Theories of Vertical Integration

by

Dennis W. Carlton

MIT Department of Economics

Cambridge, Mass.

March, 1976

Presented at the Fourth Annual Telecommunications Policy Research Conference,

Airlie House, Virginia, April, 1976.
Theories of Vertical Integration

The subject of vertical integration is a difficult one for economic theory, but one that carries with it significant public policy implications. The difficulty in studying vertical integration is that under the usual competitive assumptions that economists often postulate, there are absolutely no incentives or disincentives for vertical integration to occur. It is not hard to observe that many industries which once were composed of disintegrated competitors have substantial degrees of vertical integration. Many reasons have been advanced for vertical integration ranging from production economies of scale to transaction cost arguments. The purpose of this paper is to survey some of the recent theoretical work which has tried to advance our understanding of vertical integration. The specific models that I have been asked to discuss are those of Arrow (1974, 1975), Carlton (1975, 1976), Green (1974) and Perry (1975).

In one way or another, all models which demonstrate an incentive for firms to vertically integrate must postulate conditions other than the usual ones of competition and constant returns to scale in production. Since vertical integration occurs for such a variety of reasons, it cannot be expected that one simple model could explain all forms of vertical integration. Rather the models under discussion are all designed to focus on the effect of one particular set of assumptions about market behavior.

The models of Arrow, Carlton and Green examine the incentives for and consequences of vertical integration in markets where there is some
type of uncertainty present. The distinguishing feature of these models is that competition is taking place under uncertainty, and the assumptions that all participants have the same information or that the price can instantly adjust to clear markets are not imposed.

In the Arrow model, there is an uncertainty on the part of a producer about the input price. By integrating backward, a producer is able to obtain information about the supply of the input, and thereby obtain information about the likely price of the input. This information is useful for technology choice and provides the incentive to vertically integrate.

The models of Carlton and Green both attempt to make sense of businessmen's frequent cries that vertical integration is necessary to obtain an assured or certain supply of input. Since inputs are often sold on competitive markets, it must be that markets are not always equating supply and demand. In Carlton's model, both final product and intermediate product firms face uncertain demand and it is natural for firms to have either excess or insufficient capacity. Vertical integration is viewed as a mechanism for transferring risks from one sector of the economy to another. Vertical integration involves balancing the risk of being caught with excess capacity versus the advantage of having an assured supply of input.

In Green's model, the price of the input is exogenous and rationing occurs when demand and supply are out of balance. Final product firms which integrate are assumed to be less efficient producers of the input than are intermediate product firms. Vertical integration involves balancing
the extra costs of producing the good versus the advantage of the avoidance of rationing.

The work of Perry deals with issues much different from those of the previous models. Perry analyses markets which have no uncertainty and focuses on the effect of having noncompetitive behavior in the input market. The model from which his main results follow is one in which there is a sole purchaser (monopsonist) of the input. The monopsonist causes the well-known inefficiencies associated with monopsony. Vertical integration becomes a profitable activity which can turn efficiency losses into profits. Perry also shows that a monopsonist will have an incentive to vertically integrate so as to avoid paying rents to the competitive suppliers.

The welfare conclusions of these models are not identical. In the Arrow model, horizontal integration accompanies vertical integration, so no problems with monopoly power could arise. When no problems with monopoly power arise, complete vertical integration is desirable. In the Carlton model, vertical integration is an undesirable market structure from the point of view of static efficiency. From the point of view of a "dynamic" efficiency in the introduction of technology, vertical integration is a desirable market structure. In the Green model, the welfare implications depend on the severity of the rationing scheme. In Perry's monopsony case, complete vertical integration is desirable. The fact that the welfare implications differ between the models is not surprising since the assumptions underlying the models are so different.

The details of each of the models will now be described, and a brief conclusion will follow.
Arrow's model emphasizes the role of uncertainty in the supply of an input and most importantly the need for information by final producers as motives for vertical integration. An interesting feature of Arrow's model is that firms have an incentive to integrate backward not to obtain a source of input supply but only to obtain information about future supply conditions.

There are assumed to be some given number of input suppliers who behave competitively. The production of the input is subject to uncertainty (e.g., because of weather conditions) and the $i^{th}$ supplier produces some amount of input $x_i$ where $x_i$ is a random variable. The market for the input takes place one period after the supply conditions have been observed. Knowledge of the supply, $x_i$, of an individual firm provides information on the total amount of input to be supplied since $x_i$ is part of total supply and since $x_i$ is most likely correlated with the other random supplies of the other input suppliers. (Good weather helps everyone's crops.)

The output producers are assumed to be competitive. There is free entry into the output market, so zero profits in the output market is a condition for equilibrium. The output is produced from the input and from capital. Capital is always available at some fixed price. Since the input market clears at price $q$, there is absolutely no uncertainty in being able to obtain as much of the input as is desired. Similarly, since there is an output market which clears at price $p$, there is no
uncertainty about not being able to sell all the output that is produced. The sole uncertainty is in the prices, $q$ of the input and $p$ of the output.

If this were the entire model, there would be no incentive to vertically integrate. The input market always clears at some price $q$. From the viewpoint of a vertically integrated firm, the market price $q$ reflects the cost of using the input. The fact that a firm owns some input is irrelevant to the firm's output decisions. Owning an input supplier is no protection against high input prices. The shadow cost of the input is the same whether a firm produces the input itself or whether it purchases it on the open market.

The additional feature that makes the model interesting is the requirement that output producers must decide how much capital to purchase one period before the prices of the input and output are known. If an output producer could find out these prices, he would be able to increase his profitability by being able to choose the efficient level of capital to use. Remember that suppliers know their supplies $x_i$ one period before the input price is determined. Since input price depends on supply and since output price depends on input price, information about the supply $x_i$ of the $i^{th}$ supplier provides an output producer with some information about prices. It will be this information that provides the incentive for vertical integration.

There are two markets and hence two prices that are endogenous, the price, $p$, of the output and the price, $q$, of the input. To see how these two prices are determined, suppose that no vertical integration is allowed. In period 0, before any prices are observed, output producers must choose how much capital to employ. They choose that level of capital, such
that the expected value of output price times marginal product equals the
price of capital - where expectations are taken over the possible values
that supply, x, can take on.\textsuperscript{1} Also in period 0, the supplies \(x_1 \cdots x_n\)
are known to the individual suppliers, but not to the output producers.
In period 1, there are two markets that clear. At prices \(p\) and \(q\), the
demand for the output must equal the supply of output, and the demand
for the input must equal the supply \(x_1\).\textsuperscript{2} Notice that the decisions by
output producers about the amount of capital to employ affects the demand
for the input and hence its price.

To investigate the incentive for vertical integration, we observe
that if an output firm buys an input supplier, then the output firm will be
able to obtain some information, namely \(x_1\), about total supply \(x\), and
hence be able to form an improved estimate of future prices. As mentioned
earlier, the owning of some input, \(x_1\), does not affect the output decisions
of output producers. The output producer could always sell \(x_1\) at price
\(q\) and buy back at price \(q\) any amount of input desired. The value of
vertical integration is coming solely from the information obtained about
prices through observations on supply \(x_1\).

\textsuperscript{1}The level of capital is chosen so that \(E_x[p(x)F_y(x,y)] = p_y\), where
\(E_x\) stands for expectations over \(x\), \(p\) is the price of output, \(x\) is input,
\(y\) is capital, \(F\) is a linearly homogeneous production function, and \(p_y\)
is the price of capital. (Since \(x\) is the only variable that is random,
the output price depends on \(x\), hence we write \(p(x)\).

\textsuperscript{2}These two relations can be written as \(D(p) = F(x,y)\) and \(pF_x(x,y) = q\),
where the notation is the same as before, and \(D(p)\) is the demand curve
for output.
As the output producer obtains more information about the likely values of prices in period 1, he is able to make more informed choices about what capital stock to commit himself to in period 0. Since the information $x_1$ allows an output producer to make more informed decisions about capital stock, his expected profits increase as a result of observing $x_1$. Any input supplier would be willing to be acquired at its expected value of sales, $E(q^*x_{1})$. Any output producer would willingly pay $E(q^*x_{1})$ to acquire an input firm, since this acquisition cost can be recouped by selling the input $x_1$ at price $q$, while the additional information about supply, $x_1$, provides information on prices and causes profits to increase because the added information about prices allows more efficient utilization of capital in the production of output. The implication of this model is that there always will be an incentive for some vertical integration to exist.

Can there be a competitive market equilibrium with more than one integrated firm? Arrow shows that if such an equilibrium exists, then any input firm (which output firms competitively bid for) can be acquired at $E(q^*x_{1})$ - the expected value of the input firm's sales. If this acquisition cost were higher, then a vertically integrated firm would earn negative profits, a condition incompatible with equilibrium. If the acquisition cost were lower, no input firm would be willing to be acquired. For exactly the same reasons as described above, an output firm has an incentive to acquire input firms at $E(q^*x_{1})$ in order to use the information contained in $x_1$ to make more efficient capital decisions. Any time
there are several integrated firms in the market, there is always an
incentive for output firms to try to acquire more input firms. Equilib-
rium with more than one integrated firm simply cannot exist in this model.

The only possible equilibrium for this model occurs when there is
only one output firm who owns all the input suppliers. The output firm
can then perfectly predict total input supply and hence prices and thereby
can efficiently choose the capital stock. To avoid the acquisition of
monopoly power, Arrow postulates a perfectly elastic output demand curve
facing the output producer. In such a case, the improvement in the effi-
ciency of output production that results from perfect knowledge of supply
is passed on entirely to the initial owners of the supply firms. When
vertical integration occurs, the increased profitability of the vertically
integrated firm is driven to zero by a rise in the acquisition cost of
an individual supplier.

The socially efficient outcome is to have one firm owning all input
firms, and behaving competitively in the output market. This socially
efficient result occurs in the Arrow model for the special case of a
perfectly elastic demand for output. When the demand facing the single
vertically integrated firm is not perfectly elastic, monopoly power will
arise as vertical integration occurs. In that case, the information ad-
vantages of vertical integration will be accompanied by opportunities to
exercise monopoly power with its resulting inefficiencies.
Carlton

Carlton examines the effects of demand uncertainty, supply uncertainty and risk on incentives to vertically integrate. An important feature of this model is the transmission of uncertainty from one market to another. To understand the workings of the model, it is first necessary to explain how a single market under uncertainty operates, and then to look at the more complicated question of how markets interact under uncertainty.

For many markets, the assumptions that price instantaneously adjusts to equate supply and demand, that production occurs instantaneously and that firms can sell all they want at the market price do not precisely hold. More realistic assumptions are that to provide an effective signal to buyers, price must remain fixed for some period, that production takes time, and that firms never know exactly how much of their product will be demanded during a market period.

Imagine a number of firms competing with each other. At the beginning of each market period, before any firm can observe its random demand, firms must set price and decide how much to produce (or equivalently how much production capacity to have available). Prices do not change during the market period, firms do not trade the good amongst themselves, and recontracting markets for the good do not arise. These assumptions correspond to the observation that in the real world such phenomena often do not occur presumably because of transaction costs. A firm which does not sell out discards its excess production. (Inventory holding, provided
it is a costly activity, would leave unaltered all conclusions of the
model.) It is perfectly natural in this model for firms to be concerned
with the risk of under or overproducing.

A "good" at a particular firm has two relevant characteristics,
its price and its probability of availability. Before choosing a firm,
consumers know the price the firm charges and, through reputation, the
probability that the firm can supply the good. Consumers do not know
which firms still have the good when they visit firms during the market
period. Disappointed buyers return home dissatisfied. (Search activity
on the part of buyers, provided it is costly, would leave the conclusions
of the model unaltered.) Consumers (who we take to be identical for
this paper) have preferences which firms recognize between the price of
the good and the probability of obtaining the good. Firms compete on
the utility level (i.e., the mix of price and availability) that they
offer to consumers. Firms bid up the utility level until expected profits
are driven to zero. The higher the probability of availability, the
higher is the price needed to compensate the firm for the risk they face
of having unsold goods (or equivalently of having unused but available
capacity). Just as in usual competitive markets, in this market equilib-
rium, no firm offering less than the best deal (i.e., highest utility
level) will receive any customers.

There are several noteworthy features of competitive equilibrium
under uncertainty. First, price must exceed the constant per unit cost
of production. Revenue from sold goods must compensate not only for
the cost of production of these goods, but also for the cost of production
of the unsold good (or equivalently, compensate for unused but available capacity). Second, supply and demand will not in general be equal even in expected values since equilibrium is determined, in part, by consumers' willingness to pay for the risk of obtaining the good. Third, there will in general be a nonzero probability of not being able to obtain the good. Lastly, in the absence of insurance markets, the resulting competitive equilibrium is not Pareto-efficient, and it is possible to derive the conditions under which subsidization of the industry is desirable.

Now, let's turn to a discussion of vertical integration. There are two types of firms - output firms and input firms. The demand to any output firm is random and this generates a random derived demand for the input. The production technologies for producing the input and output both involve fixed coefficients and constant returns to scale. Production of the output can take place instantaneously. Production of the input must occur before any of the random demands can be observed. As before, prices are set at the beginning of each market period. There is a risk that input will be produced but not used by the end of the market period. (As before, allowing inventory holdings would not alter any of the model's results.)

As before, we do not allow input firms to ship the good back and forth to each other, nor do we allow recontracting markets to develop.

Output firms face the risk of being unable to obtain the input from input firms. We allow output firms the option of producing some of the
factor input for itself. We refer to the production and holding of the input by output firms as vertical integration. If an output firm produces some input for itself, it bears the risk of having unused input at the end of the market period. An output firm is not allowed to sell its inputs in the input market. This last assumption is designed to capture the notion that a vertically integrated firm is producing for its own needs to better assure itself of supply.\(^2\) The important feature of this market structure is that decisions of output firms to produce some of their own input affects the riskiness of demand that input firms see.

Suppose initially that there is no vertical integration and that the input market reaches equilibrium as discussed earlier. In deciding whether to vertically integrate, an output firm must weigh offsetting benefits and costs. Vertical integration involves a balancing of risks.

\(^2\)Let me give two possible reasons and examples to illustrate why this assumption is reasonable. First, an output firm might be distrustful of the quality of an input purchased from another competing output firm. Second, there may be transaction costs or other increases in bureaucratic costs as a company expands from being a seller of output to becoming a seller of both outputs and inputs. Complex organizations can increase costs.

This nonsharing of an input supply by output producers is quite common. As an example consider a firm's secretarial pool that is occasionally idle. It is not frequently the case that an employer will try to obtain outside typing in order to keep his secretaries fully busy. Next, consider a firm that keeps stocks of some input for production. This firm will usually not be always willing to sell its input. The firm will usually enter the input market as a seller only if its stock of input starts piling up and the firm becomes convinced that it miscalculated its needs. Only when the firm becomes sure that it has made a sizable miscalculation will it pay for the firm to incur transaction costs and enter the input market as a seller.

A model with equivalent implications can be constructed by assuming that vertically integrated firms can purchase input from each other, but the input is also demanded by other sectors of the economy, and that vertically integrated firms do not sell to these other sectors.
If an output firm produces a unit of input itself at cost, \( c \), then it assures itself of a supply and can avoid paying the higher market price, \( p \), for the input (remember \( p \) exceeds \( c \) in equilibrium). On the other hand, the output firm must bear the risk of not being able to use the input.

Strong incentives to vertically integrate arise because a vertically integrated firm is able to impose a differential risk between the use of its own input versus that of an input firm. Incentives for vertical integration come about because the output firms base their decisions to integrate on the \textit{marginal, not average}, probability of using an additional input. The way the markets operate, the price of the factor in the input market reflects not only the cost \( c \) of producing the input, but also the average probability of not being able to sell that input. When an output firm is deciding whether to hold one unit of the input itself, it is \textit{not} concerned with the average probability of being unable to use any unit of input. Rather, since the output firm will use its input holdings first, the output firm is concerned with the probability of being able to use that first unit of input. This (marginal) probability will always exceed the average probability that any unit of input is unused so that at least initially the expected benefits from vertical integration are likely to outweigh the risks. It is precisely because output firms can use their own input to satisfy their "high probability" demand and use the input market to satisfy their "low probability" demand that incentives for vertical integration occur.

Depending on the risk caused by demand uncertainty, and the wedge between market price of the input \( p \) and the cost \( c \) of producing it, three types of market structure are possible: no vertical integration, partial
integration, and complete integration. The case of partial integration occurs when an output firm produces only some of its own input and relies on the input firms for the remainder. In this case, only the "risky" demand is left for the input firms, who charge a large price premium to compensate themselves for the risk.

The welfare implications of the model derive from the assumption that vertically integrated firms do not sell their input on the input market. Any holdings of input by output firms is a less efficient way to satisfy input demand than holdings by input firms since input held by an input firm has a possibility of being used by any output firm, while input held by a vertically integrated firm can only be used by that firm. Using this result, Carlton is able to show that any vertical integration is socially inefficient, even though, as seen above, strong private incentives for vertical integration exist.1/ Although space precludes a detailed discussion, it should be mentioned that it is possible to show that the coordination ability of a vertically integrated firm makes it more likely that a vertically integrated firm, and not a nonintegrated firm, will choose to introduce socially beneficial technology. We are thus led to a view of the world where some static inefficiency in market structure must be tolerated to obtain an efficiency in the introduction of production technologies.

1/ It is worth pointing out what happens if we drop the assumption that vertically integrated firms do not sell inputs on the input market. First, the undesirable welfare implications of vertical integration disappear. Second, complete forward integration is the only equilibrium market structure since there is an advantage associated with insuring supply of and demand for the input.
Vertical integration in this model is a mechanism for transferring risk from the input market to the output market. The model shows that strong incentives from this risk transfer can occur, even when such a risk transfer is socially undesirable. Based on descriptive studies, it does appear that the risk considerations identified in the model can exert significant influence for certain industries. "...integration is a two-edged sword. Though, it reduces the risk of supply failure...the financial penalties of losses (that is risks) have increased... A way of reducing the risks...is through partial...integration: a company can produce a portion of its needs...and buy the fluctuating remainder. This has the advantage of providing full utilization of its own equipment, and allowing the suppliers to absorb the risk of fluctuations in demand. The company has to pay a premium to get someone else to absorb the risks. Tapered integration plays a large role in the industry".1/

Green

Green analyzes the consequences of having price inflexibility in the market for the input. Demand and supply of the input are stochastic and when demand exceeds supply at the exogenous price, demanders are rationed and cannot produce their desired level of output. When supply exceeds demand at the exogenous price, suppliers of the input are rationed and are unable to sell all they produce.

There are three types of firms in Green's model, an input firm, an output firm, and an integrated firm. The input firm sells its input on the input market, if it is able to. An output firm purchases input on the input market, when it is able to, in order to use its output capital to produce the output. An integrated firm produces some input for itself, possibly buys or sells some input on the input market (provided it is able to do so), and produces output with the input and its output capital. An advantage of vertical integration is that the internally supplied input can be passed directly to the output production sector of the firm without incurring any risk of rationing.

All uncertainty in the model comes from the input market. Output firms are assumed to be able to always sell all their output at the given price. Input suppliers must decide how much to supply, and input demanders must decide how much to demand before it is known how severely buyers or sellers will be rationed. Input firms and vertically integrated firms who sell input worry about not being able to sell their input.

\[1/\text{There are two types of capital, that used to produce input, and that used to produce output.}\]
Output firms and vertically integrated firms who buy input worry about not being able to obtain the input. Production of both the input and output are characterized by fixed coefficients. Hence, if an output firm is unable to obtain input, its output capital stock stands idle.

As described, vertical integration is always the desired market structure since it provides a means of obtaining the input without running the risk of being rationed on the input market and hence allows full utilization of a firm's input and output capital. Green goes on to postulate that the integrated firm has a less efficient production technology for producing both the output and the input when compared to the nonintegrated output and input firms. Green is assuming that a firm which performs several different tasks does not perform any one task as well as a firm that performs only that task. Vertical integration involves balancing the benefits of assuring input supply versus the cost of inefficient production of both inputs and outputs.

The first question Green asks is how should the vertically integrated firm compose itself. Should it only produce a little input using its input capital (and its inefficient technology), attempt to purchase some input on the input market, and then produce the output from its input and its output capital (and inefficient technology)? Or, should it produce a lot of input, try to sell some on the input market, and then use the remaining input with its output capital to produce the output? Notice that because of the assumption of a fixed coefficient technology, there is exactly one input level that is compatible with full utilization of the firm's output capital. When the vertically integrated firm cannot purchase sufficient input some of the output capital stock stands idle, while when
excess input cannot be sold the firm cannot use it to produce extra output. Before it observes the rationing that will occur in the input market, the vertically integrated firm must decide how much input capital (needed to produce the input), and how much output capital (needed to produce the output from the input) to hold and how much input to try to buy or sell on the input market.

Green is able to show that the only type of vertically integrated firm that can exist will be one that does not buy or sell in the input market. To see this point, remember that there are constant returns to scale in production. The profit function of the integrated firm depends on whether it is a buyer or seller of input since different rationing schemes with different consequences face the integrated firm under the two situations. Let the price of output capital, \( K_0 \), be \( r_0 \) and the price of input capital, \( K_1 \), be \( r_1 \), and suppose that one dollar\(^1\) is available to spend on capital (i.e., \( r_0 K_0 + r_1 K_1 = 1 \)). The total profits of an integrated firm will be a linear function only of \( K_1 \). For small \( K_1 \), the firm is a buyer of input, for large \( K_1 \), a seller and for some \( K_1^* \), it neither sells nor buys input. With an input capital of \( K_1^* \), the firm produces exactly that amount of input that is compatible with full utilization of its output capital \( K_0^* \) (recall \( r_1 K_1^* + r_0 K_0^* = 1 \)). A vertically integrated firm that neither buys nor sells the input on the market is called "balanced".

Diagrammatically, one possible configuration for the profit function is shown below.

\(^1\)Since there are constant returns to scale, and since we are interested only in the ratio of \( K_1 \) to \( K_0 \), it suffices to solve the problem for the case of one dollar.
Possible Profit Function for an Integrated Firm

There is a (usually) nondifferentiable point at $K_1^*$ where the profit function changes in response to the switch from the integrated firm being a buyer to the firm being a seller of the input. Because of the linearity of the profit function, only points $K_1$, $K_1^*$, $K_u$ can be maximum profit points and hence only these three points can represent possible configurations for a vertically integrated firm. (For the particular diagram above, $K_1^*$ is the maximum profit point.) The point $K_1$ corresponds to spending all money on output capital (i.e., $K_1 = 0$), in which case the firm acts as a pure output firm. The point $K_u$ corresponds to spending all money on input capital (i.e., $K_u = 1/r_1$), in which case the firm acts as a pure input firm. The only configuration for a truly vertically integrated firm is represented by point $K_2^*$, the point of "balanced" proportions.

If a vertically integrated firm is to exist, then it must be "balanced" - never enter the input market to either buy or sell. A "balanced" firm is completely isolated from the uncertainty and rationing in the market.
This result about the "balanced" configuration being the only form for an integrated firm depends heavily on the linearity of the profit function which in turn depends on the assumed production technology. With other production technologies, it is possible to have vertically integrated firms who take part in the input market. However, even for different technologies, as the previous diagram suggests, Green argues that it is not unlikely that vertical integration at the "balanced" point $K^*_1$ will still remain optimal.

Green then goes on to address the question of whether a market structure can exist with both nonintegrated output and input firms and vertically integrated firms. The analysis of this question becomes very complicated because of the difficulty of accounting for the changing amount supplied and demanded of the input, as the level of vertical integration increases. It is not possible to reach a general answer. However, based on the analysis of special cases, Green suggests that there is a tendency for a basic instability, either the industry becomes as vertically integrated as possible, or else total nonintegration of both output and input firms occurs. This instability of market structure can cause an industry to drastically change its composition from integrated to nonintegrated in response to small changes in the environment, especially regarding rationing probabilities.

To evaluate the social welfare aspect of vertical integration, Green uses a measure of value of output, where value is calculated at the exogenous prices. (There is of course the problem, which Green fully recognizes, that the exogenous prices may bear no relation to welfare.) The required
computation becomes very complex, and only the situations of complete vertical integration and total disintegration are analyzed. Even for these two cases, general conclusions are difficult to obtain. The main result is that there is a tendency for the system to remain integrated longer than is optimal when sellers are rationed more severely than buyers in the input market, or when sellers' rationing frequency would be improved in a nonintegrated market. Conversely, a market structure will have a tendency to remain disintegrated longer than is desirable when buyers are rationed more severely than sellers.
Perry deals with cases in which there is some noncompetitive behavior in the market for the input. A firm recognizes its influence on price. Most of Perry's analysis deals with the incentives for and consequences of vertical integration by a monopsonist. A monopsonist is the sole purchaser of an input. The input is competitively supplied at a price that rises with the quantity supplied. One possible way to think of this model is that one firm is buying mineral supplies from many mines. As more minerals are produced, production proceeds deeper down into the mine and the more costly becomes the production. In the diagram below, mc represents the rising marginal cost of production and represents the supply curve for the mineral industry. At any price p, the amount OABC represents the cost of production while ADB is pure rent.

The curve dd is the one output firm's (monopsonist's) demand for input. The efficient solution is at point B, the intersection of dd and mc.

\footnote{To keep the model simple, we ignore the dynamic consideration that today's supply decision may affect tomorrow's supply curve.}
However, if as Perry assumes, the monopsonist is not allowed to price discriminate (i.e., must pay the same price for every unit of mineral), then an efficient solution will not be reached. The reason is that when the monopsonist increases his purchases of a mineral, he must pay not only a higher price on that additional unit, but must also pay this increased price for every unit. The monopsonist thus realizes that he must pay higher than marginal cost for an additional unit. This creates an incentive for the monopsonist to restrict his purchases of the input to some $q^*$ which is less than the efficient $q^*$ (see diagram). The triangular shaped region $FGB$ is a pure efficiency loss to society that results from the monopsony behavior.

If the monopsonist owned the supply resource, he would move immediately to the efficient point $B$ since he would face the marginal cost of production as the cost of an additional unit. It would be profitable for the monopsonist to offer to buy the resource at an amount equal to the rent he is paying for the input at the inefficient point $F$. The supply owners would willingly accept this offer since their rent (i.e., revenue minus cost) remains unchanged. Backward integration is able to convert the efficiency loss triangle $FGB$ into profits for the firm.

The point, then, is that imperfect competition causes an efficiency loss. Backward integration becomes profitable because it enables the vertically integrated firm to turn this efficiency loss into profits. A similar point holds when the imperfect competition takes the form of monopoly power in the input market. The analysis of vertical integration in the presence of monopoly was first done by McKenzie (1951), Burstein
(1960) and more recently by Vernon and Graham (1971), Schmalensee (1973) and Warren-Boulon (1974). Whenever there is a market inefficiency in the input market, the value of the supply resource is worth more to the vertically integrated firm which will operate efficiently than it is to the suppliers of a disintegrated firm which is forced to operate inefficiently because of the lack of perfect competition. This discrepancy in value creates incentives for complete vertical integration to occur. Perry labels this incentive the "efficiency effect".

The interesting new aspect of the problem that Perry points out is that as vertical integration proceeds, the rents to the competitive suppliers fall. The ability of a vertically integrated firm to lower the rents paid to suppliers is called the "rent effect". The rent effect provides a separate incentive for a firm to vertically integrate fully.

To examine in detail the rent effect, assume that the demand for input is always 100 units - hence there is no possible way to have an efficiency effect in which the input demand falls short of the efficient quantity. With no vertical integration, 100 units are purchased at price p. Suppose there are 100 identical suppliers. Then each supply one unit and earn a rent ABD (see diagram below). All outcomes are perfectly efficient.
Suppose that the monopsonist owns 50 suppliers. Then, the cheapest way for the monopsonist to obtain 100 input units is to produce more than 50 units from its owned supply firms and purchase less than 50 units from the competitive suppliers. The vertically integrated firm overproduces from its own supply firms because in deciding whether to produce an additional unit from its own supply, it need only pay the marginal cost of production on that one unit. When deciding whether to purchase an additional unit from the competitive suppliers, it must pay not only the marginal cost of production on that one unit, but also a higher price for all previously purchased units. By producing itself, a firm avoids paying rents to the competitive suppliers. As the 50 competitive suppliers supply less than 50 units, their individual rents fall from the original value of $ABD$ to some $ANN$ (see diagram above).

The vertically integrated firm can thus obtain the 100 units of input at a total lower cost by purchasing the 50 suppliers at value $ABD$, and then
using these suppliers to produce more than 50 units so as to reduce the rents it pays to the remaining competitive suppliers. This incentive to vertically integrate derives solely from the ability of the vertically integrated firm to lower the rents that the competitive suppliers receive. This rent effect provides an incentive to fully vertically integrate.

Whether complete vertical integration occurs or not depends on whether the owners of supply are willing to always accept an offer to relinquish their ownership that leaves the owners in the same or better financial position than they would be if they maintained ownership. When such offers will be accepted, then complete vertical integration (i.e., the one monopolist buying out all the suppliers) will occur because of both the rent effect and the efficiency effect. Perry identifies a situation of less than complete vertical integration, namely partial vertical integration. This situation arises when some supply owners refuse to relinquish ownership for the market value of what the supply source is worth. Another way to think of partial vertical integration arising in this model is not from stubborn behavior on the part of suppliers, but rather from some government restriction on the allowable level of vertical integration.

Although Perry does not explicitly bring out the efficiency implications of this rent effect, it is important to note that with partial vertical integration, the rent effect involves efficiency effects on the supply side. In the example just presented, all supply firms are alike, but more supply is being produced in the 50 owned supply firms than in
the 50 competitive supply firms. The marginal cost of production from the
owned supply firms exceeds that from the competitive suppliers. Hence,
the 100 units are being inefficiently supplied.

In Perry's models, vertical integration usually enables the monop-
sonist to be more efficient in its output production. Even if the monop-
sonist is a monopolist in the output market, Perry shows that consumers
benefit from this increased efficiency. When there is no efficiency
effect (i.e., input demand is perfectly inelastic), then allowing either
total or no vertical integration is desirable. When there is an effi-
ciency effect (i.e., input demand depends on price), total vertical in-
tegration is the only desirable outcome. Laws prohibiting total vertical
integration are always undesirable in Perry's model because partial ver-
tical integration causes a production inefficiency associated with the
efficiency effect and a supply side inefficiency associated with the
rent effect.
Conclusion

The four models just discussed present a wide range of views about vertical integration. Given the variety of the circumstances that can lead to vertical integration, it is not terribly surprising that the models should postulate widely differing assumptions and reach different conclusions.

In most of the models, there is a tendency for the market to be one of either no vertical integration or complete vertical integration. If a little vertical integration is good, then more vertical integration is better. In the real world, there are many cases of partial not total vertical integration, in which a firm produces some of its inputs and relies on other independent suppliers for the remainder. It would be desirable to explain why partial vertical integration is one form of market structure. In neither the Arrow nor Green model does partial integration occur. In Perry's model, partial integration can occur if input suppliers are unwilling to sell their source of supply at market value. In Carlton's model, partial integration can arise under the assumption that vertically integrated firms do not sell inputs to each other (or alternatively, do not sell to other sectors of the economy). The Carlton and Perry model are in agreement that partial integration is not a desirable outcome, even though the Carlton model finds total integration undesirable (from a static viewpoint), while the Perry model finds it desirable.

Regarding the three models dealing with uncertainty, it should be mentioned that vertical integration is only one response to the market
forces. Actual ownership of the input supplier need not occur. For example, in the Arrow model the information on supply could be sold on a market. A curious feature of the Arrow model is that it can be in the suppliers' best interests to make supply information available free. The improved information causes output producers to make more efficient utilization of resources, and, at least in the case of elastic output demand, the price of the input rises.

In both the Green and Carlton models, there is no difference between a firm buying input at cost c or producing it itself at cost c. "Long-term" contracts\(^1\) that charge c per unit are indistinguishable from the arrangement in which the firm itself produces the input at cost c. With constant returns to scale, "long-term" contracts to buy at cost is a way to vertically integrate. Once such contracts are allowed, the results of the Green model would be unchanged only if it is possible to explain why an output firm (which, remember, has a fixed demand for input each period) could not sign a "long-term" contract with an input firm to always insure supply. Such a contract could be profitable for both the input and output firm and would insure that only the efficient technologies are used to produce input and output. Arguments concerning trust and enforceability of such contracts could possibly be advanced to explain why in Green's model a spot market exists but not a "long-term" contract market. Presumably, in the real world costs associated with

\(^1\) "Long-term" contracts in the models are simply a one period contract in which an output firm guarantees to accept delivery of an input at the constant cost of production, c.
contracts (Williamson (1971)) explain why in-house production often replaces contracts as the means to vertically integrate.

A similarity of the Arrow, Carlton and Green models is their emphasis on the role of decision making under uncertainty. It is the requirement that certain input decisions have to be made before the uncertain future can be observed that makes these models interesting.

All the models find incentives for vertical integration but differ considerably in their assessment of vertical integration. In the Arrow model, vertical integration is a means to obtain information. Provided such integration does not entail monopoly power, vertical integration is one mechanism to attain efficiency. In Carlton’s model, incentives for vertical integration exist because a vertically integrated firm can impose a differential risk between the use of its inputs and the use of other firms’ inputs. The firm uses its input to insure that it can satisfy low risk demand and pass off the high risk demand to someone else.

Since input firms are more efficient poolers of demand risk than are output firms (who recall do not sell their input), vertical integration is, from a static viewpoint, usually undesirable and never beneficial in Carlton’s model. From the viewpoint of introduction of technology, vertical integration can be a desirable market structure. The Green model finds an incentive to vertically integrate to insure a source of supply. The assumed inefficiency of vertically integrated firms makes welfare statements about the desirability of vertical integration depend on balancing the extra

\[1/\text{At the risk of bringing out the obvious, all these assessments are based on whether vertical integration is an efficient market structure in a partial equilibrium model. Income distribution and general equilibrium considerations are not dealt with.}\]
costs caused by inefficient production versus the benefits of having an assured supply. In Perry's model, vertical integration can occur for two reasons - the "efficiency" effect and the "rent" effect. When the efficiency effect operates, then total vertical integration is desirable. When only the rent effect operates, then there are incentives for vertical integration to exist, even though, from an efficiency standpoint, there is no difference between a world with no vertical integration and one with complete vertical integration.

Interestingly enough, there is a close analogy between the Perry and Carlton model for the special case in which there is no efficiency effect associated with vertical integration. In Carlton's model, for this special case the strong incentive to vertically integrate still exists for the same reason as outlined above - namely the ability of vertically integrated firms to use their inputs first. Thus, both Carlton and Perry find strong incentives for vertical integration, even when such vertical integration is from an efficiency viewpoint of no consequence.

Since the Green and Carlton models address a similar question, it is useful to draw out the differences in assumptions between the two models. For both Green and Carlton, input markets do not always clear. In Green's model, price is exogenous and supply and demand are out of balance because random forces from outside sectors exert influence on supply and demand. In Carlton's model, price is endogenous, and the randomness in demand is generated by the random demands of customers in the sector of concern. Green is concerned with analyzing what happens when the uncertainty of others affects a firm through the input market. Carlton
is concerned with analyzing what effect uncertainty of demand in an industry has on the input supply and the incentive to integrate for that industry. In Green's model, vertical integration is a way to bypass completely dealing with random elements. In Carlton's model, a vertically integrated firm still faces uncertain demand for its output. In Green's model, there is no risk facing the output firm so that an output firm can provide input more efficiently from a risk viewpoint to itself than can the input market. In Carlton's model, just the reverse is true. In that model, the input market pools uncertain demands from other firms (or alternatively from other sectors) and hence provides the input more efficiently from a risk viewpoint than a vertically integrated firm. In Green's model, the integrated firm has an inferior technology for producing input and output. This is not the case for Carlton's model. In both models, the vertically integrated firm is a less efficient provider of the input than the input firm for either technological reasons (Green) or risk sharing reasons (Carlton). Both models deal with the problem of assuring supply, but the two models ask different questions and not surprisingly reach different conclusions regarding the desirability of integration.

The classical assumptions of complete information, instantaneous and costless market clearing, and competitive markets with constant returns to scale in production lead to a model that uncovers no incentives or disincentives for vertical integration to occur. Models that explicitly relax these traditional assumptions must be investigated if we are to gain insight into the reasons for and consequences of vertical integration. Analyses of vertical integration in the presence of noncompetitive markets
aspects of American industry today. Moreover, studying vertical integration as a response to risk is time well spent. Chandler (1964, p. 37) in his study of vertical integration concluded that a primary force motivating vertical integration was not to lower unit costs — but specifically to obtain a more certain source of supply.

Our understanding of the incentives for and consequences of vertical integration has improved recently, but much remains to be done at both a theoretical and empirical level. The subject of vertical integration is a challenging one for economic theory and an important one for public policy.

X. Arrow, "Vertical Information and Communication," The Bell Journal of Economics, Spring, 1975


J. Green, "Vertical Integration and Assurance of Markets," Harvard Institute of Economic Research, Discussion Paper 363, October, 1974

L. McKenzie, "Ideal Output and the Interdependence of Firms," Economic Journal, December, 1951


