externalities between firms during negotiations.

**VI. Network Duopoly**

Next we turn to consider the incentives for integration by upstream duopolists. In so doing, we suppose that competition takes place between complete networks, or in
other words, that downstream firms are specific to a particular upstream firm. This allows us to focus on the impact of downstream competition on upstream competition. There are many industry examples of this form of systems competition. For example, many mobile phone networks have independent service providers that retail their products and cannot easily switch to other network operators.

In this environment, the assumed game becomes:

(i) Each upstream firm, \( j \) designates \( N_j \) potential downstream firms that it will deal with and integrates \( I_j \) of them.

(ii) The monopsonist and each independent firm engage in one-on-one negotiations over input supply terms. This bargaining is of the SZ type.

(iii) Production and downstream competition begin.

Once again we assume that upstream firms have only one opportunity to integrate downstream but that price can be renegotiated thereafter.

The final price earned in one network is now a function of own supply and supply from its competitor, \( P(N_j, N_{-j}) \). The payoff to a duopolist \( j \) given the number of downstream firms tied to the other duopolist, \( N_j \), is:

\[
v(N_j, I_j; N_{-j}) = \frac{1}{N_j - I_j} \sum_{i} d(i, N_{-j}) - \theta i di - \Delta I_j
\]

where \( \Pi(N_j, N_{-j}) \) has been spelt out explicitly. Notice that the choices of the other duopolist interact with the duopolist’s payoff only through downstream demand. In particular, as \( N_j \) rises, it reduces the marginal incentive to expand output and the marginal return to integration:

\[
\frac{\partial^2 v}{\partial N_j \partial N_{-j}} = \frac{P_2(N_j, N_{-j}) N_j}{N_j - I_j} - \frac{1}{(N_j - I_j)^2} \sum_{i} d^2 P_2(i, N_{-j}) di \leq 0
\]
\[
\frac{\partial^2 v}{\partial I \partial N_{-j}} = -\frac{P_2(I_j, N_{-j})I_j}{N_j - I_j} + \frac{1}{(N_j - I_j)^2} \int_i P_2(i, N_{-j})di \leq 0
\]

where the inequality follows if it is assumed that \( P_{12} \leq 0 \). \(^{12}\)

Thus, the converse holds true for an increase in \( I_j \), which will reduce \( N_{-j} \), as we have seen. A competitor’s decision to integrate increases the returns to expanding output and integrating. This implies that each upstream firms’ choice of integration are strategic complements. Fundamentally, the reason is that a competitor who integrates reduces its output, leaving the firm with higher effective demand and higher profits. But at the same time as profits increase, so does the bargaining power of downstream firms, and therefore the distortion in \( N \) and \( I \) increases as well.

---

\(^{12}\) The proof of these inequalities are as follows. For the first inequality,

\[
\frac{\partial^2 v}{\partial N \partial N_{-j}} = \frac{P_2(N_j, N_{-j})}{N_j - I_j} - \frac{1}{(N_j - I_j)^2} \int_i P_2(i, N_{-j})di
\]

\[
\leq \frac{P_2(N_j, N_{-j})}{N_j - I_j} - \frac{1}{(N_j - I_j)^2} \int_i P_2(N_j, N_{-j})di \text{ as } P_{12} \leq 0
\]

\[
= \frac{1}{(N_j - I_j)^2} \int_i (N_j - i)P_2(N_j, N_{-j})di \leq 0
\]

And for the second inequality,

\[
\frac{\partial^2 v}{\partial I \partial N_{-j}} = \frac{-P_2(I_j, N_{-j})}{N_j - I_j} + \frac{1}{(N_j - I_j)^2} \int_i P_2(i, N_{-j})di
\]

\[
\leq \frac{-1}{(N_j - I_j)^2} \int_i P_2(i, N_{-j})di + \frac{1}{(N_j - I_j)^2} \int_i P_2(i, N_{-j})di \text{ as } P_{12} \leq 0
\]

\[
\leq \frac{1}{(N_j - I_j)^2} \int_i ((i - I_j)P_2(i, N_{-j})di \leq 0
\]
Example 3. Returning to our example, once again assume a homogenous goods industry with linear demand. With SZ bargaining, an upstream firm and its downstream retailers agree to contractual terms that maximise their joint profits given the decision of the other upstream firm. In equilibrium, an upstream firm will not designate a downstream firm it does not intend to utilise. Hence, one can conceive of upstream firms in quantity competition with one another, acting to maximise -- not profit -- but their average profit. Given this, in equilibrium, $N_j = \min \left[ \frac{A - \theta + 2\Delta}{3}, \frac{3(A - \theta)}{2} \right]$ and $I_j = \max \left[ 0, \frac{A - \theta - 7\Delta}{3} \right]$. Industry output is, therefore, $Q = \min \left[ \frac{2(A - \theta + 2\Delta)}{3}, \frac{6(A - \theta)}{7} \right]$ compared with $\frac{2(A - \theta)}{3}$ under Cournot duopoly when upstream firms have all the bargaining power. Compared with the monopoly case, a smaller proportion of output produced is produced by integrated firms and the critical cost threshold of $\Delta$ that leads to no integration is lower.

Example 4. We can extend the above example to the case of $m > 2$ upstream firms. In this case, in equilibrium, $N_j = \min \left[ \frac{A - \theta + 2\Delta}{1+m}, \frac{3(A - \theta)}{1+3m} \right]$ and $I_j = \max \left[ 0, \frac{A - \theta - (1+3m)\Delta}{1+m} \right]$. Industry output is, therefore, $Q = \min \left[ \frac{m(A - \theta + 2\Delta)}{1+m}, \frac{3m(A - \theta)}{1+3m} \right]$ compared with $\frac{2(A - \theta)}{1+3m}$ under Cournot duopoly when upstream firms have all the bargaining power. Note that as $m$ gets large, outcomes become perfectly competitive. More importantly, however, for strictly positive $\Delta$, no integration is chosen when $m$ rises above a finite number $\frac{A - \theta - \Delta}{3}$.\textsuperscript{13}

In summary, our discussion of duopoly here demonstrates the effect of indirect competition upstream (namely, when its effect is felt only through downstream competition). Integration by one upstream firm has a positive externality on the other. Given the (social) inefficiency of integration, competition fails to restrict the otherwise socially harmful incentives to integrate.\textsuperscript{14}

VII. Conclusions and Directions for Future Research

It is not a controversial proposition in economics to suggest that vertical integration can alleviate hold-up problems. The literature on contractual incompleteness and asset ownership (beginning Grossman and Hart, 1986), is explicitly concerned with

\textsuperscript{13} Garvey and Pitchford (1995) also consider the effect of integration as competition on either side of the market varies. They demonstrate that perfect competition on either side of the market results in efficient
the effect ownership has on ex post price negotiations. They demonstrate that integration changes the outside options of parties and hence, their incentives to undertake relationship-specific investments that are themselves non-contractible.

The rationale behind vertical integration in this paper is essentially the same. Instead of having a non-contractible investment, the monopolist has market power and can choose to limit supply to downstream firms in order to raise price. This choice, however, is not contractible and any rents earned are subject to hold-up by those downstream firms. By supplying more of them, the monopolist can reduce the bargaining power of those firms. This, however, results in overcapitalisation and too much output downstream. Integration, by changing the outside options of downstream firms, raises the return to the monopolist from exercising its market power by limiting supply. Hence, even if technically inefficient, vertical integration can be a profitable strategy for the monopolist to leverage its market power downstream.

This model of vertical integration sits well with the empirical literature. One prediction of the model is that the number of downstream firms is fewer in a market with some vertical integration compared to a market with complete separation. In studies of the U.S. cable industry (where the bottleneck is downstream), Waterman and Weiss (1994) and Chipty (1995) demonstrate that the number of channels on local networks that are owned by channel providers is less than on independently owned networks. de Fontenay (1997) found that monopsony sugar mills in Honduras bought up a number of their supplying farms when a new law permitted integration. Integration significantly reduced the price paid to remaining independent farms.

integration decisions. Their analysis is, however, based on transactions costs arising from agency considerations. Moreover, they do not consider, as we do, the effect of integration on downstream prices. 14 We leave for later work the case of duopolists in direct competition for downstream suppliers.
The approach holds promise for re-considering other motives for integration besides monopoly leverage. The theory of vertical foreclosure, for one, suggests that a supplier might integrate downstream to harm his competitors (see, among others, Ordover, Saloner and Salop, 1990, and Hart and Tirole 1990). In this paper, we examined the case of competition between two networks composed of one upstream and several downstream firms. This simple case showed that the foreclosure literature has overlooked possible positive effects of integration on competitors. Future work would involve extending our framework to consider what happens when downstream firms can switch between upstream suppliers. This, however, involves departing from the Stole-Zwiebel framework as it currently stands. Another literature examines the incentives for integration, in the presence of non-contractible investment (most famously, Hart and Moore, 1990 and Grossman and Hart 1986 and Williamson 1985; for a model that considers such motives in the traditional theory of supply assurance, see Bolton and Whinston 1992). Integrating this stream of the literature with our own is also a potentially fruitful area for future research.
References


