

# *Profit maximising rigid prices and vertical integration*

Article

Accepted Version

Wadeson, Nigel (2017) Profit maximising rigid prices and vertical integration. *International Journal of the Economics of Business*, 24 (1). pp. 53-72. ISSN 1466-1829 doi: <https://doi.org/10.1080/13571516.2016.1199457> Available at <https://centaur.reading.ac.uk/65897/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1080/13571516.2016.1199457>

To link to this article DOI: <http://dx.doi.org/10.1080/13571516.2016.1199457>

Publisher: Routledge

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# **Profit Maximising Rigid Prices and Vertical Integration**

By Nigel Wadeson<sup>1</sup>

---

<sup>1</sup>Centre for Institutional Performance  
University of Reading  
PO Box 218  
Reading  
RG6 6AA  
UK

n.s.wadeson@rdg.ac.uk  
Tel +44 (0)118 3785070

## Profit Maximising Rigid Prices and Vertical Integration

### *Abstract*

This article explores profit maximising rigid pricing for a price setting firm and relates the results to vertical integration, which is an important area of corporate strategy and antitrust policy. The setting of a profit maximising rigid price is investigated in the face of a known distribution of short-run demand levels as a compromise between the flexible prices that would be appropriate in the short run at different levels of demand. The price and level of capacity are therefore set to maximise expected profits across varying levels of demand. With the help of computer simulations, it is shown that price rigidity increases the incentives for vertical integration particularly where upstream production is capital intensive, due to the increased importance of rationing. The incentives will also be particularly strong for more efficient and more capital intensive downstream production with low short-run marginal costs.

**Keywords:** price rigidity, vertical integration, restraints, rationing, capacity.

**JEL:** L10, L23, D42, D45

### **Introduction**

This article explores profit maximising rigid pricing for a price setting firm and relates the results to vertical integration. It will be shown that the capacity cost level has a key influence on the setting of a rigid price relative to the related profit maximising flexible price levels. This then determines the frequencies and severities of periods of rationing of supply relative to periods of excess supply. While purchasing from an upstream supplier allows the pooling of demand uncertainties (Carlton, 1979: 199), and hence a reduced level of upstream capacity, in periods of rationing firms with high marginal values of the intermediate product can be left short of supply.

Vertical integration and vertical relationships between firms are a key area of firm strategy and of antitrust policy. Many supply chains have been transformed in recent times by globalisation and modern information and transportation technologies, with some becoming increasingly ‘fine sliced’, pointing to a need for further research in this area. There are various established theories of vertical integration (Casson, 1984). Among them is the idea that it can result from demand variability where prices are subject to short-run rigidity. This results in downstream firms being subjected to rationing and upstream firms bearing costs of unused capacity or unsold production. Perhaps the best-known explanation for price rigidity is that of menu costs, though there are also various other theories (Blinder et al., 1998).

Both supply assurance and demand assurance have been shown to be important motives for vertical integration. For instance, shortages of supply led the Ford Motor Company to engage

in fairly extensive vertical integration by the 1920s (*The Economist*, 2009). In a study of the automobile industry, White (1971) found that supply risks motivate vertical integration. However, some modern motor manufacturers, such as Toyota, instead have close relationships with suppliers. Chandler (1976: 287-290) claimed that, with the rise of mass manufacturing, problems of ensuring adequate demand and supplies in capital intensive industries led to large costs of underutilised capacity to which firms responded by vertically integrating. In order to achieve more stable demand, a manufacturer may forward integrate into distribution or be linked to distributors by vertical relationships such as franchising contracts. A more recent study surveyed 150 senior executives and found that supply assurance measures, including forming long-term relationships with key suppliers, were considered key to flexibility by 72% of respondents (CSCMP's Supply Chain Quarterly, 2011). Some industries face significant demand variability combined with high capacity costs while the value and rapid obsolescence of products can make inventory holding expensive (Wu et al., 2005: 126-8).

As vertical integration is significant in antitrust policy it is important that there is a good understanding of efficiency-based incentives for it. Vertical integration and some vertical restraints, such as exclusive dealing arrangements, may sometimes be at least partly aimed at addressing the costs of price rigidity. The frequency and severity of periods of rationing and of excess supply arising from price rigidity can be seen as a key part of the economic performance of firms and industries.

Green (1986) showed that firms have an incentive to vertically integrate in order to avoid market rationing. In Green's model, firms are competitive and demand in the intermediate product market is subject to stochastic variations in excess demand. The decision of how much of the non-durable intermediate product to produce has to be made before demand is known in each period. Prices in the model are fixed and exogenous. Full vertical integration and full vertical disintegration are shown to be generally the only stable equilibria. Intermediate outcomes are inherently unstable.

Carlton (1979), citing an earlier discussion paper version of Green, argued that the traditional assumption that prices change instantly to clear markets is mostly unrealistic and that this is important to the question of vertical integration, citing evidence on the importance of supply assurance to vertical integration from Chandler (1964: 37, 84). He considered the case of a competitive market with identical firms, a known distribution of demand variation, and a price that is rigid in respect of that variation. The upstream firms do not know the level of demand in advance of deciding how much to produce in each period. Any unsold output is discarded. Carlton reasoned that the price must be greater than marginal cost due to the need to allow for unsold production. He demonstrated that downstream firms can have an incentive to vertically integrate production for part of their demand, as each firm requires at least that part of its demand with greater certainty than for higher levels of demand. Note that partial vertical integration seems to be fairly common (Krzeminska et al., 2013).

Carlton did not solve for the optimal price in his mathematical model. However, he supplemented it with a diagrammatic analysis consisting of iso-utility and zero-profit curves

relating to the price and the probability of satisfaction (i.e. the probability of being able to supply a customer in any period). He reasoned that the zero-profit curve must be upward sloping as a higher probability of satisfaction requires a greater level of output and hence greater costs of unsold output and a higher price. Note that he assumed that upstream firms are frequented by customers based on the combination of the probability of satisfaction and the price. He also made the key assumption that each customer frequents only a single upstream firm per time period, rather than continuing to search if unable to gain supply from the first upstream firm approached. All upstream firms are driven by competition to offer the same level of utility to customers, based on the combination of price and probability of satisfaction.

Hendrikse and Peters (1989) made a further contribution to this stream of research, following in the spirit of Green. Similar to Carlton, they found partial vertical integration to be an equilibrium outcome. However, they did so by different means, namely by assuming differing risk attitudes of market participants. They also assumed that the price is fixed, following Green, and that it is equal to the lowest reservation price of buyers in the market.

This article will address a gap in the literature in relation to the price of the intermediate product for the case of price-setting upstream firms with market power. Some of the relevant vertical integration literature takes the price as given, while Carlton provides a diagrammatic analysis of the price in a competitive market based on the assumption that each downstream firm only ever frequents a single supplier in each time period. Also, this article will focus on rationing of supply as a motive for vertical integration in addition to excess supply which was the focus of Carlton's model.

In the model presented below the price is determined endogenously and is that of a monopoly upstream firm, as the simplest case of a price-setting firm. It is, however, hoped that the results on price setting will prove relevant in non-monopoly cases in future research. The monopoly is assumed to charge a single price for all units sold. While this is essentially a simplifying assumption, the monopolist might not have good information over the efficiencies of the downstream firms with which to price discriminate and a resale market could come into existence. Also, discriminating against more capital intensive and efficient firms, based on short-run marginal values, would be unlikely to be an effective strategy as it could simply result in the relevant firms becoming more labour intensive and less efficient. Additionally, introducing a degree of competition in the upstream industry would in any case curtail or eliminate the ability to price discriminate. The monopolist could instead form special relationships with a subset of firms. For instance, it could try to avert backward integration by some downstream firms by forming longer-term contracts which involve supply priority and additional capacity investments by both the upstream and downstream firms. However, such arrangements would amount to vertical quasi-integration. A single price is therefore assumed in the model as the simplest way of exploring the setting of rigid prices in the absence of vertical quasi-integration. The model assumes that in each period demand is observed before the decision over the level of production is made. Therefore, if demand is not high enough there is unused capacity rather than unsold production.

Demand variability is also part of transaction cost theory. However, transaction cost theory could be argued to be less relevant in situations where demand variations for an intermediate product have a known distribution. Unexpected shifts in the distribution might be more likely to result in problems of opportunistic haggling due to the need to revise the terms of contracts between upstream and downstream firms. However, transaction cost theory could help to explain the nature of the contract used in a vertical relationship and why vertical integration might sometimes be preferred to vertical quasi-integration.

Peak-load pricing is a further body of literature that concerns pricing and rationing with demand variability (Carlton, 1977; Crew et al., 1995). However, it deals with social welfare maximising pricing rather than profit maximising pricing which is of concern here.

Note that the more general literature on rigid prices focuses its attention on why a price may be rigid (or ‘sticky’) and on testing for rigidity rather than on the level at which a rigid price will be set. Some literature has also explicitly modelled price inertia. For instance, Taylor (1980) assumed that prices are set according to competitor prices and the level of excess demand. However, the pricing studied in this article is not that of a process of readjustment towards a conventional equilibrium level. Instead the price represents a compromise made to maximise profits over varying levels of demand.

## **The Model**

It is assumed that an upstream firm sets a long-run equilibrium price and capacity level. Assume, for simplicity, that the firm is a monopoly. It is further assumed that the firm’s price is fixed in the face of short-run fluctuations in demand, as each shock is relatively short-lived. Each shock lasts for a single period and the resulting demand variations have a known distribution. Therefore, the equilibrium price and capacity levels reflect the distribution of the short-run variations. Assume that output either cannot be stored between periods or that it is too expensive to do so. However, the effects explored in the model will still have force so long as the stock mechanism is costly to use.

Assume that the downstream firms are competitive price takers. Because the downstream firms are price takers, there is no issue of double marginalisation, or of two-part tariffs resulting from it, in the model. However, assume that the downstream firms vary in their cost functions. This will be of significance in terms of the marginal values that the downstream firms place on supplies of the intermediate product. Assume also that they differ in terms of their distributions of quantities demanded across different time periods, as will be enlarged on below.

Consider first the upstream firm’s choice of its capacity level,  $Q_{cap}$ . Assume that the firm decides how much to produce at the beginning of each time period after the level of demand in that period has become known. The cost of each unit of capacity per period is constant at  $c_{cap}$ . The marginal cost of producing a unit of output using an available unit of capacity is the constant value  $c_{prod}$ . The price is  $P$  ( $P > c_{prod} + c_{cap}$ ). It is assumed that this is the only price charged for all units sold and that there is no fixed-fee component of the pricing structure.

It is useful to assume that demand has a symmetrical triangular distribution. This is intended to be representative of what might be seen as more realistic distributions, in as much as the density has a peak and tails, and has the advantages that it is amenable to the mathematics that follows and that negative values can be fully excluded. The use of a triangular distribution would also facilitate the investigation of the effects of skewness in future research.

The quantity demanded is  $q_t = q(P, s_t)$ , which is assumed to be continuous and strictly decreasing in  $P$ , where  $s_t$  is the shock to demand in time period  $t$ . This is a derived demand and reflects the fact that the downstream firms are price takers in both the intermediate market and the market for their output. It is assumed that rationing of consumers is always due to lack of upstream rather than downstream capacity. Assume that demand pooling efficiencies mean that upstream capacity is always less than total downstream capacity, as will be explored later, and that a consumer who cannot be served by one downstream firm in any time period can continue search for supply across other firms.

For any given price, the quantity demanded is distributed between a maximum level of  $q_{max} = q_{max}(P)$  and a minimum of  $q_{min} = q_{min}(P)$ . The mean of the distribution is denoted  $\bar{q} = (q_{min} + q_{max})/2$ . The probability distribution function for the quantity demanded is  $F(q_t, P)$ :

$$F(q_t, P) = 2 \left( \frac{q_t - q_{min}}{q_{max} - q_{min}} \right)^2, \text{ where } q_t \leq \bar{q} \quad (1.1)$$

$$F(q_t, P) = 1 - 2 \left( \frac{q_t - q_{max}}{q_{max} - q_{min}} \right)^2, \text{ otherwise} \quad (1.2)$$

The distribution is further illustrated in Figure 1.

[Figure 1 near here]

Expected profits are composed of expected revenues net of the short-run production costs for periods when demand is at or below capacity, plus those for periods when it exceeds capacity, less the total cost of capacity:

$$E(\pi) = \left( \int_{q_{min}}^{Q_{cap}} f(q_t, P) q_t dq_t + \left( 1 - F(Q_{cap}, P) \right) Q_{cap} \right) (P - c_{prod}) - Q_{cap} c_{cap} \quad (2)$$

Maximising expected profits involves maximising with respect to both price and capacity. The profit maximising price will be considered later. The simplest way of solving for the profit maximising level of capacity, in terms of the price, is by setting the marginal expected benefit of capacity per period ( $MEB_{cap}$ ) equal to the marginal cost of capacity per period<sup>i</sup>.  $MEB_{cap}$  is the product of the probability that demand is at least as great as  $Q_{cap}$  multiplied by the margin of the price over the short-run marginal cost of production:

$$MEB_{cap} = \left( 1 - F(Q_{cap}, P) \right) (P - c_{prod}) \quad (3)$$

Setting this equal to the marginal cost of capacity,  $c_{cap}$ , gives:

$$F(Q_{cap}, P) = 1 - \frac{c_{cap}}{P - c_{prod}} \quad (4)$$

Solving for the capacity level,  $Q_{cap}$ , for the assumed distribution gives:

$$Q_{cap} = q_{min} + (q_{max} - q_{min}) \sqrt{\frac{P - c_{cap} - c_{prod}}{2(P - c_{prod})}} \quad , \text{ where } Q_{cap} \leq \bar{q} \quad (5.1)$$

$$Q_{cap} = q_{max} - (q_{max} - q_{min}) \sqrt{\frac{c_{cap}}{2(P - c_{prod})}} \quad , \text{ otherwise} \quad (5.2)$$

The determination of the level of capacity involves a trade-off between the cost of capacity and the margin between the price and the marginal cost of production that determines how close the capacity level is either to the minimum or the maximum possible quantity demanded. A low cost of capacity relative to the mark-up means that the optimum level of capacity is close to  $q_{max}$ . A high cost of capacity relative to the mark-up means that the optimum level of capacity is close to  $q_{min}$ .

It can also be seen that the condition  $Q_{cap} \leq \bar{q}$  is equivalent to:

$$c_{cap} \geq \frac{1}{2}(P - c_{prod}) \quad (6)$$

There is rationing of supply in time periods when the quantity demanded exceeds  $Q_{cap}$ . The expected excess demand per period,  $E(Q_{ration})$ , is:

$$E(Q_{ration}) = \int_{Q_{cap}}^{\frac{q_{min}+q_{max}}{2}} \frac{4(q_t - q_{min})}{(q_{max} - q_{min})^2} (q_t - Q_{cap}) dq_t + \int_{\frac{q_{min}+q_{max}}{2}}^{q_{max}} \frac{4(q_{max} - q_t)}{(q_{max} - q_{min})^2} (q_t - Q_{cap}) dq_t \quad , \text{ where } Q_{cap} \leq \bar{q} \quad (7.1)$$

$$E(Q_{ration}) = \int_{Q_{cap}}^{q_{max}} \frac{4(q_{max} - q_t)}{(q_{max} - q_{min})^2} (q_t - Q_{cap}) \quad , \text{ otherwise} \quad (7.2)$$

Reductions in  $Q_{cap}$  within the probability distribution of  $q_t$  increase expected rationing. Substituting for  $Q_{cap}$  with the values given in equations 5.1 and 5.2 then gives:

$$E(Q_{ration}) = (q_{max} - q_{min}) \left( \frac{1}{2} - \left( 1 - \frac{P - c_{cap} - c_{prod}}{3(P - c_{prod})} \right) \sqrt{\frac{P - c_{cap} - c_{prod}}{2(P - c_{prod})}} \right) \quad , \text{ where } Q_{cap} \leq \bar{q} \quad (8.1)$$

$$E(Q_{ration}) = (q_{max} - q_{min}) \frac{1}{3\sqrt{2}} \left( \frac{c_{cap}}{P - c_{prod}} \right)^{\frac{3}{2}} \quad , \text{ otherwise} \quad (8.2)$$

Expected rationing is increasing in the range of demand variation, in the cost per unit of capacity, and in the marginal cost of production. Also, a higher price, relative to the other factors, results in a lower expected quantity of rationing, assuming that a higher price does not lead to a greater range of demand variation.



The ratio of the expected quantity of rationing to the expected quantity demanded per time period, a measure of supply unreliability, is positively related to the proportionate level of demand variability, as well as the cost of capacity and the short-run marginal cost of production.

The mark-up that results from market power is already well-recognised as being important in terms of double marginalisation, excessive substitution of the intermediate product in downstream production (Warren-Boulton, 1974), and a downstream firm having an incentive to backwardly integrate simply in order to avoid paying a high price. However, in existing literature it has not been clear how the mark-up relates to theories of vertical integration based on rigid pricing. It has been shown above that it leads to increased capacity and reduced levels of rationing for a given level of demand variability facing the upstream firm, a given unit cost of capacity, and a given price. However, it remains to be seen how the price is determined. This will be explored later.

Now consider the incentive for a downstream firm to internalise production of the intermediate product. Assume that the firm takes current upstream supply conditions as given. Assume, for convenience, that the marginal value of the intermediate product for downstream firm  $i$ ,  $v_i$ , is constant for each unit of its demand and that it is equal to the price in the downstream market,  $p_d$ , minus the short-run marginal cost of production (excluding the cost of purchasing the intermediate product),  $c_i$ ;  $v_i = p_d - c_i$ .

A U-shaped long-run average cost curve could have been assumed, which would help to explain why the industry is competitive, but it is a useful simplifying assumption to have a constant short-run unit cost. As the model involves a hard short-run capacity constraint, the conventional assumption of production rising in the short-run to the point where price equals short-run marginal cost is not relevant. It is, however, similar to a case where short-run marginal cost rises sharply beyond some point.

Assume that competition in the downstream industry results in the downstream price being determined by the constant short-run marginal cost of labour intensive downstream firms (again excluding the cost of purchasing the intermediate product),  $c_L$ , and by the price of the intermediate product. However, some downstream firms are more efficient and more capital intensive. It is these more efficient and capital intensive firms that are of interest in the present context in terms of vertical integration, as will be seen.

$$p_d = c_L + P \tag{9}$$

And so the marginal value of supplies to firm  $i$ , up to the point where it is itself at capacity or has satisfied the demand that it faces, depends on its production cost advantage over less efficient and more labour intensive firms:

$$v_i = P + c_L - c_i \tag{10}$$

Assume that, for any given level of market rationing of the intermediate product, firm  $i$  will face a given proportion,  $z_i$ , of that rationing on average. Backward integration allows the firm

to escape rationing and leads it to adjust its capacity level. The change in the firm's total cost of capacity if it vertically integrates is  $c'_{icap}(Q_{icap})$ , where  $Q_{icap}$  is its new capacity level. This includes both its cost of capacity for the intermediate product and the change in the cost of capacity for its downstream production. The latter is the result of the adjustment to the firm's capacity level caused by integration. The incentives regarding capacity levels both before and after vertical integration are considered below.

The expected value of the effect on sales of the adjustment in the firm's capacity level following vertical integration is  $E'_{Si}(Q_{icap})$ . This depends on both the capacity adjustment and the frequency of use of each unit of capacity concerned, which depends on the probability distribution of demand facing the firm. The firm's average level of output per period before vertical integration is  $\bar{q}_i$  and its constant short-run marginal cost of producing the intermediate product is  $c_{il}$ .

The incentive for downstream firm  $i$  to internalise production of the intermediate product can then be represented by the following expression:

$$(z_i E(Q_{ration}) + E'_{Si}(Q_{icap}))(v_i - c_{il}) - \bar{q}_i(c_{il} - P) - c'_{icap}(Q_{icap}) \quad (11)$$

It would be possible to model a case where a downstream firm's capacity is exactly matched to a fixed proportion of upstream capacity. The downstream firm would have a fixed proportionate market share and the demand faced by it would therefore vary according to overall downstream market demand variations. At times of rationing, the proportion of the available supply of the intermediate product allocated to the downstream firm would be equal to its fixed market share. In such circumstances, a fixed proportion of upstream capacity is effectively dedicated to the downstream firm.

However, it is more realistic to assume instead that the demand faced by any downstream firm is affected by firm-level variations in demand in addition to those concerning overall market demand. This means that extra capacity continues to have a positive value to the downstream firm even after its capacity level exceeds that part of the upstream firm's capacity available to it at times of rationing. So the downstream firm has an incentive to invest in some level of additional capacity.

Assume a degree of randomness in the spreading of demand across different downstream firms. Some consumers regularly frequent the same firm whenever they want to buy. However, their loyalty does not extend to being willing to pay a higher price than that charged by other downstream firms, or at least not by more than an insignificant amount. Other consumers randomly select which downstream firm to visit for each purchase. Consumers continue to other firms if they have not already been offered available supply. Variations in each downstream firm's demand for the intermediate product in each time period therefore reflect a combination of the effects of changes in overall downstream market demand, variations in how random visitors' demand is spread across different firms, and variations in factors specific to the firm's regular customers.

Assume that the demand facing downstream firm  $i$  can be expressed as follows:

$$q_{it} = \alpha_i q_t (1 + e_{it}) \quad (12)$$

In this expression  $e_{it}$  represents the firm-level shock to the demand faced by firm  $i$  at time  $t$ , where  $\sum_i e_{it} = 0$ .  $q_t$  is overall market demand and  $\alpha_i$  represents the firm's proportionate market share when  $e_{it} = 0$ . For simplicity, assume that spillovers of demand from firms that have reached capacity are included in the  $e_{it}$  variable. It is therefore necessary to treat  $e_{it}$  and  $q_t$  as jointly distributed because such spillovers will be greater during periods of high industry-level demand. Spillovers provide some reduction in the impact of firm level shocks.

Consider first firm  $i$  without vertical integration. Outside of times of rationing it will sometimes receive more demand than  $\alpha_i Q_{cap}$  due to positive firm-specific demand shocks.

Its marginal expected benefit of expanding capacity beyond  $\alpha_i Q_{cap}$  is then:

$$\int_{\frac{Q_{cap}}{1+e_{imax}}}^{Q_{cap}} \int_{\frac{Q_{cap}}{q_t}}^{e_{imax}} f(q_t, e_{it}) (c_L - c_i) d e_{it} d q_t \quad (13)$$

The lower limits of the integrals are such that  $q_{it} > \alpha_i Q_{cap}$ . The expression demonstrates that the presence of firm-level demand shocks creates an incentive to invest in additional capacity, over and above the level  $\alpha_i Q_{cap}$ . There is therefore a resulting mismatch in capacity levels between the upstream and downstream firms. Some downstream firms will therefore have idle capacity at times of rationing of the intermediate product. Note that it can be seen in the expression that this pertains to the more efficient and capital intensive firms in the presence of significant firm-specific shocks. There is therefore a greater incentive for vertical integration in order to address idleness of capacity imposed by rationing where downstream production is relatively efficient and capital intensive.

Now consider the vertically integrated firm  $i$ 's marginal expected benefit from investing in capacity beyond a given level,  $Q_i$ , with its intermediate product capacity matched to its downstream production capacity. The limits of the inner integral are set such that  $q_{it} > Q_i$ .

$$\int_{\frac{Q_i}{\alpha_i(1+e_{imax})}}^{q_{imax}} \int_{\frac{Q_i}{\alpha_i q_t}}^{e_{imax}} f(q_t, e_{it}) (P + c_L - c_i - c_{ii}) d e_{it} d q_t \quad (14)$$

Note that the downstream price,  $p_d$ , has replaced the intermediate product price,  $P$ , in contrast to equation 5, but with  $P$  reintroduced in expanding  $p_d$  as per equation 9. It is the marginal value of the intermediate product to the integrated firm that determines the level of capacity and this depends on its production cost advantage,  $c_L - c_i$ . This supplements the incentives for more efficient firms to undertake vertical integration; increasing their capacity levels to take account of the freedom from rationing that vertical integration affords.

These effects represent a different argument to that explored in Carlton's (1979) model which demonstrated a motive for partial vertical integration relating to the firm's high probability demand. This was based on the problem of excess supply; the probability of not being able to sell units of the intermediate product. Periods of excess demand in the intermediate product market will sometimes be more important than periods of excess supply in leading to either vertical integration or vertical quasi-integration.

The incentive to backwardly vertically integrate is reduced by pooling efficiencies relating to those parts of the downstream firm's variability of demand that are not common to the market as a whole. Such variations are pooled across many downstream firms in the demand facing the upstream firm which is therefore left with the variation in demand for the market as a whole. However, the use of vertical quasi-integration maintains pooling efficiencies.

Forward vertical integration by the upstream firm is another possibility. This could combine diversified demand sources with more efficient levels of capacity and supply reliability. However, it would also depend on the ability of the monopoly to operate efficiently in the downstream industry in which it may be in other ways disadvantaged, such as due to downstream diseconomies of scale. In some cases the monopoly might become a franchisor rather than engaging in fully-fledged vertical integration.

There is the possibility of some other form of price structure being offered to at least a subset of downstream firms to help address the problem of inefficient allocations under rationing and to induce the upstream firm to increase its capacity for the sake of downstream firms with higher marginal values of the intermediate product. This is subject to the problem that, in any single time period, there is asymmetric information between the upstream firm and the downstream firms over the total quantity of demand faced by the upstream firm. This is due to demand faced by the downstream firms being partly determined in each time period by firm-specific shocks rather than industry-level demand alone. Therefore, in any single period, the downstream firms cannot be certain of the degree of supply priority that they are actually being given.

One strategy would be to offer more reliable supply to the more efficient and capital intensive downstream firms as a way of preventing backward integration. However, to be fully effective, the fees charged to any particular downstream firm would need to depend on the distribution of that firm's demand, in terms of both the mean value and the variability. It would also need to allow for different levels of downstream firms' marginal values. There could therefore be significant bargaining issues involved in negotiating the level of a fixed fee, or some more complex arrangement. Such negotiations could be impeded by a danger for the downstream firms that in volunteering to pay more they declare themselves as having low production costs and so open themselves up to greater rent extraction by the monopolist.

Such an arrangement, in relying on contracting that depends on an agreed degree of reliability of supply being delivered over an extended period of time in the face of a pattern of demand from the downstream firm that may or may not turn out to be as expected by either side when it was agreed, would require continuing relationships between the upstream and downstream firms and contractual safeguards to facilitate adaptation to changing underlying conditions which would amount to vertical quasi-integration. The need for such safeguards would be strengthened by the additional capacity investments needed. A formal example of such a relationship is a just-in-time supply contract. Such a contract, by its nature, has to exist for a significant period of time to be effective and so is a fairly large departure from the pure market.

We are therefore left with the following factors (in addition to differences in unit costs of depending on whether production is upstream or downstream) concerning the incentives for either vertical integration or vertical quasi-integration:

1. It allows for the profitable expansion of both upstream and downstream capacity levels. Capacity can be set at a level that reflects downstream marginal value. Absent of vertical integration and quasi vertical integration, the upstream firm invests in a capacity level that instead reflects the price in the intermediate product market. For some downstream firms, marginal value may be much greater than price. Restricted upstream capacity then leads to lower levels of downstream capacity due to costs of rationing of the intermediated product. With vertical integration, upstream capacity is matched to downstream capacity and both are expanded.
2. It gains the advantage of removing costs of idleness of downstream capacity, in periods of adequate downstream demand, that result from the downstream capacity level being greater than the upstream capacity level. The capacity mismatch happens where additional capacity investments are made downstream due to firm-specific variations in demand.
3. It avoids the problem of some portion of rationed supplies of the intermediate product being allocated to firms with lower marginal values.
4. Backward vertical integration sacrifices the advantages of pooling downstream firm-specific variations in demand, though these can be retained if there is quasi-integration instead.
5. The frequency and severities of rationing help to determine the incentive to integrate and depend on the pricing of the upstream firm. The pricing and the resulting rationing will now be considered.

The expected profit of the upstream firm can be re-expressed as:

$$E(\pi) = E(q_t - Q_{ration})(P - c_{prod}) - Q_{cap}c_{cap} \quad (15)$$

Substituting the values given in equation 5 for  $Q_{cap}$  and the values given in equation 7 for  $E(Q_{ration})$  then gives:

$$E(\pi) = q_{min}(P - c_{cap} - c_{prod}) + (q_{max} - q_{min}) \frac{\sqrt{2} (P - c_{cap} - c_{prod})^{\frac{3}{2}}}{3 \sqrt{P - c_{prod}}} , \text{ where } Q_{cap} \leq \bar{q} \quad (16.1)$$

$$E(\pi) = \bar{q}(P - c_{prod}) + (q_{max} - q_{min}) \frac{\sqrt{2} c_{cap}^{\frac{3}{2}}}{3 \sqrt{P - c_{prod}}} - c_{cap}q_{max} , \text{ otherwise} \quad (16.2)$$

Assume now that mean demand follows the equation  $\bar{q} = a - bP$  and define  $r$  ( $r \leq 2$ ) as the range of variation of quantity demanded as a proportion of the mean quantity demanded:

$$r = \frac{q_{max} - q_{min}}{a - bP} \quad (17)$$

It will help in understanding the profit maximising rigid price,  $P^*$ , to compare it with prices that react with full flexibility to short-run conditions but with the same level of capacity,  $Q_{cap}$ , imposed. Say that there is multiplicative uncertainty, where the shock to demand in an individual time period,  $t$ , is  $s_t$  ( $-r/2 \leq s_t \leq r/2$ ), and  $q_t = (a - bP_t)(1 + s_t)$ . The assumption of multiplicative uncertainty, as opposed to additive uncertainty, means that any fluctuations in demand are themselves sensitive to the price. This form of uncertainty is therefore assumed on the grounds that it seems appropriate in the current context. The profit maximising flexible price in cases where the resulting quantity demanded does not exceed capacity can be found by maximising the following short-run surplus:

$$(a - bP_t)(1 + s_t)(P_t - c_{prod}) \quad (18)$$

This gives the following flexible price:

$$P_{1t} = \frac{1}{2b}(a + bc_{prod}) \quad (19)$$

This gives a quantity traded that is increasing in  $s_t$ :

$$q_{1t} = \frac{1}{2}(a - bc_{prod})(1 + s_t) \quad (20)$$

For higher values of  $s_t$  price is set to match the quantity demanded to the level of capacity,  $Q_{cap}$ :

$$(a - bP_{2t})(1 + s_t) = Q_{cap} \quad (21)$$

This gives the following profit maximising flexible price:

$$P_{2t} = \frac{1}{b}\left(a - \frac{Q_{cap}}{1 + s_t}\right) \quad (22)$$

Note that this price increases with  $s_t$  at a decreasing rate. The associated constant quantity traded is:

$$q_{2t} = Q_{cap} \quad (23)$$

The price  $P_t$  is set at  $P_{1t}$  so long as the quantity demanded does not exceed  $Q_{cap}$  which is the case when the following condition holds:

$$s_t \leq \frac{2Q_{cap}}{a - bc_{prod}} - 1 \quad (24)$$

Otherwise it is set at  $P_{2t}$ . A price of  $P_{2t}$  has the advantage compared to the rigid price,  $P^*$ , that it avoids rationing and so ensures that the output is allocated efficiently, with demand at that

price being fully satisfied. Also,  $P^*$  is above  $P_{1t}$  and so the rigid price is inefficiently high at lower levels of market demand, resulting in excess supply.

## Simulations

Figure 2 was produced by brute-force enumeration, performed by a purpose-written computer programme, to establish the profit maximising value of the rigid price  $P^*$  for many values of the range of demand variability,  $r$ .

[Figure 2 near here]

In Figure 2 increasing demand variability leads to an increase in the rigid price, compromising more strongly between the profit maximising flexible prices for lower and higher levels of demand, so that price rigidity imposes increasing inefficiencies on both the upstream and downstream firms. Capacity falls in this example, tending more towards  $q_{min}$  than  $q_{max}$  (see equation 5), and expected rationing increases. Expected profits fall. The capacity level becomes increasingly inefficient from the point of view of customers with a relatively high valuation of the product. The greater incidence of rationing also increases the losses incurred by higher-valuation downstream firms caused by the rationing mechanism failing to allocate available supplies efficiently. Increased demand variability therefore increases the incentives for vertical integration.

Each graph in Figure 3 was produced by mathematically solving for one value of  $P_{1t}$  and many values of  $P_{2t}$ , while the value of  $P^*$  was again found by brute force enumeration. The figure illustrates the fact that  $P^*$  is a compromise price between the range of the flexible  $P_t$  prices. Rationing of supply occurs to the right of the point where  $P_t$  and  $P^*$  are equal. To the left of this point there is excess supply in the form of idle capacity. In each graph, the severity of rationing or excess supply is increasing in the vertical distance between  $P^*$  and  $P_t$ . In Figure 3b  $P_t$  is always equal to  $P_{2t}$ . In Figures 3a and 3c it is equal to  $P_{1t}$  to the left of the kink and equal to  $P_{2t}$  to the right of the kink. The value of  $P_{1t}$  is the same in all three graphs.

It can be seen from the graphs in Figure 3 that the frequencies and severities of rationing and of excess supply vary substantially across the different cases. The rigid price does not simply even out periods of rationing and of excess supply.

[Figure 3 near here]

What factors determine the weighting of the value of the rigid price,  $P^*$ , relative to the range of values of  $P_t$ ? The height of the  $P_{2t}$  prices above the  $P_{1t}$  prices is a key factor, together with how early the point is reached where  $P_t = P_{2t}$ . The  $P_{1t}$  prices equate short-run marginal cost to marginal revenue. Raising the rigid price a little above  $P_{1t}$  therefore has a fairly small effect on the surpluses earned in time periods with the relevant values of  $s_t$  (i.e. left of the kink) because marginal revenue and marginal cost are still fairly close in value. However, the  $P_{2t}$  prices are set at a higher level due to restricted capacity. Where, for a given value of  $s_t$  to the right of the kink, the  $P_{2t}$  price is substantially above the  $P_{1t}$  price then raising the rigid price

even a little above the  $P_{2t}$  price has a much more significant effect on the surplus earned because marginal revenue is then significantly greater than marginal cost.

Differentiating expression 18 with respect to  $P_t$  gives  $(\bar{q} - b(P_t - c_{prod}))(1 + s_t)$ , which is zero where  $P_t = P_{1t}$ . This shows how the magnitude of the negative impact on the surplus generated in a time period where demand is below capacity is increasing in the excess of the current price over  $P_{1t}$ . In the figure, this holds to the left of where the rigid price and the flexible prices cross.

For instance, a lower level of capacity resulting from a higher unit cost of capacity causes  $P_{2t}$  to become the relevant flexible price earlier into the range of  $P_t$  prices, as can be seen by comparing the different graphs in Figure 3. High costs of capacity therefore mean that the rate of loss of revenues involved in raising the value of the rigid price,  $P^*$ , above values of  $P_{2t}$  is high on average. This results in  $P^*$  being set lower and so crossing  $P_t$  further to the left, so that rationing is more frequent and, on average, more severe than if  $P^*$  were set higher.

The variation in the effect of a rise in the rigid price on the surpluses earned in time periods when there is adequate capacity to supply the quantity demanded, in other words where the rigid price is at or above  $P_t$ , is in contrast to its fixed and unmitigated positive effect at times of rationing of supply. This is because, when there is rationing, the quantity sold remains fixed as the price is raised until the price reaches the point where rationing has been eliminated. So, in time periods where the value of  $s_t$  is to the right of the kink and the rigid price,  $P^*$ , is below the flexible price  $P_{2t}$ , an increase in the rigid price to bring it closer to  $P_{2t}$ , but without crossing above it, has a fixed effect for any given level of capacity.

A further factor is that the weighting is partly determined by the differences in the levels of output concerned at the different values of  $s_t$ . The level of output is at capacity during periods of rationing and is increasing in  $s_t$  at other times. This helps to weight the rigid price towards higher levels relative to the range of  $P_t$  prices when  $P_{1t}$  is the relevant flexible price over a significant part of the range of levels of demand, as determined by inequality 24; in other words, when the level of capacity is well into the range of possible levels of quantity demanded. A higher cost of capacity, lowering the level of capacity, results in  $P_{1t}$  ceasing to be the relevant flexible price further to the left. Hence, a higher cost of capacity again results in a rigid price being weighted lower relative to the range of values of  $P_t$ , crossing  $P_t$  further to the left, resulting in more frequent and severe rationing.

Note that the shape of the probability density is one factor that does systematically increase the weightings of flexible prices that are associated with values of  $s_t$  closer to zero. This is because of its central peak. The strength of this affect is therefore decreasing in the degree of spread of the distribution; in other words, with the degree of demand variability.

Figure 3b represents a case with a higher cost of capacity than Figure 3a and hence a lower level of capacity. It can be seen that the price is weighted lower into the range of  $P_t$  prices. The point where  $P^*$  is equal to  $P_t$  is therefore further to the left and so rationing is more frequent. Figure 3c further illustrates the same point, but with a lower cost of capacity so that



the price is weighted further to the right into the range of  $P_i$  prices and the frequency of rationing is lower.

A higher unit cost of capacity in upstream production therefore leads to a price of the intermediate product that results in more frequent excess demand. Downstream firms face more frequent rationing that is more severe on average. This therefore increases incentives for either vertical integration or vertical quasi-integration in relation to more efficient and more capital intensive downstream production.

This approach to examining the rigid price could be taken further by incorporating choices over backward vertical integration explicitly into the demand function. Rather than simply making the long-run demand function more elastic, this would involve explicitly incorporating the level of reliability of supply into it. The results would depend on the distribution of downstream firms' preferences concerning the trade-off between the price level and the degree of supply reliability. This is something that the monopoly might not have good knowledge of and so its managers would then have to use their judgment to estimate the effects. It would also depend on the strength of the monopoly's cost and product superiority advantages. Note that in more competitive industrial structures firms might have little choice but to set their prices low, in order to be able to sell at times of market excess supply, irrespective of longer-run implications for demand, except where they can form longer-run relationships with individual customers. Without such relationships, there could be a risk of a low price, high rationing industry outcome. Some relationships would then be of a deeper nature in order to further address the problems associated with rigid pricing.

## **Conclusion**

The model and simulations explored profit maximising rigid pricing by an upstream price-setting firm. It was argued that, while purchasing from the upstream firm results in efficiencies from the pooling of demand variations, at times of higher market demand downstream firms that have high marginal values for the intermediate product (due to efficiency and capital intensity) can find themselves subjected to rationing. Vertical integration results in the integrated firm ensuring that it can produce enough of the upstream product for itself to match its downstream capacity level. It also has an incentive to expand its overall capacity level. The advantages of vertical integration can alternatively be captured through quasi integration while still gaining pooling advantages.

The focus is different to Carlton's (1983) model which concentrates on excess supply and resulting costs of unsold production. His model relies on the insight that the market price of the intermediate product must exceed its constant marginal cost of production, as it must also cover the costs of unsold production, to demonstrate an incentive for vertical integration (p. 197). It demonstrated a motivation for vertical integration based on high probability demand and excess supply rather than on excess demand. The current model is also in contrast to Green's (1986: 180) model which treats the price as exogenous. Additionally, both Green and Carlton focussed on competitive upstream markets, whereas here the upstream firm has market power.

In the model, the profit maximising rigid price is a compromise between the flexible prices that would be appropriate in the short-run at different levels of demand. However, the price does not systematically even out periods of rationing and excess supply. Rather, the characteristics of the case concerned determine the frequencies and severities of rationing across different time periods. They can therefore produce more frequent rationing, more frequent excess supply, or an intermediate outcome. With the help of simulations, it was shown that a higher unit cost of capacity results in the upstream firm setting the rigid price so that rationing is more frequent and more severe on average. This keeps the upstream firm working more consistently close to full capacity. Where the upstream firm has a low unit cost of capacity the rigid price is set at a level that results in more frequent excess supply so that the firm can profit more from periods of high demand. These results were explained in relation to the range of quantities sold in different time periods when demand is within capacity, to the excess of marginal revenue over marginal cost where flexible prices are set to prevent the quantity demanded from exceeding capacity, and to the fixed effect on the surplus earned of a higher price in a period of rationing.

A key implication of the model, therefore, is that price rigidity increases the incentives for vertical integration particularly where upstream production is capital intensive, due to the increased importance of rationing. The incentives will also be particularly strong for more efficient and more capital intensive downstream production with low short-run marginal costs.

With rigid prices, the price mechanism is less effective than might be thought at allocating the output of a low capacity level (due to frequent rationing) and at keeping high capacity levels well-utilised (due to frequent excess supply). The upstream firm's profits decline and its price rises as demand variability increases, with the rigid price being pushed further above the marginal cost of production. Both private performance and social efficiency therefore suffer as demand variability increases.

The centrality of the cost of capacity in the model and of firm efficiency, combined with price-setting power, in generating incentives for vertical integration accords with empirical evidence showing that vertical integration is more common in capital intensive and more concentrated industries and among more productive firms (MacDonald, 1985; Hortaçsu and Syverson, 2007). The capital intensity of upstream firms is important in terms of their decisions over their capacity levels and prices. The capital intensity of downstream firms is important in terms of the costs to them of having their production processes disrupted by unavailable supplies. Vertical integration will sometimes be a necessary accompaniment to investment in more capital intensive downstream production.

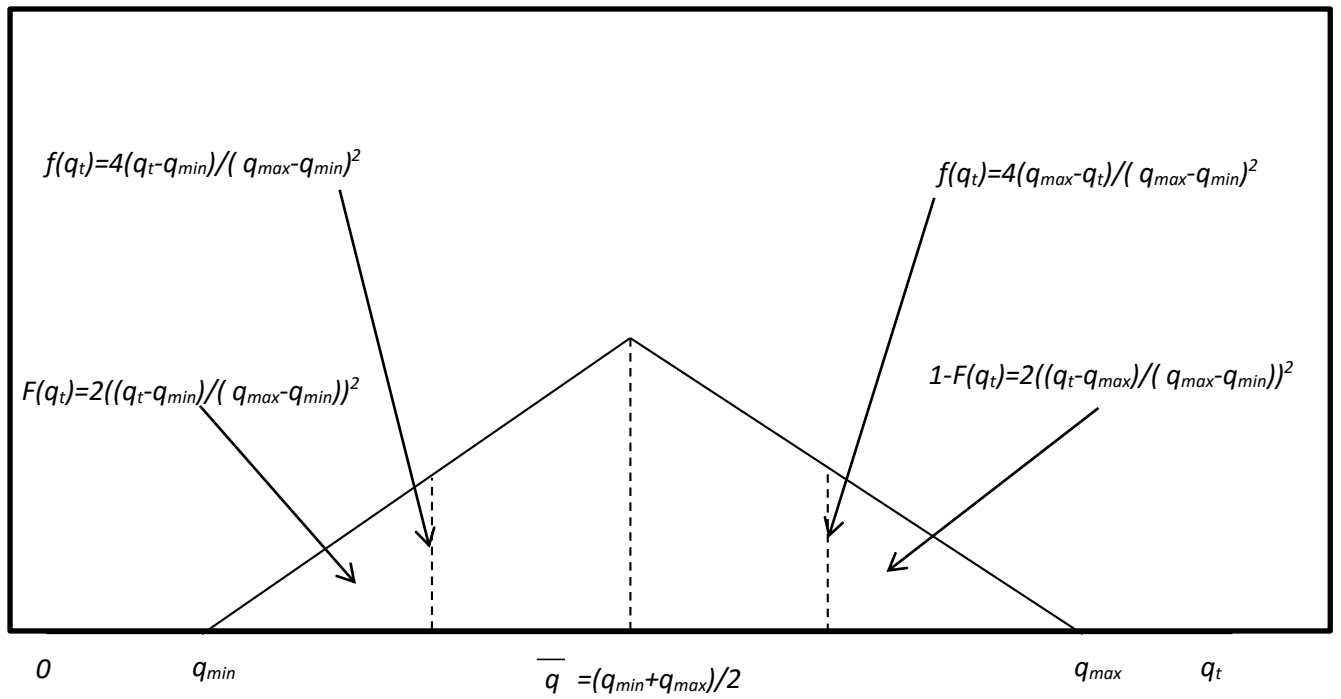
Considering more competitive upstream industries, an alternative to vertical integration is for the upstream and downstream firms to form relationships so that the access given to upstream firms' capacities can be better matched to the needs of particular customers. Each firm could then serve a mix of higher and lower priority customers, so allowing for better capacity usage than when just selling to high priority customers. Some downstream firms might be willing to pay higher prices and to give consistent trade to a supplier in order to gain privileged access

to upstream capacity. On the other hand, a downstream firm might be given a lower price for guaranteeing some minimum level of demand in each period and then charged a fee for the guarantee of supplies to satisfy higher levels of demand. Others might choose instead to buy at low prices, incurring high costs of searching for available supplies across different suppliers when necessary, though at times being unsuccessful in their searches. Without relationships between firms, price competition might result in a low price, low industry capacity outcome. Such relationships would also be deepened in order to further address problems caused by price rigidity.

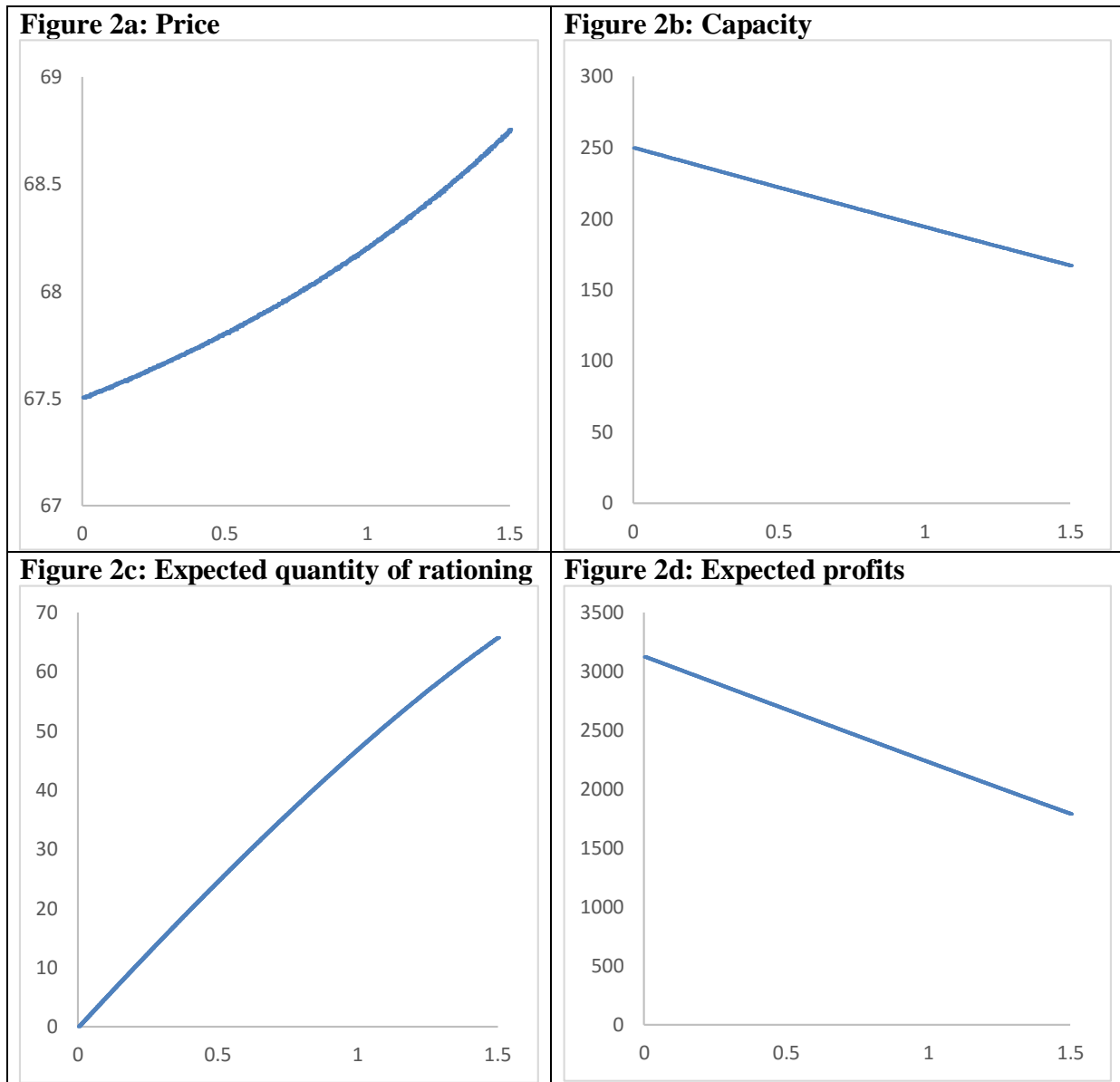
The motives for vertical integration and for relevant vertical restraints that may result from relationships between firms have implications for antitrust policy towards vertical foreclosure. A firm that is vertically integrating or seeking prioritised supply from its suppliers and exclusive dealing or other arrangements such as dedicated retail floor space with transaction specific investments may be significantly foreclosing its competitors. However, the anticompetitive effects need to be weighed against the factors explored in this article; that there can be significant efficiency gains in terms of gaining more stable demand and more reliable supply.

Further work could investigate the effects of alternative industry structures, such as the effects of competition in the upstream industry. It could also investigate the effects of endogenising variables treated as exogenous in the model in order to control its scope.

**Figure 1: The symmetrical triangular distribution**

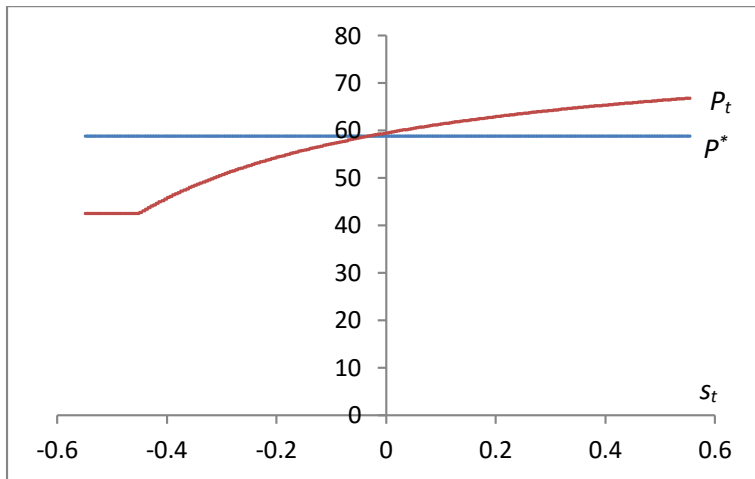


**Figure 2: How price, capacity, expected rationing, and expected profit vary with demand variability ( $r$ )**

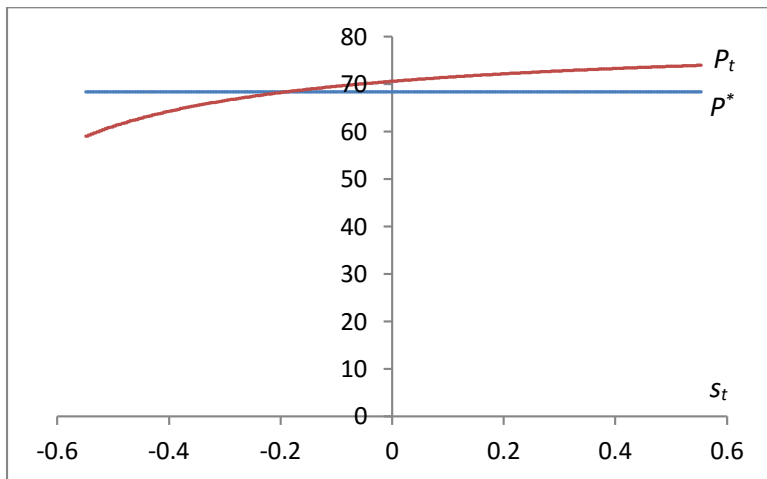


**Figure 3: The rigid price,  $P^*$ , as a compromise price**

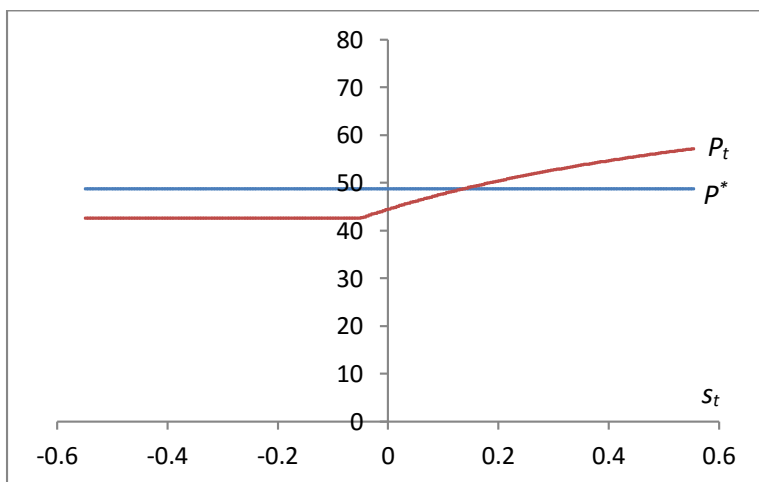
**Figure 3a: With parameter values as per figure 2 ( $c_{cap}=30, r=1.1$ )**



**Figure 3b: With a higher cost of capacity ( $c_{cap}=50$ )**



**Figure 3c: With a lower cost of capacity ( $c_{cap}=10$ )**



## Acknowledgement

I am grateful to the journal's anonymous referees for their help with this article.

## References

- Blinder, Alan S., E. R. D. Canetti, D. E. Lebow and J. B. Rudd. 1998. *Asking About Prices: A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.
- Carlton, Dennis W. 1977. "Peak Load Pricing with Stochastic Demand." *The American Economic Review* 67 (5): 1006-1010.
- Carlton, Dennis W. 1979. "Vertical Integration in Competitive Markets under Uncertainty." *The Journal of Industrial Economics* 27 (3): 189-209.
- Casson, Mark C. 1984. "The Theory of Vertical Integration: A Survey and Synthesis." *Journal of Economic Studies* 11 (2): 3-43.
- Chandler, Alfred. 1964. *Strategy and Structure: Chapters in the History of American Industrial Experience*. Cambridge, MA: MIT Press.
- Chandler, Alfred. 1976. *The Visible Hand: The Managerial Revolution in American Business*. Cambridge: Harvard University Press.
- Crew, Michael A., Chitru S. Fernando and Paul R. Kleindorfer. 1995. "The Theory of Peak-Load Pricing: A Survey." *Journal of Regulatory Economics* 8 (3): 215-48.
- Green, Jerry K. 1986. "Vertical Integration and Assurance of Markets." In *New Developments in the Analysis of Market*, edited by Stiglitz, J.E. and G.F. Mathewson. Cambridge, MA: MIT Press.
- CSCMP's Supply Chain Quarterly, 2011. "More Companies adopt Supply-Assurance Programs." *CSCMP's Supply Chain Quarterly* Quarter 3.  
[http://www.supplychainquarterly.com/news/201103forward\\_supplyassurance/](http://www.supplychainquarterly.com/news/201103forward_supplyassurance/)
- Hendrikse, George and H. Peters. 1989. "A Note on Partial Vertical Integration." *Economics Letters* 30 (3): 249-252.
- Hortaçsu, A. and C. Syverson. 2007. "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." *Journal of Political Economy* 115 (2): 250-301.
- Krzeminska, Anna, Glenn Hoetker and Thomas Mellewig. 2013. "Reconceptualizing Plural Governance." *Strategic Management Journal* 34 (13): 1614-1627.
- MacDonald, James M. 1985. "Market Exchange or Vertical Integration: An Empirical Analysis." *Review of Economics and Statistics* 67 (2): 327-331.

Peters, Michael. 1984. "Bertrand Equilibrium with Capacity Constraints and Restricted Mobility." *Econometrica* 52 (5): 1117-1127.

Taylor, John B. 1980. "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88 (1): 1-23.

The Economist, 2009. "Vertical Integration: Moving on Up." *The Economist*, March 27. <http://www.economist.com/node/13173671>

Warren-Boulton, Frederick R. 1974. "Vertical Control with Variable Proportions." *Journal of Political Economy* 82 (4): 783-802.

White, Lawrence. 1971. *The Automobile Industry Since 1945*. Cambridge, MA.: Harvard University Press.

Wu, S. David, Murat Erkoç and Suleyman Karabuk. 2005. "Managing Capacity in the High-Tech Industry: A Review of Literature." *Engineering Economist* 50 (2): 125-158.

---

i

$$\begin{aligned} \frac{\partial E(\pi)}{\partial Q_{cap}} &= \left( \frac{\partial}{\partial Q_{cap}} \left( \int_{q_{min}}^{Q_{cap}} f(q_t, P) q_t dq_t \right) + \frac{\partial}{\partial Q_{cap}} \left( (1 - F(Q_{cap}, P)) Q_{cap} \right) \right) (P - c_{prod}) - c_{cap} \\ &= \left( f(Q_{cap}, P) Q_{cap} - f(Q_{cap}, P) Q_{cap} + (1 - F(Q_{cap}, P)) \right) (P - c_{prod}) - c_{cap} \\ &= \left( (1 - F(Q_{cap}, P)) \right) (P - c_{prod}) - c_{cap} \end{aligned}$$