Another Look at Market Power in Antitrust

by

Richard Schmalensee

WP #1238-81   July 1981

Another Look at Market Power in Antitrust

Richard Schmalensee*

The recent Harvard Law Review article on "Market Power in Antitrust
Cases" by Professors Landes and Posner (hereinafter referred to as LP)
is an important contribution that deserves careful study.[1] My aim here
is not to quarrel with their basic analytical approach, but rather to point
out some limitations of their analysis and some improper implications for
antitrust policy that might be drawn from it.

Part I deals with the theoretical analysis of market power. I attempt
to clarify the definitions and implications of the measures that LP employ.
In addition, I point out some limitations of the particular market model they
use and some confusion that may arise because of their failure to deal explicitly
with the difference between short-run and long-run analysis. Part II examines
the use of market share as a measure of market power. Even though, as LP note,
the computation of market share is "the standard method of proving market power
in antitrust cases,"[2] it is not a very reliable method in many situations.
Attempts to adjust market shares to reflect departures from "standard" market
conditions or to compute shares when markets are ill-defined can lead
to improper inferences about market power. Part III briefly examines
some alternative approaches to detecting market power that deserve attention
because they may be more reliable under some conditions.

I. Market Power in Theory

As LP note, a firm has market power if it has the ability "to raise
price above the competitive level without losing so many sales so rapidly
that the price increase is unprofitable and must be rescinded."[3] We do
not live in a perfect world, and absolutely perfect competition is rarely
encountered outside textbooks. Almost all firms have some market power,
but most have very little. The relevant question in antitrust cases is

thus not the presence or absence of market power, but rather its _importance_. I begin by examining this question in the static, single-firm model used by LP, then discuss a pair of important complications, and conclude this Part by considering the linkage between single-firm analysis and market parameters.

A. A Single-Product Firm in the Short Run

Following LP, let us consider a single-product firm facing a well-defined demand curve and maximizing its profits in the short run.[4] Then, as they show, at the profit-maximizing point, the following condition holds:[5]

$$L = (P - MC)/P = 1/\epsilon^f , \tag{1}$$

where P is price, MC is marginal cost, ϵ^f is the elasticity of the _firm's_ demand curve, and the first equality defines L, the Lerner index of monopoly power. In Part I of their article, LP treat L as "a precise economic definition of market power" and examine its theoretical determinants by relating the firm demand elasticity, ϵ^f, to other quantities, such as market share. In Part II, however, where they focus on applications of theory, they shift their focus to the deadweight loss caused by monopoly pricing. Because of its greater usefulness, deadweight loss seems the more important of the two concepts.

The deadweight loss caused by monopoly is shown by the sum of cross-hatched areas in C and E in Figure 1, which is essentially identical to Figure 2 in LP. D is the firm's demand curve, and MC is its marginal cost schedule. (The other labeled areas are discussed in Section A of the Appendix.) Deadweight loss is the dollar value of the loss to society caused by the monopolist's failure to increase output from Q^m, assumed to be the profit-maximizing level, to Q^c, the competitive level.[6] If

the MC and firm demand curves are linear, as drawn, it is shown in Section A of the Appendix that deadweight loss is given by the relation

$$DW = L(P^m Q^m)K/2 , \tag{2}$$

where K is between zero and one and equals one when MC is constant.[7]

Equation (2) makes clear the main difference between L and DW: the latter automatically reflects the amount of commerce affected by the market power considered. L is a dimensionless quantity, whereas DW is measured in dollars of loss. It is thus quite appropriate for LP to turn from L to DW when considering the degree of market power sufficient to warrant antitrust concern,[8] since a small firm with a high L may cause insubstantial social loss. In order to possess substantial market power, equation (2) indicates that a firm must have <u>both</u> a high L (or, equivalently, face a low ε^f) <u>and</u> significant dollar revenues.

Note that the <u>firm's</u> revenues appear in (2), while LP assert that "the relevant sales volume is not the defendant's, but the market's."[9] This assertion is not correct in general; Section A in the Appendix shows that (2) is valid for the dominant firm -- competitive fringe model that LP develop and employ. Since defining "the relevant market" is usually difficult and sometimes impossible,[10] this result is reassuring. In principle, at least, one need not go beyond information on a single firm in order to assess the significance of its market power.

B. Multiple Products

Single-product firms of the sort analyzed above are very common in textbooks but very uncommon in reality. LP do not deal explicitly with the use of L or DW in the multiple-product case.[11] If a multiple-product firm has any monopoly power at all, the L's for all of its products will generally be different. Depending on the demand relations

among its outputs, some L's may even be negative. That is, it may even
be optimal to sell some products below marginal cost in order to stimulate
the demand for others.[12] In order to use the apparatus developed by LP
and discussed above in the multiple-product case, one must be able to say
something in principle about aggregation of L's or (more importantly) DW's
across products.

In general, this is a very difficult problem. Section A of the
Appendix considers a tractable special case: a two-product monopoly with
constant marginal costs, linear demands, and no "income effects." In this
case aggregation is very simple in principle. First one uses equation (2)
with $K = 1$ (because marginal costs are constant) to compute DW for each
product. Then one adds these to obtain the total DW for the firm. Alter-
natively, one can calculate the weighted average L, using dollar revenues
as weights, and treat it as applying to the entire firm. Multiplication
of this average L by one-half of total dollar revenues, as in equation (2),
then yields the firm's total deadweight loss. This special case pro-
vides aggregation rules that should not be grossly misleading in more
general settings: compute average L as a weighted average of individual
product L's using dollar revenue weights, and multiply by the firm's
total revenue to obtain total firm DW. Note that these rules do not in-
volve attempting to define or estimate any sort of aggregate demand
elasticity for all the firm's products taken together.

One serious difficulty must be mentioned, however. When the firm's
products are either substitutes or complements, the second equality in
equation (1) no longer holds at the profit-maximizing point. The optimal
markup over marginal cost depends on the cross-elasticities of demand
among the firm's products. In carrying out the computations discussed
in the preceding paragraph, one cannot use readily available elasticity

information to form judgements about the size of product-specific L's.

Under some conditions, discussed in Part III, Section A, _infra_, one can

use information on excess profits instead of data on product-specific

prices and marginal costs.[13]

C. Dynamics: Short-Run and Long-Run Analysis

The formal analysis in Section A, above, is static and timeless, as

is the corresponding analysis in LP. Such analysis is strictly valid only

if demand and cost curves do not shift over time in response to the seller's

actions. Under this extreme assumption, DW gives the net cost to society

per period caused by the monopoly's exercise of its market power. In order

to obtain the net present value of those costs, one would simply capitalize

the stream of deadweight losses by dividing DW by an appropriate rate of

interest.[14]

In fact, one usually expects demand curves to change over time in

response to price changes. It is customary in economic analysis to employ

a distinction between short-run and long-run reactions to price changes.

In the short run, all investments in such long-lived assets as plant and

equipment are taken as fixed, while in the long run all assets are assumed

to adjust in response to market conditions. Thus in the short run the

set of producers of any particular product is taken as fixed, unless there

exist non-producers who could begin production without making significant

new investments. In the long run, firms are assumed to enter any markets

in which they expect to earn adequate returns. Fixed costs are thus en-

countered only in short-run analysis, while the concept of "barriers to

entry" refers to the long run.[15] One generally expects a price increase

to cause a smaller sales reduction in the long run than in the short run,

both because buyers can adapt more completely to the change in the long

run and because new sellers may enter the market in response to the higher price. This means that the individual firm demand curve is more elastic in the long run than in the short run, since quantity is then more responsive to price.

All else equal, one might thus expect a given firm's market power and DW to decline over time, since higher demand elasticities give rise to lower values of DW, as LP demonstrate.[16] To see what this implies for the net cost of market power, suppose that a particular firm's measured DW at time zero (now) is DW^S. (This might be measured by equation (2).) Assume also that in the long run this loss is expected to decline to DW^L, and the gap between current and long-run DW is expected to decline at a rate γ. Then the flow of losses per period due to market power at time t can be written as

$$DW(t) = DW^L + [DW^S - DW^L]e^{-\gamma t} .$$

If the interest rate is r, the net present value (in dollars) of this expression is given by[17]

$$TDW = \frac{1}{r}[\frac{r(DW^S) + \gamma(DW^L)}{r + \gamma}] . \qquad (3)$$

Equation (3) serves to emphasize once again that the measurement of market power is inherently a <u>quantitative</u> exercise: the relevant question in a dynamic setting is the <u>magnitude</u> of TDW. Equation (3) also has a number of interesting qualitative implications. First, suppose that $\gamma = 0$ or that $DW^L = DW^S$, so that the short run and long run are identical. Then $TDW = DW^S/r$, as asserted above. On the other hand, suppose that γ is very large so that adjustment to the long run, via new entry for instance, is very rapid. Then TDW is approximately equal to DW^L/r; only the long-run

deadweight loss matters. Now as a matter of definition, if there are no
barriers to entry or imitation, no seller can hold its price above marginal
cost in the long run. Under these conditions, $DW^L = 0$, and thus TDW is
approximately zero if adjustment to the long run is very rapid. In general,
equation (3) makes clear that in order to assess the net cost of market
power over time, one needs to know the present cost (in dollars per year),
DW^S, the long-run cost, DW^L, and the rate at which market power is eroded
over time, γ. Unless γ is either zero or infinite, both DW^S and DW^L affect
the net cost to society of market power; both the long run and the short
run matter.

The distinction between short-run and long-run analysis does not
appear explicitly in the LP discussion, and this leads to some needless
confusion.[18] For instance, LP find "puzzling" the Cellophane Court's
definition of market power as "the power to control prices or exclude
competition."[19] But there is no problem with this definition. Market power
has two conceptually distinct dimensions.[20] The power to control price in
the short run, short-run market power, leads to a high value of DW^S, while
DW^L can be positive only if the firm has long-run market power that can
be used to exclude competition. Confusion also arises when LP state that
"The supply elasticity of the competitive fringe is determined by both
the ability of existing firms to expand output and the ability of new firms
to enter the market."[21] The first of these is all that matters in the short
run, while the conditions of entry are a key element in determining the
long-run elasticity, so that it is not clear what sort of elasticity LP
have in mind.

The most serious confusion of this sort, however, arises in LP's dis-
cussion of the implications of finding low profits together with high values
of the Lerner index.[22] They essentially argue that if L is large but profits

are low because of fixed costs, firms do not possess significant market power. They apparently have in mind a situation in which there is free entry into the business of developing new products, so that no individual product has associated with it any significant degree of <u>long-run</u> market power.[23] (In terms of equation (3), $DW^L = 0$.) A firm that has developed a desirable product will generally have some short-run discretion over the price it charges, so that L and DW may be substantial for such products in the short run. With free entry into product development, the monopoly profits generated by successful new products can be expected on average just to compensate for the costs of research and development, which are fixed (and sunk) once the products are ready to market. Thus one can have short-run monopoly power with no long-run power and no excess profits on average.

The absence of long-run power does not automatically imply that no antitrust concern is warranted. One must consider both the dynamics involved, in terms of product lifetimes, and the sort of remedy that is contemplated. If the industry described above produces popular songs, for instance, the effective lifetime of an individual product is probably so short on average that TDW is very small for any individual song. (In terms of equation (3), we assume $DW^L = 0$ and associate a short product lifetime with a large value of γ.) On the other hand, if the products are patented drugs with expected lifetimes at least as long as the patent grant, the TDW for any particular drug may be sizeable even though DW^L is zero. (A long lifetime, especially with patent protection, corresponds to a low γ in equation (3).) Even if long-run power to exclude competitors is negligible or absent, short-run power over price may decay so slowly as to be a subject of concern. In the very long run, after all, all

patents expire, but it is not obviously absurd to be concerned with the
conduct of any particular patent monopoly. As always in this area, the
questions are fundamentally quantitative.

Short-run market power may also be relevant if one is considering
relief designed to alter the prevailing mode of conduct in an industry
of the sort described above. There may be little reason to be concerned
with the pricing of any _particular_ popular song, for instance, but the
aggregate deadweight loss associated with _all_ popular songs may not be
negligible. Further, even though the market power associated with any
one song may vanish very rapidly, the industry will be marked by the con-
tinued exercise of short-run power even in long-run equilibrium. Of course,
any restriction on the ability to exercise such power (such as a restriction
on the use of tying arrangements or price discrimination) will reduce the
rewards to innovators and thus tend to reduce innovation. But just as it
is not obvious in general either that the optimal patent lifetime is in-
finite or that patent holders should be immune to antitrust prosecution,
so it is not obvious that reductions in the incentive to innovate are
always undesirable. There is a tradeoff in such cases between static
efficiency losses (as measured by DW and TDW) and long-run gains from
innovation. And the tradeoff is not the same in all cases.

D. From the Firm to the Market

All of the analysis so far relates to a single firm. This is because
all that matters in principle in assessing a single firm's market power
is the demand and cost conditions under which it operates. Since firms'
demand curves are rarely directly observable,[24] it is usually necessary
to infer their characteristics from other data. LP make a number of
specific assumptions that enable them to relate a single-product firm's
demand elasticity, ε^f in equation (1), to parameters such as market share.[25]

They assume a well-defined market in which all firms produce identical products and in which the firm of interest is a price-maker, facing a large number of price-taking, perfectly competitive rivals, each with a tiny share of the market. Though they describe the model that results from these assumptions as an "example",[26] the formulae it implies are employed extensively and serve to structure their entire discussion.

In light of its importance to their analysis, it is worth emphasizing that the dominant-firm/competitive-fringe model is indeed only an example. In particular, it rests on the structural assumption that the firm under study has a much larger share of the relevant market than any other single seller, so that it is not directly useable in oligopolistic markets. It also assumes that no single new entrant can noticeably affect market conditions, so that it has little to say about entry deterrence or predatory conduct. The assumption of product homogeneity rules out advertising and all other forms of non-price competition. If any of the assumptions underlying the LP "example" are inconsistent with the facts of any particular situation, it makes little sense to use that model as a tool of analysis or to take seriously the formulae it implies. An alternative model that is more consistent with the facts should be employed instead.

Exclusive focus on the LP example would tend to understate a firm's market power in many cases, for two reasons. First, the assumption that all firms in "the market" produce perfect substitutes will overstate the constraints that rival production imposes on the firm studied if that firm in fact produces a product that is differentiated from others. Since some differentiation exists in most markets, the LP model almost always understates market power, though with slight differentiation the understatement is trivial. The second reason for likely understatement of market power is the LP assumption that all the firm's rivals behave perfectly competitively. If any of those rivals also have market power, however, exercise of that power will tend to raise price and thus

to increase the Lerner index of the firm in question. The concept of elasticity of supply, which appears frequently in the LP discussion, is strictly relevant only under perfectly competitive behavior. Under any other assumption about rival behavior, one needs to know more than cost curves and entry conditions in order to predict rival response to changes in the price or output of the firm being studied.[27]

II. Market Share and Market Power

As LP state, "The standard method of proving market power in antitrust cases involves first defining a relevant market in which to compute the defendant's market share, next computing that market share, and then deciding whether it is large enough to support an inference of the required degree of market power."[28] This method of proof is only a diagnostic technique, which, like any other diagnostic technique applied under imperfect information, may lead to the wrong conclusion. Professor Fisher has recently stated the basic point of this Part well:

> I do not believe that the question of what is the
> relevant market is fundamentally the right question
> to ask, even though answering it in a sensible way can
> be an aid to analysis. The fundamental question is
> that of the constraints on power. Focusing on the
> question of the relevant market can often lead to
> losing sight of that fact.[29]

The market share approach to diagnosing market power has a number of intrinsic weaknesses, some of which LP note but some of which they gloss over. In addition, some of their suggested modifications of this approach may not systematically enhance its resolving power.

A. The "Standard Market" Fallacy

Implicit in the near-universal use of the market share approach in antitrust proceedings is the belief that in almost any case one can find a good approximation to a "standard", textbook market, one that passes the usual definitional tests and has the additional property that a firm's share thereof is a good measure of its market power. It is important to

be very clear that <u>there is absolutely no support in theory or in fact</u>
<u>for such a belief.</u>[30]

Almost a half-century ago, Joan Robinson proposed that monopoly
should be treated as the central case in economic analysis, recognizing
that firms differ in the extent of their monopoly power.[31] The analysis
here and in LP is in that tradition. In relating her conception to what
is still the standard textbook usage of monopoly, she described the market
share approach very clearly:

> The reader may object that there is clearly some sense
> in which Messrs. Coats have got a monopoly of sewing
> cotton, and in which a Bedfordshire market gardener has
> not got a monopoly of brussels-sprouts. . .All that
> "monopoly" means, in this old-fashioned sense, is that
> the output of the individual producer happens to be
> bounded on all sides by a marked gap in the chain of
> substitutes. Such a gap in nature provides us with a
> rough-and-ready definition of a single commodity -
> sewing cotton or brussels-sprouts - which is congenial
> to common sense and causes no trouble. When a single
> producer controls the whole output of such a commodity
> the plain man's notion of monopolist and the logical
> definition of a monopolist as a single seller coincide.[32]

The market share approach searches for "a marked gap in the chain of sub-
stitutes" for a particular firm's output, uses the gap settled upon to
construct a "rough-and-ready" definition of "the relevant market," and
infers market power if the firm in question controls all or a substantial
fraction of the market output thus defined.

The first and most basic problem with this approach is discussed
by LP: the share of any market thus defined may give a seriously in-
correct indication of the magnitude of market power.[33] There is no theo-
retical reason or factual basis for supposing that all markets defined by
"a marked gap in the chain of substitutes" have similar demand elasticities.
Thus firms that are monopolists according to "the plain man's notion"
thereof can have very different values of L. Since "markets" thus defined

can also differ dramatically in size, firms that are monopolies in the

"old fashioned sense" can produce dramatically different values of DW. If

the dominant firm -- competitive fringe model used by LP is applicable,

the L's and DW's of firms with the same "market share" can also differ

substantially because of differences in fringe supply elasticities.[34]

Two additional sets of problems with the market share approach deserve

discussion. LP suggest that this approach can be modified in certain ways

if it yields shares that are misleading, in the sense that they suggest

(by implicit reference to "standard" markets) more or less power than the

firm actually possesses. Their proposed modifications may be very informa-

tive in some cases, but they can be very misleading in others and should

not be applied mechanically. It is safer to reinterpret market shares in

light of information about elasticities. A second set of problems arise

when products are differentiated and markets are not well defined. While

these may be more important in many cases than difficulties caused by

atypical supply and demand elasticities, LP do not discuss them systemat-

ically.[35]

B. Adjusting Shares to Reflect Elasticities

For most of their discussion, LP assume that there is no debate

about the definition of "the relevant market." I will maintain that

assumption in this Section, and I will also assume that their dominant

firm -- competitive fringe model is descriptively valid for the situation

being analyzed. In that model, the dominant firm's demand elasticity,

which determines L via equation (1), is given by[36]

$$\epsilon^f = \epsilon^m/S + \mu^s(1 - S)/S , \qquad (4)$$

where ϵ^m is the <u>market</u> demand elasticity, S is the dominant firm's market

share, and μ^s is the elasticity of supply of the competitive fringe. Holding the two elasticities on the right of equation (4) constant, larger values of S imply smaller values of ε^f, which in turn (from equation (1)) imply larger values of L. If those two elasticities were the same in all markets, that is, if all markets were "standard markets", and if the dominant firm -- competitive fringe model were valid for all markets, one would only need to know S in order to compute L. Combining equations (1) and (4) and solving for L, one obtains[37]

$$L = S/[\varepsilon^m + (1 - S)\mu^s] . \tag{5}$$

Multiplication of an estimate of L derived explicitly or implicitly from equation (5) by one-half of the dominant firm's dollar revenue would then produce an estimate of DW, the net cost of the dominant firm's market power in the period corresponding to the revenue figures employed. (See equation (2). This assumes MC is approximately constant.)

As LP clearly recognize, all markets are not "standard."[38] The demand and supply elasticities on the right of equation (5) may vary considerably across well-defined markets that are "bounded on all sides by a marked gap in the chain of substitutes." LP mention two different approaches to dealing with this problem. If market power is to be defined in terms of market shares, they discuss modifications in the computation of market shares designed to reflect elasticity differences.[39] Alternatively, they suggest that it might be desirable "not to define market power in terms of specific market shares at all, but instead to interpret the market share statistics in each case by reference to qualitative indicia of the market elasticity of demand and the supply elasticity of the fringe firms."[40] While either of these approaches is acceptable in principle, the second has the great advantage of focusing attention on

the relevant questions. Share adjustment procedures mechanically applied
may not always operate in the right direction, they may not produce
adjustments of the appropriate magnitude, and they tend to focus attention
on peripheral issues. Let me illustrate these problems with two of LP's
specific proposals.

1. Excess Capacity -- LP suggest that in general, "the sum of
the capacity, or potential output, of competitors and the current output
of the firm in question should be the denominator in computing the firm's
market share."[41] The stated rationale for this is that "The greater the
difference between capacity and current output, the greater is the supply
elasticity of competing firms, and therefore the greater is the constraint
that these firms place on a firm that tries to raise price above marginal
cost."[42] They qualify this rule as follows: "When the incremental cost
of converting excess capacity to output is greater than the marginal
cost of the last unit actually produced, only so much of the excess
capacity as can be converted to output without increasing marginal cost
should be included in the computing market share."[43] There are at least
three reasons why this "general rule" should not be employed in antitrust
cases, even when the dominant firm -- competitive fringe model is appli-
cable.

First, it is very difficult to define "capacity" in a universally
satisfactory way, and it is comparably difficult to obtain good estimates
under any definition selected.[44] LP attempt to deal with this problem
by means of the qualification quoted above, but this is unsatisfactory
on at least two counts. If firms have constant marginal costs over some
range of outputs that includes current output, as they assume, supply
elasticity is infinite over that entire range. Thus the variation in
excess capacity caused by variations in current output over that range

provides absolutely no information about supply elasticity. Moreover, if such a range of constant marginal costs exists, it is generally incompatible with the price-taking competitive fringe behavior that underlies equation (5).[45] It is thus both more common and more sensible to assume that competitors' marginal cost schedules are rising over the relevant range. But this is exactly what makes "capacity" difficult to define clearly or to measure accurately, thus making it very hard to interpret estimates of "excess capacity" in individual industries. It is easy to imagine such definition and measurement issues consuming enormous amounts of time and money in antitrust cases, diverting attention from the more fundamental question of competitive supply elasticity.

Second, the adjustment that LP propose does not necessarily work in the right direction in all cases. Section B of the Appendix shows that well-behaved, non-pathological short-run supply curves exist for which increases in output raise the elasticity of supply. This means that the elasticity of supply may be lower the greater is excess capacity (i.e., the lower is current output) in some cases. The qualitative rationale for the use of capacity information quoted above is thus not universally correct, though it may hold in many cases.

Finally, even if the LP adjustment is generally in the right direction, there is no reason to suppose that it is generally of the correct magnitude. Section B of the Appendix shows that even in a simple case in which supply elasticity falls when output increases (as LP assume), the quantitative relation between supply elasticity and excess capacity depends on all the parameters of the cost function. Thus the extent of excess capacity (however measured) by itself provides essentially no quantitative information about the elasticity of supply in any particular case. Even if there exist some situations in which application of the LP adjustment

produces adjusted market shares that are good indicators of market power, there will surely exist others in which the adjusted shares will considerably understate the leading firm's market power, along with still other cases in which this adjustment will not lower that firm's share enough and thus leave market power substantially overstated.[46]

In view of these difficulties, it is surely more sensible to define market share in the usual fashion and to employ quantitative or qualitative information on competitive supply elasticity to interpret that share, using equation (5) explicitly or implicitly, than to shift attention to the definition and measurement of "excess capacity" and to compute an "adjusted" market share that may be even less informative than the original share.

2. <u>Market Definition</u> -- LP also suggest that the market can be re-defined to reflect variations in the market elasticity of demand from case to case.[47] They argue that if a particular well-defined product has good substitutes, its (market) demand elasticity is likely to be high, relative to some standard,[48] and they seem to endorse "the usual approach in antitrust cases: before market shares are computed, commodities that are very good substitutes for each other are aggregated into a single product."[49] Here again it seems preferable to use elasticity information to interpret ordinary market shares rather than to try to incorporate such information indirectly into the market definition process. The three basic objections raised above to use of excess capacity figures also apply here.

First, this approach focuses attention on the wrong question. Arguments about market definition are usually couched in terms of substitutability among commodities or, more technically, in terms of cross-elasticities of demand. But the real issue is the own-price demand elasticity facing the firm being studied. LP recognize this and argue that ambiguity "could

be avoided by using elasticity of demand instead of cross-elasticity of demand as the ruling concept in antitrust cases."[50]

Second, the proposed adjustment does not go in the right direction in all cases. Suppose that one begins analysis by looking for the first "market gap in the chain of substitutes" around the outputs of the firm of interest. This will usually produce a narrow market definition including sellers of very close substitutes. The proposed adjustment (and standard practice) would then be to expand the definition of the market if the original "gap" is not judged to be sufficiently large, or, in other words, if other commodities are judged to be sufficiently close substitutes in consumption for the commodities in the original market. But it is not true that commodities with good but imperfect substitutes always have unusually high demand elasticities.

Section C in the Appendix develops the following relation for the case of two substitute goods:

$$\bar{\varepsilon} = w^1 \varepsilon^1 + w^2 \varepsilon^2 - \sigma , \qquad (6)$$

where $\bar{\varepsilon}$ is an aggregate market demand elasticity for both goods, defined in the Appendix, ε^1 and ε^2 are the market demand elasticities for goods 1 and 2, respectively, w^1 and w^2 are the shares of the two goods in their total revenue, and σ is the appropriate weighted average of the cross-price elasticities. (There is in general no such thing as the cross-price elasticity, as the Appendix demonstrates.) If these two goods, taken together, do not have any especially close substitutes, $\bar{\varepsilon}$ will be some finite number. (One might think of processed lemon juice and fresh lemons, for example.[51]) Equation (6) indicates then that if σ is exceptionally large, indicating that the two goods are very close substitutes in con-

sumption, both ϵ^1 and ϵ^2 will probably be large and at least one must be. But in general one cannot be sure that either ϵ^1 or ϵ^2 is unusually large just because σ is non-negligible. Even if one knows that $\epsilon^1 = \epsilon^2$, for example, $\sigma = 2$ is compatible with $\epsilon^1 = 2.5$ ($\bar{\epsilon} = .5$) and with $\epsilon^1 = 5.0$ ($\bar{\epsilon} = 3.0$). These two market elasticities suggest very different interpretations of particular market share values. The point here is that one simply cannot infer much about a product's own-price elasticity of demand from information about its cross-price elasticities with respect to one or two other products.[52]

Finally, the proposed adjustment may produce very misleading quantitative results. The consequences for the computed market share of a firm selling product 1 of concluding that products 1 and 2 are sufficiently close substitutes that they should be aggregated into a single market depend _entirely_ on the sales volume of product 2.[53] But this need have no relation at all to the market demand elasticity of product 1. Thus even if product 1 does have an unusually high market demand elasticity in large part because product 2 is a good substitute for it, there is no guarantee at all that this will be accurately reflected in "adjusted" share computations. If product 2 is unimportant, aggregation may produce a share that is misleadingly large if interpreted in the context of a "standard" market. On the other hand, if product 2's revenues greatly exceed product 1's, aggregation may incorrectly suggest that even a monopoly of product 1 would have negligible market power.

It is thus both more natural and more reliable to focus _directly_ on the issue of demand elasticity than to haggle at length about market definition and about whether particular products are or are not "close substitutes."[54]

C. Product Differentiation and Ill-Defined Markets

The market share approach to measuring market power depends on the

implicit assumption that "marked gaps in the chain of substitutes" gen-
erally occur in convenient places. That is, it is assumed not only that
the gaps that separate a firm and its most direct rivals from other
products are sufficiently wide that all other products may be neglected
in the analysis of market power but also that all products within "the
relevant market" as defined by those gaps are very close substitutes.[55]
If the latter condition is not satisfied, very little meaning attaches
to computed market shares.

It is easy to think of examples in which "marked gaps in the chain
of substitutes" do not occur in convenient locations. It seems clear,
for instance, that small four-bit microprocessors are not especially
close substitutes for the largest mainframe computers, though some sub-
stitutability undoubtedly exists. On the other hand, there are no obvious
"marked gaps" in the array of computers of intermediate power across
which little substitution takes place. Similar problems seem likely to
exist in automobiles (with, say, a stripped down Chevrolet Chevette at
one extreme and a fully-equipped Mercedes 450 SEL at the other), in
cameras (consider a Kodak Instamatic and a Hasselblad), and in other lines
of business. The problem can be encountered in geographic contexts as
well. Along Interstate Route 55 between Chicago and St. Louis there are
(or at least there used to be) a large number of gasoline stations, dis-
tributed more or less evenly. If one attempted to measure the share of
"the relevant market" accounted for by, say, a hypothetical group of
colluding stations near Springfield, one would not find a "marked gap in
the chain of substitutes" by proceeding either north or south. But it
would clearly make little sense to compute the ratio of the sales of the
hypothetical colluding group to total gasoline sales between Chicago and
St. Louis.

When a sizeable number of differentiated products are available, "marked gaps in the chain of substitutes" are unlikely to occur in convenient places, and the market share approach to measuring market power is not likely to perform reliably. Regardless of the number of brands considered, if all plausible definitions of "the relevant market" require grouping together products that are significantly differentiated, essentially nothing can be said about the relation between shares of such markets and firms' demand elasticities.[56] There is no general, universally applicable model of the competitive relationships among differentiated products. Under differentiation, a low market share does not establish that market power is negligible, since competition may be "localized", so that a particular firm or brand has only a few effective rivals even though a large number of broadly similar brands may be marketed, or firms may have long-run market power by virtue of a membership in "strategic groups" protected by "mobility barriers."[57] A large share of a market with differentiated products provides evidence of substantial market power only if the market definition is not excessively narrow. In short, if significant differentiation can be demonstrated, market share computation should generally not be taken very seriously.

LP discuss the implications of product differentiation when relating their views on geographic market definition to those of Areeda and Turner.[58] LP argue convincingly that if the products involved are identical and if only transportation costs and tariffs impede trade, then "if a distant seller has some sales in a local market, all its sales, wherever made, should be considered a part of that local market for purposes of computing the market share of a local seller."[59] On these assumptions, Areeda and Turner do not disagree.[60] Indeed, this is basically an application of

Alfred Marshall's classic dictum that "the more nearly perfect a market
is, the stronger is the tendency for the same price to be paid for the
same thing at the same time in all parts of the market: but of course
if the market is large, allowance must be made for the expense of de-
livering the goods to different purchasers; each of whom must be supposed
to pay in addition to the market price a special charge on account of
delivery."[61] If knowledge of the price in the "local" (or domestic)
market permits one accurately to predict price in the "distant" (or for-
eign) market by adding or subtracting transportation and tariff costs,
the two areas are effectively linked on the supply side and should be
treated as a single market.

LP and Areeda and Turner agree that foreign production should not
be used to compute the market shares of domestic producers if domestic
and foreign products are strongly differentiatied, but they disagree
about the appropriate tests for strong differentiation in this context.[62]
Areeda and Turner would exclude the output of foreign producers if the
product in question is regularly exported from as well as imported into
the United States or if the sum of the foreign price and the transporta-
tion and tariff costs of importation exceed the domestic price. LP con-
tend that substantial two-way trade argues instead for use of a narrower
market definition, and I would agree.[63] They seem to suggest that the
second of the Areeda-Turner tests should be amended to require a substantial
excess.[64] I again agree but would add that a substantial difference of
either sign should serve to rebut the presumption that the domestic and
foreign market are effectively welded together by trade. LP themselves
would infer strong differentiation if foreign sellers had only a negligible
share of the domestic market or if foreign sellers were observed to make
substantial sales in spite of a serious cost disadvantage (including

transportation and tariffs) in the U. S. market. Since the issue is
whether the foreign and domestic markets are tightly linked, the Areeda-
Turner price comparison is more natural than the LP test involving costs,
and it is likely to be simpler to perform as well.

In addition to these tests, courts should be prepared to consider
other evidence bearing on the question of differentiation, such as studies
of the correlation of domestic and foreign price movements over time.
Both LP and Areeda and Turner agree on the relevant question; neither of
their proposals should be interpreted as restricting the set of ways one
might attempt to answer it in any particular case.

III. Other Indicators of Market Power

If the central LP argument, that market share is of interest only
to the extent that it provides information about market power, is widely
understood and accepted, they will have performed a valuable service.
It is an immediate corollary of that argument that even though market share
has been traditionally relied upon by the courts as the best indicator of
market power, other evidence deserves equal standing, as it may be at least
as informative in some cases.[65] Two broad categories of non-share evidence
deserve at least brief discussion here: evidence on profitability and
evidence derived from patterns of conduct.[66]

A. Persistently High Profitability

Under the kinds of simple assumptions made by LP in their formal
analysis and in Part I above, the excess profits earned by a firm
exercising market power is directly proportional to the deadweight loss
(DW) it inflicts on society. As Section A of the Appendix shows, under
linear demand and constant marginal cost, DW is equal to exactly half of

the firm's excess profits. To see what this implies algebraically, suppose a firm employs capital K^m, earns a rate of return r^m on those assets, and could earn a rate of return r on investments of comparable risk elsewhere in the economy.[67] Then the firm's excess profits are equal to $(r^m - r)K^m$, and deadweight loss is given by

$$DW = [(r^m - r)/r^m][r^m K^m]/2 \ . \tag{7}$$

The first term on the right of equation (7) corresponds to L in equation (2); it measures the percentage deviation from the competitive norm. The second term is total accounting profit; it automatically factors in a measure   of substantiality just as total dollar revenue did in equation (2). As Section A of the Appendix demonstrates, this relation between excess profit and deadweight loss also holds in at least one special multiple-product case. In simple models, then, profitability information is exactly as informative about DW as information about price-cost margins or firms' demand elasticities. Moreover, profitability is considerably easier to use (at least in principle) when a firm sells multiple products; total excess profit is a convenient and directly relevant aggregate.

There are two serious problems with the use of this approach, however, that must be emphasized.[68] First, it is very difficult in practice to measure actual profitability, and it may be even more difficult to measure excess profits.[69] There are no simple, universally applicable techniques for obtaining accurate estimates of these quantities, though there are reasons to hope for progress in this area.[70] Second, substantial excess profits can arise in the short run even under perfect competition. Such profits provide socially essential signals to guide the flow of investment funds in competitive economies. One thus cannot use excess profits to

establish short-run market power, though their presence may serve to

confirm the existence of such power established by other means. Persistent

excess profits, however, provide a good indication of long-run power; they

show clearly that there is some impediment to effective imitation of the

firm in question.[71] In addition, as equation (7) indicates, information

about excess profits derived from market power can be used to produce rough

estimates, at least, of the deadweight loss caused by the exercise of that

power.

B. Patterns of Conduct

Evidence that competitors have conspired to fix price or divide

markets is treated as very good evidence that they have market power.[72]

In other contexts, other kinds of evidence about firms' market conduct

may provide useful information about their market power.

It is a standard textbook proposition that "for a seller to practice

price discrimination profitably," it "must have some control over price --

some monopoly power."[73] If the same product is sold to different customers

at different prices even though costs are known to be the same, and if it

is reasonable to assume that no sales are made below cost, one can obtain

a lower bound on the extent of market power by using the lowest price as

an estimate of marginal cost. In general, however, it is hard to go from

the fact of price discrimination to estimates of the importance of market

power.[74] Thus evidence on price discrimination is probably most useful in

cases in which only some minimum quantum of market power is required.[75]

In a similar vein, one can argue that proof of predatory conduct should

suffice in principle to establish market power. Let us follow Bork and

define predation as

> a firm's deliberate aggression against one or more rivals
> through the employment of business practices that would

not be considered profit-maximizing except for the
expectation either that (1) rivals will be driven
from the market. . .or (2) rivals will be chastened
sufficiently to abandon competitive behavior the
predator finds inconvenient or threatening.[76]

Assuming that firms rarely engage in strategies with negligible chances

of success, a firm that predates on this definition has (or at least

thinks it has) the ability to affect market conditions materially, since

without a material effect from the exit or passivity of the prey, preda-

tion would almost never appear to be profitable.[77] This implies the

possession of some degree of market power but does not lead directly to

any estimate of its importance. A more serious problem is the difficulty

of establishing directly the motives and expectations that underlie firm

conduct, especially conduct alleged to be predatory in intent and expected

effect.[78] The point here is simply that as a logical matter, if those

motives and expectations are established in some particular case, and if

the expectations of the effects of predation cannot be shown to be totally

unrealistic, some market power has been proven. Whether that should be

taken as proving enough market power to meet the threshold requirements

in monopolization cases is quite another question, however, and one that

does not have an obvious general answer.

Finally, the methods by which a firm makes price and output decisions

may provide direct evidence of the presence of market power. In some cases,

if data are unusually abundant and of unusually high quality, skillful

econometric analysis may permit rigorous testing of hypotheses about mar-

ket power or collusive behavior.[79] In other settings, documentary evidence

of recognition of market power in price setting and other marketing de-

cisions, coupled with failure of the market to reject those decisions,

provides evidence of some market power.[80] Unless this sort of evidence

is unusually strong, it may be difficult to use it to establish enough

market power for a monopolization case, though such proof may be adequate

where less power is required.[81]

IV. Conclusions

Even though I have taken issue with a number of LP's specific assertions and proposals,[82] it should be clear that I endorse their basic approach to the analysis of market power. The deadweight loss (DW) associated with a firm's price and output decisions is in principle a good measure of the short-run importance of its market power. Fundamentally, this measure depends on the firm's net demand curve, not on its share of the relevant market or any other aggregate. The dominant firm -- competitive fringe model used by LP to relate firm and market parameters is directly useable only under particular structural conditions. One expects market power to erode over time, and information on the likely speed and extent of erosion should be used to assess the total social cost imposed by the exercise of such power (TDW).

Computation of market share can provide information about the importance of market power, but markets differ considerably, and shares should be interpreted in light of evidence on market demand elasticities and other conditions. Mechanical adjustments to ordinary share computations can be misleading. In particular, I do not think that excess capacity estimates should be used to adjust market shares. In situations in which "gaps in the chain of substitutes" do not occur in convenient places, whether because of product differentiation or for other reasons, the market share approach may be vary unreliable because a "standard", textbook market may not exist. Other approaches to proving the existence of substantial market power are in principle no less valid than the market share approach. Depending on the facts of the case at hand, data on profitability or on patterns of conduct may be more informative than market shares.

APPENDIX

A. Properties of the Deadweight Loss Measure

In the linear demand and marginal cost case depicted in Figure 1, the formula for the area of a triangle yields

$$DW = (1/2)(P^m - MC^m)(Q^c - Q^m)$$

$$= (1/2)[(P^m - MC^m)/P^m][P^m Q^m][(Q^c - Q^m)/Q^m]$$

$$= L^m(P^m Q^m)K/2 , \qquad (8)$$

where $K = (Q^c - Q^m)/Q^m$. This is equation (2) in the text. If MC is constant, $Q^c = 2Q^m$, and $K = 1$. If marginal cost is rising, Q^m exceeds $Q^c/2$, and K is less than one. In the linear case, K can be written as a function of the elasticity of output with respect to marginal cost (this would be an elasticity of supply under competition) and the firm's elasticity of demand, both evaluated at monopoly equilibrium; it is increasing in the first of these arguments and decreasing in the second.

Suppose that the firm depicted in Figure 1 is a textbook monopoly, so that the firm demand curve shown gives total market demand. One can think of the net benefit to consumers of each addition to the monopoly's output, starting from zero, as the difference between the maximum price they would be willing to pay for that unit, as given by the height at the demand curve, and the price they actually must pay. Then net benefit, or consumers' surplus, is given by the area A under monopoly and by the sum of areas (A + B + C) under competition. With output Q^m, the monopoly's revenue, $P^m Q^m$, is equal to area (B + D + F + G), and costs are G. (This neglects fixed costs for simplicity; their addition would change nothing.)

Profits with output Q^m are thus given by (B + D + F), while with output Q^c profit equals (D + E + F). Define net surplus as the sum of consumers' surplus and monopoly profit. We can then compute the cross-hatched dead-weight loss area as the difference between net surplus at output Q^c and net surplus at output Q^m:

$$[(A + B + C) + (D + E + F)] - [(A) + (B + D + F)] = C + E.$$

Similarly, one can compute the change in deadweight loss associated with a change in the monopoly's price or output as the sum of the induced <u>reductions</u> in consumers' surplus and monopoly profits. If, for instance, price were lowered by one (small) unit below P^m, Figure 1 shows that consumers' surplus (area A) would increase by approximately Q^m. (Formally, the derivative of consumers' surplus with respect to P^m is exactly equal to $(-Q^m)$.) This gain would be <u>subtracted</u> from the reduction in the monopoly's profits to give the net change in deadweight loss. (That change would be negative in this case; a price reduction would lead to an output increase and a rise in net surplus.) We could clearly compute the total deadweight loss at price P^m by adding the reductions in consumers' surplus and monopoly profit as price rises from P^c to P^m or, alternatively, by summing the increases in surplus and profit as price falls from P^m to P^c.

Now suppose that we are instead dealing with a "dominant" firm that has a substantial market share but faces competition from a set of price-taking firms producing the identical product. This is the assumption made by LP in their formal analysis. In this case, the firm demand curve shown in Figure 1 is a net demand schedule, computed by subtracting competitive supply from market demand at each price level. Changes in the dominant firm's price now induce changes in suppliers' profits, along with

changes in consumers' surplus and the dominant firm's profits. In order

for equation (2) in the text to be correct, we must be able to use the

firm's demand curve to compute changes in all three quantities as price

moves from P^c to P^m or vice versa. In fact we can do this exactly as

above.

The change in the firm's profits caused by, say, a small reduction

in P^m clearly must be computed using the firm's net demand curve. From

the preceding paragraph, we thus need only show that the increase in

consumers' surplus plus competitors' profits caused by a small unit decrease

in P^m is equal to Q^m. To see that this is indeed the case, consider Fig-

ure 2, which shows the supply curve of the competitive producers along

with the market demand curve. At price P^m, the dominant firm sells Q^m,

the difference between market demand and competitive supply, Q^s. That

part of the cross-hatched area above the P^m line gives consumers' surplus;

the rest gives profits of the competitive producers. If P^m is lowered to

$P^m-\Delta$, the increase in the sum of consumers' surplus and competitor's

profits is given by the shaded area, which is approximately equal to $Q^m\Delta$.

(Formally, the derivative of consumers' surplus plus competitors' profits

with respect to P^m is $(-Q^m)$, exactly as in the monopoly case.) Thus the

firm demand curve is the relevant one for these computations, and equation

(2) holds for the dominant firm -- competitive fringe model, as asserted

in the text.

Let us now consider a monopoly selling two products, with unit sales

Q_1 and Q_2, having constant marginal costs v_1 and v_2, respectively. If

market demands for these products are linear and there are no income

effects, then sales with prices P_1 and P_2 are given by

$$Q_1 = b_1 - c_1 P_1 - dP_2, \quad \text{and} \quad Q_2 = b_2 - c_2 P_2 - dP_1 ,$$

where b_1, b_2, c_1, and c_2 are positive constants, and d is a constant that is positive if the goods are complements and negative if they are substitutes.[83] Under these assumptions, one can write consumers' surplus in both markets as

$$V(P_1,P_2) = a - b_1 P_1 - b_2 P_2 + (c_1/2)(P_1)^2 + (c_2/2)(P_2)^2 + dP_1 P_2 ,$$

where a is a constant of integration.[84] The monopoly's profits are given by

$$\Pi(P_1,P_2) = (P_1 - v_1)Q_1 + (P_2 - v_2)Q_2 .$$

As above, we can compute the deadweight loss associated with charging prices P_1^m and P_2^m instead of the competitive prices v_1 and v_2 as the difference between net surplus under competition and monopoly:

$$DW = [V(v_1,v_2) + \Pi(v_1,v_2)] - [V(P_1^m,P_2^m) + \Pi(P_1^m,P_2^m)]$$

$$= (c_1/2)(P_1^m - v_1)^2 + (c_2/2)(P_2^m - v_2)^2 + d(P_1^m - v_1)(P_2^m - v_2)$$

The demand functions imply that the differences between monopoly and competitive quantities are given by

$$Q_1^c - Q_1^m = c_1(P_1^m - v_1) + d(P_2^m - v_2) , \quad \text{and}$$

$$Q_2^c - Q_2^m = c_2(P_2^m - v_2) + d(P_1^m - v_1) .$$

Solving for the price/cost differences, substituting into the expression above for DW and simplifying, one obtains after considerable algebra,

$$DW = (1/2)[(P_1^m - v_1)(Q_1^c - Q_1^m) + (P_2^m - v_2)(Q_2^c - Q_2^m)] . \tag{9}$$

This is of the form of equation (8), above. Solving the monopoly's profit maximization problem, one finds that $Q_1^m = Q_1^c/2$ and $Q_2^m = Q_2^c/2$. Substituting for the Q^c's into equation (9), one obtains

$$DW = (1/2)[L_1(P_1^m Q_1^m) + L_2(P_2^m Q_2^m)] , \tag{10}$$

where the L's are the Lerner indices for the two products, as defined by equation (1) in the text. Equations (1) and (10) establish that DW is equal to one-half of the monopoly's (excess) profit here, as in the single-product case with linear demand and constant marginal cost. Comparing equation (2) in the text, it is seen that total deadweight loss caused by exercise of the firm's market power is given by simply applying equation (2) to each market separately and adding the results, as asserted in the text. But it is important to note that the second equality in equation (1) does _not_ hold in the multiple-product case. The optimal markup over marginal cost in each market is determined by both own-price responsiveness of demand (measured here by c_1 and c_2) _and_ cross-price responsiveness (measured here by d).

B. Excess Capacity, Supply Elasticity, and Share Adjustment

Let the supply function of a set of price-taking competitive firms (the competitive fringe of suppliers in the LP dominant firm -- competitive fringe model) be $Q^S(P)$, where P is the market price. In the short run, it is usually assumed that higher values of P will induce more production, so that $Q^S(P)$ is an increasing function. (Excess capacity, with which we are concerned here, has no meaning outside the short run.) As LP state, the elasticity of supply is defined by

$$\mu^S = (dQ^S/dP)(P/Q^S) ,$$

where dQ^S/dP is the derivative of Q^S with respect to P.[85] For a given

level of capacity, excess capacity increases when Q^S decreases. Since

P and Q^S move together along the supply curve, excess capacity also in-

creases when P decreases. In order for the LP assertion that elasticity

of supply increases with excess capacity to be correct,[86] it thus must

be the case that μ^S increases when P decreases. Differentiating the

definition of μ^S above, this in turn requires the following expression

to be negative:

$$d\mu^S/dP = (d^2Q^S/dP^2)(P/Q) + (\mu^S/P)(1 - \mu^S) ,$$

where d^2Q^S/dP^2 is the second derivative of Q^S with respect to P. If the

supply curve is approximately linear at any point, the first term on the

right of this equation will be very small. Then increases in P will in-

crease μ^S if μ^S is less than one. Since LP assume $\mu^S < 1$ in some of their

examples,[87] they apparently agree that there is nothing pathological or

abnormal about inelastic supply. But if supply is inelastic, increases

in excess capacity lower the elasticity of supply unless d^2Q^S/dP^2 is

negative and sufficiently large, and a priori this derivative is as likely

to be positive as negative. In general, then, increases in excess capacity

may either increase or decrease the elasticity of supply.

In order to examine the quantitative properties of the LP proposal

to add competitors' excess capacities to market sales for purposes of

defining "adjusted" market shares, let us consider a special case in which

increases in excess capacity do increase the elasticity of supply. Suppose

that a perfectly competitive firm's total cost is given by

$$TC = F + AQ + BQ^2 , \qquad (11)$$

where F, A, and B are positive constants and Q is the firm's output. If
we define capacity, $Q*$, as the output level at which average cost (TC/Q)
is minimized,[88] it is easy to show that $Q* = \sqrt{F/B}$. Recognizing that as
long as market price exceeds A the firm's supply curve is equal to its
marginal cost curve, obtained by differentiating equation (11), one can
write the firm's elasticity of supply as

$$\mu^S = 1 + (Z + 1)(A/2\sqrt{FB}) , \tag{12}$$

where $Z = (Q* - Q)/Q$ is a measure of the extent of excess capacity. If
all of a given firms' rivals are identical to the competitive firm con-
sidered here, equation (12) will hold for the elasticity of competitive
fringe supply. Note that increases in Z increase μ^S in this example.
But note also that knowing Z tells one essentially nothing about the value
of μ^S. Even in this special case one would have to know all three param-
eters of the cost function (F, A, and B) in order to know whether μ^S was
particularly large or small. Thus even if μ^S is clearly an increasing
function of excess capacity for some industry, one would need to have con-
siderable quantitative information about that industry's cost structure
in order to know whether any given level of excess capacity implies an
unusually high or low value of the elasticity of supply relative to other
industries.

We can use the example of elasticity of supply variations to explore
the basic quantitative properties of approaches that mechanically adjust
market shares to take into account deviations of individual markets from
average or "standard" conditions. As equation (5) in the text makes clear,
the relation between the dominant firm's market share and the Lerner index
depends on elasticities of market demand and competitive fringe supply.

In determining threshold values of S under various statutes, one can imagine that the courts have in mind average values of these elasticities. Let those averages be $\bar{\varepsilon}^m$ and $\bar{\mu}^s$, respectively. The idea of share adjustment then amounts to searching for an S* such that

$$L = S*/[\bar{\varepsilon}^m + (1 - S*)\bar{\mu}^s] ,\qquad(13)$$

where L is the firm's <u>actual</u> Lerner index. That is, one searches for an S* to which one can apply thresholds or other standards based on average or "standard" market conditions.

Now suppose that a particular market has $\varepsilon^m = \bar{\varepsilon}^m$ but it is expected that μ^s may differ from $\bar{\mu}^s$ because of excess capacity held by the dominant firm's rivals. LP would then recommend computing

$$S* = S/[S + (1 - S)(1 + Z)] ,$$

where Z is defined, as above, as the ratio of excess capacity to current output for the competitive fringe. On the interpretation of share adjustment advanced in the preceding paragraph, this is the correct adjustment if and only if substitution of this S* into equation (13) yields the firm's actual L. Using equation (5) in the text, this is seen to be true if and only if the following relation holds:

$$\mu^s = \bar{\mu}^s + Z(\bar{\mu}^s + 1) .\qquad(14)$$

(This makes clear that the definition of $\bar{\mu}^s$ must implicitly assume zero excess capacity.) There are two reasons why this equation (14) will not be correct in general. First, there is no reason to expect it to hold for a particular industry even if $\bar{\mu}^s$ is defined as the elasticity of supply

for _that_ industry when $Z = 0$. Indeed, it does not hold for the quadratic cost example studied above (equation (11)) under those assumptions. Second, that example makes it clear that one cannot compute a reliable estimate of the supply elasticity of a particular set of firms knowing only their excess capacity and the average or normal supply elasticity for some broad set of firms or markets; the details of the particular cost functions involved matter. Similar quantitative problems arise for other mechanical rules for adjusting market shares. Even if they go in the right direction they may produce misleading magnitudes.

C. Own-Price and Cross-Price Elasticities

Consider two products, which will be identified by super-scripts 1 and 2, that are related in demand. One can then write the demand function as

$$Q^1 = Q^1(P^1, P^2) \quad \text{and} \quad Q^2 = Q^2(P^1, P^2) \, ,$$

where the Q's are unit sales and the P's are prices. The two own-price (market) demand elasticities are defined as in LP as follows:[89]

$$\varepsilon^1 = -(\partial Q^1/\partial P^1)(P^1/Q^1) \quad \text{and} \quad \varepsilon^2 = -(\partial Q^2/\partial P^2)(P^2/Q^2) \, . \tag{15}$$

Similarly, the two cross-price elasticities are given by

$$\sigma_2^1 = (\partial Q^1/\partial P^2)(P^2/Q^1) \quad \text{and} \quad \sigma_1^2 = (\partial Q^2/\partial P^1)(P^1/Q^2) \, . \tag{16}$$

There is no reason to expect σ_2^1 and σ_1^2 to be equal. If there are no income effects of price changes, it follows that $\partial Q^1/\partial P^2 = \partial Q^2/\partial P^1$.[90] In this case one has

$$w^1\sigma_2^1 = w^2\sigma_1^2 \, ,$$

where the w's are revenue shares:

$$w^1 = P^1Q^1/(P^1Q^1 + P^2Q^2) \quad \text{and} \quad w^2 = 1 - w^1 = P^2Q^2/(P^1Q^1 + P^2Q^2) \ . \tag{17}$$

If one defines σ as the revenue-weighted average of the two cross-price elasticities, it follows that

$$\sigma = w^1\sigma_2^1 + w^2\sigma_1^2 = 2w^1\sigma_2^1 = 2w^2\sigma_1^2 \ . \tag{18}$$

If it is necessary to summarize the degree of substitutability between two products by a single number, σ would seem a sensible quantity to employ. It will generally be positive if the two products are substitutes and larger the more readily they are substitutable in consumption.[91]

When a set of products are imperfect substitutes (or imperfect complements), there is no unambiguously correct method of defining the aggregate demand elasticity for the set as a whole. The procedure adopted here in order to derive equation (6) in the text focuses on the response of total revenue to proportional price changes. Consider a single product with market demand function Q(P) and define the revenue function

$$R(\lambda;P) = (\lambda P)Q(\lambda P) \ .$$

One can verify the following relation

$$1 - [(dR(1;P)/d\lambda)/R(1;P)] = -(dQ/dP)(P/Q) = \varepsilon^m \ ,$$

where $(dR(1;P)/d\lambda)$ is the derivative of R with respect to λ evaluated at $\lambda = 1$. The value of ε^m depends in general on P.

Similarly, in the case of two products we can define the revenue function

$$R(\lambda;P^1,P^2) = (\lambda P^1)Q^1(\lambda P^1,\lambda P^2) + (\lambda P^2)Q^2(\lambda P^1,\lambda P^2) \ .$$

By analogy with the single-product case, we can define the demand elasticity for the pair of products, $\bar{\varepsilon}$, as follows:

$$\bar{\varepsilon} = 1 - [(dR(1;P^1,P^2)/d\lambda)/R(1;P^1,P^2)] \ ,$$

where again the derivative is evaluated where $\lambda = 1$. Carrying out the differentiation and employing the definitions (15) - (18) above, one obtains

$$\bar{\varepsilon} = w^1(\varepsilon^1 - \sigma_2^1) + w^2(\varepsilon^2 - \sigma_1^2) \ . \tag{19}$$

Use of the definition of σ yields equation (6) in the text. Note that changes in λ correspond to equal percentage changes in P^1 and P^2. Consideration of other sorts of changes would lead to different definitions of the aggregate elasticity. Note also that the value of $\bar{\varepsilon}$ depends in general on the values of both P^1 and P^2.

Footnotes

1. W. M. Landes and R. A. Posner, <u>Market Power in Antitrust Cases</u> 94 Harv. L. Rev. 937 (1981). Though the discussions there and here speak almost exclusively of "market power", it should be noted that economists generally consider that term to be synonomous with "monopoly power" and "economic power".

2. <u>Id</u>. 938.

3. <u>Id</u>. 937. To simplify the discussion, I deal explicitly only with the problem of measuring a single firm's market power, though as Landes and Posner note, the same concepts and techniques may be applied to a group of firms alleged to be acting jointly or proposing merger.

4. LP do not explicitly state that their analysis is concerned with the short run, but their references to fixed costs (<u>Id</u>. 939) indicate that it is at least in part, since no costs are fixed in the long run. I have more to say about the relation between short-run and long-run analysis in this context in Section C, <u>infra</u>.

5. <u>Id</u>. 939-40, 983-85. I have changed notation somewhat.

6. Actually, this measure is exact only if the commodity involved is sufficiently unimportant that changes in its price do not cause noticeable changes in consumers' real incomes and if it is considered appropriate to work with unweighted sums of losses incurred by

affected firms and household. The deadweight loss measure is widely

employed in applied work as a useful approximation, however, For

discussions of the issues involved in its use, <u>see</u> Harberger, <u>Three</u>

<u>Basic Postulates for Applied Welfare Economics: An Interpretive</u>

<u>Essay</u>, 9 J. Econ. Lit 785 (1971), Willig, <u>Consumer's Surplus Without</u>

<u>Apology</u>, 66 Amer. Econ. Rev. 589 (1976), and Hausman, <u>Exact Consumer's</u>

<u>Surplus and Deadweight Loss</u>, 71 Amer. Econ. Rev. forthcoming (1981).

7.   This can be shown to be a generalization of the formulae presented

by Landes and Posner, <u>supra</u> note 1, at 954 (note 32), 993 (eq. 20)

for linear demand and constant marginal cost.

8.   <u>Id</u>. 952-55.

9.   <u>Id</u>. 953.

10.   See Part II, Section 6, <u>infra</u>.

11.   Landes and Posner, <u>supra</u> note 1, at 943, do note that for the related

case of a firm selling a single product at different prices to dif-

ferent buyers, "The Lerner index for the price-discriminating firm

will be a range of numbers rather than a single number." But they

do not say how such a range should be interpreted or summarized.

12.   See, for instance, E. M. Singer, <u>Antitrust Economics</u> 177-82 (1968)

and R. Schmalensee, <u>Monopolistic Two-Part Pricing Arrangements</u>,

<u>12 Bell</u> J. Econ. forthcoming (1981).

13.   As Section A of the Appendix shows, total deadweight loss in the

special two-product case examined there is equal to one-half of the

firm's excess profits, so that <u>if</u> profit data can be used, they pro-

vide automatic aggregation of deadweight losses. The usefulness of

profit data is explored in Part III, Section A, <u>infra</u>.

14.   On the computation of present values, see, for instance, R. Brealey

and S. Myers, <u>Principles of Corporate Finance</u> 10-41 (1981).

15. In some situations the concept of "barriers to mobility" introduced
by R. Caves and M. E. Porter, From Entry Barriers to Mobility
Barriers, 91 Quart. J. Econ. 241 (1977), is more relevant, but I
generally restrict my attention to situations in which only "barriers
to entry" matter.

16. Landes and Posner, supra note 1, at 991-96.

17. Here e is the basis of the system of natural logarithms. On this
sort of present value calculation, see Brealey and Myers, supra
note 14, at 36-38.

18. Landes and Posner are of course aware of the dynamic phenomena dis-
cussed here. They suggest use of longer-run elasticities in computing
DW in some situations than in others, depending on the nature of the
offence alleged and the delays to be expected in the judicial process.
Landes and Posner, supra note 1, at 959. While their observations in
this regard seem generally sensible, I think it is preferable to
consider separately the net loss involved, TDW, and the costs and
delays likely to be involved in attempting to mitigate it, since
they are conceptually distinct.

19. Id. 977. The quotation is from United States v. E. I. du Pont de
Nemours & Co., 351 U. S. 377, 391 (1956).

20. I have discussed these two dimensions of market power elsewhere;
see Schmalensee, On the Use of Economic Models in Antitrust: The
ReaLemon Case, 127 U. Pa. L. Rev. 994, 1005-09 (1979).

21. Landes and Posner, supra note 1, at 945 (note 19).

22. Id. 956-57.

23. This sort of situation is also discussed (at greater length) by
F. M. Fisher, Diagnosing Monopoly, 19 Quart. Rev. Econ. Bus. 7,
10-12 (1979). Fisher's prescriptions are essentially identical to
those of LP.

24. Indeed, in a situation in which oligopolistic interdependence is recognized, the firm's demand curve is not even well-defined. See, for instance, J. M. Henderson and R. E. Quandt, Microeconomic Theory 199 (3d. ed. 1980). Even in such situations, however, it is often useful to think of firms as behaving as if they faced well-defined demand curves and then to inquire as to the properties (such as elasticities) of those curves. The discussion here and in LP is consistent with this sort of approach. See Landes and Posner, supra note 1, at 951.

25. Id. 944.

26. Id.

27. It is worth noting that the assumption of perfectly competitive rival behavior likely produces an upward bias in the enhancement of market power predicted by LP's equation (4). Id. 972. (LP do note a number of other problems with such predictions. Id. 973-74.) That equation assumes that the larger firm, with a 20% pre-merger market share in their example, behaves as a dominant firm, while the smaller merger partner, with a 10% pre-merger share in the example, behaves perfectly competitively. But if shares are of the same order of magnitude, there is no obvious reason to expect such sharp qualitative differences in pre-merger behavior. Surely no such differences would be expected if the two shares were, say, 16% and 14%, for instance. If the smaller firm exercises some market power before the merger, the difference between the pre- and post-merger prices will generally be less than equation (4) indicates. That equation makes most sense when the smaller merger partner has a negligible pre-merger share, in which case almost any formal or informal model would predict a

negligible impact of the merger on price.

28. <u>Id</u>. 938.

29. Fisher, <u>supra</u> note 23, at 16.

30. I do not mean to suggest that LP share this belief, as they clearly do not. <u>See</u> Landes and Posner, <u>supra</u> note 1, at 947.

31. J. Robinson, <u>The Economics of Imperfect Competition</u> (1933).

32. <u>Id</u>. 5-6.

33. Landes and Posner, <u>supra</u> note 1, at 947-51.

34. <u>Id</u>. 948-50.

35. They do consider differentiation in their discussion of geographic market definition, which I examine in Section C, <u>infra</u>. <u>Id</u>. 963-72.

36. This is LP's equation (2) with some changes in notation. <u>Id</u>. 945.

37. This is LP's equation (3) with some changes in notation. <u>Id</u>. 945.

38. <u>Id</u>. 947-51.

39. <u>Id</u>.

40. <u>Id</u>. 958.

41. <u>Id</u>. 949. See also <u>Id</u>. 966.

42. <u>Id</u>. 949.

43. <u>Id</u>. 949-50.

44. Four different definitions of "capacity" are discussed by G. J. Stigler, <u>The Theory of Price</u> 156-58 (3d. ed. 1966). On some of the empirical problems involved in measuring "capacity", <u>see</u> A. Phillips, <u>An Appraisal of Measures of Capacity</u>, 53 Amer. Econ. Rev. Proc. 275 (1963).

45. Suppose that for a price-taking competitive firm marginal cost is rising except when output, q, satisfies $\underline{q} \leq q \leq \bar{q}$. Let marginal cost equal a constant, v, for all outputs in that range. Then if price is above v by any amount at all, the firm's optimal output

exceeds $\bar{q}$. Similarly, if price is below v, no matter how slightly, the firm optimally produces less than $\underline{q}$. Only in the unlikely razor's edge case in which price exactly equals v will q be between $\underline{q}$ and $\bar{q}$. And in that case, all outputs in that range yield identical net profit. Note also that "excess capacity" is a short-run concept with no meaning in the long run.

46. These quantitative problems are explored further in Section B of the Appendix.

47. Landes and Posner, _supra_ note 1, at 947-48.

48. _Id._ 961 (note 43).

49. _Id._ 948.

50. _Id._ 961 (note 43).

51. _See_ Schmalensee, _supra_ note 20, for a discussion of this example.

52. This is in fact clear from the development in Landes and Posner, _supra_ note 1, at 961 (note 43), and it would have been even more apparent if they had not imposed the assumption of constant real income.

53. _See_ the discussion of _Alcoa_ in _Id._ 978-79 or the discussion of ReaLemon in Schmalensee, _supra_ note 20, at 998, 1013-16.

54. LP and I may not have any real disagreement on this point: _see_ Landes and Posner, _supra_ note 1, at 958-59, 978-79. My aim here is mainly to clarify and support what I think is our common position on this issue. It would seem that we are also in general agreement with Fisher, _supra_ note 23, at 12-16.

55. Thus Fisher (_Id._ 16) notes that in antitrust, "The temptation is to regard products which are in [the relevant market] as all counting equally and products which are out [of the relevant market] as not counting at all."

56. If one regards products 1 and 2 in equation (6) as in the same
    "relevant market" even though σ is finite (so that the products are
    not perfect substitutes), it is plain that knowing w^1 provides
    essentially no information about ε^1, even if $\bar{\varepsilon}$, which would corre-
    spond to the market demand elasticity, is known.

57. On localization, see Schmalensee, _Entry Deterrence in the Ready-to-
    Eat Breakfast Cereal Industry_, 9 Bell J. Econ. 305 (1978), and _The
    New Industrial Organization and the Economic Analysis of Modern
    Markets_, in W. Hildenbrand (ed.), _Advances in Economic Theory_ (forth-
    coming). On strategic groups and mobility barriers, see Caves and
    Porter, _supra_ note 15.

58. Landes and Posner, _supra_ note 1, at 965, 969-70.

59. _Id_. 963. _See_ also their qualifications to this proposition, _Id_. 965,
    967. Their assertion that it is "unlikely that firms would risk
    giving imports a foothold in their market merely to improve their
    litigating position" (_Id_. 965, note 46), suggests another qualifi-
    cation. If distant or foreign sellers must make special investments
    in order to expand their output in the local market (the costs of
    establishing or expanding a network of dealerships, for instance),
    inclusion of the total output of distant sellers in "the relevant
    market" will tend to over-state the short run constraints those
    producers place on the power of local sellers. If the costs are
    small relative to the output expansion involved or if required
    investments could be easily liquidated, the over-statement will be
    minor. Subject to these same assumptions and qualifications, their
    treatment of exports (_Id_. 968-69) also appears correct.

60. 2 P. Areeda and D. Turner, _Antitrust Law_ ¶523 (1978).

61. A. Marshall, Principles of Economics 325 (8th ed. 1920). See also the general discussion of market definition in Id. 323-30.

62. Compare Landes and Posner, supra note 1, at 965 with Id. 969 or Areeda and Turner, supra note 60.

63. Landes and Posner, supra note 1, at 970.

64. Id.

65. The courts have, of course, generally been willing at least to consider other evidence; see for instance the discussions of price discrimination and product uniqueness in United States Steel Corp. et al v. Fortner Enterprises, Inc. 429 U. S. 610, 617, 619-21 (1977).

66. This discussion is motivated in part by the considerable weight apparently given to information about profitability by Landes and Posner, supra note 1, at 957, discussed in Part I, Section, supra.

67. In the absence of the measurement problems mentioned in the next paragraph, r^m would be the firm's accounting rate of return, and r would be its cost of capital. On this last concept, see Brealey and Myers, supra note 14, at ch. 9.

68. Fisher, supra note 23, at 19-22 provides a clear discussion of some of the pitfalls encountered in using profitability information in this context.

69. In addition to the measurement problems discussed by Fisher (Id.), inflation can make conventional accounting statements almost meaningless. On this problem, see J. B. Shoven and J. I. Bulow, Inflation Accounting and Nonfinancial Corporate Profits: Physical Assets, 2 Brookings Pap. Econ. Act. 557 (1975), and Inflation Accounting and Nonfinancial Corporate Profits: Financial Assets and Liabilities, 3 Brookings Pap. Econ, Act. 15 (1976). On some of the difficulties

involved in estimating a firm's cost of capital or normal rate of
return, see Brealey and Myers, supra note 12, at ch. 9, and Schmalensee,
Risk and Return on Long-Lived Tangible Assets, 9 J. Fin. Econ. forth-
coming (1981). In general, see Areeda and Turner, supra note 60,
at ¶512.

70. For some interesting recent work in this general area, see E. B.
Lindenberg and S. A. Ross, Tobin's q and Industrial Organization,
59 J. Bus. 1 (1981).

71. Areeda and Turner, supra note 60, at 332-34. As they note, in order
to establish power over price, one must also show that excess returns
are not entirely attributable to ownership of a naturally unique
and long-lived asset that provides lower cost but does not confer
market power. The best (and perhaps only) example of such an asset
is a small but unusually rich deposit of some mineral. As Areeda
notes, such competitive explanations for high profits as "superior
production resources or managerial skill. . .become less likely as
the period of high profits is more prolonged." P. Areeda, Antitrust
Analysis 38 (2d ed. 1974).
Also, to the extent that resources are expended to obtain or maintain
market power or simply wasted by managers not subject to the discipline
of effective competition, deadweight loss estimates based on observed
profits understate the costs of market power; see R. Posner, The Social
Costs of Monopoly and Regulation, 83 J. Polit. Econ. 807 (1975).

72. See, for instance, A. D. Neale, The Antitrust Laws of the U. S. A.
37 (2d. ed. 1970), or Areeda and Turner, supra note 60, at 345 (note 2).

73. F. M. Scherer, Industrial Market Structure and Economic Performance
315 (2d. ed. 1980).

74. Areeda and Turner, supra note 60, ¶514 provide a good discussion of
these problems.

75. This is consistent with the Court's apparent willingness to rely on such evidence in tying cases. United States Steel Corp. et al v. Fortner Enterprises, Inc. 429 U. S. 610, 617 (1977).

76. R. H. Bork, The Antitrust Paradox 144 (1978). For a rigorous development of the implications of this definition, see J. Ordover and R. Willig, An Economic Definition of Predation, unpublished manuscript.

77. For an application of this reasoning to a particular case, see Schmalensee, supra note 20, at 1016, 1029-31. In the dominant firm -- competitive fringe model used by LP, all rivals of the firm studied are assumed to be too small to affect market conditions, so that this model cannot be used to analyze situations in which a firm's profits would be noticeably affected by the entry or exit of any single rival.

78. Thus in most cases it is probably most sensible to employ something like the two-tier approach advocated by P. L. Joskow and A. K. Klevorick, A Framework for Analyzing Predatory Pricing Policy, 89 Yale L. J. 213.

79. Examples of such analyses include P. W. MacAvoy, Price Formation in Natural Gas Fields (1962), E. Applebaum, Testing Price Taking Behavior, 9 J. Econometrics 283 (1979), F. M. Gollop and M. J. Roberts, Firm Interdependence in Oligopolistic Markets, 10 J. Econometrics 313 (1979), D. Carlton, W. Landes, and R. Posner, Benefits and Costs of Airline Mergers: A Case Study, 11 Bell J. Econ., 65 (1980), and T. F. Bresnahan, Competition and Collusion in the American Automobile Industry: The 1955 Price War, unpublished manuscript.

80. By this I mean simply that if a firm sets prices assuming it has market power and is then forced to lower them because of inadequate sales, the market's rejection of its decisions shows its lack of power.

81. Areeda and Turner, supra note 60, at 345 reach this same conclusion regarding other sorts of evidence on conduct.

82. I must disagree with one more dictum. Landes and Posner, supra note 1, at 176, assert that "It should always be open to a defendant in an antitrust case to rebut an inference of market power based on market share by showing that its market share is the result of low prices." Except for cases in which mergers, regulatory limits on entry, or blatantly exclusionary practices are important, firms generally acquire large market shares because either cost advantages or product superiority permit them to sell at a lower quality-adjusted price than their closest rivals. If one followed the LP recommendation, most "clean-handed" monopolists could thus rebut the inference that they had any market power at all. Evidence on the sources of market power is much more appropriately used to judge the legality of the acquisition and exercise of that power; it would cause considerable confusion if "market power" were to be re-defined (as LP implicitly suggest) so that "clean-handed" firms could never possess it.

83. Equality of the cross-price coefficients is a consequence of the assumption of no income effects; see, for instance, H. Varian, Microeconomic Analysis 84-99 (1978). Landes and Posner, supra note 1, at 961 (note 43) make this same assumption when they hold real income constant.

84. This is the indirect utility function corresponding to the demand functions given above; See Varian, supra note 83, at 89-90. In order for this function to be well-behaved, we need $c_1 c_2 > d^2$. This suffices to show that the first expression for DW in the next paragraph is always non-negative and is zero only when $P_1 = v_1$ and $P_2 = v_2$.

85. This is identical to the definition given by Landes and Posner, supra note 1, at 984, except that they use the partial derivative of Q^S with respect to P instead of the ordinary derivative that appears here. Since Q^S is a function of only one argument, the ordinary derivative is more appropriate, but their definition is perfectly correct.

86. Id. 949.

87. See, for instance, Id. 951.

88. This is the easiest of the four notions of "capacity" discussed by Stigler, supra note 42, to employ here.

89. Landes and Posner, supra note 1, at 983-84.

90. See note 83 supra.

91. The presence of income effects in general gives rise to technical problems in the definition of terms like "substitutes" and "complements". It is possible in theory for σ_2^1 and σ_1^2 to have opposite signs, though this is unlikely to be encountered in antitrust contexts. In any case it seems sensible to focus on the weighted average of these two elasticities, σ. For a discussion of some of the technical issues that arise in this context, see P. A. Samuelson, Complementarity -- An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory, 12 J. Econ. Lit. 1255 (1974).
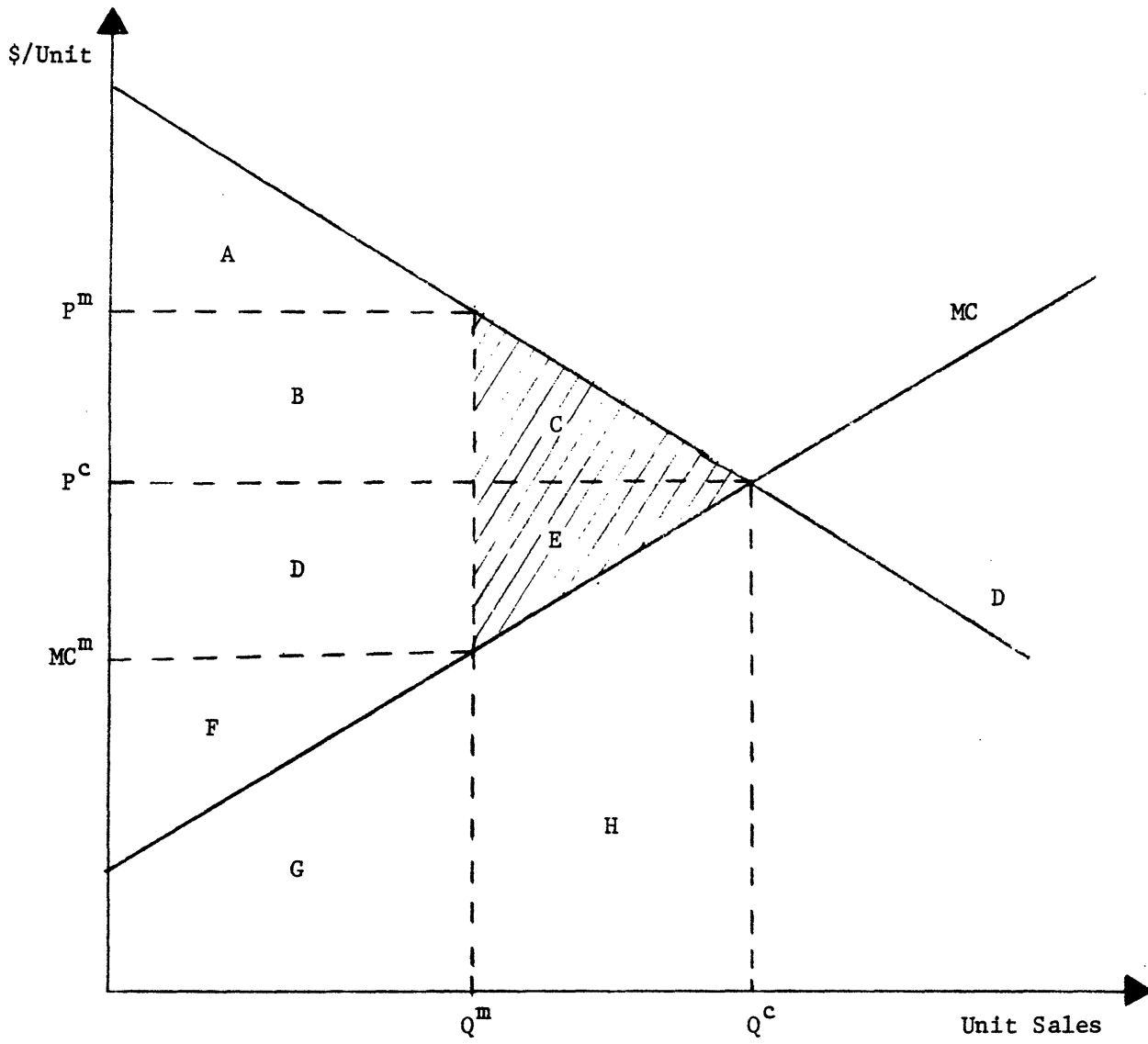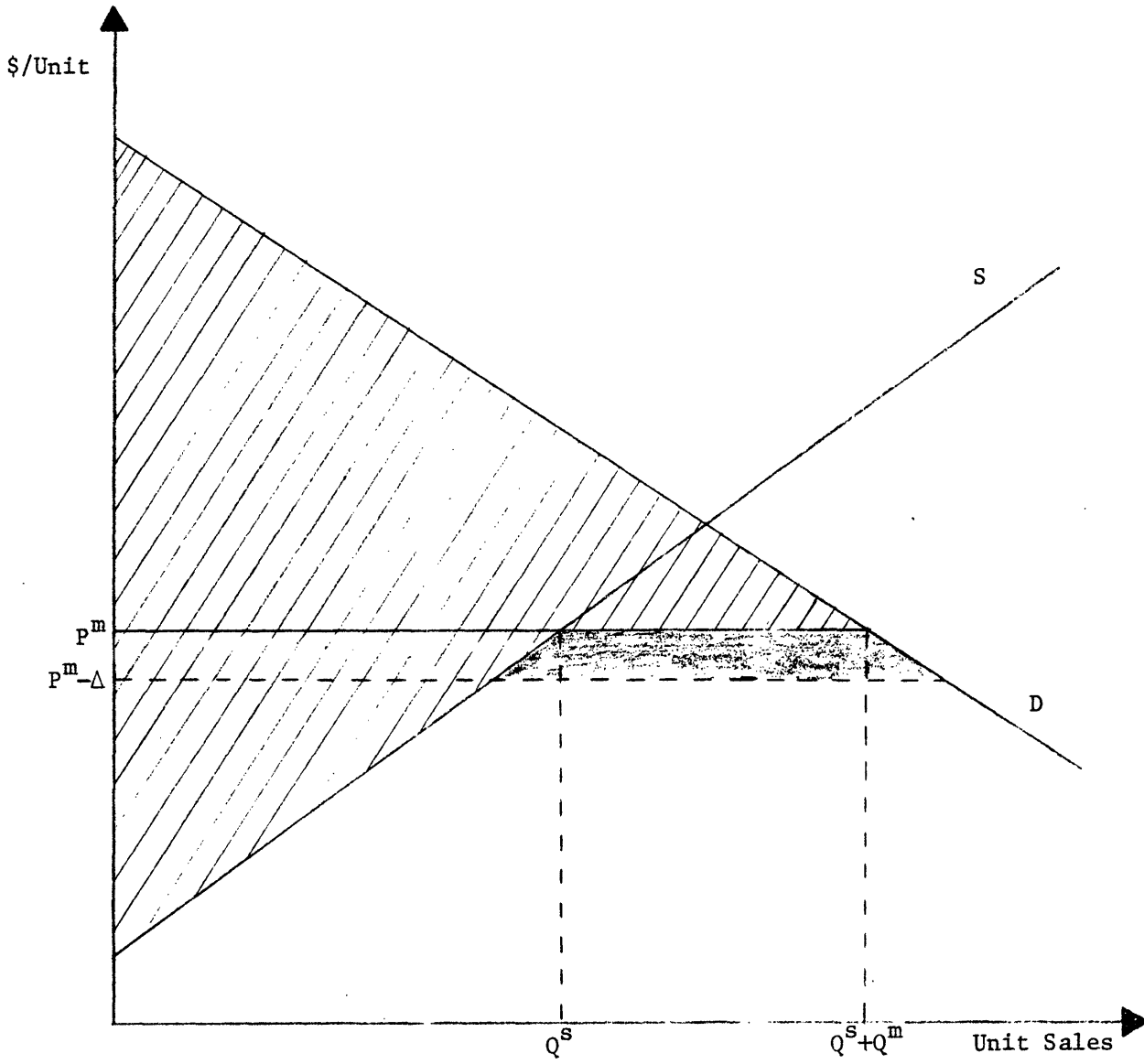
Figure 1

Figure 2