

VERTICAL INTEGRATION AND DIFFERENT INPUTS

M.E. Porter
A.M. Spence

NUMBER 120

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK
COVENTRY

VERTICAL INTEGRATION AND DIFFERENT INPUTS

M.E. Porter
A.M. Spence

NUMBER 120

November, 1977.

This paper is circulated for discussion purposes only and its contents should be considered preliminary.

VERTICAL INTEGRATION AND DIFFERENT INPUTS

M. E. Porter and A. M. Spence
Harvard University*

Introduction

Vertical integration decisions, which determine the boundaries of the firm, are among the largest investment decisions firms make and can have a major influence on firm success. In some cases the failure to integrate can result in the inability to obtain the desired inputs at reasonable prices. In other instances, an integrated firm may be severely affected by economic fluctuations due to overhead that the unintegrated firms do not bear. Research on vertical integration has stressed the balance between the inflexibility of integration and the cost reduction and/or potential increase in market power that integration may allow.

In case studies, however, executives often report that their motivation for integration is to obtain products with a desired set of attributes, ones that are not offered on the market. Desired attributes can include product design, durability, performance limits, adherence to tolerances or specifications, timeliness and reliability in delivery, and the amount of customer service. For example, IBM is reported to have integrated backward into products that its equipment uniquely required, while purchasing more standardized products on the open market.¹ Polaroid integrated backward into batteries for its instant developing film because it could not obtain batteries with the required specifications and reliability from its suppliers.²

The general point that emerges is that in a variety of product dimensions, firms seem to find it desirable to integrate to obtain inputs that are tailored to the particular needs of their manufacturing and distribution requirements, and that these inputs are difficult or impossible to obtain in the market. The purpose of this paper is to identify the economic basis for

this type of vertical integration. We will discuss by means of a model and examples the structural determinants of the incentive to integrate for these reasons. In addition to establishing the structural determinants of the need to integrate to obtain desired inputs, we will provide a rationale for the observed differences in the level of vertical integration of firms within a given industry and for changes in the level of integration of given firms over time. Industry and firm characteristics associated with these differences will be identified. Finally, we will assess some of the welfare aspects of vertical integration to obtain inputs with the desired characteristics.

1. Integration to Obtain Inputs: Basic Theory

The first question we must answer is "Why might a firm find it necessary to integrate to obtain the inputs it requires?" The most basic answer to this question is much the same as the explanation of the failure of a monopolistically competitive consumer goods industry to supply the full range of products. The benefits to the buyers exceed the revenues (under the price system) to the sellers. Thus revenues may fail to cover costs, while the benefits if they could be appropriated would justify the costs. Vertical integration eliminates revenues as a test of desirability by making the buyer and seller the same unit.

To illustrate, let us assume a world with no uncertainty where the input in question has a marginal cost of c and a fixed cost of F . These cost conditions are assumed to be equally applicable to both the unintegrated upstream firm and downstream firm that integrates backward. The downstream firm is the sole demander of the particular variety of the input in question; it has an inverse demand $S(x)$. There are no substitute inputs available at equal or lower costs.

The input product cannot be profitably supplied if the average cost $c + \frac{F}{x}$ exceeds the inverse demand, $S(\bar{x})$. That is, there will be no market for the input if for all x

$$S(x) < c + \frac{F}{x}$$

That does not imply, however, that the net benefits of the input are always less than zero.

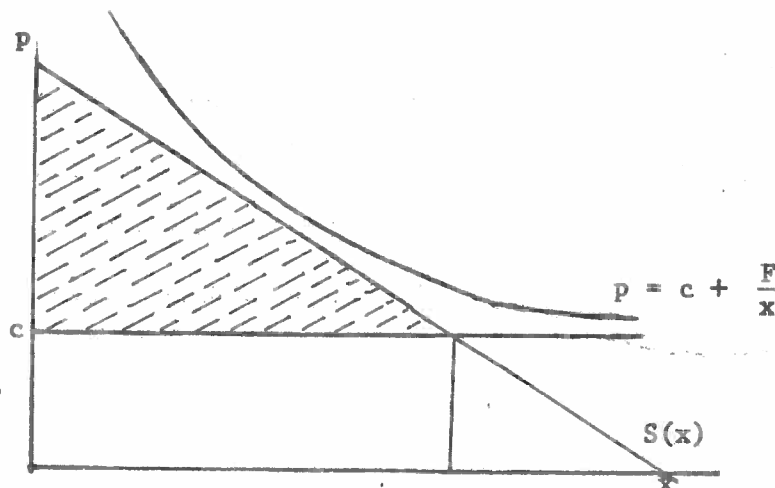
The profits of the upstream firm are $\pi(x) = S(\bar{x}) - c\bar{x} - F$.

The net benefits to the downstream are

$$\begin{aligned} \text{NB} &= \int_0^x S(v)dv - cx - F \\ &= \left[\int_0^x S(v)dv - xS(x) \right] + [xS(x) - cx - F] \\ &= \left[\int_0^x S(v)dv - xS(x) \right] + \pi(x) \end{aligned}$$

Provided the demand is downward sloping (i.e., not infinitely elastic), the first term is positive. This term is the profit of the downstream firm when it is not integrated. [This follows from the fact that the inverse demand for an input is equal marginal revenue of that input, for a profit maximizing firm.] Thus, while the second term may be negative for all x , the sum of the two is not necessarily negative. Assuming NB is positive, the benefits of the product can be obtained by vertical integration. Figure 1 illustrates the problem.

Figure 1



Provided the shaded area is greater than F , the input is worth more than it costs to the downstream firm, even though it could not be simply priced to make money by an independent seller.⁴

Although something of a simplification, the example illustrates a number of points. Increasing returns may make it necessary to integrate to obtain an input with the desired attributes. Integration also yields benefits because with increasing returns, the market price must exceed the marginal cost of the input. The integrated firm can attach the appropriate marginal cost shadow price to the input while the nonintegrated firm cannot. This second point has been discussed by Schmalensee (1973) and Perry (1975) who argued that a firm will integrate to avoid a divergence between a market price of an input and its marginal cost.⁵

If there were N firms each with inverse demands of $S(x)$ for the same input, the market inverse demand in Figure 1 would shift up. The prospects for the market improve because there are more firms among whom to distribute the overhead, fixed costs, F . Integration under these circumstances becomes less attractive.

2. The Case of a Standardized Product Alternative

There remain a number of questions associated with market interaction and the integration decision which we can illustrate by relaxing some of our earlier assumptions. In practice, the downstream firm will often face a tradeoff between producing an input which maximizes benefits and buying an

imperfect substitute. The downstream firm faces a cost of producing the unique input internally, which we assume is not supplied by the market as above, and compares this cost to the availability and market prices of imperfect substitutes. The interaction can become somewhat complicated. For example, as the demand for inputs by the individual firm grows, the benefits of integration (and producing the unique input) increases, but the market price of the imperfect substitute falls because the demand, over which to spread fixed costs, has increased. It is not, therefore, clear whether industry growth is likely to be accompanied by integration or disintegration.⁶

The question of imperfect substitutes is usefully thought of in the following way. The possible inputs are points in a space of attributes, such as plants have geographic locations. The downstream firms may prefer inputs with different attributes, and may integrate to obtain these inputs. But an upstream firm may supply a "compromise" product, one which is no firm's first choices because its fixed costs can be spread over a number of downstream firms. Intuitively, a compromise product is likely to be sold to the set of firms whose first choices form a cluster in the attribute space. Firms whose first choices are relatively isolated may integrate.

The fully general analysis of integration, starting from the space of attributes, is analytically too complicated to generate intuitive insights. Thus to examine the structural influences on vertical integration decision in this context, we consider a model of the following type. There are N similar downstream firms; they may buy a standardized input, which can be thought of (in the spirit of the literature on television

programming) as a common denominator product which approximates the desired attributes of the group of firms though not maximizing the net benefits to any of them. ⁴⁷ The unit cost of the standardized input is one, and the fixed cost F. The demand for this product by the representative firm is Ap^{-n} , where A is a scale factor and n is the elasticity. The elasticity is assumed to be greater than one. The benefits of the product to the buying firm when the price is p are, therefore,

$$\frac{A}{n-1} p^{1-n} .$$

Note that the derivative of the benefit function is minus the demand.

Each firm has an alternative input which is tailored to its specific needs. The units of this unique input are scaled so that its marginal cost is also one, and its fixed cost is assumed to be F. Thus the cost function of the unique input is assumed to be equal to that of the standardized input, and invariant to whether it is produced by an outside supplier or within an integrated firm. In particular, the outside supplier cannot produce the tailored product at a lower cost, as a result of producing either the standardized product or other tailored products. This is a special assumption, and one that will be relaxed later. The firm specific input has a demand θAp^{-n} where $\theta > 1$, and thus the benefits and the demand are greater at each price for the firm specific alternative to the common denominator product. For the moment, the firms are assumed to be the same size. Later we shall relax this assumption to allow the downstream demands to be distributed with respect to size.

Given a market price of p, the firm will compare the benefits of buying from the market,

$$\frac{A}{n-1} p^{1-n}$$

with the net benefits from integration. The benefits from integrating and producing the specialized product are

$$\frac{\theta A}{n-1} p^{1-n} - Ap^{-n} (p-1) - F$$

when the internal shadow price is p . These benefits are maximized when $p = 1$, i.e., the internal shadow price is equal to the marginal cost. The result is a total benefit from integrating of

$$(2) \quad \frac{\theta A}{n-1} - F.$$

The firm will integrate and produce the product that is tailored to its needs if

$$\frac{\theta A}{n-1} - F \geq \frac{A}{n-1} p^{1-n}.$$

Thus the highest market price consistent with a preference for no integration is

$$(3) \quad \hat{p} = \left[\theta - \frac{F(n-1)}{A} \right]^{\frac{1}{1-n}},$$

that being the price at which (1) and (2) are equal.

The term in square brackets will be denoted by ϕ :

$$\phi = \theta - \frac{F(n-1)}{A}$$

The parameter ϕ figures prominently in what follows, and therefore deserves some comment here. Suppose the number of downstream firms was very large, so that the average and marginal costs of the upstream firm were virtually the same

Then the market price for the standardized product would be one (or very close to one) and the surplus for the typical downstream firm would be $\frac{A}{n-1}$. The surplus for the integrated downstream firm is

$$\frac{\theta A}{n-1} - F .$$

The parameter ϕ is the ratio of the latter to the former:

$$\phi = \left(\frac{\theta A}{n-1} - F \right) / \left(\frac{A}{n-1} \right)$$

Thus ϕ is a summary measure of the pure advantage or disadvantage of integrating to produce the specialized input. The actual ratio of benefits, when the market price is p , would be ϕp^{n-1} , a number that increases with p . The parameter ϕ captures the combined effects of θ , F , A and n on the relative merits of integrating. This is useful since it reduces the number of parameters that one needs to keep track of, and makes diagrammatic treatment of the subject possible.

Let us return to the integration decision. From (3), the highest price the downstream firm(s) will pay for the standardized input before integrating is

$$\hat{p} = \phi \frac{1}{1-n} .$$

If $\phi > 1$, then since $n > 1$, $\phi \frac{1}{1-n} < 1$, and integration will surely result because the market price cannot fall short of marginal cost. This will occur if the advantage of the firm specific product outweighs the cost of absorbing all the overhead.

If the downstream firms buy from the market at price p , the profits of the seller are:

$$\begin{aligned} \pi &= NA(p - 1)p^{-n} - F \\ &= NA \left((p - 1)p^{-n} - \frac{F}{NA} \right) \end{aligned}$$

where N is the number of downstream firms. Noting from (3) that

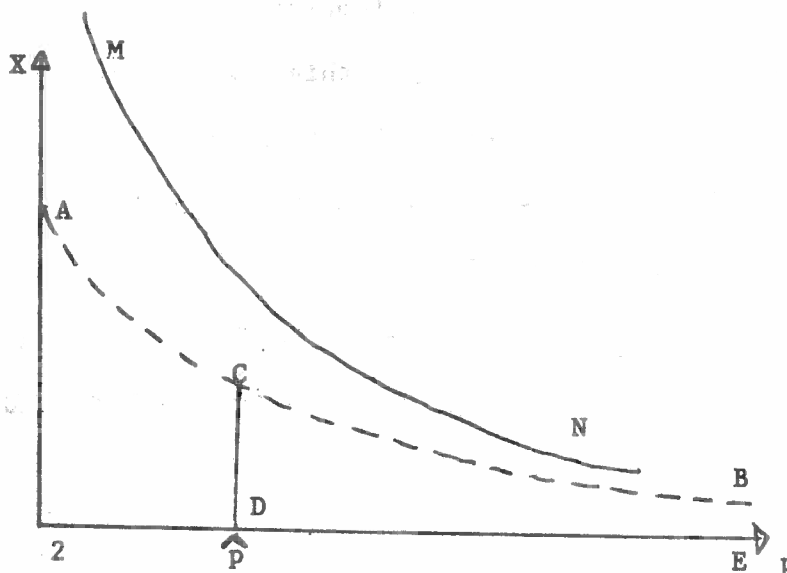
$$\frac{F}{A} = \left(\frac{\theta - \phi}{n - 1} \right),$$

profits can be written

$$\pi = \left[(p - 1)p^{-n} - \left(\frac{\theta - \phi}{N(n - 1)} \right) \right] NA.$$

The central issue is whether an upstream firm can supply the standardized product at a price low enough to prevent integration, and still earn non-negative profits. The problem can be depicted in a manner analogous to Figure 1. In Figure 2, the line AB is the demand function of the representative firm for the standardized product. However, because of the availability of the firm specific product, demand is truncated at $\hat{p} = \phi^{\frac{1}{1-n}}$. Thus the upstream firm faces an effective demand from individual downstream firms of ACDE. Let $D(p)$ be this effective demand.

Figure 2



An upstream firm can supply the market and earn non-negative profits if at some price p

$$(p-1)ND(p) > F,$$

or

$$(4) \quad D(p) > \frac{F}{N(p-1)}$$

The line MN depicts the right hand side of (4). As drawn, it is everywhere above the ACDE. Thus integration would be the outcome.

Recall that ϕ is the ratio of the profits from integration to the profits from the standardized good, if the latter had a price of one. As ϕ increases \hat{p} falls. That clearly increases the chances of integration by reducing the effective demand for the standardized product. The effect of a shift in the elasticity, with ϕ held constant, is more complicated, as we shall see shortly.

We turn now to the conditions under which integration will occur. It is useful to begin by noting that the profit maximizing price for the upstream firm, if integration were disallowed, would be

$$p^* = \frac{1}{n-1}.$$

There are two cases that require attention.

Case A: If $p^* < \hat{p}$: In this case the standardized product will dominate integration if $\pi(p^*) > 0$. That is, if the reservation price is above the monopoly price, then the market can provide the standardized product if profits are non-negative at the monopoly price.

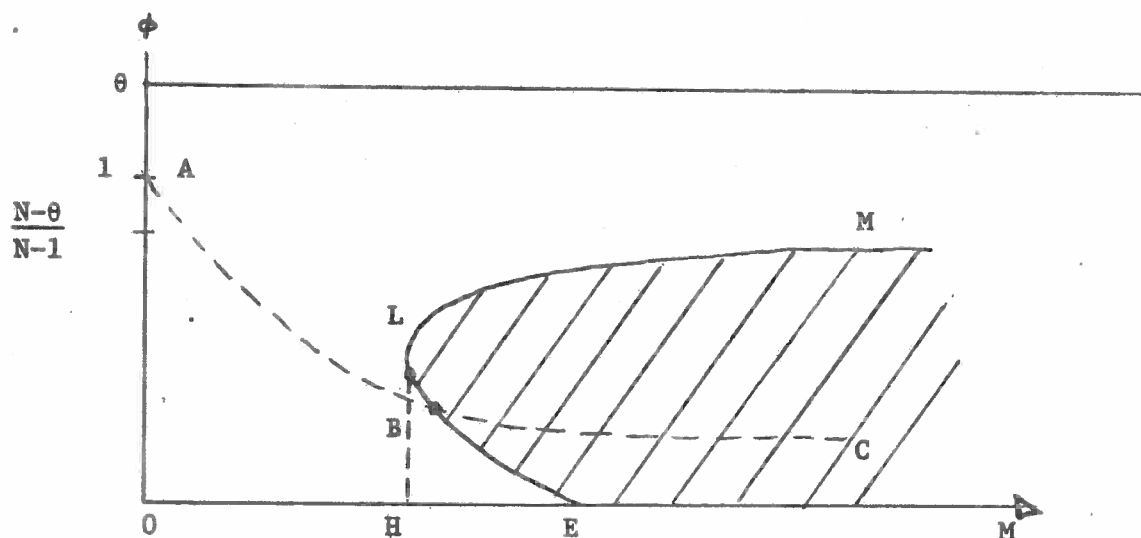
Case B: $p^* > \hat{p}$: In this case, the reservation price is below the monopoly price. The standardized product will be preferred to integration if

$$\pi(\hat{p}) > 0.$$

These two cases divide the space of possibilities into two parts, the market solution and the integration solution.

It is useful to take the parameters of the model to be ϕ , a measure of the natural advantage of the firm specific input, and the elasticity of demand. We have plotted the regions in which integration occurs or the standardized product is supplied by the market, in the two dimensional space with n and ϕ on the axes. While the details of the argument are not especially illuminating (and have been relegated to our appendix), the qualitative results are interesting. Figure 3 is an example of the kind of picture that emerges. The picture is drawn for a fixed number of firms. Some of its characteristics change as N changes, as will be discussed below.

Figure 3



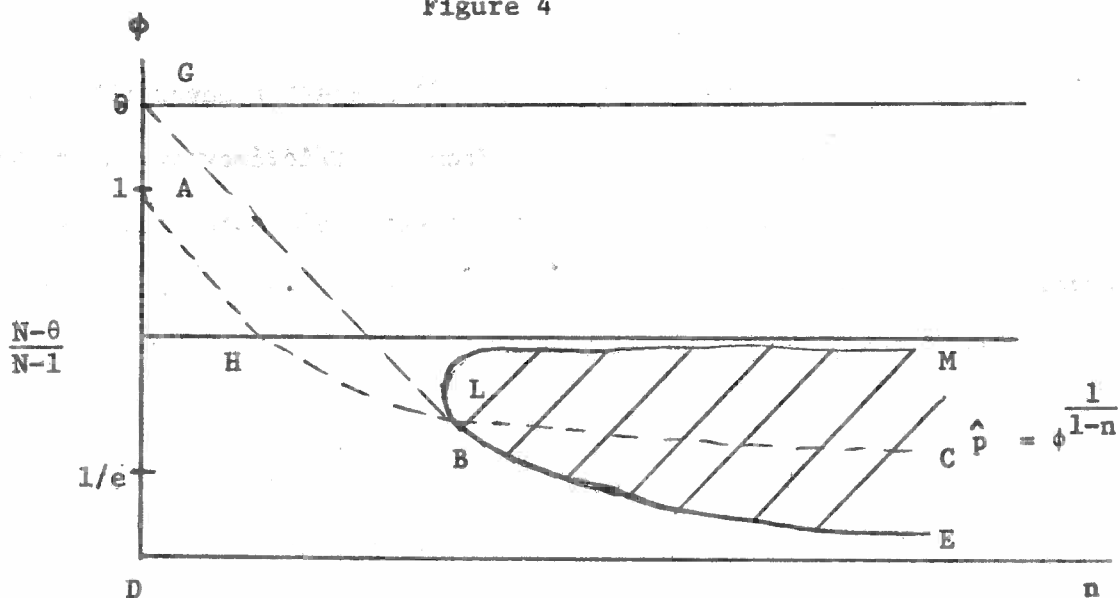
The points in the shaded area correspond to the parameter values at which the standardized product will prevail. All the remaining points below $\phi = \theta$, will result in integration. Note that θ is the maximum possible value of ϕ . It occurs, for example, when $F = 0$.

The Figure illustrates several points. Integration occurs either when ϕ is relatively large or when the elasticity, n , is small. The maximum value of ϕ that is consistent with the market outcome is $\frac{N - \theta}{N - 1}$, and then the market solution occurs only for very large elasticities. For n in the interval HE, it is possible for the market to pass through successive stages of integration, disintegration, and integration again, as ϕ rises. The parameter ϕ rises, for example, when A , the demand scale parameter, rises. Therefore, increases in firm size do not have an unambiguous effect on the integration decision.

The boundary of the "market" region actually has two parts. On the segment BE, the monopoly price for upstream firms is below the downstream firm's reservation price (case 1 above), and profits are zero at the monopoly price. On BLM, the monopoly price is above the reservation price (case 2 above), and profits are zero at the reservation price. The line ABC is the set of points where the monopoly price and the reservation price are equal. Below ABC, the downstream firms would cheerfully accept the monopoly price and the standardized product. But in the non-shaded area below ABC, the upstream firm cannot earn positive profits even at its profit maximizing price. Above ABC, $\hat{p} < p^*$, so that the downstream firms will integrate if the price is above \hat{p} . In the non-shaded area above ABC, profits for the upstream firm are negative at \hat{p} . Thus integration occurs in some cases because the upstream firm cannot make profits at any price, and in other cases because integration is attractive enough to preclude upstream profits at a price that is low enough to deter integration. Qualitatively, these are quite different reasons for integration.

When the number of firms is relatively small, the picture changes to that in Figure 4.

Figure 4



The principal difference between Figures 3 and 4 is that the market region does not touch the horizontal axis. This means that for ϕ sufficiently small, the market cannot earn positive profits for any elasticity. The reason is relatively straightforward. The profitability of the upstream firm is determined largely by the ratio F/A of fixed costs to individual downstream firm size, and N the number of downstream firms. The upstream firms will be unprofitable when F/A is high or N is low or both. Now $\phi = \theta - \frac{F(n-1)}{A}$. Thus, when ϕ is low, F/A is high. Thus when ϕ is low, and N is low, the upstream firm is unprofitable. The elasticity plays a role in this. For a given ϕ and N , the chance of a positive profit rises with n . But a high elasticity cannot compensate completely for high fixed costs in relation to sales. ⁸

In the special case where $N = 1$, the market is never the outcome, given this set of considerations. From an efficiency point of view as well as in the interest of profits, an integrated monopoly is preferable. For in the case of the monopoly, the benefits of distributing fixed costs among several firms is not an issue.

It is worth noting also that the preceding model can be applied to the case in which the downstream firms integrate to produce the same product as the upstream firm. That is, the input they produce internally is not unique but is the standardized product. In the model, one simply sets $\theta = 1$, so that the firm specific products are undifferentiated from what we have been calling the standardized product. Setting $\theta = 1$ will weaken the case for integration, but not eliminate it. Even with $\theta = 1$, the integrated downstream firm still has the benefit of "buying" the input at a marginal price of one, the marginal cost.

Welfare in the Symmetric Case

The standardized, compromise product, if appropriately priced at marginal cost, yields total net benefits of

$$\frac{NA}{n-1} - F.$$

The firm specific input generate benefits of

$$\frac{NA\theta}{n-1} - NF.$$

Using the fact that $\phi = \theta - \frac{F(n-1)}{A}$, the compromise product is preferable

if

$$\phi < \frac{N-\theta}{N-1}$$

As can be seen from Figures 3 and 4, this region contains the region where integration does not occur. Therefore, the market failure consists of integration occurring under conditions where it should not occur. From the figures one can see that this failure is more likely the lower is the downstream elasticity of demand for the input.

The welfare maximizing outcome would obtain if the standardized input were offered with a fixed charge of F/N , and a margin price of one (the marginal cost). This points up a familiar argument in vertical integration theory. A suitable contract, in this case a set of two part tariffs, would in theory accomplish the desired objective. Later, we shall see that, when the downstream firms are of different size, the appropriate pricing is discriminatory.

The reason vertical integration is excessive primarily for low demand elasticities is that when elasticities are low, the upstream firm, using the price system, appropriates a lower fraction of the benefits that it does in the case of high elasticities. This is the analogue of a result in the theory of monopolistic competition.⁹ As a result, with low elasticities the upstream seller has more difficulty achieving positive profits.

3. Downstream Firms of Different Sizes

If the downstream firms are distributed with respect to the size of their input demands, then there are some additional descriptive and normative issues that are of interest. The impact of having a size distribution of downstream firms is not difficult to discover. As compared with the symmetric case, the market with a non-degenerate size distribution will have smaller demands for the standardized product at low prices, and larger demands at high prices.

To illustrate, let us suppose that the parameter A in the input demand function is distributed according to the cumulative distribution, $G(A)$. For any \bar{A} , $G(\bar{A})$ is the number of firms with a demand parameter, A , equal or less than \bar{A} . Let N be the total number of firms and \bar{A} be the mean size. At a market price of p , a firm of size A will select integration if

$$A > \frac{F(n-1)}{\theta \cdot p^{1-n}}$$

Thus at price p , the market demand is

$$D(p) = p^{-n} \int_0^{\left[\frac{F(n-1)}{\theta \cdot p^{1-n}} \right]} AdG(A) .$$

Let

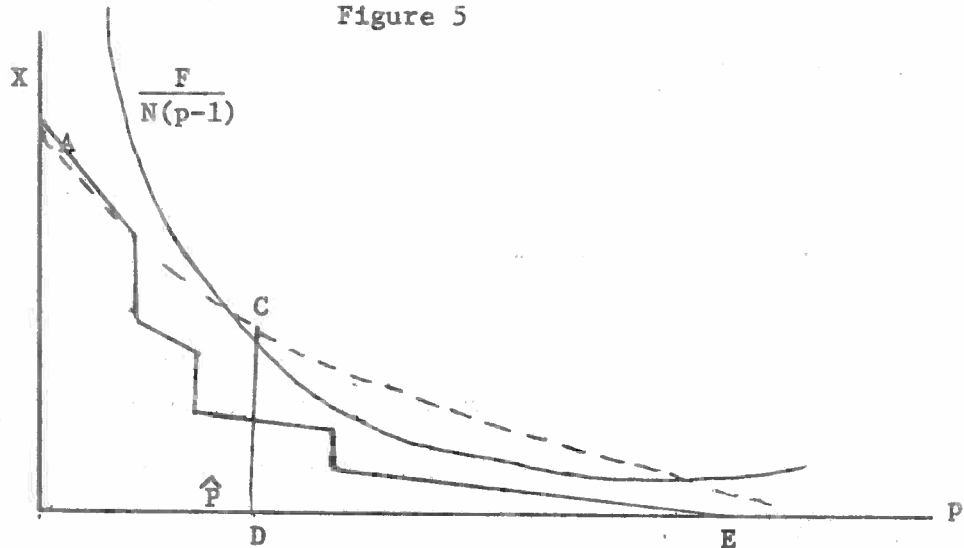
$$B(p) = \frac{1}{N} \int_0^{\left[\frac{F(n-1)}{\theta \cdot p^{1-n}} \right]} AdG(A) .$$

The condition for the market in the standardized good to be profitable is that

$$p^{-n} B(p) > \frac{F}{N(p-1)} .$$

Let $\hat{p} = \left[\theta - \frac{F(n-1)}{\bar{A}} \right]^{\frac{1}{1-n}}$, the analogue of the reservation price in the symmetric case. If all firms were of the mean size, then $B(p) = \bar{A}$ for $p \leq \hat{p}$, and $B(p) = 0$ for $p > \hat{p}$. In the non-symmetric case, large-r firms have lower reservation prices. Thus, the industry demand curve has a sequence of discontinuities at the reservation prices of the various firms. The comparison is depicted in Figure 5.

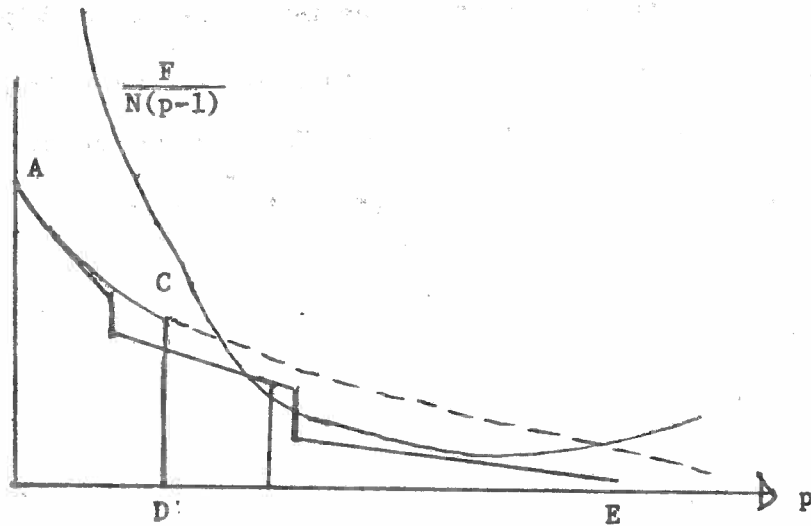
Figure 5



The line ACDE is the market demand when all N firms are of size \bar{A} , the mean. When they are not, the demand is the line with downward steps. Thus demand is lower for $p < \hat{p}$, and higher for larger values of p . As drawn, the line $\frac{F}{N(p-1)}$ goes underneath ACDE, so that the symmetric case would produce no integration. With the asymmetric case, integration will result in this example.

It is clear that the outcome could go the other way, but with an important qualification. Figure 6 depicts a case in which the standardized product is sold to smaller firms, where integration would have occurred with equal sized firms.

Figure 6



It is important to note, however, that in this equilibrium the large firms have integrated, and only the smaller firms buy from the market. The market price is high, reflecting the reduced demand associated with the integration of the large firms. Thus a size distribution, even when it salvages the market, may not salvage the entire market, but only a fraction of it.

With a size distribution, the larger firms will generally integrate. The smaller ones may be forced to integrate. Very small firms may find integration unprofitable. If $\frac{A \theta}{n-1} - F < 0$, or $A < \frac{F(n-1)}{\theta}$, then integration is not a viable strategy. Such firms will buy from the market if there is a market, or go out of business.

Thus, there is an externality associated with the integration decision of the larger firms. Integration increases the costs of smaller firms and reduces their profitability. Integration by large firms therefore can increase the height of the entry barriers, by raising the price of the standardized good, thereby increasing the minimum profitable scale.

To summarize, the greater the variance in the size distribution, the more likely is integration. The integration may be partial, that is, undertaken only by large firms. This sort of integration may drive small firms out of business in the manner described above, thereby increasing industry concentration. Or it may lead to a market in the standardized product serving small firms, who will operate at a cost disadvantage with respect to larger integrated firms. Whether integration will increase concentration or leave an unintegrated fringe with a cost disadvantage will depend on the extent of returns to scale. In assessing the competitive consequences of the integration by larger firms, it is also important to note that the larger firms are making inputs tailored to their needs, which may imply quality differences between the products of large and small firms relegated to the standardized input.

Welfare When Firms Are of Different Sizes

The welfare analysis of the case where firms are of different sizes is as follows. Large firms, those for which

$$A > \frac{F(n-1)}{\theta - 1}$$

should and will integrate. There is no welfare problem with respect to them, and we can, conceptually, simply remove them from the market.

Let us assume that all firms have $A \geq \frac{F(n-1)}{\theta}$, so that if the compromise product is not offered, all firms can survive by integrating. If the compromise product is offered at a price of one, and firms purchase it, the net benefits are

$$\int \frac{A}{n-1} dG - F = \frac{N\bar{A}}{n-1} - F,$$

If all firms integrate, then net benefits are

$$\int \left[\frac{\theta A}{n-1} dG - F \right] dG = \frac{\theta N \bar{A}}{n-1} - NF.$$

Thus, the compromise product yields a larger total surplus if

$$(5) \quad \frac{\bar{A}N}{n-1} - F > \frac{\theta N \bar{A}}{n-1} - NF$$

or

$$N > \frac{\theta - \phi}{1 - \phi},$$

where

$$\phi = \theta - \frac{F(n-1)}{\bar{A}}.$$

This is exactly the same condition as in the symmetric case. From a welfare point of view, integration should occur either completely or not at all, and under the same conditions that apply to the case when every firm is average in size.

Against that standard, there is too much integration. In particular, large firms are likely to integrate when it is not desirable from an efficiency standpoint. With a size distribution of firms, there is a tendency for integration in large firms to occur even when none is desirable.

As in the symmetric case, a suitable set of two part prices will generate the desired result. However, in the non-symmetric case, the fixed costs of the compromise product have to be distributed among downstream firms in such a way as to leave them all better off than they are when integrated. If that cannot be done, then complete integration is the surplus maximizing outcome.¹⁰ When the compromise product yields the higher surplus, it can be sustained on the market only with discriminatory two-part tariffs. To see this, let $Q(A)$ be the fixed charge to a firm of size A .

We look for a $Q(A)$ such that

$$(6) \quad \frac{A}{n-1} - Q(A) \geq \frac{A}{n-1} - F$$

for all A , and

$$\int Q(A)dG(A) = F$$

The inequality indicates that all firms prefer the standardized product (at a fixed charge of $Q(A)$ and a marginal price of one) to the tailored product and integration. The equality indicates that the fixed charges to firms cover the fixed component of production costs.

Suppose that (5) holds, and let

$$\delta = \frac{\bar{A}}{n-1} - \frac{F}{N} - \frac{\theta\bar{A}}{n-1} + F > 0.$$

Let

$$(7) \quad Q(A) = \frac{A}{n-1} - \frac{\theta A}{n-1} + F - \delta$$

Since $\delta > 0$, the condition (6) holds. Then, multiplying (7) by $dG(A)$ and integrating, and using the definition of δ , we have

$$\int Q(A)dG(A) = \frac{\bar{A}N}{n-1} - \frac{\theta N\bar{A}}{n-1} + (F - \delta)N = F$$

Therefore, when the market is preferable to integration, there is a discriminatory way of distributing the fixed costs so as to induce all downstream firms to buy the standardized product, and conversely.

There is another rather important welfare question that can be asked. Suppose that the upstream industry or firm has to charge a single price that is high enough to break even. If we could decide on the fraction of firms that are allowed to integrate, would we select a fraction that is larger or smaller than the fraction of integrated firms that characterizes the market outcome? This is a second best question. We take as given the price system, and the requirement upstream firm earnings be non-negative. With those constraints, we compare the "optimal" amount of integration with what actually occurs.

In the symmetric case, this issue arises in a trivial form. Given the constraints, firms integrated if the standardized product can be supplied profitably only at prices that are high enough to make integration desirable. In the nonsymmetric case, however, this is not true. Large firms integrate, ignoring the cost they impose on smaller firms in terms of a higher price for the compromise product. Therefore, generally the fraction of firms that are integrated is too high in the market outcome. This is not to say that the optimal amount of integration is zero. On the contrary, there are benefits from large firms integrating. The proposition is only that the market tends to exceed the optimal amount of integration.

The remainder of this section gives an argument for this proposition, using the previous model. The reader who is convinced by the argument above may wish to pass over the technicalities.

Let \tilde{A} be the smallest integrated firm. We consider \tilde{A} a decision variable. Non-integrated firms pay a price $p(\tilde{A})$ that is just sufficient to make upstream profits zero. This price is defined implicitly by

$$(p-1)p^{-n} \int_0^{\tilde{A}} AdG(A) = F .$$

It is easily verified that $dp/d\tilde{A} < 0$; the price falls as the number of unintegrated firms increases.

The total net benefits, given \tilde{A} , and the implied price are

$$(8) \quad T(\tilde{A}) = \int_{\tilde{A}}^{\infty} \left[\frac{\theta A}{n-1} - F \right] dG(A) + \int_0^{\tilde{A}} \frac{A}{n-1} p^{1-n} dG(A).$$

For this argument, let us assume $G(A)$ is differentiable, and let $g(A) = G'(A)$. This makes the argument easier by allowing us to differentiate. Differentiating (8), we have

$$(9) \quad T'(\tilde{A}) = g(\tilde{A}) \left[F - \frac{\tilde{A}(\theta - p^{1-n})}{n-1} \right] - \frac{F}{n-1} \frac{dp}{d\tilde{A}}.$$

The market equilibrates, when the marginal integrated firm is just indifferent between integrating and buying the compromise product. This condition,

derived earlier, is that

$$p = \left[\theta - \frac{F(n-1)}{\tilde{A}} \right] \frac{1}{1-n}, \text{ which implies}$$

that the term in square brackets is zero. Therefore, at the market outcome

$$T'(\tilde{A}) = - \frac{F}{n-1} \frac{dp}{d\tilde{A}}$$

which is positive. Therefore, at the market outcome the surplus would be increased by an increase in \tilde{A} , that is, a reduction in the fraction of firms that are integrated. The externality referred to above is captured precisely by the second term in (9). The first term reflects the net benefits of integrating for the marginal firm. At an equilibrium it is zero. That leaves the negative price effect of integrating, which is ignored by the integrating firm.

4. Extensions

Two extensions and modifications of these results are worth noting. First is the case where the supplier, or upstream firm, can supply multiple varieties of inputs. Thus far we have assumed independence in the cost of producing different varieties, or that the firm must incur new fixed costs F to produce a different variety. If by virtue of supplying the standardized input (or a specialized input), the multiproduct supplier can produce a firm-specific input (or a second firm specific input) with a marginal cost less than c or additional fixed cost less than F (the fixed cost of the integrated firm), then downstream firms are less likely to need to integrate to achieve the greater benefits of the specialized input. That is, increasing returns to producing several products works in the direction of "solving" the problem with the market noted in Section 1. If the reduction in overhead of the multiproduct supplier is proportional to the total volume produced by the supplier, complicated interactions similar to the previous case occur. But integration is less likely.

The qualitative predictions of the model are similarly modified in the case where there is a "clustering" of the firm specific inputs in the space of attributes as noted earlier. If the multiproduct returns to scale are greater for inputs that are close to each other in the attribute space, integration is less likely for the firms with "clustered" input attribute needs. Firms with attribute preferences that are remote from those of other firms will integrate. The alternative is to accept a distinctly second-best input. This latter strategy is undertaken only when the costs of integration are severe.

It is clear at this point that there are a wide variety of possible outcomes, depending upon the distribution of preferences with respect to attributes, and the multiproduct structure of costs of the upstream industry. The forces identified in the simple models, however, continue to operate.

5. The Capacity Decision

In this section, we seek to show that when the upstream supplier faces increasing returns and has a capacity decision to make, it will, under fairly general conditions, carry less capacity than the integrated firm would have. The integrated firm would carry more capacity and put a shadow price on the input that causes its division producing the input to lose money. The reason is that the upstream firm will set a price above marginal cost, and hence downstream consumption of the input will be lower than in the integrated condition.

Let $D(p)$ be the demand for the good by the downstream firm when the price is p . Let x be the amount of the input produced, k the capital used to produce it, r the unit cost of capital per period, and $c(x,k)$ be the variable cost function. Total costs are

$$c(x,k) + rk$$

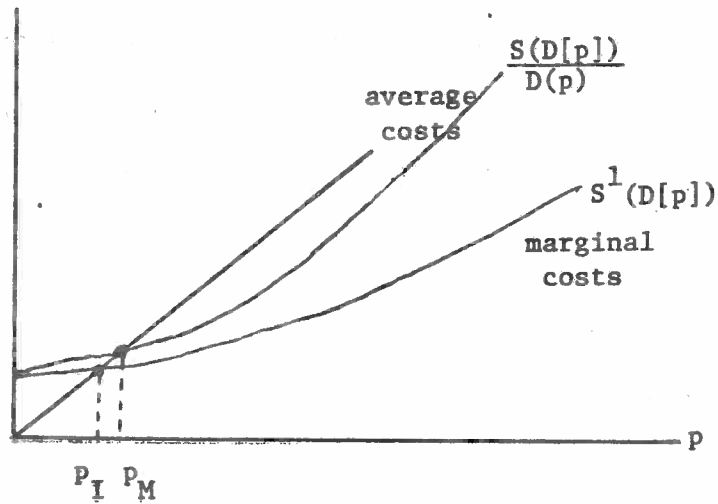
and the long run cost curve is

$$S(x) = \min_k [c(x,k) + rk].$$

The optimal amount of capital increases with x .

Now if $S(x)$ exhibits declining average costs, then the market price in the disintegrated case will be above the optimal shadow price in the integrated case. That follows from the fact that with integration, the optimal price is marginal cost, which is below average cost. In Figure (7) the situation is depicted.

Figure 7



Average and marginal costs are drawn as a function of p . Recall that $D'(p) < 0$. The integrated firm sets the price at p_I where it equals marginal cost. The market sets price at p_M where it equals the average cost. With declining average costs, $p_I < p_M$. It follows easily then that the integrated firm uses more of the input than it would when purchasing from an upstream firm. And the latter carries less capacity than would the integrated downstream firm.

6. Uncertainty and Capacity

One of the attributes of a product is availability. The demand by the downstream firm for an input is often somewhat random, and upstream suppliers may carry insufficient capacity to meet the demand when the latter is high. Firms are thought to integrate to ensure the timely delivery of sufficient quantities of the input when it is needed. A version of this argument does withstand rigorous scrutiny, but it requires some care. An integrated firm will not necessarily carry enough capacity to meet the demand under all circumstances, for capacity is costly. The qualitative proposition

must therefore be that an independent upstream supplier will carry less capacity than would an integrated downstream firm. This is true under some circumstances, as the following example illustrates.

Let the downstream firm face random demand. Let $\theta_1 B(x)$ be the revenues that the downstream firm can generate with x units of the input when demand is low. Let $\theta_2 B(x)$ be the same thing when demand is high. We assume that $\theta_1 < \theta_2$ which is the operational meaning of demand being low or high. The fraction of periods in which demand is low is α_1 , while α_2 is the fraction of the time it is high. Given θ_1 , the demand for the input satisfies

$$\theta_1 B'(x_1) = p$$

where p is the price of the input. This follows from the fact that the downstream firm, faced with a price of p , will maximize $\theta_1 B(x) - px$, the net benefits of buying the good x .

Let k be capacity measured in units of the input and let r be the unit cost of capacity. It is clear that the integrated firm, or an upstream supplier will produce at capacity when demand is high. For this example, we assume the marginal cost of the input is zero. The assumption of a positive marginal cost contributes nothing to the example

Let us consider the integrated firm. Its expected profits are maximized when

$$\alpha_1 \theta_1 B'(x_1) = 0$$

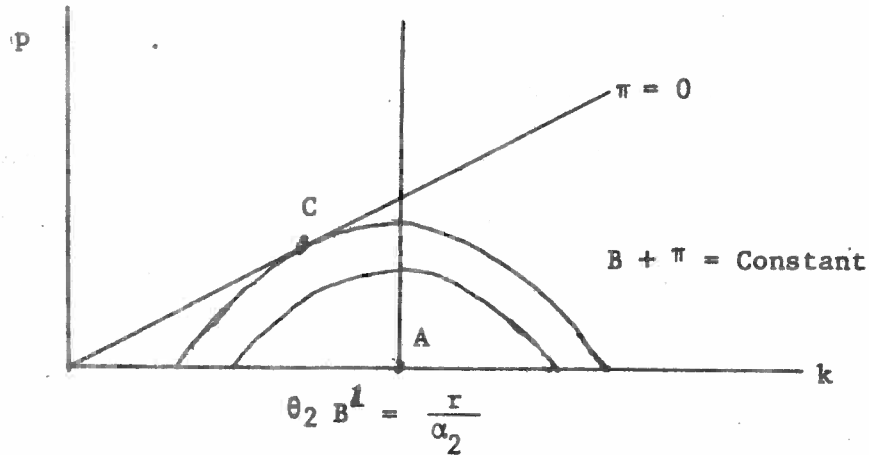
and

$$(10) \quad \alpha_2 \theta_2 B'(k) = r.$$

A useful interpretation is as follows. The division of the integrated firm producing the input sets k to satisfy (10). It establishes a shadow price of zero (the marginal variable cost), and sells as much as its capacity will permit to the downstream division at that price.

Figure 8 illustrates the solution. On the axes are capacity and the price facing the downstream division.

Figure 8



The contours are constant profit countours for the integrated downstream firm. Point A is the optimum.

Without integration, an upstream firm with capacity k and a price of p has profits of

$$\pi = (\alpha_1 x_1(p) + \alpha_2 k) p - rk .$$

Such a firm will act so as to maximize the benefits to the downstream firm, subject to the constraint $\pi \geq 0$. The benefits to the downstream firm are

$$M = \alpha_1 \theta_1 B(x_1(p)) + \alpha_2 \theta_2 B(k) - (\alpha_1 x_1 + \alpha_2 k) p .$$

The upstream firm acts so as to maximize $M(p,k)$ subject to $\pi(p,k) \geq 0$. This is equivalent to maximizing $B + \pi$ subject to $\pi = 0$. Thus in Figure 8, the outcome without integration occurs where the line $\pi = 0$ is tangent to an iso-profit contour for the integrated firm. That is the point C in the diagram. The disintegrated outcome will have lower capacity than the integrated firm provided $\pi = 0$ is upward sloping.

The fact that it is upward sloping is argued in the following way. When $k = 0$, then $p = 0$. For $k > 0$, p must be positive. Thus the line $\pi = 0$ must start out with a positive slope. Suppose it went flat at some point. That would occur when $\pi_k = \alpha_2 p - r = 0$. But when $\pi = 0$, then $(\alpha_2 p - r) = -\frac{\alpha_1 p x_1(p)}{k}$. Provided that $x_1(p)$ is not zero, this cannot happen. We will return to that case in a moment. Excluding it, the contour $\pi = 0$ is positively sloped. The result is that the upstream firm will carry less capacity than the integrated firm. It will therefore run out of the supplied input more quickly than the downstream firm would like.

There is an interesting special case where $x_1(p) \equiv 0$. In that case the upstream firm sets a price of $p = \frac{r}{\alpha_2}$ and holds capacity sufficient to supply demand at that price. In this case, the result is optimal for the downstream firm. The reason is that if $x_1(p) \equiv 0$, the fact that the internal shadow price of zero is irrelevant. The integrated firm establishes capacity at the level where its expected marginal benefits $\alpha_2 \theta_2 B^1(k)$ equal its marginal cost r . But if it faces a market price of $p = \frac{r}{\alpha_2}$, it demands x_2 satisfying $\theta_2 B^2(x_2) = r/\alpha_2$ when demand is high. This demand is exactly equal to the capacity held by the upstream firm.

There is an interesting interpretation of this special case. The fact that demand for the input could be zero with probability α_1 , might result from there being a probability of α_1 that a superior input becomes available. This is generally thought to be an argument for integration, because it represents a risk to the upstream firm. This is incorrect. Provided the probability of there being a superior input is the same for the

upstream and downstream firms, the risk of investing in capacity is the same for the upstream firm and the integrated firm. The discovery of a superior input will make the integrated firm's capacity as valueless as that held the unintegrated upstream firm.

To sustain this argument, one could argue (perhaps plausible) that the upstream firm perceives β_1 to be higher than the perception of the downstream firm, or that the upstream firm is more risk-averse. The argument here was conducted with the assumption of risk neutrality. Or one could argue that the cost of changing over to another activity are different for the upstream and downstream firms. The point we wish to make is simply that the presence of the risk of a superior substitute becoming available does not, by itself, create a presumption in favor of integration.

7. Empirical Implications

We conclude by drawing together some empirical implications of the models presented above and related observations. These can be grouped into three areas:

- (1) What are the structural features of industries in which we will observe integration rather than the market supplying inputs, and what is the nature of the inputs that will be supplied?
- (2) What factors cause firms within an industry to have differing levels of vertical integration?
- (3) What factors cause the level of integration in a given industry (or firm) to change over time?

With respect to the first question, it is clear that integration is more likely in industries which demand unique inputs.¹¹ These would be industries demanding inputs made to extremely rigid tolerances, produced from unusually high quality materials made to unusual specifications, or requiring

extremely reliable delivery. Industries such as aircraft and precision instruments produce many of their own components, for example, while industries such as recreational vehicles and mobile homes merely assemble commonly available panels, appliances, fixtures, windows and the like.

Integration is even more likely in industries meeting such specifications where each firm has differing preferences from other firms with respect to input characteristics. Using our earlier terminology, integration is more likely in industries where input demands of firms do not cluster. Since the demand for unique inputs is presumably related to the presence of differentiated outputs, we would expect integration to be positively related to physical product differentiation in the downstream industry. Each medical instrument producer makes many parts for his particular designs, for example, while cleaning powder manufacturers purchase standard chemical inputs in the market.

Second, we would expect integration to occur in industries and for inputs where firm demands for inputs are moderate or large relative to the increasing returns in input production. Where firm input demands are small, the sharing of fixed costs by purchasing a standardized product on the market loom more important. Many recreational vehicle producers fabricate their own metal panels and fasteners, for example, while none except General Motors produce their own engines and chassis involving significant scale economies. Third, we would expect integration where the downstream industry is concentrated. With many firms in the downstream industry, sharing of fixed costs favors the market. Fourth, integration is more likely if the elasticity of demand for the input is small. Fifth, we would expect

to observe integration in those industries where there are few economies of multiproduct production of different varieties of unique inputs, since multiproduct production economies increase the attractions of the market relative to integration. These would be inputs where raw materials, set-up, tooling, production machinery and/or distribution arrangements are specific to the particular variety of input being produced. Sixth, we expect integration in those industries where specialist upstream firms have relatively few cost advantages over in-house production of the inputs by downstream firms. Inputs requiring production knowhow far removed from that of the downstream industry are more likely to be produced in the market by specialist firms. Examples of these would be inputs drawing on a different basic technological discipline than do the products of the downstream industry, such as the manufacture of automobile tires which involve little of the electro-mechanical and metal working technology of the automobile itself. Tires are one of the few areas auto manufacturers have not integrated into.

A related set of implications apply to static differences in the level of integration of firms in a given industry. Larger firms will be more likely to be integrated than small firms, particularly if the firm size distribution has high variance. This is commonly observed: General Motors and Ford are considerably more integrated than Chrysler and American Motors; Cessna is more integrated than Piper Aircraft or Beech Aircraft. Firms with lower elasticities of input demand will be more likely to be integrated. Firms demanding "distant" inputs from those of other firms in the attribute space will tend to be integrated, while those firms demanding inputs that cluster in attribute space will more likely buy an

"average" standardized product on the market. Firms demanding inputs subject to few multiproduct production economies with other varieties of those inputs will be more likely to integrate, while those demanding inputs with low change-over costs from other varieties will more likely buy on the market. Note that clustering and multiproduct economies of scale are conceptually different, though in practice they may be positively correlated.

There are also a series of dynamic implications of the models, some of which have been touched on. Growth of an industry and the firms in it tends to promote integration by reducing the impact of overhead cost in producing the unique varieties firms demand. However, growth in the industry also increases the potential size of the market for the standardized product, lowering its cost. The outcome is complicated, in general, as has been discussed. On balance, we observe a net tendency for integration to increase over time in many industries with growth, such as in mini-computers, snowmobiles and light aircraft.

A number of other phenomena are likely to be occurring over time in an industry which bear on vertical integration, however. First, a common empirical observation is that downstream sellers' products often become less differentiated as an industry matures.¹² If input demands tend to cluster over time in an industry, this favors the market. Going in the other direction is the diffusion of knowledge about producing inputs. If knowledge about producing inputs becomes more widespread over time, cost advantages held by specialized producers over potentially integrating downstream firms diminish and integration becomes more likely. This occurs in part because the numbers of people with the production know-how increases over time. This seems to have been the case in aerosol packaging, when initially independent specialist aerosol fillers dominated the market when aerosol technology was being developed, but end product marketers now produce most of their need internally.

Footnotes

1. G. W. Brock, *The U.S. Computer Industry*, Ballinger, Cambridge, MA 1975, page 31.
2. Polaroid/Kodak, Intercollegiate Case Clearinghouse, #4-376-266, Boston, MA 1975.
3. Let $f(x_1, \dots, x_n)$ be the production function for the downstream firm, where x_i are the inputs. Let S be the output price and w_i be input prices. The firm maximizes $sf - \sum w_i x_i$, by setting

$$sf_i = w_i$$

for $i=1, \dots, n$. It follows that for any i

$$\begin{aligned} sf &= \int_0^{x_i} sf_i(k_1, \dots, v_i, \dots, x_n) dv_i \\ &= \int_0^{x_i} w_i(x_1, \dots, v_i, \dots, x_n) dv_i \end{aligned}$$

Thus, the integral under the inverse demand for any of the inputs equals the revenues. Thus

$$\int_0^{x_i} w_i(x_1, \dots, v_i, \dots, x_n) dv_i = w_i x_i$$

represents profits exclusive of deductions for the costs of other inputs.

4. It may be that the integrating downstream firm may not be able to produce the input as cheaply, for a given output level, as the upstream specialist firm can. This would reduce the likelihood of integration and create a disparity between social and private benefits.
5. R. Schmalensee, "A Note on the Theory of Vertical Integration," Journal of Political Economy, March/April 1973. M.K. Perry, "Theory of Vertical Integration by Imperfectly Competitive Firms," Stanford Center for Research in Economic Growth, 1975.

Footnotes cont.

6. A related point was discussed by G. Stigler, "The Division of Labor is Limited by the Extent of the Market," Journal of Political Economy, June 1951. Stigler pointed out that there is a trade-off between the returns achievable by a single supplier serving many firms, and continuous process increasing returns achievable by integrated firms. He argued that a market could, in the course of growing, integrate and then disintegrate. A similar effect will be observed here.
7. There could be more than one "standardized" input available; this does not change the basic nature of the result.
8. Formally the argument is as follows. The profits at the monopoly price of $n/n-1$ are

$$\pi = \frac{NA}{n-1} n^{-n} - F.$$

Since $F = \frac{(\theta - \phi)A}{n-1}$, the profits can be rewritten

$$\pi = \frac{A}{n-1} [N(1-1/n)^n - (\theta - \phi)].$$

These are positive if the term in square brackets is positive. The term $(1 - 1/n)^n$ is increasing in n and has a maximum value of $1/3$ at $n = \infty$. Thus profits are always negative if

$$\frac{N}{e} < \theta - \phi,$$

or

$$\phi < \theta - N/e.$$

For N sufficiently small, the right hand side is positive. In that case, when ϕ is below that value, the standardized good cannot be sold profitably even when the elasticity is very high. However, for larger N , the left hand side is negative. That is the case in Figure 3.

Footnotes cont.

9. For a discussion, see M. Spence, "Product Selection, Fixed Costs, and Monopolistic Competition," Review of Economic Studies, Vol. XLIII (2), June, 1976, pp. 217-235.
10. Such an overhead reduction is plausible given any excess capacities in the multiproduct supplier.
11. As with all these empirical implications, we hold constant the variety of other factors affecting vertical integration such as the basic product configuration and technology shape of the cost curve in the upstream industry, market power considerations and the like.
12. See Theodore Levitt, "Exploit the Product Life Cycle," Harvard Business Review, Vol. 43(6) Nov./Dec. 1965, pp 81-94 and other literature on the product life cycle.

*This research was supported by the Division of Research at the Harvard Graduate School of Business and by the National Science Foundation. The authors are particularly grateful to the members of the Social Science Research Council's Summer Workshop on Oligopoly Theory, held at the University of Warwick, July 1977. The development of these ideas benefited from their comments and criticisms.

Appendix: Derivation of the Properties of Figures

This appendix develops the properties of the symmetric model, described and depicted in the text. The integrated firm derives benefits $\frac{\theta A}{n-1} - F$. The unintergrated firm derives benefits of $\frac{A}{n-1} p^{1-n}$ when facing a price of p . These are equal when $p = \phi^{1/(1-n)}$ when $\phi = \theta = F(n-1)/A$. We call this the reservation price, \bar{p} . A firm supplying the market at a price of p has profits

$$(A.1) \quad \pi = NAp^{-n}(p-1) - F.$$

The profit maximizing price is $\hat{p} = \frac{n}{n-1}$.

There are two cases: corresponding to when the reservation price is above or below the profit maximizing price.

Case 1: If $\bar{p} > \hat{p}$ or equivalently $\phi < (1-1/n)^{n-1}$, then the market will prevail if profits are positive at the profit maximizing price, or

$$(A.2) \quad \phi > \theta - N(1-1/n)^n.$$

Case 2: If $\bar{p} < \hat{p}$ or

$$(A.3) \quad \phi > (1-1/n)^{n-1}$$

then the market will prevail if profits are positive at the reservation price. This can be written

$$(A.4) \quad (N + \frac{1}{n-1})\phi - \frac{\theta}{n-1} > N\phi^{\frac{n}{n-1}}$$

It remains to derive the properties of these regions. The boundary in $n-\phi$ space where $\bar{p} = \hat{p}$ is defined by (A.3) in its equality form. The

right hand side is 1 when $n=1$ and declines monotonically to $\frac{1}{e}$ at $n=\infty$.

To see this, note that

$$(A.5) \quad \log \left(1 - \frac{1}{n}\right)^{n-1} = (n-1) \log \left(1 - \frac{1}{n}\right) \\ (n-1) \left(-\frac{1}{n} + \frac{1}{2n^2} - \dots\right)$$

As $n \rightarrow \infty$, the right hand side approaches minus one. As to the slope, we differentiate $(n-1) \log \left(1 - \frac{1}{n}\right)$, to obtain

$$(A.6) \quad \log \left(1 - \frac{1}{n}\right) + \frac{1}{n}$$

This is negative because $-\log \left(1 - \frac{1}{n}\right) = \int_{1-1/n}^1 \frac{dx}{x} > \frac{1}{n}$.

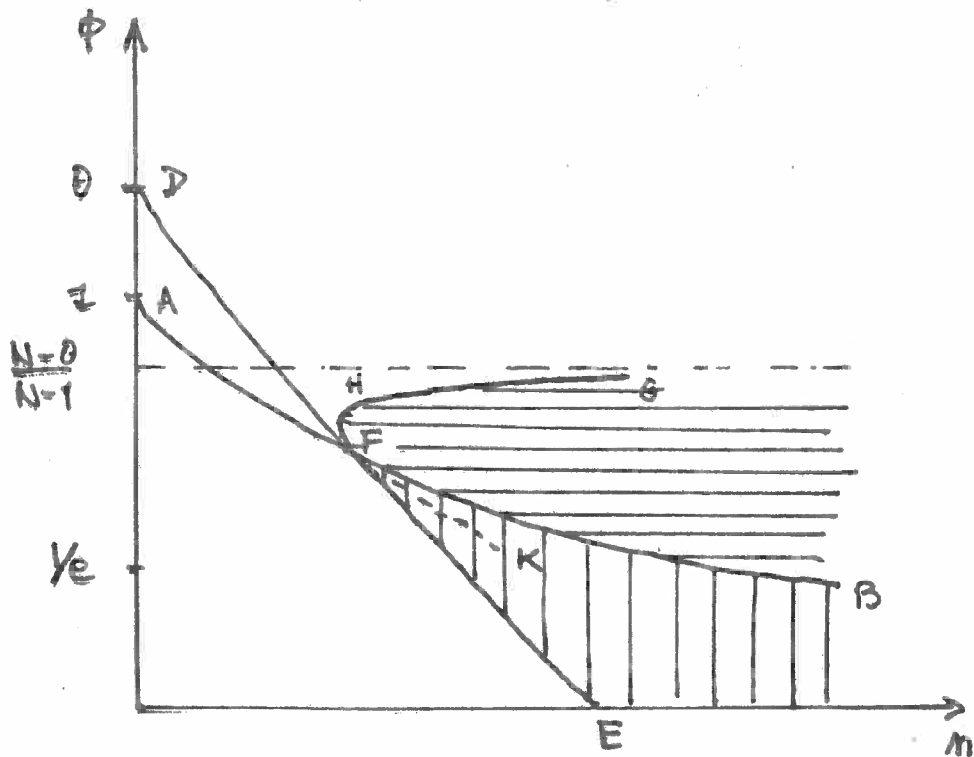
The curve is shown in Figure A.I as the line AB.

The equality version of A.2 has the θ when $n=1$. It declines monotonically to $\theta - \frac{N}{e}$. The analysis is similar to that carried out for A.3. The extent of this decline clearly depends on the size of N . Figure depicts the case where $\theta < \frac{N}{e}$. Note however that $\theta - \frac{N}{e} > 0$ is possible. And if $\theta - \frac{N}{e} > \frac{1}{e}$, then the curve would remain above AB. In that case, i.e. $N < \theta e - 1$, there are no realizations of case 1. Figure 3 in the text depicts the case $\theta < \frac{N}{e}$. Figure 4 corresponds to $\theta e > N > \theta e - 1$. In Figure A.I shows this line as DE. The shaded area BEF is the section where the market prevails, when the reservation price is above the monopoly price.

Turning to case 2, the market boundary is defined by the equality version of (A.4). This is a more difficult relation analytically.

We know the market cannot prevail at prices below 1. From the main body of the paper, integration is preferred at a price of one when $\phi > (N-0)/(N-1)$. As n becomes large, if ϕ is fixed, the ratio $\frac{F}{A}$ must fall to zero. Thus profits will approach zero at a price of one.

Figure A.1



Hence the boundary approaches $\phi = (N-\theta)(N-1)$, asymptotically as $n \rightarrow \infty$. For n small enough, A.4 will have no solution because for $0 < \phi < 1$, the left hand side of (A.4) is negative. Recall that $\phi > 1$ assures integration. As n rises, a solution appears, and then two solutions. The result is the curve GHFK in Figure A.I. The horizontal shaded area corresponds to parameters where the market will prevail.

A final remark is that as N becomes small, the market areas GFB and FBE move to the right. When N is 1, there is no market region. The reason is that when $N = 1$, the market has no advantage, it not being possible to distribute fixed costs over firms.