

ECONOMIES AS AN ANTITRUST DEFENSE: THE WELFARE TRADEOFFS

By OLIVER E. WILLIAMSON*

Suppose that a merger (or other combination) is proposed that yields economies but at the same time increases market power. Can the courts and antitrust agencies safely rely, in these circumstances, on a literal reading of the law which prohibits mergers “where in any line of commerce or any section of the country, the effect of such acquisition may be substantially to lessen competition, or to tend to create a monopoly,”¹ or does this run the risk of serious economic loss? In the usual merger where both effects are insubstantial this problem is absent.² But in the occasional case where efficiency and market power consequences exist, can economies be dismissed on the grounds that market power effects invariably dominate? If they cannot, then a rational treatment of the merger question requires that an effort be made to

* The author is associate professor of economics at the University of Pennsylvania.

¹ Public Law 899, Sec. 7, 38 Stat. 731, as amended; 15 U.S.C. 18.

² Donald Dewey has observed in this connection that most mergers “have virtually nothing to do with either the creation of market power or the realization of scale economies” [9, p. 257]. Jesse Markham agrees that since 1930 monopolization has not been a principal merger objective, but finds that “some mergers have undoubtedly come about as adjustments to major innovations . . . : the first great wave of mergers followed a period of rapid railroad building, and the wave of the 1920s came with the rise of the motor car and motor truck transportation and a new advertising medium, the home radio” [22, pp. 181–82]. It might be useful briefly to summarize some of the ways in which efficiencies might result from combination. These would include miscalculation, shifts in demand, technological developments, displacement of ineffective managements, and mixtures thereof.

As an example of miscalculation consider two firms that have entered a market at an efficient plant scale but have incorrectly estimated the volume necessary to support an efficient distribution system. Combination here could lead to efficiencies but might also have some market power effects (reducing competition between the two but possibly enhancing their competitive position with respect to their rivals). A significant, persistent decline in demand might produce a condition of excess capacity in which combination would permit economies but would also have market power consequences. As discussed in Section III, an increase in demand might induce a change from job shop to assembly line type operations with vertical integration consequences. Technological developments may similarly provide opportunities for a significant reorganization of resources into more efficient configurations—the electronic digital computer being a recent example. Finally, merger may be the most expeditious way of displacing an inefficient by a more efficient management—but the benefits here may only be of a short-run variety. A manifestly inefficient management would, hopefully, be displaced by other means if, by reason of the market power consequences of a combination, the merger route were closed.

A merger can, of course, produce diseconomies as well. What I have previously characterized as the “control loss” phenomenon appears to be an increasing function of firm size [31]. See also Parts 7 and 8, Section II, *infra*.

establish the allocative implications of the scale economy and market power effects associated with the merger.

The initial indication of the Supreme Court's view on this question came on the occasion of the first merger case to come before it under the 1950 amendment to Section 7 of the Clayton Act. In a unanimous opinion, the Court took the position in *Brown Shoe* that not only were efficiencies no defense, but a showing that a merger resulted in efficiencies could be used affirmatively in attacking the merger since small rivals could be disadvantaged thereby [6, p. 374]. Opportunities to reconsider this position have presented themselves since, *Procter & Gamble* being the most recent.

Justice Douglas, in delivering the opinion of the Court, observed that Procter & Gamble "would be able to use its volume discounts to advantage in advertising Clorox," and went on to state that "economies cannot be used as a defense to illegality. Congress was aware that some mergers which lessen competition may also result in economies but it struck the balance in favor of protecting competition" [10, pp. 1230-31]. Although reference to congressional intent may relieve the Court of the responsibility for making tradeoff valuations, this does not fully dispose of the issue. What tradeoff calculus did Congress employ that produced this result?

In a concurring opinion to the Clorox decision, Justice Harlan provides the first hint that efficiencies may deserve greater standing. At least with respect to conglomerate or product-extension mergers "where the case against the merger rests on the probability [as contrasted, apparently, with a certainty] of increased market power, the merging companies may attempt to prove that there are countervailing economies reasonably probable which should be weighed against the adverse effects" [10, pp. 1240-41]. But inasmuch as the economies in Clorox were in his opinion merely pecuniary rather than real, which distinction is of course appropriate, he concluded that Procter's efficiency defense was defective [10, p. 1243].

Even if Justice Harlan's position were the prevailing one, it is clear that economies would be an acceptable antitrust defense for only a restricted set of structural conditions. Since the relevant economic theory, although widely available, has never been developed explicitly on this issue, such a result is not unexpected. Indeed, lacking a basis for evaluating net effects, for the Court to hold that the anticompetitive consequences of a merger outweigh any immediate efficiency advantages is only to be expected. An institution acting as a caretaker for the enterprise system does not easily exchange what it regards as long-term competitive consequences for short-term efficiency gains.

The merits of the Supreme Court's position on mergers are at the

heart of the recent Bork and Bowman v. Blake and Jones debate [2, 3, 4, 5]. Although this dialogue deals directly with the critical issues, its failure to produce a consensus is at least partly due to the fact that essential aspects of the relevant economic model were not supplied. Lacking a tradeoff relation, Bork is forced to assert that "Economic analysis does away with the need to measure efficiencies directly. It is enough to know in what sorts of transactions efficiencies are likely to be present and in what sorts anticompetitive effects are likely to be present. The law can then develop objective criteria, such as market shares, to divide transactions [into those predominantly one type or other]" [5, p. 411]. But this obviously leaves the mixed cases, which are the hard ones, unresolved. Blake and Jones, by contrast, conclude that "claims of economic efficiency will not justify a course of conduct conferring excessive market power. The objective of maintaining a system of self-policing markets requires that all such claims be rejected" [3, p. 427]. But what are the standards for "excessive" market power and "self-policing" markets? And are these really absolute or do they reflect an implicit tradeoff calculation? And if it is the latter, should we (if we can) make this tradeoff explicit?

Indeed, there is no way in which the tradeoff issue can be avoided. To disallow tradeoffs altogether merely reflects a particularly severe a priori judgment as to net benefits. Moreover, it is doubtful that a goal hierarchy scheme of the sort proposed by Carl Kaysen and Donald Turner has acceptable properties. As they formulate the problem, higher level goals strictly dominate lower level goals, so that only when the latter are available without sacrifice in the former is lower level goal pursuit allowed [16, pp. 44–45]. Inasmuch as they rank efficiency and progressiveness above reductions in market power, an absolute defense would appear to obtain when, for any structural condition present or prospective, it could be shown either that economies have not yet been exhausted or that discreteness conditions (indivisibilities) would not efficiently permit a separation [16, pp. 44–46, 58, 78]. But this may be to construe their intentions too narrowly; for it is with antitrust actions that result in *substantial* efficiency losses [16, pp. 44, 133] and involve *too great* a sacrifice in performance [16, p. 58] that they are especially concerned. Although these distinctions are important, they are not ones for which goal hierarchy analysis is well suited to deal. Tradeoff analysis, by contrast, is designed to cope with precisely these types of issues.

The relevant partial equilibrium model with which to characterize the tradeoffs between efficiency and price effects together with a representative set of indifference relations are developed in Section I of this paper. A variety of essential qualifications to this naive model are then

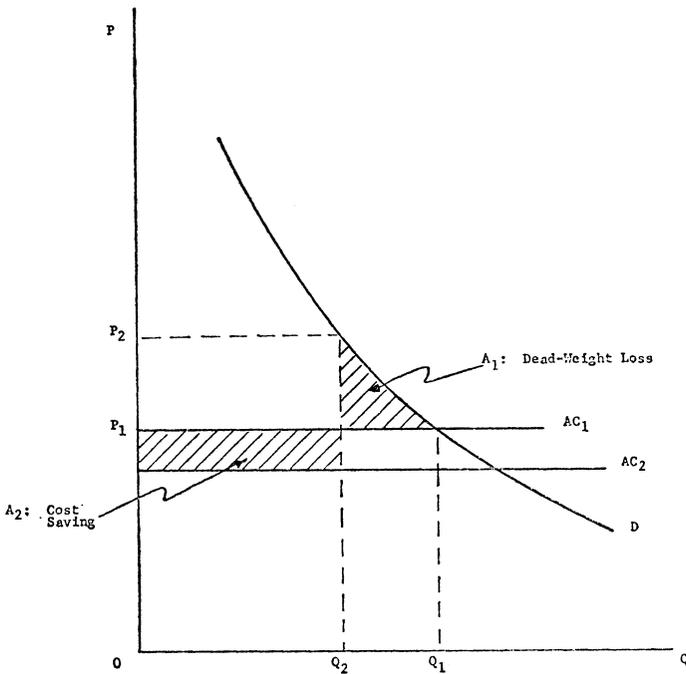


FIGURE 1.

presented in Section II. Extensions of the argument, which is developed initially in horizontal merger terms, to deal with questions of dissolution as well as vertical and conglomerate mergers, are given in Section III. The conclusions follow in Section IV.

I. *The Naive Tradeoff Model*

The effects on resource allocation of a merger that yields economies but extends market power can be investigated in a partial equilibrium context with the help of Figure 1. The horizontal line labeled AC_1 represents the level of average costs of the two (or more) firms before combination, while AC_2 shows the level of average costs after the merger. The price before the merger is given by P_1 and is equal to k (AC_1), where k is an index of pre-merger market power and is greater than or equal to unity. The price after the merger is given by P_2 and is assumed to exceed P_1 (if it were less than P_1 the economic effects of the merger would be strictly positive).³

³ This is a simple but basic point. It reveals that market power is only a necessary and not a sufficient condition for undesirable price effects to exist. It would be wholly irrational to regard an increase in the price to average cost ratio ($P_2/AC_2 > P_1/AC_1$) as grounds for opposing a merger if, at the same time, the post-merger price were less than the pre-merger level ($P_2 < P_1$).

The net welfare effects of the merger are given (approximately) by the two shaded areas in the Figure. The area designated A_1 is the familiar dead-weight loss that would result if price were increased from P_1 to P_2 , assuming that costs remain constant. But since average costs are actually reduced by the merger, the area designated A_2 , which represents cost savings, must also be taken into account. The net allocative effect is given by the difference, $A_2 - A_1$, of these two areas.⁴

The area A_2 is given by $(AC_2 - AC_1)Q_2$, or $[\Delta(AC)]Q_2$, while A_1 is given approximately by $\frac{1}{2}(P_2 - P_1)(Q_1 - Q_2)$, or $\frac{1}{2}(\Delta P)(\Delta Q)$. The net economic effect will be positive if the following inequality holds:

$$(1) \quad [\Delta(AC)]Q_2 - 1/2(\Delta P)(\Delta Q) > 0.$$

Dividing through by Q_2 and substituting for $\Delta Q/Q$ the expression $\eta(\Delta P/P)$, where η is the elasticity of demand, we obtain:

$$(2) \quad \Delta(AC) - 1/2(\Delta P)\eta \frac{\Delta P}{P} > 0.$$

Finally, dividing through by $P_1 = k(AC_1)$ we have as our criterion:

$$(3) \quad \frac{\Delta(AC)}{AC} - \frac{k}{2} \eta \left(\frac{\Delta P}{P}\right)^2 > 0.$$

If this inequality holds, the net allocative effect of the merger is positive. If the difference is equal to zero the merger is neutral. If the inequality is reversed the merger is negative.

In words, the inequality shown in (3) says that if the decimal fraction reduction in average costs exceeds the square of the decimal fraction increase in price multiplied by one-half k times the elasticity of demand, the allocative effect of the merger is positive. Setting k equal to one (which it will be if the pre-merger market power is negligible), the cost reductions necessary to offset price increases for various values of the elasticity of demand are shown in Table 1.

For example, if price were to increase by 20 per cent, then running across the row $[(\Delta P/P) \times 100] = 20$ we observe that if η is 2 a cost reduction of 4 per cent will be sufficient to offset the price increase, while if η is 1 only a 2 per cent cost decrease is needed to neutralize the price effect, and if η is $\frac{1}{2}$, a cost reduction of 1 per cent is sufficient. More generally it is evident that a relatively modest cost reduction is usually

⁴ My use of dead-weight loss is somewhat restrictive. Inefficiency is also a dead-weight loss. For convenience of exposition, however, I refer to the Marshallian triangle as the dead-weight loss and compare this to the cost saving (efficiency) aspects of a merger. Estimating the value of consumers' surplus by the Marshallian triangle follows the common (and broadly defensible) practice of suppressing the income effects associated with a price change. The net social benefit associated with a particular cost-price configuration is defined as total revenue plus consumers' surplus less social cost, where social and private costs are assumed to be identical (externalities and producers' surplus are both assumed to be zero).

TABLE 1—PERCENTAGE COST REDUCTIONS $[(\Delta(AC)/AC)\times 100]$ SUFFICIENT TO OFFSET PERCENTAGE PRICE INCREASES $(\Delta P/P\times 100)$ FOR SELECTED VALUES OF η

η [($\Delta P/P$) $\times 100$]	2	1	1/2
5	.25	.12	.06
10	1.00	.50	.25
20	4.00	2.00	1.00
30	9.00	4.50	2.25

sufficient to offset relatively large price increases even if the elasticity of demand is as high as 2, which is probably a reasonable upper bound. Indeed, if a reduction in average costs on the order of 5 to 10 per cent is available through merger, the merger must give rise to price increases in excess of 20 per cent if $\eta \cong 2$, and in excess of 40 per cent if $\eta \cong \frac{1}{2}$, for the net allocative effects to be negative. Moreover, it should be noted, if the merger reduces average costs by x per cent and the post-merger price increases by y per cent, the post-merger price to average cost differential slightly exceeds $x+y$ per cent. Thus, expressing price with respect to the post-merger level of average costs yields an even greater differential than is reflected by the relations stated above. The naive model thus supports the following proposition: a merger which yields non-trivial real economies must produce substantial market power and result in relatively large price increases for the net allocative effects to be negative.

II. Qualifications

Our partial equilibrium analysis suffers from a defect common to all partial equilibrium constructions. By isolating one sector from the rest of the economy it fails to examine interactions between sectors. Certain economic effects may therefore go undetected, and occasionally behavior which appears to yield net economic benefits in a partial equilibrium analysis will result in net losses when investigated in a general equilibrium context. Such a condition has been shown to exist in an economy in which monopoly exists in many sectors. Thus, whereas partial equilibrium analysis indicates that an increase in the monopoly price in any one sector invariably yields a loss, viewed more generally such an isolated price increase may actually lead to a desirable reallocation of resources.⁵ Conceivably, therefore, a merger that has monopoly power and cost-saving consequences could yield benefits in *both* respects—although it is probably rare that operational content can be supplied

⁵ This is the familiar "second-best" argument. For a discussion of second-best qualifications in treating the monopoly problem, and references to this literature, see Ferguson [11, pp. 16-17, 49-51].

to this qualification. But were there no other considerations, such bias as our partial equilibrium construction produces would be to underestimate the net economic gains of combination.

This does not, however, exhaust the range of qualifications. Among the other factors that can or should be taken into account are inference and enforcement expense, timing, incipiency, weighting, income distribution, extra-economic political objectives, technological progress, and the effects of monopoly power on managerial discretion.

A. Inference and Enforcement Expense

The relevant effects are those which take the form of real rather than pecuniary economies. Also, since evaluating a claim that economies exist will itself absorb real resources, it seems reasonable to impose a requirement that the net gain exceed some threshold value before such a defense will even be entertained. This, in conjunction with qualifications B through D below, would appear to meet Donald Turner's point that if economies are to be invoked as a defense "the law might well require clear and convincing evidence that the particular merger would produce substantial economies that could not be achieved in other ways" [27, p. 1328]. As the tools for assessing economies are progressively refined (and the incentive to make such improvements is obvious once an efficiency defense—even in principle—is granted), this threshold level should be reduced accordingly.

Operationally it may be essential to express the value of the threshold as a function of the ease with which economies can be established. Economies that have a highly speculative aspect should be required to reach a higher minimum level than those which are more objectively specified. (Thus if economies in both production and distribution expenses are claimed, and if the former are better specified than the latter, distribution economies would have to reach a higher threshold than would production economies to be admissible.) Since the ease with which exaggerated claims are detected varies directly with the degree of distortion attempted, and since evidence of distortion seriously debilitates a defense, adjusting the threshold in this way will tend to protect the enforcement agencies against grievously inflated efficiency claims.

Bork, apparently, would resist the argument that the defendants should bear the burden of proof on efficiencies since many efficiencies may be difficult to establish [5, p. 410]. But if efficiencies are to be a defense at all, it is clear that the companies—which are, presumably, sensitive to the relevant economies in proposing the merger in the first place—must be prepared to make the case for them in court. They have the data and these must be supplied. Otherwise the mixed case which involves both scale economy and market power effects can only be handled arbitrarily—and this is satisfactory to no one.

B. *Timing*

Significant economies will ordinarily be realized eventually through internal expansion if not by merger. Growth of demand can facilitate this internal adjustment process; the necessity for part of the industry to be displaced in order that efficient size be achieved is relieved in a growing market. Thus, although a merger may have net positive effects immediately (cost savings exceed the dead-weight loss), when allowance is made for the possibility of internal expansion these effects can become negative eventually (the cost savings persist, but these could be realized anyway, and the dead-weight loss could be avoided by prohibiting the merger).

Designating the dead-weight loss effects of the merger by $L(t)$ and the cost savings by $S(t)$, the argument would be that the value of $S(t)$ falls while $L(t)$ persists over time. Thus, taking the discounted value of net benefits (V) we have:

$$(4) \quad V = \int_0^T [S(t) - L(t)]e^{-rt} dt,$$

and if initially $S(t)/L(t) > 1$, but eventually $S(t)/L(t) < 1$, this can easily become negative. Consider, for example, the case where $S(t) = \bar{S}$ for a period of length T' and then becomes zero, while $L(t) = \bar{L}$ indefinitely. Using a social discount rate of 10 per cent, what initial combinations of \bar{S}/\bar{L} and T' would leave us just indifferent over the allocative effects of a merger? For \bar{S}/\bar{L} of 3, indifference occurs at a value of T' of 4 years; any value of T' less than 4 years would reveal that the scale economies can be realized by internal expansion in a sufficiently short interval that the merger should be disallowed, while any value of T' that exceeds 4 years would show that net gains are available by approving the merger. For \bar{S}/\bar{L} of 2, the corresponding value of T' is 7 years, while for \bar{S}/\bar{L} of 1.5, the value of T' increases to 11 years. The necessary qualifications to our earlier results are thus obvious: only if \bar{S}/\bar{L} is relatively large, or T' reasonably long, should a merger which results in eventual net losses be approved.

By contrast with a growing market, to force economies to be realized by internal expansion in a static market is generally without merit. The market power effects will occur here anyway, and the internal expansion route merely delays and may upset the market adjustment.

The above results are merely illustrative. More generally, equation (4) calls attention to the importance of considering the shape of the time stream of benefits and costs that a merger produces. Thus it is not sufficient to justify a merger on the basis of merely potential economies. Not only is it relevant to consider whether the merger would produce net benefits, but whether the timing is such as to maximize these gains. If a

merger is proposed that promises potential economies, but these will not be realized for some time, it may be better to delay the combination. Such might be the case in circumstances where the existing plant has not exhausted its useful life and has limited value in other uses; in this situation investment in the new facilities may not be economical immediately. For the merger to occur much earlier than the indicated economies will be realized would permit whatever market power effects as the merger produces to take effect at an earlier time than is clearly most beneficial.

Plausible as this last argument may appear, it raises a serious question of how extensive a "management" function the enforcement agencies should play in merger matters. It is an easy step from the suggestion that a proposed merger should be delayed until maximum net gains are realized to the proposition that the enforcement agencies should "arrange" optimal firm pairings. Both of these, however, are much more ambitious tasks than merely testing whether the net gain associated with a proposed combination is positive. Subject possibly to occasional exceptions where the social net benefit calculus identifies a distinctly superior timing or combination from that which has been proposed privately (and assuming that the change can be implemented), the simple requirement that discounted net gains be positive is probably a sufficient test. Otherwise, mergers are too complex to postpone casually; and the enforcement agencies are not designed (nor should they be redesigned) to function in a brokerage capacity.

C. *Incipiency*

It is likewise vital to consider not merely the market power effects of any single merger taken in isolation, but whether the merger is representative of a trend. If a series of such mergers can reasonably be expected, the judgment of whether to permit any given combination should properly be cast in an industry context—in which case the anticipated economy and market power effects throughout the industry should be examined. Since, if economies are available by combining one pair of firms they will often be available more generally, this may frequently be an important consideration. The notion of incipiency thus has special relevance in administering the law on mergers where economies are claimed.

This proposition might usefully be contrasted with that of Bork and Bowman [2, p. 594]:

The difficulty with stopping a trend toward a more concentrated condition at a very early stage is that the existence of the trend is *prima facie* evidence that greater concentration is socially desirable. The trend indicates that there are emerging efficiencies or economies of scale—whether due to engineering and production developments or to

new control and management techniques—which make larger size more efficient. This increased efficiency is valuable to the society at large, for it means that fewer of our available resources are being used to accomplish the same amount of production and distribution. By striking at such trends in their very earliest stages the concept of incipency prevents the realization of those very efficiencies that competition is supposed to encourage.

Their evaluation of the social desirability of a trend suggests a certain insensitivity to the relevant scale economy-market power tradeoff considerations, and they appear to read the significance of a trend somewhat too loosely. That a trend necessarily implies emerging efficiencies is incorrect: it may also indicate an emerging awareness that market power advantages might be realized through a series of combinations.⁶ Moreover, whereas they seem to suggest that to disallow a merger is to prevent the realization of scale economies altogether, ordinarily it is not a question of whether economies will be realized but when and with what market power effects. Thus, while Bork and Bowman may be correct in charging that scale economy justifications have not been given sufficient weight in the recent enforcement of the merger law, they are also guilty of a certain heavy-handedness in their own treatment of the incipency question.

D. *Weighting*

The economies that a merger produces are usually limited strictly to the combining firms. But the market power effects of a merger may sometimes result in a price increase across a wider class of firms. Where this occurs, a weighting factor should be introduced into expression (3) to reflect this condition. The criterion becomes:

$$(3') \quad \left(\frac{Q_2}{Q_T}\right) \frac{\Delta(AC)}{AC} - \frac{k}{2} \eta \left(\frac{\Delta P}{P}\right)^2 > 0,$$

where Q_2 is the output of the merging firms and Q_T is the total quantity of industry sales for which the price increase becomes effective.

E. *Income Distribution*

An additional qualification to our analysis involves income distribution effects. The rectangle in Figure 1 bounded by P_2 and P_1 at the top and bottom respectively and O and Q_2 on the sides represents a loss of consumers' surplus (gain in monopoly profits) that the merger produces. On the resource allocation criteria for judging welfare effects advanced above, the distribution of these profits becomes a matter of indifference.

⁶ This is George Stigler's point in his treatment of "Monopoly and Oligopoly by Merger" [24]. Bork concedes this possibility in his response to Blake and Jones [5, p. 412]; but his principal emphasis, which is probably correct, is that a trend signals emerging economies.

For specific welfare valuations, however, we might not always wish to regard consumer and producer interests symmetrically—although since, arguably, antitrust is an activity better suited to promote allocative efficiency than income distribution objectives (the latter falling more clearly within the province of taxation, expenditure, and transfer payment activities), such income distribution adjustments might routinely be suppressed. If they are not, the tradeoff between efficiency gains and distributive losses needs explicitly to be expressed. Thus, while economies would remain a defense, any undesirable income distribution effects associated with market power would be counted against the merger rather than enter neutrally as the naive model implies.

Inasmuch as the income redistribution which occurs is usually large relative to the size of the dead-weight loss, attaching even a slight weight to income distribution effects can sometimes influence the overall valuation significantly. Thus, expressing the dead-weight loss ($L = \frac{1}{2} (\Delta P) (\Delta Q)$) as a ratio of the income distribution effect ($I = (\Delta P)Q$), and substituting into this ratio the expression for the elasticity of demand (η), the fraction $L/I = \frac{1}{2} (\Delta P/P) \eta$ obtains. It is therefore obvious that, except where the elasticity of demand is “high,” the dead-weight loss as a fraction of the income distribution effect is relatively small—certainly less than unity. Hence if, as is probably common, the income redistribution which results when market power is increased is regarded unfavorably, an appropriate weighting of this factor will, at least occasionally, upset a net valuation which on resource allocation grounds is positive.

Note in this connection that the transfer involved could be regarded unfavorably not merely because it redistributes income in an undesirable way (increases the degree of inequality in the size distribution of income), but also because it produces social discontent. This latter has serious efficiency implications that the above analysis does not take explicitly into account. This same point also appears to have gone unnoticed in the entire Bork and Bowman v. Blake and Jones exchange [2] [3] [4] [5]. Distinguishing social from private costs in this respect may, however, be the most fundamental reason for treating claims of private efficiency gains skeptically.

F. *Political Considerations*

Combinations which involve firms that are already very large in absolute terms might be resisted on grounds that these raise extra-economic problems of political significance. There is not, however, any obvious way in which to integrate these into the analysis. Rather, although the political implications of control over wealth are a matter for serious concern, these are separable from the economic problems posed by control over markets; a different calculus is required to deal

with each. The necessary political judgment, ideally, is one for Congress to make. Possibly, as Carl Kaysen has suggested, this would take the form of a prohibition against expansion by merger of the largest 50 or 100 corporations [17, p. 37].

The issue here reaches beyond the social discontent matter raised above. Thus, whereas social discontent can be reduced, in principle at least, to efficiency-equivalent (net value product) terms, the political implications of the control over wealth involve a judgment of how the quality of life in a democracy is affected by size disparities. The latter is less easily (or even appropriately) expressed in efficiency terms. The issue is nevertheless important, and failure to deal with it may be unresponsive to the position taken by Blake and Jones. Inasmuch as several of the counterexamples that they pose in their critique of Bork and Bowman appear deliberately to have been selected from the giant firm universe [5, pp. 425–27], possibly it is mergers within this subset that concern them most. Should economies be allowed as a defense, therefore, the rule proposed by Kaysen would limit such a defense in a way which would presumably relieve this aspect of their concern.

G. *Technological Progress; and*
H. *Managerial Discretion*

The highly conjectural nature of qualifications G and H makes it unclear at this time what weight ought to be assigned to them. It is at least arguable that the prevailing uncertainties are too great to give any effect to these two factors at this time. They are, nevertheless, potentially of such significance that to dismiss them may run the risk of serious error. In consideration of this potential importance, additional research which would permit us better to evaluate their actual significance would seem warranted. The manner in which each would influence the estimate of net effects is sketched out below.

Consider technological progress first. Such increases in market power that result in predictable effects on technological progress should, if they can easily, be taken into account. The present evidence, while hardly abundant, suggests that, as a general rule, the research and development expenditures of the four largest firms in an industry are neither as large proportionately nor as productive as those of their immediately smaller rivals.⁷ But this fails to answer the question of what

⁷ With respect to size, Mansfield found that the ratio of innovations to firm size reached a maximum at about the sixth largest firm for the petroleum and coal industries, and at a much lower rank for steel [20, p. 566]. Elsewhere Mansfield reports that the largest firms in petroleum, drugs, and glass spent somewhat less on R&D, relative to sales, than did somewhat smaller firms; in chemicals they spent somewhat more; in steel they spent less, but the difference was not statistically significant [21, p. 334]. Scherer concludes from his study of patent behavior in a group of 448 firms selected from the Fortune list of the largest 500 industrial corporations in 1955 that “the evidence does not support the hypothesis that corporate bigness is especially favorable to high inventive output” [23, p. 1114]. Turning to productivity, Mansfield con-

market structures most enhance progressiveness. The evidence on this latter is somewhat mixed.⁸ It seems unlikely, however, that subsequent investigation will upset the basic proposition that progressiveness is promoted by at least some elements of competition at virtually every stage of an industry's development—if for no other reason than that competition tends to assure that variety in research approaches will be employed. Local or regional monopolies may provide partial exceptions (since here the requisite variety will be available nationally, although the rate at which innovations are implemented may nevertheless lag if competitive pressures are lacking), but monopoly, or near-monopoly, would not seem to be the perfect instrument for technical progress in industries for which the relevant market is national.

Lacking additional evidence, it would not seem injudicious to assume that mergers between relatively small-sized firms rarely have negative (and may frequently have positive) effects on progressiveness, whatever the condition of concentration. This judgment probably holds for most mergers involving lower-middle sized firms as well. Thus it is mainly in the relatively large firms, particularly those in moderately to highly concentrated national markets (which, of course, are also ones where market power effects may be important), that the effects of a merger on technological progress deserve special attention.

Whether the effects be positive or negative, the necessary extension to the model is identical. Assume therefore that a merger is proposed involving a large firm in a concentrated industry, and that while it yields economies it also predictably decreases the rate of progressiveness. Holding constant for the moment the effects on price, how large a change in the rate of technical progress would be required to offset the available economy of scale advantage? To obtain a crude estimate of this, let θ be the ratio of the immediate post-merger to pre-merger average costs (so that $1 - \theta$ is the immediate decimal fraction reduction in average costs), g_1 be the rate of productivity increase in the absence of the merger and g_2 the rate if the merger is approved (where $g_1 \geq g_2$), $Q(t)$ be the output in period t , and let r be the social discount rate. Then the merger will have neutral effects if the discounted value of costs under

cludes that “in most industries, the productivity of an R&D program of given scale seems to be lower in the largest firms than in somewhat smaller firms” [21, p. 338]. Comanor found that diseconomies of scale in the pharmaceutical industry were encountered at even moderate firm sizes [8, p. 190]. For a recent review of this literature, see Johnson [15, pp. 169–71].

⁸ Hamburg [13, Ch. 4] and Horowitz [14, pp. 330–01] report a positive correlation between R&D expenditures and industrial concentration. Scherer finds a much weaker but slightly positive association [23, pp. 1119–21]. Kendrick concludes from an examination of [Terleckyj] data that there is no significant correlation between productivity changes and industrial concentration [18, p. 179]. Stigler found in an earlier study “hints that industries with lower concentration had higher rates of technological progress” [26, p. 278], while I, using Mansfield's data, found a negative correlation between the proportion of innovations introduced by the four largest firms and industrial concentration [30].

each condition is the same. This requires that the equality given below should hold:

$$(5) \quad \int_0^{\infty} [(AC)Q(t)e^{-\rho_1 t}]e^{-r t} dt = \int_0^{\infty} [\theta(AC)Q(t)e^{-\rho_2 t}]e^{-r t} dt$$

Assuming that output increases exponentially at the rate α , the critical value of g_2 is given by:

$$(6) \quad g_2 = \theta g_1 - (1 - \theta)(r - \alpha)$$

If, for example, the values of θ , g_1 , and $r - \alpha$ were .90, .03, and .07 respectively, the critical value of g_2 would be .02. Were g_2 to fall below this value, an indicated economy of 10 per cent would not be sufficient to offset the cumulative productivity loss associated with the merger, to say nothing of the market power effects that the merger produces. If indeed the selected values of g_1 and $r - \alpha$ are at all representative, a predictable decrease in the rate of productivity advance by one-third or more would thus be sufficient to disallow a merger for which an efficiency advantage as large as 10 per cent could be expected.⁹

Consider now the managerial discretion argument. Here the direction of the effect is not so much a matter for dispute as is its quantitative significance. The argument is that market power provides a firm with the opportunity to pursue a variety of other-than-profit objectives. Although this is an "old" argument, its persistence at least suggests the possibility that it may not be without merit.¹⁰ Whether qualitatively there is anything to it turns essentially on the behavioral proposition that where competition in the product market presents no significant threat to survival, the resources of the firm are absorbed in part as corporate consumption activities by those members of the firm who are knowledgeable of discretionary opportunities, powerfully situated, and disposed to be assertive [29, 32]. Its quantitative significance rests on a judgment over whether the conspicuous evidence is sufficiently strong.¹¹

If indeed a predictable relaxation in the least-cost posture of a firm which has acquired market power through merger can be made, the

⁹ If the beneficial economies of scale are available only to the combining firms, while the negative progressiveness effects are felt throughout the industry, the above results underestimate the extent of economies necessary to produce indifference.

¹⁰ As Arthur Hadley observed in 1897, "The tendency of monopoly to retard the introduction of industrial improvement is . . . a more serious thing than its tendency to allow unfair rates. This aspect of the matter has hardly received proper attention. We have been so accustomed to think of competition as a regulator of prices that we have lost sight of its equally important function as a stimulus to efficiency. Wherever competition is absent, there is a disposition to rest content with old methods, not to say slack ones. In spite of notable exceptions this is clearly the rule" [12, p. 383].

¹¹ This presently is the weakest part of the argument. For a recent survey of the data, see [19].

estimated cost savings that appear in equation (4) should be adjusted accordingly. Economies which are available in theory but, by reason of market power, are not sustainable are inadmissible.

III. *Extensions*

Although the foregoing analysis has been concerned exclusively with horizontal mergers, the argument applies generally to problems in which market power-efficiency tradeoffs exist. Dissolution, vertical mergers, and conglomerate mergers can all be treated within this general framework.

A. *Dissolution*

The argument here is perfectly straightforward. It is simply not sufficient in a monopolization case for which dissolution is the indicated relief that (1) a persistent monopoly condition ($P_1 > AC_1$) exist, and (2) a reduction in price following dissolution ($P_2 < P_1$) be expected. It is necessary in addition that the gains realized by the price reduction be sufficient to offset any losses in economies that result. The relevant test is that shown in equation (3)—modified, as may be necessary, by the qualifications discussed in Section II above.

B. *Vertical Mergers*

It is important to note in dealing with vertical mergers that the conventional analysis of vertical integration, which takes a historical definition of an industry as given, often leads to incorrect results. The logical boundaries of a firm are not necessarily those which have been inherited but rather are defined by the condition that the firm be unable to arrange a transaction internally more cheaply than the market.¹² This is not something which is given once-for-all but depends both on technology and the extent of the market. Thus what may be regarded as “vertical integration” under a historical definition of an industry might, in many instances, more accurately be characterized as a reorganization into a more efficient configuration. For example, as technology evolves processes that are more fully automated or as demand for a commodity increases sufficiently to warrant continuous processing techniques, combinatorial economies may result by serially linking activities within a single firm that had previously been done in separate specialty firms.¹³ A transformation of this sort accomplished in part

¹² As Ronald Coase has pointed out, “a firm will tend to expand until the costs of organizing an extra transaction within the firm become equal to the costs of carrying out the same transaction by means of an exchange on the open market or the costs of organizing in another firm” [7, p. 341].

¹³ Stigler argues that increasing the extent of the market will often lead to dis-integration of manufacturing processes since now the market will be sufficient to support a specialized firm [25, pp. 188–90]. Although this may often occur, there is also the countervailing tendency

through vertical mergers is probably common in the production of commodities which shift from sequential job shop to continuous assembly line type operations.

That vertical integration can produce real economies is a result of the fact that the market does not perform its exchanges costlessly. Going to the market involves search costs, contracting costs, misinformation costs, delay costs, transfer costs, interface costs, etc.,¹⁴ and these must be balanced against the costs of organizing a transaction internally. Where the former exceed the latter, "vertical integration" is indicated. But of course this is vertical integration in only an apparent sense: in fact it represents a rationalization of the firm into an optimum economic unit.

The historical organization of an industry can ordinarily be presumed to reflect adequately basic efficiencies where significant market or technological developments have been lacking. And even where such recent changes have occurred, an efficiency defense is not automatic. Furthermore, if an efficiency defense can be supplied, any market power consequences that a vertical merger produces need also to be considered.¹⁵ Again the basic tradeoff calculation is that given by equation (3)—modified as necessary by the qualifications discussed in Section II.

C. Conglomerate Mergers

The principal ways in which conglomerate mergers can produce efficiencies have been given previously by M. A. Adelman [1, pp. 241–42] and Turner [27, pp. 1323–39, 1358–61]. The ways in which conglomerate mergers may produce market power are also discussed by Turner. All that remains, essentially, is to deal with the tradeoff question. Again the rules for estimating net benefits are substantially those given above.

IV. Conclusions

Most mergers produce neither significant price nor efficiency consequences, and where this is true the analysis of this paper has limited relevance. Where both occur, however, and if without merger the transition to an efficient industrial configuration is apt to be both painful and delayed, an efficiency defense deserves consideration. This does not of

to maintain or extend integration where coordination among the parts in the face of market uncertainties is critical—as it often is where assembly line operations are employed. See Coase [7, p. 337].

¹⁴ Coase discusses some of these [7, pp. 336–37]. (For an early example in which the costs going to the market were examined in a common law proceeding, see *Hadley v. Baxendale*.) In addition, if suppliers possess market power, going to the market may involve pecuniary expenses that could be avoided by integrating backward into supply activities.

¹⁵ Stigler identifies barriers to entry that take the form of increased capital and/or knowledge requirements as potential anticompetitive consequences of a vertical merger [25, p. 191].

course mean that the mere existence of economies is sufficient to justify a merger. But since a relatively large percentage increase in price is usually required to offset the benefits that result from a 5 to 10 per cent reduction in average costs, the existence of economies of this magnitude is sufficiently important to give the antitrust authorities pause before disallowing such a merger. There are, as indicated in Section II, a variety of qualifications that may upset this general conclusion in any particular case, but absent these and the result clearly holds.

It might be objected that the courts do not possess the expertise to make the types of judgments described. This is typically true. But that does not mean that an analysis of these effects should be not performed by the Antitrust Division or Federal Trade Commission before deciding to challenge a merger. The enforcement agencies can obtain, at reasonable cost, the necessary expertise to make these evaluations.¹⁶ Only after they are convinced that such economies as may exist are not sufficient to justify a merger should a case go forward. Although possibly this extends the responsibility of the enforcement agencies beyond those that are clearly intended, the alternative is scarcely acceptable. For if neither the courts nor the enforcement agencies are sensitive to these considerations, the system fails to meet a basic test of economic rationality. And without this the whole enforcement system lacks for defensible standards and becomes suspect.

Once economies are admitted as a defense, the tools for assessing these effects can be expected progressively to be refined. Since such refinements will permit both the courts and the enforcement agencies to make more precise evaluations, the threshold value under which an economies defense will be allowed can be reduced accordingly. Thus even if initially only a few mergers for which mixed effects are present are able to pass an appropriately qualified tradeoff test because of high threshold requirements, this proportion can be expected to increase as research results and analytical aids for evaluating scale economies accumulate. As an interim gain, solemn references to early oratory might finally be displaced in favor of analysis in the continuing dialogue on antitrust enforcement.

¹⁶ That the enforcement agencies are sensitive to scale economy considerations is evidenced by the recent Federal Trade Commission merger guidelines "Enforcement Policy with Respect to Mergers in the Food Distribution Industries," issued January 3, 1967. See especially pages 6-9.

Justice Brennan observed in the Philadelphia National Bank merger that "a merger the effect of which 'may be substantially to lessen competition' is not saved because, on some ultimate reckoning of social or economic debits and credits, it may be deemed beneficial. . . . [Such] is beyond the ordinary limits of judicial competence" [28, p. 371]. My point is that, at least with respect to efficiencies, such reckoning need not and indeed should not be beyond the competence of the antitrust agencies. It is here that the first critical decision of whether to file suit is made.

REFERENCES

1. M. A. ADELMAN, "The Antimerger Act, 1950-60," *Am. Econ. Rev., Proc.*, May 1961, 51, 236-44.
2. H. M. BLAKE AND W. K. JONES, "In Defense of Antitrust," *Columbia Law Rev.*, March 1965, 65, 377-400.
3. ———, "Toward a Three-Dimensional Antitrust Policy," *Columbia Law Rev.*, March 1965, 65, 422-66.
4. R. H. BORK AND W. S. BOWMAN, "The Crisis in Antitrust," *Columbia Law Rev.*, March 1965, 65, 363-76.
5. R. H. BORK, "Contrasts in Antitrust Theory: I," *Columbia Law Rev.* March 1965, 65, 401-16.
6. *Brown Shoe v. United States*, 370 U.S. 294 (1962).
7. R. H. COASE, "The Nature of the Firm," *Economica*, N.S., IV, 1937, 386-485. Reprinted in G. J. Stigler and K. E. Boulding, eds., *Readings in Price Theory*, Homewood, Ill., 1952, pp. 331-51.
8. W. S. COMANOR, "Research and Technical Change in the Pharmaceutical Industry," *Rev. Econ. and Stat.*, May 1965, 47, 182-90.
9. D. DEWEY, "Mergers and Cartels: Some Reservations about Policy," *Am. Econ. Rev., Proc.*, May 1961, 51, 255-62.
10. *Federal Trade Commission v. Procter & Gamble Co.*, 87 S. Ct. 1224 (1967).
11. C. E. FERGUSON, *A Macroeconomic Theory of Workable Competition*. Durham, N. C., 1964.
12. A. T. HADLEY, "The Good and Evil of Industrial Combination," *Atlantic Monthly*, March 1897, 79, 377-85.
13. D. HAMBURG, *R & D: Essays on the Economics of Research and Development*. New York 1966.
14. I. HOROWITZ, "Firm Size and Research Activity," *So. Econ. Jour.*, Jan. 1962, 28, 298-301.
15. R. E. JOHNSON, "Technical Progress and Innovation," *Oxford Econ. Papers*, July 1966, 18, 158-76.
16. C. KAYSSEN AND D. F. TURNER, *Antitrust Policy: An Economic and Legal Analysis*. Cambridge, Mass., 1959.
17. C. KAYSSEN, "The Present War on Bigness: I," in *The Impact of Antitrust on Economic Growth*, Fourth National Industrial Conference Board Conference on Antitrust in an Expanding Economy, New York 1965, pp. 31-38.
18. J. W. KENDRICK, *Productivity Trends in the United States*. Princeton, N. J., 1961.
19. H. LEIBENSTEIN, "Allocative Efficiency versus 'X-Efficiency'," *Am. Econ. Rev.*, June 1966, 56, 392-415.
20. E. MANSFIELD, "Size of Firms, Market Structure, and Innovation," *Jour. Pol. Econ.*, Dec. 1963, 71, 556-76.
21. ———, "Industrial Research and Development Expenditures: Determinants, Prospects, and Relation to Size of Firm and Inventive Output," *Jour. Pol. Econ.*, Aug. 1964, 72, 319-40.
22. J. W. MARKHAM, "Survey of the Evidence and Findings on Mergers,"

- in *Business Concentration and Price Policy*, Princeton, N. J., 1955, pp. 141–82.
23. F. M. SCHERER, “Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions,” *Am. Econ. Rev.*, Dec. 1965, 55, 1097–1125.
 24. G. J. STIGLER, “Monopoly and Oligopoly by Merger,” *Am. Econ. Rev., Proc.*, May 1950, 40, 23–24, reprinted in R. B. Heflebower and G. W. Stocking, eds., *Readings in Industrial Organization*, Homewood, Ill., 1958, pp. 69–80
 25. ———, “The Division of Labor is Limited by the Extent of the Market,” *Jour. Pol. Econ.*, June 1951, 59, 185–93.
 26. ———, “Industrial Organization and Economic Progress,” in L. D. White, ed., *The State of the Social Sciences*, Chicago 1956, pp. 269–82.
 27. D. F. TURNER, “Conglomerate Mergers and Section 7 of the Clayton Act,” *Harvard Law Review*, May 1965, 78, 1313–95.
 28. *United States v. Philadelphia Nat'l Bank*, 374 U.S. 312 (1963).
 29. O. E. WILLIAMSON, *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm*. Englewood Cliffs, N. J. 1964.
 30. ———, “Innovation and Market Structure,” *Jour. Pol. Econ.*, Feb. 1965, 73, 67–73.
 31. ———, “Hierarchical Control and Optimum Firm Size,” *Jour. Pol. Econ.*, April 1967, 75, 123–38.
 32. ———, “A Dynamic-Stochastic Theory of Managerial Behavior,” in A. Phillips and O. E. Williamson, eds., *Prices: Issues in Theory, Practice and Public Policy*, Philadelphia (forthcoming).