THE GAINS AND LOSSES FROM INDUSTRIAL CONCENTRATION*

SAM PELTZMAN
University of Chicago and
National Bureau of Economic Research

No field in the industrial organization literature has been as well plowed as the relationship between concentration and profitability. Weiss,[1] in his latest review of this literature, discusses over forty such studies since 1951, and this is not a complete census. The reader has by this late date earned the right to demand strong justification for a new entrant. Mine is simply that, despite its bulk, the literature fails to inform us how to interpret its main findings.

Those findings are well known: with few exceptions, market concentration and industry profitability are positively correlated. Since the correlation is usually weak, the literature has tended to become a search for more complex and/or accurately specified relationships. I eschew that approach to focus on a more basic question: if concentration and profitability are indeed related, what market process produces the relationship? The traditional answer has been that high concentration facilitates collusion and hence super marginal-cost pricing, for which some profitability measure is a proxy. Unfortunately, this answer does not logically follow from the usual evidence, so its acceptance by economists and practitioners of antitrust policy is little more than an act of faith.

Any profitability measure implies a corresponding difference between price and average cost. As a matter of simple arithmetic a causal relationship running from concentration to profitability can operate either through an effect on price (the usual interpretation), or on average cost, or, of course,

* This study was supported by a grant from the National Science Foundation to the National Bureau of Economic Research for research in law and economics. The article is not an official Bureau publication since it has not yet undergone the full critical review accorded Bureau publications, including review by the Bureau's Board of Directors. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

I am indebted to Gerald Dwyer and Stephen Entin for their research assistance, and to Gary Becker, Yale Brozen, Harold Demsetz, George Stigler, and Lester Telser for helpful comments on an earlier draft.

[1] Leonard W. Weiss, The Concentration-Profits Relationship and Antitrust, in Industrial Concentration: The New Learning 184 (Harvey J. Goldschmid, H. Michael Mann, & J. Fred Weston eds. 1974).

both. Acceptance of the pure collusion interpretation of the evidence has so far hinged on the largely untested assumption that concentration has little or no connection with the implied average cost measure. However, indirect tests of the assumption, most notably by Demsetz,[2] imply that it may not be useful, and his finding is one motive for the present work.

In essence, this paper will try to decompose the concentration-profits relationship into separate concentration-price and concentration-cost relationships. By doing this, I hope to shed light on some of the allocative and distributive issues that, I suspect, give the subject its intrinsic interest, but which have not yet been confronted empirically: Does high concentration save or waste resources? Does it lead to higher prices? Who gains and loses from a social policy hostile to high concentration? Since the unique aspect of the paper is its focus on a concentration-cost relationship, most of the analytical effort is spent here. I review the theory underlying such a relationship and develop a model designed to estimate its importance. Subsequently, I try to estimate how much of the usual profit-concentration relationship is due to cost effects and how much to price effects. The main conclusion is that, while price effects are not absent, the cost effects so dominate as to cast doubt on the efficacy of any general legal rule hostile to industrial concentration.

I. Market Structure, Costs, and Prices

The possibility that market structure and costs are related has of course long been recognized. An unconcentrated industry in which a technological advance produces "natural monopoly" or oligopoly cost conditions will become both more concentrated and more efficient over time. The process by which the old technology is rendered economically obsolete will also entail a fall (or at least no increase) in price. The price decline need not be great enough to eliminate producer rents, either because the associated increase in concentration permits collusion or because the new technology diffuses slowly enough to leave room in the market for a fringe of old-technology firms. Whichever force operates, there is a clear dilemma for antitrust policy: attempts to thwart increased concentration will merely waste resources without benefiting consumers.[3]

The concentration-profitability literature has so far given little weight to "natural oligopoly" interpretations of the data. Two reasons for this neglect are important enough to affect the structure of this study:

[2] Harold Demsetz, Industry Structure, Market Rivalry, and Public Policy, 16 J. Law & Econ. 1 (1973).

[3] See Oliver E. Williamson, Economics as an Antitrust Defense: The Welfare Tradeoffs, 58 Am. Econ. Rev. 18 (1968), for a discussion of this kind of problem.

1. A "natural oligopoly" interpretation is asymmetric. Some technological progress can be scale-reducing. If this sort of change diffuses slowly enough, or cannot be implemented as efficiently by all firms, the large firms will be the marginal firms and small firms will earn rents. Thus, we should observe unusually low (or declining) as well as high concentration associated with unusually high rates of return. By and large, such a U-shaped concentration-profits relationship has not been found. One inference could be that rents to size-specific technical change are unimportant empirically. However, it would be just as easy to conclude that large size-related economies are simply more prevalent. Or, to the extent that such economies are specific to a few organizations, the large size-related economies will dominate in the usual data: Three clever firms producing one thousand cars per week at half of General Motors' unit cost will have a trivial impact on automobile prices or the measured efficiency of the automobile industry. I shall attempt to disentangle the possibilities empirically rather than intuitively. My model permits any kind of change in market structure to reflect size-related technological change and leaves the importance of the relationship to be determined empirically.

2. The empirical literature on economies of scale seems to conflict with a "natural oligopoly" interpretation. A common finding of this literature is that of long-run constant costs at the firm level over a wide range of output, wide enough to encompass many existing-firm sizes and a large fraction of industry output. To illustrate, Bain[4] and Stigler[5] both find that the minimum-efficient-size steel firm produces something like 2 per cent of national output. Smaller firms had less than 20 per cent of national capacity in 1951.[6] While some efficiency might be gained by a decline of this inefficient fringe, any substantial change in market structure would likely involve a reallocation of output among efficient-sized firms.

However, that sort of inference might only mean that the economies-of-scale paradigm is not very useful rather than that market structure and efficiency are unrelated. Indeed, the imprecision of scale-economy rationalizations of market structure can itself be of help in formulating such a relationship. If there is no unique efficient firm size, but only a wide band encompassing many existing firms, then any of these firms can grow to the upper end of the band before it incurs size diseconomies. This kind of expansion by a firm becomes likely, instead of merely possible, if the firm discovers a lower cost technology which is not immediately available to others. The potential profits from the cost advantage will then attract capital

[4] Joe S. Bain, Economies of Scale, Concentration, and the Condition of Entry in Twenty Manufacturing Industries, 44 Am. Econ. Rev. 15, 23 tab. II (1954).

[5] George J. Stigler, The Economies of Scale, 1 J. Law & Econ. 54, 58 (1958).

[6] *Id.*

and permit the firm to grow at least to the upper end of the band, though, to get there, the firm may also have to reduce prices. Thus, the fortunate firm, or firms, become big instead of merely "average" (concentration increases), and resource costs and prices are lowered by this unusual growth.

This argument, which can be found in Demsetz[7] and McGee,[8] is sketched in Figure I.

Let P_1L be the long-run supply curve of a competitive industry with demand D. Let P_1RM be the long-run firm supply curve, so, following the scale-economies literature, firm size is indeterminate in the range OB. We observe an actual (or average) firm output of OA. The industry is in equilib-

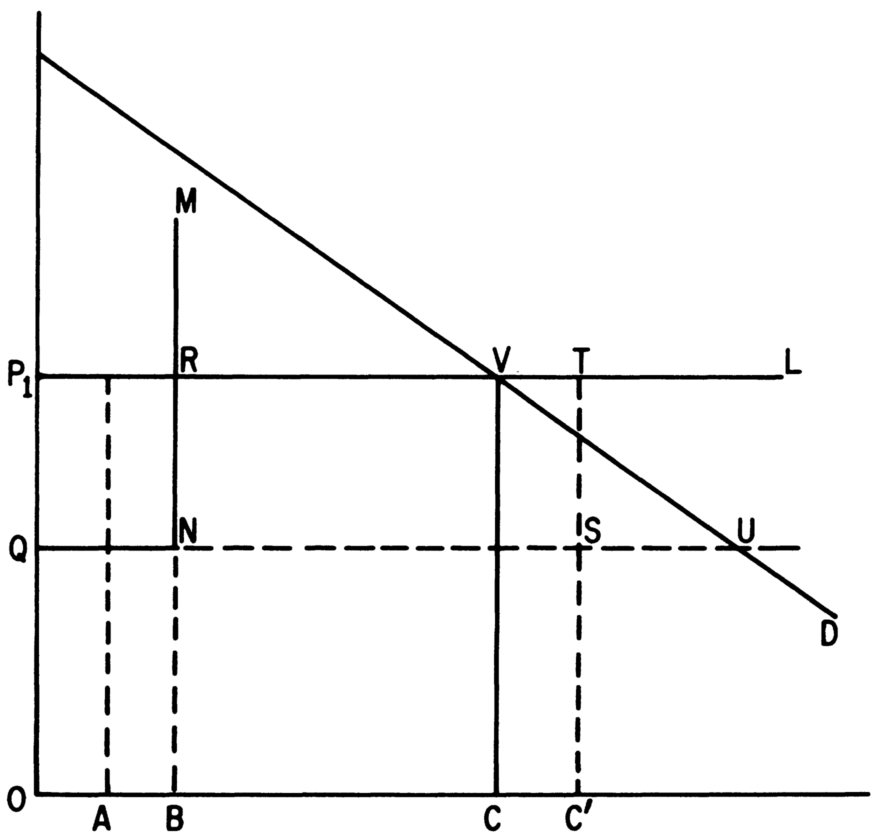

FIGURE I

[7] Harold Demsetz, *supra* note 2.

[8] John McGee, In Defense of Industrial Concentration 41-52, 75-79 (1971).

rium with a price of OP_1, zero rents, and a four-firm concentration ratio of, say, $4(OA/OC)$. Now, one representative firm discovers a way to lower marginal costs to $QNRM$. In this perfectly-competitive-industry example, it cuts price trivially and expands from OA to OB. There are now positive rents for this firm (and for the industry aggregate) equal to the resource cost saving P_1QNR, and concentration has increased by (AB/OC). Statistically, the increased concentration will be correlated with the increased rents, but there is a more substantive connection: the increased concentration is the mechanism by which part of the resource cost saving $(AB \times NR)$ is realized.

If enough firms make the same cost-reducing discovery, consumers will share the resource cost saving. For example, if (QS/QN) firms make the discovery, industry supply becomes $QSTL$. Price falls, but there are still positive rents and concentration increases (so long as demand is sufficiently inelastic to keep $CC'/OC < AB/OA$). If the discovery is sufficiently general, of course, the rents will disappear. However, with an industry supply like QU, the correlation between efficiency and concentration can still hold. As new firms catch on, each grows toward maximal size, and eventually this maximal size becomes the typical firm size. Of course, with a sufficiently elastic demand, this firm growth need not imply increased concentration. But, given the very large maximal sizes usually encountered in the scale economies literature, increased concentration would be the expected outcome.

While the constant-returns-to-scale case is important empirically and useful didactically, some of the ambiguities of the more conventional (by textbook standards) diminishing-returns case merit elaboration. This requires distinguishing between increasing costs at the firm and industry level. If the industry is constant cost, the preceding analysis needs no qualification: in the long run a uniquely efficient firm will expand as much as others contract, and it will collect as rent all of its differential productivity. The more interesting case analytically is increasing costs for both the firm and industry. This implies rents for some factors, and hence an equilibrium excess of marginal cost over average cost (net of rents) for intra-marginal firms. This means, in turn, that any efficiency-induced output expansion has two offsetting effects on a firm's average cost: the efficiency lowers the level of its average cost curve, but the output expansion causes a move up along the curve. Under some supply and demand conditions, the latter effect can dominate, so that average cost increases.[9] To see just what these conditions are, write the total industry supply (S^T) curve

$$S^T(p) = \overline{S}(p) + S^A(p, \lambda), \quad (1)$$

[9] This does not, of course, mean that efficiency, appropriately defined, deteriorates, but only that average cost is an inappropriate efficiency proxy in this case.

where

p = price,
$\bar{S}$ = the aggregate of all firms but one (A), and
λ = shift parameter.

Then suppose A becomes more efficient, so its supply curve shifts rightward (that is, λ increases and $S_\lambda{}^A > 0$, the subscript denoting a partial derivative). Assume for now that the resulting increase in A's output also increases any measure of industry concentration. Let us then see what is required for average cost to decline while profitability increases.

Since costs and rents (R) must exhaust industry receipts, industry unit costs (C) can be written

$$C = p - \frac{R^T}{Q} = p - \left(\frac{\bar{R} + R^A}{Q}\right), \tag{2}$$

where

Q = industry quantity sold.

The effect of A's efficiency (the increase in λ) on C can be shown to be

$$\frac{dC}{d\lambda} = -\frac{R_\lambda{}^A}{Q} + S_\lambda{}^A \cdot \frac{R^T}{Q^2}\left[\frac{1}{1 + \dfrac{E_S}{E_D}}\right], \tag{3}$$

where

$E_{D,S}$ = absolute value of the industry demand and supply elasticities.[10]

The first term on the right-hand side of (3) is the pure "efficiency" effect, and were this a constant-cost industry ($R^T = 0, E_S = \infty$) that would tell the whole story: industry unit costs and unit rents would change by equal and opposite amounts. The second term, the "output" effect, can offset the first if (a) the supply shift $S_\lambda{}^A$ is large enough, and if (b) industry supply is sufficiently inelastic (other firms save few resources by cutting their output) or demand

[10] This follows from

$$\frac{dC}{d\lambda} = \frac{-R_\lambda{}^A}{Q} + S_\lambda{}^A \frac{R^T}{Q^2}\left[1 + \frac{S_p{}^T}{S_\lambda{}^A}\frac{dp}{d\lambda}\right],$$

when we impose a supply-demand equilibrium condition on $dp/d\lambda$. This condition comes down to

$$\frac{dp}{d\lambda} = \frac{-S_\lambda{}^A \cdot p/Q}{E_D + E_S}.$$

sufficiently elastic (so output expands enough to make the diminishing returns important).[11]

There is a further ambiguity in the relationship between unit cost and unit rent (that is, profitability). Since

$$\frac{d(R^T/Q)}{d\lambda} = \frac{dp}{d\lambda} - \frac{dC}{d\lambda}, \quad (4)$$

a negative relationship between the two requires that price fall by less than unit cost. Whether or not this holds again depends on supply and demand elasticities, but here it turns out that more elastic demand, on balance,[12] favors an increase in rents.

All the preceding results—and ambiguities—would hold if A were initially a small firm, except, of course, that concentration could decrease. More generally, where differences in firms' costs underlie changes in their market shares, one ought to expect any change in market structure to promote efficiency. However, with constant returns-to-scale, there is a clear bias toward increased concentration as the main source of lower costs. So long as a firm's superior technology simply lowers the level of its horizontal marginal cost curve, the firm will expand to maximal efficient size.

Demsetz[13] tests for this bias by comparing rates of return of large and small firms by industry. He finds no difference in these rates of return in low-concentration industries, so small firms do not seem to have a cost advantage there. However, in highly concentrated markets, the large firms have the higher rates of return, so Demsetz concludes that they have lower costs, and, by inference, that this cost advantage is the source of their large size.

Even if one accepts this inference, the results can be consistent with either competitive or oligopolistic pricing. For example, an industry whose supply

[11] For two familiar cases—a constant shift in S^A, and a constant percentage shift—$dC/d\lambda$ is always < 0. In the former, $S_\lambda{}^A = 1$ and $R_\lambda{}^A = P$, so (3) is

$$\frac{dC}{d\lambda} = \frac{-C}{Q} - \left(\frac{E_S}{E_S + E_D}\right)\frac{R^T}{Q^2} < 0.$$

In the latter case,

$$\frac{S_\lambda{}^A}{S^A} = \frac{R_\lambda{}^A}{R^A} = \frac{1}{\lambda},$$

and

$$\frac{dC}{d\lambda} = \frac{-S^A}{\lambda Q}\left[\frac{R^A}{S^A} - \frac{R^T}{S^T} + \frac{R^T}{S^T}\left(\frac{E_S}{E_S + E_D}\right)\right].$$

If A is initially an "average" firm, so that its unit rents equal R^T/S^T, this too is negative.

[12] It attenuates the price decrease, but also retards the fall in unit cost.

[13] Harold Demsetz, *supra* note 2.

schedule passes through V in Figure I, while some firms have marginal costs like QNM, could be characterized as competitive, while the cost difference generates both high concentration and producer rents. However, in another industry the aggregate marginal cost of the superior firms could pass through U, while a collusive agreement among them keeps the price at P_1. In that case, a smaller firm could survive, and the same disparity among rates of return would be observed. Indeed, once one allows for both the traditional connection between concentration and collusion and for differences among firm costs, still less benign solutions become consistent with Demsetz's results. For example, let the long-run supply from less efficient firms be increasing. Then let the process described by Demsetz and McGee generate increased concentration which incidentally decreases the cost of collusion. The large firms may now find it in their interest to set a price above the previous competitive price P_1, even though they must yield some market share to do this. The marginal firm in this case would be both "small" and earning a "competitive" rate of return, and this result would also be consistent with the Demsetz data.[14]

My intent here is not to catalog possibilities, but to indicate that there is insufficient evidence for a conclusion that the effects of concentration are either wholly beneficial or costly.

If an eclectic integration of the prevailing theory cannot therefore be ruled out, it becomes a useful framework with which to empirically evaluate the main costs and benefits. In the remainder of the paper I try to do just this. Specifically, I ask the following questions:

1. How important is the relationship between market structure and costs?
2. How much of any resulting change in costs is translated into price changes?
3. How important is the relationship between market structure and collusion, and how much are prices thereby affected?

While answers to these questions would be useful in clarifying an important academic literature, they also have important policy implications. The merits of an anticoncentration policy can hinge on whether the collusion effects of concentration outweigh the cost effects.

The discussion so far can be summarized symbolically as follows. The prevailing view on the effects of market structure on the price of any good would be

$$p = p(X, C, MS), \tag{5}$$

[14] See Harold Demsetz, Two Systems of Belief About Monopoly, in Industrial Concentration: The New Learning 164, 178-79 (Harvey J. Goldschmid, H. Michael Mann, J. Fred Winston eds. 1974).

where

p = price of the good
X = set or index of demand shifters
C = index of supply shifters, which can be subsumed under the rubric of "costs"
MS = some measure of market structure, that is, the number and size distribution of firms, which is a proxy for the cost of collusion.

Putting aside important qualifications about the effects of varying demand elasticities, the measurement and relevance of the various components of C, and so forth, low-cost collusion is assumed to lead systematically to an increase in the ratio of price to either marginal or average cost, so that

$$\frac{\partial p}{\partial MS} > 0 \quad \text{(where higher } MS \text{ implies lower cost collusion).}$$

The underlying theory permits (5) to be applied across isolated markets for a homogeneous good, or to a particular market over time. For my purposes, it will be useful to treat the variables in (5) as (logarithmic) time derivatives rather than levels. The essential eclecticism is then introduced by a companion function for costs:

$$C = C(Y, MS), \tag{6}$$

where

Y = set of exogenous determinants of the cost index, for example, factor prices.

The MS term in (6) is meant to summarize what is really a two-way (and nonmonotonic) relationship: changes in MS (in either direction) both cause costs to decline and are induced by changes in costs. Once a relationship like (6) is admitted, the total effects of MS on p become more complex. Specifically,

$$\frac{dp}{dMS} = \frac{\partial p(\cdot)}{\partial MS} + \frac{\partial p(\cdot)}{\partial C}\frac{dC}{dMS}. \tag{7}$$

In the particular case of increased concentration $\frac{dC}{dMS} < 0$, the prevailing view that the first term on the right-hand side is positive can be correct even though the total derivative is negative. Moreover, if some of the effect of a cost change shows up in producer rents ($\partial p/\partial C < 1$), the prevailing view can be wrong, but measured ***profitability*** will be positively related to MS.

The next section specifies (6), which is subsequently estimated, and the results are used to estimate (5) and (7). These, in turn, provide the answers to the questions I have posed.

II. The Relationship Between Costs and Market Structure

For simplicity, classify the firms in any industry into two groups: type L firms are or will become the largest in the industry; type M firms are all the others. The industry's cost per unit of output (C) at any point in time is then

$$C = sL + (1 - s)M, \tag{8}$$

where

L, M = the unit cost of a firm in each group, which is assumed to be the same for all firms in the group.

s = type L's share of industry output (for example, a four-firm concentration ratio, if the type is defined as the largest four).

Since we will be interested in percentage changes over time in C ($\dot{C}$), note that

$$\dot{C} = \frac{1}{C}\ [s\,\dot{L}\,L + (1 - s)\,\dot{M}\,M] + \left[\frac{L - M}{C}\right]\frac{ds}{dt}. \tag{9}$$

The last term on the right-hand side of (9) captures the effect of market share changes on efficiency. It says that if, for example, the type L are more efficient ($L < M$) and their market share increases, the resource cost of industry output is thereby reduced. Now, express each type's unit cost change as

$$\dot{L} = r + \ell \tag{10}$$

$$\dot{M} = r + m, \tag{11}$$

where r is the sum of all forces changing costs which are common to the two types such as secular productivity growth and factor price changes, and ℓ, m summarize forces peculiar to each group. This allows (9) to be expressed as

$$\dot{C} = r + \ell + (1 - s)\frac{M}{C}\,\delta + \left[\frac{L - M}{C}\right]\frac{ds}{dt}, \tag{12}$$

where

$$\delta = m - \ell.$$

This says that the level, as well as the change, of market shares matters. For example, if type L becomes relatively more efficient over time ($\ell < 0$, $\delta > 0$), then industry costs will grow more slowly the greater type L's market share.[15]

Further rearrangement of the terms leads to the following alternative expressions, both of which will be useful subsequently:

$$\dot{C} = r + \ell + \left[\frac{\frac{ds}{dt} + \frac{1}{D}(1 - s)\delta}{s + \frac{1}{D}} \right], \text{ or} \tag{13}$$

$$\dot{C} = r + \ell + \left[\frac{-\frac{ds}{dt} + \left(\frac{D' + 1}{D'}\right)(1 - s)\delta}{(1 - s) + \frac{1}{D'}} \right] \tag{13'}$$

where

$D = L/M - 1$
$D' = M/L - 1.$

The meaning of these equations can be grasped by focusing on (13) when $D > 0$. In that case, small firms have a cost advantage. If $ds/dt < 0$, industry efficiency will improve, because the more efficient firms gain market share ($\dot{C}$ declines). The degree of improvement is greater, the larger the small firm advantage (that is, the smaller the $1/D$ term in the denominator—assume $\delta = 0$ for the moment), and the larger the share of output due to the more efficient firms (that is to say, the smaller s is). Equation (13′) applies symmetrically where the large firms have the cost advantage.

It is now necessary to specify the link between market structure changes (ds/dt) and cost differences (D or D'). I assume the following simple relationships.

$$G_L - G_M = \alpha_1 D', \text{ and} \tag{14}$$

$$G_M - G_L = \alpha_2 D, \tag{14'}$$

where

α_1, $\alpha_2 > 0$, and are constants
G_i = output growth rate of type i firms.

Each equation is applicable for D' or $D > 0$. These say that the type of firm with a cost advantage grows faster over time in proportion to its cost advantage. Adjustment costs affect the size of the proportionality constants, α_1 and

[15] There is another sense in which the level of market share matters, which turns out to be empirically important: the bracketed term gets larger, the larger the share of the most efficient type firm, because C gets smaller. Thus, for a given increase in that type's share, the cost effect is larger, the larger its market share.

α_2, and these are allowed to differ—mergers may have different costs than divestitures. For example, notice that the adjustment process is only indirectly affected by differences in the rate of change of firm costs (δ), which will affect D or D', but, for simplicity, I make no allowance for forecasting. That is, if δ is expected to be maintained, the future values of D will change, and this could affect the response to the current D. The process described by (14) and (14′) is, however, completely myopic in this regard.

Since

$$\frac{ds}{dt} = s(1 - s)[G_L - G_M], \tag{15}$$

the adjustment process can be expressed

$$Z = s(1 - s)\alpha_1 D', \text{ if } ds/dt > 0 \tag{16}$$

$$Z = s(1 - s)\alpha_2 D, \text{ if } ds/dt < 0 \tag{16'}$$

where

$Z = |ds/dt|$.[16]

The next step is to introduce the adjustment process into (13) and (13′) by solving (16) and (16′) for D and D' and replacing these in the former equations. This yields

$$\dot{C} = r + \ell + \left\{\frac{-Z^2/(1 - s) + [Z + \alpha_1 s(1 - s)]\delta}{Z + \alpha_1 s}\right\} \text{ for } ds/dt > 0 \tag{17}$$

and

$$\dot{C} = r + \ell + \left\{\frac{-Z^2/s + \alpha_2(1 - s)^2\delta}{Z + \alpha_2(1 - s)}\right\} \text{ for } ds/dt < 0. \tag{17'}$$

Next, I introduce an assumed relationship between δ—the differential between large and small firm cost changes—and market growth. The motivation for so doing is the empirical relationship between market structure and growth. For example, Nelson[17] reports a significant negative correlation between the 1935-54 change in industry concentration ratios and growth in value added. This is also present in the 1947-67 sample I shall use subsequently.[18] In a model which purports to link market structure to differential costs, such an empirical regularity must logically be cost related. Therefore,

[16] Note that there is an implicit conjecture here that Z is correlated with $s(1 - s)$, because the αD terms affect only relative firm growth rates directly. In fact, in the sample we shall use subsequently, the correlation of the 1947-67 Z with the 1947 $s(1 - s)$ is a significant +.22.

[17] Ralph L. Nelson, Concentration in the Manufacturing Industries of the United States 51-56 (1963).

[18] The simple correlation is −.23.

assume that rapid market growth reduces the small-firm cost change relative to that of large firms, perhaps because it permits the smallest firms to adjust cheaply to minimum efficient size more quickly or creates favorable "learning curve" effects for them.[19] To incorporate this relationship most simply, let

$$\ell = ag \tag{18}$$

$$m = bg, \tag{19}$$

so

$$\delta = (b - a)g, \tag{20}$$

where

a, b = constants
g = growth in demand for output.

The sign of a and b need not be specified, but if growth is advantageous for small firms, $(b - a)$ should be negative.[20] If that is true, then (17) and (17′) imply that, given large-firm cost changes (ℓ), growth reduces industry costs. This occurs generally because growth promotes lower small firm costs. In the specific case of increasing concentration, there is another force at work, captured by the $z\delta$ term in (17). Increased concentration in the face of rapid growth (declining small firm costs) would imply unusually low-cost large firms, and an unusual decline in industry costs as they increase their market share.

To allow for empirical implementation of (17) and (17′) by conventional techniques, I shall use three-term Taylor expansions about $z = 0$, which capture most of the important nonlinearities in the model.[21] They can be written as the single equation

$$\dot{C} = r + ags + bg(1 - s) \tag{21}$$

$$+ (b - a)g\left[\frac{Z^2M_i}{\alpha_i^2K_i} - \frac{ZM_i}{\alpha_i}\right] - \frac{Z^2}{\alpha_i s(1 - s)} + \textit{Remainder Term},$$

[19] That is, the small firm in a growing industry could accumulate a given volume of output (and experience) as quickly as a larger firm in a declining industry.

[20] There is an ambiguity here in that the empirical regularity we are seeking to incorporate applies to small firms generally, while the m variable applies to firms which end up small whatever their original size. I treat the effect of this ambiguity later, but it is essentially forced by the available data.

[21] The important nonlinearities arise from the dual role of Z as an indicator and implementer of cost changes. Ignoring the growth interaction implied by (20), this feedback leads to $\dot{C}_{zz} < 0$. The growth interaction complicates this: C_{zz} becomes uncertain for $ds > 0$, because higher Z implies loss of the growth benefits on small firm costs. These benefits depend on the level of small firm costs, and this makes $\dot{C}_{zg} \lesseqgtr 0$ for $ds \gtreqless 0$. For example, if $ds > 0$, $dZ > 0$ implies an increase in the ratio M/L (or D'). Since the beneficial effect of g on $\dot{C}$ is amplified at higher levels of M and diluted by the positive ds, this implies further that $\dot{C}_{zg} < 0$. These results for $\dot{C}_{zz}$ and $\dot{C}_{zg}$ hold both for (17) and (17′) and their Taylor expansion (21).

where, if $ds/dt > 0$, then $i = 1$; and if $ds/dt < 0$, then $i = 2$; and

$$M_1 = -1$$
$$M_2 = +1$$
$$K_1 = s$$
$$K_2 = (1 - s).$$

Finally, to obtain an empirical counterpart to r, the cost changes common to all firms, assume for simplicity that each industry's total output Q can be described by a Cobb-Douglas production function

$$Q = F(t, \phi) \prod_i I^{a_i}, \quad \Sigma a_i = 1, \tag{22}$$

where

I = input amount
a_i = constants, and
F = input productivity shift function, depending on secular productivity growth (t) and other forces (ϕ).

Then assume, again for simplicity, that if some firms become more efficient, this efficiency is not specific to any one input. At the industry-aggregate level, efficiency and its subsequent spread through any change in market structure can then be included among the other forces (ϕ) in (22). As a further simplification, assume that the competitive profit-maximizing conditions for a firm can be used in approximating industry unit cost changes, so

$$\dot{C} = -\dot{F} + \sum_i a_i \dot{p}_i, \tag{23}$$

where

p_i = input prices, and now
a_i = (constant) input cost shares.

Then assume a constant percentage secular growth (γ_i) of industry input productivity, so that $\dot{F}$ can be decomposed

$$\dot{F} = \dot{F}(\phi) + \gamma_i. \tag{24}$$

We can now relate (23) and (24) to (21): r, summarizing the productivity and input price trends common to all firms, is simply $-\gamma_i + \Sigma a_i \dot{p}_i$, and everything else on the right-hand side (r.h.s.) of (21) equals $-\dot{F}(\phi)$.

In the next section, I estimate the scheme in (21), (23), and (24).

III. Empirical Results: Costs and Market Structure

The basic cost relationships just developed are estimated here for a subset of four-digit Standard Industrial Classification (SIC) manufacturing industries. The subset consists of those industries for which a meaningful change in unit cost (essentially industry expenditures deflated by an output index) can be computed for 1947-67. Since this is the longest period practicable, we face the risk that the basic cost relationships will be obscured by technological change. However, a greater risk would be for short-period changes in market structure to hide the more fundamental relationship we seek to measure, a relationship which, after all, purports to rationalize a nontrivial part of the existing variety of market structure. Even twenty years is not a long enough period to permit as much change in this variety as we might like.

Many current four-digit industries had to be dropped because of changes in classification between 1947 and 1967,[22] or because reliable output indexes were unavailable from 1947.[23] To limit potential measurement error, "industries" with low or changing coverage or specialization ratios were also deleted.[24] Finally, two industries—drugs and ballpoint pens—which experienced profound technological change in this period were omitted. They met all the formal tests for inclusion, but their measured productivity growth was so atypical as to obscure some of the results and raise questions about the comparability of their earlier and later outputs. This left a sample of 165 industries.[25]

Cost data available at the four-digit level include only labor and raw material expenditures by Census establishments. Expenditures typically incurred by administrative offices rather than plants (for example, advertising) and capital costs are excluded. A modest adjustment for the former deficiency is discussed later. I assume that capital costs are proportional to an industry's gross book value of plant and equipment, for which data are available or can be estimated.[26] The factor of proportionality is derived from

[22] The main changes in classification occur in 1958. In some cases, the post-1958 SIC combines pre-1958 industries. These were retained in the sample only if pre-1958 concentration ratios could be reliably inferred for the post-1958 industry. In practice, this means that the merged industries have firms so small that none could conceivably be among the four largest prior to 1958, because the total output of the merged industry's largest four is much less than the average output of the merging industry's largest four.

[23] U.S. Bureau of the Census, Census of Manufacturers: 1954 Indexes of Production (vol. 4, 1958), singles out industries with unreliable output indexes. I deleted such industries where the 1954 index (1947 = 100) was more than 50 per cent different from that of its two-digit class.

[24] Low specialization means that the plants in the SIC industry produce substantial amounts for other markets, thus calling into question the relevance of the market definition. Low coverage means that plants elsewhere produce a large part of this industry's output, thus calling into question the meaning of narrowly based market structure measures. I deleted industries where the 1947 or 1967 product of the specialization and coverage ratio was under .6 or where either changed by over .1 between these years.

[25] A list of these industries is available on request. They account for about half of U.S. manufacturing sales.

[26] Gross book value (GBV) data are available from U.S. Bureau of the Census, Annual Survey of Manufactures: 1954 (1956), but the coverage expands over time. Where gaps had to

Berndt and Christensen's[27] estimate of annual capital cost shares for all manufacturing. I simply choose the proportionality factor that yields for my sample the Berndt-Christensen cost shares for 1947 and 1967.[28]

To calculate the cost-share-weighted input price changes on the r.h.s. of (23), distinguish three inputs: labor, raw materials, and capital. The input cost shares are set at their 1947-67 means, and the input price changes are calculated from the following data:

1. Labor—payroll per employee. (Compensation per man hour is not consistently available for 1947.)

2. Raw materials. Purchase price indexes are not available by industry. However, the major change here is a decline in the relative price of agricultural to manufactured materials between 1947 and 1967. Therefore, separate raw material price indexes for each industry are estimated using 1963 Input-Output data[29] on the direct and indirect purchases of agricultural products per dollar of purchases for each SIC industry as weights.[30]

3. Capital. Since industry-specific data are unavailable, it is assumed that the change in the rental price of capital is the same for each industry.

Finding empirical counterparts to the crucial market structure variables on the r.h.s. of (21) poses a major difficulty. The relevant theory applies to firms classified by terminal relative size regardless of initial size. Unfortunately, corresponding published data do not exist. We have only conventional concentration ratios, which do not reveal where today's largest firms

---

be filled, a log linear form of $GBV_{t-1} = f(GBV_t, X)$ was first estimated, where X is a vector which includes capital expenditures between $t - 1$ and t (so f is an inverse form of the relationship running from GBV_{t-1} to GBV_t), industry sales (which is used as a size deflator), and Internal Revenue Service data on assets for the three-digit industry superclass. The missing GBV_{t-1} are then filled in using the regression weights and known values of the independent variables. The entire set of GBV_{1947} had to be estimated in this fashion, since pre-1954 GBV data are unavailable. IRS data are from U.S. Internal Revenue Service, Source Book of Statistics of Income, various years (unpublished).

[27] Ernst R. Berndt & Laurits Christensen, The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-68, 1 J. Econometrics 81 (1973).

[28] None of the results reported subsequently are substantially different if capital costs are excluded. The correlation between the capital cost inclusive and exclusive cost changes is +.98.

[29] U.S. Dep't Comm., Office of Bus. Econ., 3 Input-Output Structure of the U.S. Economy: 1963, Total Requirements for Detailed Industries (Supp. to Survey of Current Bus., 1969).

[30] The industry price indexes are:

$$Index_i = PM + AG_i(PA - PM),$$

PM = Manufactured products wholesale price index,

PA = Farm products wholesale price index,

AG_i = 1963 direct and indirect purchases of agricultural products per dollar of SIC industry i's total purchases of materials.

For 1957 = 100: $PM_{47} = .759$, $PA_{47} = 1.099$, $PM_{67} = 1.072$, $PA_{67} = 1.004$.

came from. Therefore, I must make the strong assumption that the large firms in 1967 are the same as (or at least dominated by) those in 1947. In keeping with the simple size dichotomy embodied in (21) and to make the results of this study comparable with the bulk of the literature, Z is then defined simply as the absolute 1947-67 change in an industry's four-firm concentration ratio (and s is the average of the two concentration ratios). The reader should, however, be aware that the model predicts cost effects for turnover of large firms as well as for net changes in their relative size.[31] Finally, g is defined as the log of 1967 divided by 1947 sales net of the log change in the manufactured goods Wholesale Price Index. This variable's accuracy as a demand change proxy depends on an implicit assumption of unit elastic demand for each industry's product.

I want to test Demsetz and McGee's conjecture[32] that the process described by (21) is more important for increases in concentration than decreases. The formal model does not allow for this except through differences in α. Therefore, I simply estimate separately the effects of increases and decreases in concentration, and see whether these effects are indeed different. The basic regression is given in Table 1 for the values of α_i which minimize residual variance (see note to Table 1).

The results lend some support to the basic model and uncover a few puzzles. For example, the model predicts a coefficient of $+1$ for the first two-factor price variables, and indeed both are insignificantly different from 1. However, the results show neither the expected difference between large and small firm cost changes (the coefficient of $G2$ should be less than that of $G1$), nor the relationship of the $CR1$ coefficients to those of $G2$ and $G1$ (the former should equal the difference of the latter two). And the coefficients of the $CR2$ terms deviate from their theoretical value of -1. In spite of these drawbacks, a market share effect on costs does show through, and as I demonstrate later, it is empirically important. Moreover, this effect is asymmetric in just the way implied by Demsetz and McGee. Decreases in concentration do reduce costs, as predicted, and this does not depend on the aberrant negative α.[33] However, the effect is statistically insignificant and only a fraction of the effect of similar increases in concentration. Essentially, only increases in concentration matter very much and their effects, more-

[31] In addition, if my implicit assumption about the importance of adjustment costs is wrong, effects of intraperiod changes in concentration will not be captured. For example, both a brief rise in concentration and a subsequent fall which offsets it might be cost-induced, and these effects are assumed away here.

[32] Harold Demsetz, *supra* note 2; John McGee, *supra* note 8, at 75-79.

[33] Much the same cost effect was found by assuming any of a wide range of positive values for α for that subsample. See subsequent discussion.

TABLE 1
REGRESSION ESTIMATE OF EFFECTS OF CHANGES IN CONCENTRATION ON CHANGES IN UNIT COST, 1947-67
FOUR-DIGIT SIC INDUSTRIES

Independent Variables			
Symbol	Description	Coefficient	t-ratio
	Input Cost Shares (a_i) × Input Price Changes ($\dot{p}_i$)		
1. *LAB*	i = labor	1.211	4.639
2. *RMT*	i = raw materials	.991	1.480
3. *CAP*	i = capital (cost share only)	.980	2.053
	Growth in Demand:		
4. *G*1	gs	.079	.901
5. *G*2	$g(1 - s)$	.116	2.150
	Change in Concentration:		
6. *CR*1+	$g\left[\frac{z^2M_i}{\alpha_i^2K_i} - \frac{zM_i}{\alpha_i}\right]$, for increases in concentration (0 otherwise)	−2.245	−3.245
7. *CR*2+	$\frac{z^2}{\alpha_i s(1 - s)}$, for increases in concentration (0 otherwise) (α = .856 for (6) and (7))	−.006	−.026
8. *CR*1−	Same as (6), for decreases in concentration	−.645	−1.546
9. *CR*2−	Same as (7), decreases in concentration (α = −1.046 for (8) and (9))	−.484	−1.006
Constant	—	−.328	−1.686
R^2	Coefficient of Determination	.342	
SE	Regression Standard Error × 100	20.760	

Note: The dependent variable is the log difference of 1967 and 1947 unit costs. Unit costs are total industry costs (labor + raw materials + capital) deflated by an output index. Total costs are from U.S. Bureau of the Census, Census of Manufactures: 1947 & 1967 (with capital costs estimated—see text). Production Indexes are from the U.S. Bureau of the Census, Census of Manufactures: Indexes of Production (various years). The 1947 value is set at 100, and the 1967 values derived by successive multiplication of cross-weighted 1954, 1958, 1963, and 1967 intercensus output ratios.

The factor cost shares are averages of 1947 and 1967 values. See text for sources of factor price changes. Since the capital cost price change is assumed to be a constant, only the capital factor share is entered on line 3.

Growth in industry sales and concentration ratios are from U.S. Bureau of the Census, Census of Manufactures (various years).

The α_i are estimated by running the regression on the components of *CR*1 and dividing the coefficient of gZM_i by that of gZ^2M_i/K_i. (Since this uses 2 degrees of freedom, t-ratios are exaggerated by about 1 per cent, given the sample size of 165.)

over, are consistent with all of the nonlinearities emerging from the basic model.[34] This consistency, though, hides a qualitative discrepancy between

[34] Specifically, at the sample means, we obtain the following signs for partial derivatives from the regression (with the signs derived from (17) for $\delta = (b - a)g < 0$ in parentheses):

$$\dot{C}_z < 0 \quad (< 0)$$
$$\dot{C}_{zz} > 0 \quad (?)$$
$$\dot{C}_g < 0 \quad (< 0)$$
$$\dot{C}_{zg} < 0 \quad (< 0).$$

the empirical results and the basic thrust of the model. Empirically, the main link between market structure and costs comes via the term ($CR1+$) in which market-structure change and growth interact. An increase in concentration in a nongrowing market has trivial cost effects. In the model, this interaction term is supposed to have effects proportional to differences in firm cost growth rates. However, the coefficient of this term in Table 1 is much larger than any plausible difference in cost growth rates, so it seems clear that the model incompletely specifies the interaction of growth and market structure with costs.[35] The size of this interaction is a puzzle that demands further work, but here I shall merely draw out its empirical implications.

Tables 2 and 3 show the estimated reduction in unit costs implied by Table 1 for several combinations of the relevant growth and market share variables. These are chosen so as to range roughly one standard deviation on either side of the sample means. (No "low growth" calculations are shown, because these would uniformly yield trivial cost savings.) Since risk of error increases with distance from sample means, it is best to focus on the upper left-hand-corner entries in both tables. These point to two conclusions: (1) Market share effects can be substantial. With total factor productivity growing at around 2 per cent per year in manufacturing, about one-fifth of this growth can be attributed to postwar market structure changes for a typical industry with increased concentration. (2) The effect is much larger for increases in concentration than for decreases. A given increase in concentration lowers costs roughly two to three times as much as does an equal decrease.

It is worth examining the sensitivity of these results to the model specification. Note, for example, the implausible negative value of the adjustment coefficient, α, which underlies the estimates for decreases in concentration. Forcing a more, or indeed any, plausible value of α on the estimates turns out to make little difference. I estimated the equation in Table 1 constraining both of the α to take on various values from -30 to $+30$. Except in the neighborhood of $\alpha_i = 0$ (where the $CR1$ variable essentially reduces to a single term and the explanatory power of the regression deteriorates noticeably), the results are very much the same: the mean cost changes are always within a percentage point of the values in Tables 2 and 3, and the regression standard error is also virtually unchanged.

Another aspect of the model that merits examination is the complexity of the $CR1$ variables. These variables dominate the main empirical result, and

[35] The importance of the growth interaction is clear when the regression in Table 1 is estimated without some of the other nonlinearities. No explanatory power is gained by adding only the change in concentration, or $CR2$, to the first five variables. However, most of the Table 1 results are reproduced when linear growth interaction terms (gZ_i) are added. See subsequent discussion.

TABLE 2
REDUCTION IN UNIT COSTS, 1947-67, FOR VARIOUS COMBINATIONS OF CONCENTRATION, INCREASES IN CONCENTRATION, AND SALES GROWTH (NATURAL LOGS × 100)

A. Average Sales Growth (g = .48)

Increase in Concentration Ratio 1947-67 ($Z+$)	Concentration Ratio (CR)		
	Average (CR = 36.3)	Concentrated (CR = 60)	Unconcentrated (CR = 15)
Average increase ($Z+$ = 8.8)	8.0	9.2	3.5
Large increase ($Z+$ = 16)	9.9	13.9	(4.8)
Small increase ($Z+$ = 2)	2.4	2.4	2.1

B. High Sales Growth (g = 1)

$Z+$ \ CR	Average	Concentrated	Unconcentrated
Average increase	16.6	19.2	7.3
Large increase	20.4	29.0	(10.2)
Small increase	4.9	5.0	4.4

Note: ( ) = increase in unit costs.

The entries show the estimated (continuously compounded) percentage decrease in the unit costs of an industry with the specified characteristics compared to an industry with unchanged concentration. The estimates are derived by calculating the $CR1+$ and $CR2+$ (see Table 1) implied by these characteristics, multiplying by the coefficients of these variables and summing. The characteristics are chosen to range roughly a standard deviation either side of the means for the subsample of industries with increasing concentration. The relevant means (standard deviations) for this subsample (89 industries) are:

$$Z+ = 8.8\ (7.1)$$
$$CR = 36.3\ (21.3)$$
$$g = .480\ (.562).$$

(Industries with unchanged concentration are included in both subsamples.)

they imply a relationship between costs and the level, as well as the rate, of change in concentration. In Tables 2 and 3, this level effect is almost always such that higher concentration reduces costs, holding constant the change in concentration. This relationship is most pronounced where concentration increases. This result merits skepticism, first, because it depends partly on the way $CR1$ is constructed and, second, because the underlying model predicts it unambiguously only when concentration > .5.[36]

[36] There are two types of cost change-concentration level relationships embodied in the model. One is described in note 15 *supra*.

The other resides in the market share adjustment process ((16) and (16′)). If we hold Z constant, as is done along the rows of Tables 2 and 3, and the adjustment coefficient is also constant, then the large firm-small firm cost difference (the D' or D in ((16) or (16′)) is implicitly a function of s, and the extent of this difference obviously affects $\dot{C}$. Specifically, the same Z

TABLE 3
REDUCTION IN UNIT COSTS, 1947-67. VARIOUS COMBINATIONS OF CONCENTRATION, DECREASES IN CONCENTRATION AND SALES GROWTH (NATURAL LOGS × 100)

A. Average Sales Growth (g = .727)

Decrease in Concentration, 1947-67 ($Z-$)	Concentration Ratio (CR) Average (CR = 45.8)	Concentrated (CR = 70)	Unconcentrated (CR = 20)
Average decrease ($Z-$ = 8.1)	2.9	3.1	2.1
Large decrease ($Z-$ = 14)	4.2	4.8	1.7
Small decrease ($Z-$ = 2)	0.9	0.9	0.8

B. High Sales Growth (g = 1.25)

$Z-$ \ CR	Average	Concentrated	Unconcentrated
Average decrease	5.9	6.4	5.0
Large decrease	9.9	11.3	6.9
Small decrease	1.5	1.6	1.5

Note: See Note to Table 2. Means (standard deviations) of variables for the 77 industries in the decreasing-concentration subsample are:

$Z-$ = 8.1(7.1)
CR = 45.8 (23.8)
g = .727 (.535).

As a crude check on the validity of both the theoretical underpinning of the $CR1$ variable and its empirical implications, I estimated a regression in which the four market share variables (6-9) in Table 1 were replaced by four simpler terms which separated the concentration change from level effects: a growth interaction term (gZ) and the concentration ratio (s) were entered separately for each subsample. The essential results were:

(1) The explanatory power of this regression is slightly smaller (R^2 = .33 vs. .34) than in Table 1, lending slight support to the more complex formulation shown there.

(2) Coefficients of gZ and s are both significantly negative ($t = -2.2$ and -2.3) when concentration is increasing and insignificantly negative when it is falling. This corroborates the basic result of Table 1 and the general pattern of results in Tables 2 and 3.

---

implies a higher D', the further s is from .5. Thus, the relationship between C and s is reinforced for $s > .5$, but offset for $s < .5$.

(3) The magnitude of the cost changes implied by this regression around the sample means tends to be larger than in Tables 2 and 3. For example, the changes corresponding to those in the first row of Table 2 would be 10.7, 15.8, and 6.2; for the first row of Table 3, we get 5.5, 8.0, and 2.8, implying that both the concentration change and level effects in those tables may be conservative. But the patterns here and in the tables are sufficiently similar to support the amalgamation of effects into the single $CR1$ variable.

So far I have been lumping together productivity and input price effects on costs. To explore any interaction between market structure and input prices, as well as to check on the reasonableness of the preceding results, I estimated market structure effects on two productivity measures. The first is an estimate of $\dot{F}$ in (23), derived by imposing the Cobb-Douglas restriction that the coefficient of each weighted-factor price change is unity. This is essentially an estimate of total factor productivity. The second is a conventional labor productivity (change in output per worker) estimate, which is motivated in part by the measurement error in the nonlabor factor prices. When these were regressed on the last six variables in Table 1,[37] the pattern in that table was repeated: only increased concentration significantly raised productivity. The magnitudes of productivity improvement are also familiar. At the sample means, these were:

1) Total factor productivity: 8.6 per cent, if concentration increases; 3.1 per cent, if concentration falls.

2) Labor productivity: 11.4 and 3.5 per cent respectively.

Finally, I examined the sensitivity of the results to the time period over which they are estimated. The relevant issue here is how long it takes before long-run effects dominate. A shorter time period will, for one thing, make the production function-cost curve relationship in equations (22) and (23) inappropriate. The more substantive risk of specification error arises from the need to distinguish the transitory from the permanent changes in market structure, upon which rests the theoretical link to cost changes. Over short periods, concentration changes will be dominated by forces—like differences between the shape of large and small firm short-run marginal costs—which are ignored by our theory. Moreover, firms which are expanding rapidly to take advantage of their lower long-run costs can incur a short-run adjustment cost penalty. All this suggests that if we focus on too short a time period, the market structure-cost relationship will be unreliable and attenuated.[38]

[37] I added a term—change in the ratio of book value of assets for noncapital costs—to the labor productivity regression to standardize for input mix.

[38] One also has to be mindful of the reciprocal nature of this relationship. A firm—and therefore the industry in which it is classified—can become more efficient today, while its cost advantage is only subsequently translated into larger market share.

TABLE 4
COST REDUCTION (LOGS) PER PERCENTAGE POINT CHANGE IN CONCENTRATION, ESTIMATED AT SAMPLE MEANS, VARIOUS SUBPERIODS, 1947-67

	Industries with Increasing Concentration		Industries with Decreasing Concentration	
Period	Cost Reduction $\div \bar{Z}+$	t-ratio for Coeff. of $CR1+$	Cost Reduction $\div \bar{Z}-$	t-ratio for Coeff. of $CR1-$
1947-67	.91	3.25	.36	1.55
1947-63	.47	2.03	.08	1.28
1954-67	.55	2.29	.26	1.85
1947-58	.60	2.39*	(.12)	1.28
1958-67	.03	.393	.43	1.54

Source: See text and Tables 1, 2, 3. Figures for 1947-67 cost reduction are the upper left-hand-corner entries in Tables 2 and 3 divided by the associated Z. All other cost reductions are derived analogously from regressions on subperiod data.

Columns headed "t-ratio . . ." show absolute ratio of the coefficient of the $CR1$ variables to its standard error in the relevant regression.

( ) = cost increase.

* = coefficient of $CR2$ also had $t > 2$. All other t-ratios for $CR2$ are less than 2.

Theory, though, gives no guidance on what is "too short" concretely. Consequently, Table 4 shows the sample-mean cost changes derived from replications of the regression in Table 1 on data from various subperiods. For ease of comparison, the cost changes are shown per percentage point change in concentration. The t-ratio for the coefficient of the growth interaction variable ($CR1$) is also given, since it turns out to be as concise a summary test of the significance of market structure effects on costs for any subperiod as it is for the whole period.

The general pattern observed in the full period tends to hold for the subperiods: changes in concentration are associated with cost reductions, and they are more pronounced when concentration increases. However, the subperiod effects tend to be smaller than the full-period effects. This indicates that the underlying process generating the cost reductions takes considerable time indeed—at least two decades—to work itself out or that it can be partly obscured by impermanent changes in market structure.[39]

In broad summary, then, the main result of this section is that long-period changes in market structure are accompanied by increased efficiency. This efficiency gain is most pronounced where concentration is high and rising and where demand is growing. In the next section, I discuss the implications

[39] One reason for this is that more general long-run equilibrium cost changes occur slowly. The subperiod regressions underlying Table 4 tend to be characterized by insignificant or implausible negative coefficients for the factor-share-weighted input price change terms, which is a symptom of incomplete adaptation to these price changes. Like the market share effects, these input price effects also tend to be more erratic over the two shorter subperiods.

of this result for output prices and for the lengthy literature on concentration and profitability.

IV. Market Structure, Prices, and Profitability

The existing literature on profitability and concentration provides a convenient starting point for our analysis. An almost universal conclusion is that high concentration and high profitability go hand-in-hand. Since the data used here share a common source with Collins and Preston's[40] contribution, their results provide a useful starting point. Their profitability measure is the "price-cost margin" (M), which is essentially:

$$\frac{\text{Revenues} - \text{Costs}}{\text{Revenues}} = \left(\frac{\text{price} - \text{unit cost}}{\text{price}}\right).$$

They make no explicit adjustment for capital costs, so their costs are essentially plant payroll plus material costs.[41] Their sample, like mine, is drawn from the four-digit SIC universe. Their essential result is the regression reproduced on line 4 of Table 5, in which M is made dependent on the four-firm concentration ratio and the ratio of gross book value of fixed assets to industry sales. (The latter is meant to adjust for capital costs.) Their results can be compared with those of similar regressions for each census year for the sample used in this study (lines 1-6). The pattern is clear: the coefficient of concentration is almost always significantly positive and on the order of .1. (The generally superior results for the capital intensity variable in my sample, while encouraging, need not concern us here.) The concentration effect seems weaker in the two earliest samples, but confidence in the basic result is greatly strengthened when the equation is estimated in first differences (lines 7-9). Here the already weak level relationship survives the noise introduced by differencing, and its magnitude consistently duplicates that in the Collins and Preston data. This comparability of level and change effects is extremely important, since it permits comparison between the main body of my results (which are necessarily estimated in changes) and that of the literature typified by Collins and Preston's study (which invariably employs levels).

A typical inference drawn from results such as those in Table 5 is that they signify inefficiency in concentrated markets. The problems with this inference may be seen with the aid of the following simplified linear representation of equations (5) and (6). Let the analogue to (5) be

$$P = C + as, \tag{25}$$

[40] Norman R. Collins & Lee E. Preston, Concentration and Price-Cost Margins in Manufacturing Industries (1968).

[41] See *id.* at 119 app. A for qualifications.

TABLE 5
REGRESSIONS OF PRICE-COST MARGIN ON CONCENTRATION AND CAPITAL INTENSITY. FOUR-DIGIT SIC INDUSTRIES. VARIOUS YEARS

Year of Census	Coefficients (t-ratios) of *CR*	*GBV*	R^2	S.E. × 100	*N*
1. 1967	.112 (3.9)	.097 (4.3)	.21	8.2	165
2. 1963	.118 (4.3)	.103 (4.2)	.23	8.0	165
3. 1958	.099 (3.6)	.069 (2.4)	.14	8.0	165
4. 1958	.122 (>2.6)	.011 (<1.6)	.13	*	288
5. 1954	.069 (2.6)	.086 (2.5)	.10	7.8	165
6. 1947	.024 (1.0)	.122 (3.9)	.10	7.4	165
Differences					
7. 1967–1947	.103 (2.4)	.018 (.8)	.04	5.9	165
8. 1967–1958	.137 (2.9)	−.031 (1.7)	.07	3.7	165
9. 1958–1947	.120 (2.5)	−.025 (.8)	.04	5.2	165

Note: Source for all data: U.S. Bureau of the Census, Census of Manufactures (various years) and *id.*, Annual Survey of Manufactures (various years). Dependent variable (M) = (value added − payroll costs)/value of shipments for four-digit SIC industry.

CR = four-firm concentration ratio.

GBV = gross book value of depreciable assets/value of shipments.

Line four is reproduced from a regression which includes an additional variable designed to measure the geographic extent of the market. The coefficient of this variable was insignificant at the .1 level. The simple regression of M on CR yielded a coefficient of .125 ($t > 2.6$), $R^2 = .12$. See Norman R. Collins & Lee E. Preston, Concentration and Price-Cost Margins in Manufacturing Industries 99 (1968).

N = sample size.

* = not reported.

Lines 7, 8, 9: Arithmetic changes in M between years indicated are regressed on changes in CR and GBV.

where a is a positive coefficient measuring the impact of collusion, which increases with concentration (s), while the analogue to (6) is

$$C = C_0 - bs, \tag{26}$$

where b is another positive coefficient, and C_0 is a constant. Here, all markets are initially atomistic ($s \to 0$). Then some firms in some markets discover lower costs and gain market share. If these firms are sufficiently few, (26) will approximate the cross-market deviations of C from C_0. Now, even if (25) and (26) hold simultaneously (see below), the reduced form typically estimated in the literature will entirely conceal the process in (26). The "margin" (m) in this context can be defined as $P - C$, and (25) implies that

$$m = P - C = as. \tag{27}$$

That is, a regression estimate of (27) reveals only the collusion effect, when the reduced form for P is

$$P = C_0 + (a - b)s. \tag{28}$$

The net allocative effect of concentration depends on the relative magnitudes of a and b. Matters become more complicated if P responds differently to the components of C in (26). For example, suppose that low-cost firms fail to capture all the market because they eventually run into scale diseconomies, so that their marginal cost exceeds average cost. Then P will not fall by the whole bs term in (26). Approximate this, by rewriting (28),

$$P = 1 \cdot C_0 - kbs + as, \tag{29}$$

where $k < 1$ is a constant. In this world, the reduced form (27) becomes

$$m = [a + (1 - k)b]s. \tag{30}$$

Note that both (27) and (30) imply the same sign for the coefficient of s, and that the sign in (30) is positive even if there is no collusion ($a = 0$). Thus, the conventional finding of a positive sign is consistent with an entirely noncollusive process. There are then two main empirical problems that have to be confronted: (1) essentially, what is the relative magnitude of a and b, and (2) is a positive?

The qualitative answer to the first question can be gleaned by looking behind the results in Table 5. Using Collins and Preston's result on line 4 and assuming that something like (28) rather than (30) holds (namely, that $k = 1$), their definition of the margin implies

$$\frac{dM}{ds} = .122 = (1 - M)\left[\frac{d\ell nP}{ds} - \frac{d\ell nC}{ds}\right] \tag{31}$$

so that the relevant total derivative is

$$\frac{d\ell nP}{ds} = \frac{.122}{(1 - M)} + \frac{d\ell nC}{ds}, \tag{31'}$$

where the first term on the r.h.s. is the assumed "collusion" component. In these samples, $(1 - \bar{M}) \approx .8$, and the top left-hand entry in Table 4 is an estimate of the second r.h.s. term when there is an increase in concentration. Thus, in this case, an estimate of (31′) yields

$$\frac{d\ell nP}{ds} \approx .15 - .91 = -.76. \tag{31''}$$

This result—that the cost effect dominates strongly over any collusion effect—will survive subsequent refinements. So a major inference of the literature needs to be reversed. In fact, increased concentration signifies a net improvement in efficiency, and this is a substantial multiple of any collusion effects.

To get at price effects more directly, we want to estimate the structural equation (5). In the present context, estimates of (log) changes in industry price indices are required to do so. The Census' industry "unit value" indices provide such estimates, but their use entails a major statistical problem. The unit cost variable is industry costs deflated by an output index. But the same output index is often used to deflate industry sales in order to estimate the price index.[42] Thus any measurement error in the output index will be shared by both price and unit cost changes, and straightforward OLS estimate of (5) will yield biased and inconsistent coefficients.

To overcome this problem, a two-stage procedure is used in which the predicted values of the unit cost variable from the Table 1 regression are used as regressors explaining price changes, thereby "purging" the cost variable of the measurement error it shares with price changes. The resulting estimate of the price change structural equation is in Table 6. In addition to the change in unit cost, it includes the following independent variables:

1. The change in concentration, to capture the partial (namely, costs-held-constant) effect of market structure on prices.

2. Growth in sales, which is a proxy for growth in demand, and which should increase price if most markets have long-run increasing costs or if adjustment to equilibrium takes over twenty years.

3. A correction (*OCST*) for costs excluded from the Census. The establishment basis of census reporting means that the Census' cost measure excludes items like advertising and central headquarters overhead. These excluded costs ought to affect price changes when they do not change proportionately with the included costs. However, Internal Revenue Service data from tax returns include total deductions by item for three-digit industries. The largest item ("cost of sales and operations") corresponds roughly to the costs measured by the Census. Therefore, it is possible to compute a proxy for the ratio of total costs to costs measured by the Census from the IRS data at the three-digit level. The log change in this ratio (*OCST*) is then entered for each four-digit industry falling within any three-digit class.[43]

[42] And the rest of the time the output index is estimated by deflating industry sales by the price index.

[43] More precisely, our main cost variable is census costs plus estimated capital (interest plus depreciation) costs. In computing *OCST*, therefore, the denominator includes depreciation, interest, and .2 times stockholders' equity (roughly the postwar average pre-tax return on equity in manufacturing) as well as "cost of sales" and operations. The numerator is reported total deductions plus the imputed cost of equity. IRS data are from U.S. Internal Revenue Service, *supra* note 26.

Unless *OCST* is a perfect proxy, it should have a positive coefficient below that of *COST*.[44]

The results are consistent with prior expectations, and all of the coefficients are considerably greater than their asymptotic standard errors.[45] It is especially interesting that a 95 per cent confidence interval for the *COST* coefficient barely overlaps unity. Therefore, at least over two decades, sellers appear to retain some small part of any unusual productivity gains (and bear part of atypical cost increases). This coupled with the finding that high and rising concentration is conducive to lower costs may help explain part of the observed correlation between concentration and profitability. But there is more to this story, since the *DCR* coefficient in Table 6 is also positive. We can get at the *net* effect of an increase in concentration by using the information in Tables 2 and 6 to evaluate the total derivative in (7). For the "average" case of increasing concentration ($DCR = +8.8$), the approximate total effect is:

$$\begin{aligned} &\text{coefficient of } DCR \times +8.8 && = .212 \times 8.8 \\ &- \text{coefficient of } COST \times && \quad - .934 \times 8.0 \\ &\quad \text{cost reduction if} && \\ &\quad DCR = +8.8 \text{ (see Table 2)} && \\ & && = +1.9\% - 7.5\% \\ & && = -5.6\%, \end{aligned}$$

44 Let the true relationship be

$$\dot{P} = a\dot{C}^*,$$

where $\dot{C}^*$ = change in total costs. Let C = costs included in *COST*, so

$$C^* \equiv R \cdot C,$$

where $R = C^*/C$. However, we know only the proxy for $\dot{R}$, *OCST*. If

$$\dot{R} = b + d(OCST),$$

the estimate of $\dot{P}$ is

$$\dot{P} = ab + ad(OCST) + a\dot{C}.$$

If there is no measurement error, $b = 0$, $d = 1$, and $ad = a$. But since *OCST* is not a perfect proxy, $d < 1$ and $E(\widehat{ad}) < a$.

So long as $d > 0$, we want to take account of any market share effects on the costs not in C. For example, if the share of central office overhead in total costs grows with concentration, part of the previously calculated cost reduction would be offset. However, regressing *OCST* on the market share variable in Table 1 yielded insignificant (and numerically trivial) effects of both increased and decreased concentration.

45 The coefficient of *GRO* may be partly spurious, since the dependent variable is *GRO*-change in output. If *GRO* is deleted, the remaining coefficients are virtually unchanged and their standard errors increase by about one-fifth.

TABLE 6
EFFECT OF UNIT COST AND CONCENTRATION CHANGES ON PRICE CHANGES, 1947-67. FOUR-DIGIT SIC INDUSTRIES

Independent Variables			Coefficient ÷
Symbol	Description	Coefficient	Standard Error
DCR	Change in four-firm concentration ratio	.212	3.818
COST	Log change in unit cost	.934	21.767
GRO	Log change in total revenues	.050	4.630
OCST	Log change of ratio of "IRS" to "Census Costs"	.323	3.773
Constant	—	1.200	0.911

Note: Dependent variable is log change of 1947-67 industry price index. Census unit value indexes are used where available, otherwise industry value of shipments is deflated by an output index. (The latter procedure introduces error where there is net accumulation or depletion of inventories.) Source: Census of Manufactures, various years.

COST is predicted value from Table 1 regression. See text.

See text for description of *GRO* and *OCST*. Sources: *GRO*—U.S. Bureau of the Census, Census of Manufactures: 1947 & 1967; *OCST*—U.S. Internal Revenue Service, Source Book of Statistics of Income: 1947 & 1967.

Sample size: 165. All variables × 100.

or −.64 per cent per percentage point increase in concentration. This last figure is directly comparable to, and not very different from, (31″), which now enables us to understand the process underlying the main result of the concentration-profitability literature. Briefly, more concentration raises profitability, not because prices rise, but because they fall by less than costs. If we ignore doubt about the significance of the cost effect when concentration falls, a similar calculation yields a price reduction of 4.4 per cent for the average ($DCR = -8.1$) case (or .55 per cent per point reduction in concentration). The two effects are roughly comparable, because the weak cost effect is reinforced by the pure price effect in the latter case.

These results pose an immediate question about the meaning of the pure price effect. Is it plausible to attribute that effect to collusion? Recall that an alternative interpretation would rely on rents to differential efficiency, which could be consistent with a competitive process. These alternatives can be distinguished by estimating the effect of *DCR* separately for the rising and falling concentration subsamples. Since costs decline for both types of change, the "rent" interpretation implies an offsetting *DCR* effect for both types. In particular, this means that the coefficient of a decrease in concentration should be negative (or, since the cost effect is weak, at least not positive), which offsets some of the tendency of the cost reduction to lower prices.

This "rent" interpretation is not, however, borne out empirically. When the *DCR* variable in Table 6 is bifurcated, with each new variable equal to the *DCR* with a common sign and zero otherwise, the coefficients of both are positive and virtually identical to the value in Table 6. This means that when concentration falls, prices decline by more than costs and measured industry profitability falls. This process seems inconsistent with a pure "rent" interpretation, so the asymmetry between the profitability effects of increasing and decreasing concentration renders the "cost of collusion" interpretation more plausible.[46]

Noncollusive interpretations cannot, however, be entirely ruled out. For example, when an output expansion by an efficient firm imposes losses on other firms, there is no need for the industry in the aggregate to earn rents from this efficiency. And the likelihood of negative aggregate rents is greater if the efficient, growing firm is small initially than if it is large, simply because the small firm has fewer inputs on which it can earn the efficiency rents that might outweigh everyone else's loss. Unless the small firm gets to be sufficiently large, concentration will decline. Thus, coexistence of declining rents with declining concentration can be consistent with competition.[47]

[46] George J. Stigler, A Theory of Oligopoly, 72 J. Pol. Econ. 44 (1964), argues that the Herfindahl (H) index is superior to the concentration ratio as a proxy for the cost of collusion. Since the Herfindahl index is unavailable for our sample, this argument cannot be directly tested. However, Ralph L. Nelson, *supra* note 17, at 111-30 app., provides H for a sample of four-digit industries. I regressed the log of H on the log of concentration and its square, and found an essentially constant elasticity of about 1.8. This crude empiricism and Stigler's theory suggest that raising concentration to a power and substituting the change in this variable for *DCR* in Table 6 will improve the results, if the "cost of collusion" interpretation is correct. I attempted this transformation for powers ranging from .25 to 2, and the results are encouraging for Stigler's model. The coefficient of this transformed variable rose steadily relative to its asymptotic standard error as the power was increased. At a power = 2, this ratio was 4.97 (compared to 3.82 for power = 1), and the coefficient is $.25 \times 10^{-2}$.

Empirically, this means that the total price effect of a change in concentration is essentially invariant to the level of concentration (*s*), because the differing price and cost effects cancel. To illustrate, for $DCR = +8.8$, and the sample average CR (= 36.3), the partial price effect is $.25 \times 10^{-2}(40.7^2 - 31.9^2) = 1.6\%$, and the total price effect is -5.9%. For $CR = 60$, this calculation is $.25 \times 10^{-2}(64.4^2 - 55.6^2) = 2.6\%$, but note from Table 2 that the cost effect is also larger (9.2% v. 8.0%), and the total effect $(2.6 - .94(9.2) = -6.0)$ is the same.

[47] I am indebted to Yale Brozen for pointing this out. A simple numerical example may clarify his argument: Consider a five-firm industry where firm *A* initially has 60 per cent of the market and *B*, . . . , *E* each have 10 per cent. Let "*CR*" then be 60 per cent. Initially $P = C = 1$, and there are zero rents. Now let any one of these firms (a) discover a way to lower C to .8, (b) cut P to .9 and (c) add 20 points to its market share. The efficient, growing firm then gets rent per unit of its output = .1, while all other firms suffer a loss of .1 per unit. If *A* is the efficient firm, *CR* will increase to 80, and industry rents per unit will be $+.06 = .80(.1) + .20(-.1)$. If *B* is the efficient firm and gains sales proportionately from other firms (including *A*), *CR* will decline to 46⅔; *B* is bigger (30 per cent of the market), but still not as big as *A*. In this case, unit rents are $-.04 = .30(+.1) + .70(-.1)$. There would be increasing rents together with decreasing concentration if *B* obtains between 50 and 60 per cent of the market and thus replaces *A* as the dominant firm. The essential logic of the example is that the firm discovering the efficiency can apply it (and earn rents) to all its output, not just the output it adds; and *A* has the larger output base.

Even if the results in Table 6 are consistent with some collusion, they may overstate its importance. The theory which permits some rents to efficient firms implies that concentration-induced cost changes have a smaller effect on price than industrywide changes in, say, wage rates. The positive coefficient of *DCR* may, in part, be correcting for our failure to allow for such differential effects by lumping all sources of cost change into the one *COST* variable. To test this, I implemented something like (29) by breaking *COST* into the component due to change in concentration and that due to all other forces. The coefficients of these were .75 and .98, respectively, while that of *DCR* declined to .16. The implied total effect on price of an average increase in concentration becomes −4.5, instead of −5.6 per cent. While this procedure does not allow a test of significance, the difference in the cost coefficients is consistent with some rents for innovating firms even in the absence of collusion.

The direct effect of concentration on price seems to have a shorter gestation period than the cost effect. This is evident in Table 7, which gives the coefficient of *DCR* in subperiod estimates of the regression in Table 6. These estimates are uniformly positive and close to the full period estimate, indicating that prices adjust completely to a change in concentration within a decade and that temporary and permanent changes have equally powerful price effects. This pattern may help explain the survival of the erroneous conventional wisdom about concentration. Consider a merger which permanently increases concentration and reduces collusion costs. This permanence may hinge on efficiencies which, however, take a long time for the merged firm to implement. Thus, the immediate and perhaps most easily detectable effect of the merger may well be an increase in price.

Since efficiency effects take hold so gradually, it would be desirable to observe the full-price effect of changes in concentration over periods even longer than two decades. The only data currently available for this are crude, but they are suggestive. For a handful of four-digit industries from 1939 to 1967, output indexes can be pieced together. Since sales data are not

TABLE 7
PARTIAL PRICE EFFECT OF CHANGE IN CONCENTRATION, VARIOUS SUBPERIODS, 1947-67

Period	Coefficient of *DCR*	Coefficient/ Standard Error
1947-67	.212	3.818
1947-63	.233	4.130
1954-67	.251	4.219
1947-58	.159	2.628
1958-67	.286	4.346

uniformly available, deflating value added by output is as close as we can come to a price index. In Table 8, the change in this "unit value added" is shown for 24 industries which experienced a large (10 or more percentage points) change in concentration from 1939 to 1967, and whose growth in output or value added over the period was at least half that of all manufacturing or of their two-digit groups. On average, these industries' "price" performance is about 20 per cent better than that of either their two-digit groups or of all manufacturing. While it is hardly uniform, this superior performance characterizes virtually all the large deviations. Like the previous results, the degree of superiority is similar regardless of the direction of change in concentration (but here it is more reliable for decreases). Since the average change in concentration here is about 15 percentage points, the average price effects are about double those for the 1947-67 sample. Not too much can be made of this result, but it hints at the danger of ignoring the longer-run consequences of a change in concentration.

The Role of the Number of Firms in an Industry. The empirical work has so far focused on concentration, since this allows comparability with a large literature. However, another structural characteristic, the number of firms, merits examination, for it may affect both costs and prices. Telser[48] has shown that, holding concentration constant, the price-cost margin increases with the number of firms. While this may disappoint Cournot's descendants, Telser suggests that it is consistent with an alternative, competitive disequilibrium explanation: high margins attract entrants. But our theory of structure-related efficiency, in its broadest form, raises yet another alternative by positing a relationship between efficiency and inequality, of which concentration is just one indicator. Thus, consider the case where concentration increases even though the number of firms also increases, so smaller firms are losing market share. This means that the discrepancy in size between the largest firm and the "typical" firm grows wider than it would if both were gaining market share. On the other hand, when concentration declines in the face of an exit of firms, size discrepancy narrows more than otherwise. In either case, the unusually rapid growth of one type of firm ought to be related to an unusual cost advantage. To test this, I added two terms to the regression in Table 1: the log change in number of firms if concentration increased (zero otherwise), and the same variable for industries with decreasing concentration (zero otherwise). On the preceding argument, these terms should have negative and positive coefficients respectively. They do, though the effects are not overly powerful (both elasticities were around .1, with t-ratios of about 1.6).

To ascertain the competitive effects of a change in firm number, I then

[48] Lester G. Telser, Competition, Collusion, and Game Theory 330-36 (1977).

TABLE 8

CHANGE IN UNIT VALUE ADDED, 1939-67. INDUSTRIES WITH LARGE CHANGE IN CONCENTRATION

Increasing Concentration				Decreasing Concentration			
Group: SIC and Industry	(1) Log Change in Unit Value Added (× 100)		(2) (1) − Average Change for Group	*Group:* SIC and Industry	(3) Log Change in Unit Value Added (× 100)		(4) (3) − Average Change for Group
Food	111.1		—	*Food*	111.1		
2043 Cereals		110.2	−0.9	2046 Corn Refining		91.5	−19.6
2071 Candy		119.2	8.1	2073 Chewing Gum		59.6	−51.5
2072 Chocolate		128.0	16.9	*Printing*	89.6		
2082 Beer & Ale		46.4	−64.7	2753 Engraving		64.4	−25.2
2087 Syrup		21.1	−90.0	*Chemicals*	26.3		
2098 Macaroni		120.5	9.4	2813 Industrial Gases		−40.6	−66.9
Chemicals	26.3		—	*Petroleum*	103.7		
2844 Toiletries		57.3	31.0	2951 Asphalt Paving		38.3	−65.4
Metal Products	122.5		—	*Rubber*	103.1		
3496 Collapsible Tubes		88.7	−33.8	3021 Rubber Shoes		125.5	22.4
Machinery	74.4			*Stone, Clay, Glass*	105.1		
3555 Printing Mach.		85.6	11.2	3291 Abrasives		102.0	−3.1
Electrical Eqpt.	72.0			3293 Gaskets		115.2	10.1
3633 Washers & Dryers		65.9	−6.1	*Metal Products*	122.5		
Instruments	108.8			3425 Handsaws		132.9	10.4
3871 Watches		56.5	−52.3	3481 Tacks & Nails		90.0	−32.5
Miscellaneous	100.2*			*Instruments*	108.8		
3953 Markers		53.5	−46.7	3843 Dental Eqpt.		88.4	−20.4
				Miscellaneous	100.2*		
				3953 Pins & Needles		63.0	−37.2
Column				Column			
Average	100.2*	79.4	−18.2	Average	100.2*	77.5	−23.2
Standard Error		9.9	11.0	Standard Error		13.5	8.5

Note: Columns (1) and (3) are logarithms of a 1967 index of value added per unit of output (1939 = 1.00). Value added per unit is value added deflated by an index of output.
Value added and output indexes are from U.S. Bureau of the Census, Census of Manufactures (various years).
"Group" refers to the two-digit SIC class.
1947 output indexes were unavailable for the following industries: 2844, 3496, 3555, 3953, and all "decreasing concentration" industries, except 2046, 2813, 3021. In these cases, the two-digit industry output indexes were substituted.
* = total manufacturing.

added this variable to the regression in Table 6. On the basis of Telser's tentative explanation for this result, more firms would be attracted by rising prices, holding costs constant. However, the coefficient of the log change in number of firms is virtually zero (.003, $t = 0.2$). Thus my data hint that the main role played by the number of firms is on the cost side of the profitability equation.

V. Conclusions

Most practitioners have chosen to interpret the profitability-concentration relationship as evidence for collusion. A minority has emphasized the concentration-efficiency nexus. The evidence here is consistent with an eclectic view, but one in which efficiency effects predominate. An important implication of this finding is that, for all its bulk, the concentration-profitability literature is incomplete. Since it has largely been motivated by a collusion model, most of its growth has been elaboration of that theme. However, attention to the efficiency effects of concentration may yield the larger research payoff. For example, one major task is to separate the symptomatic from the causal elements in the statistical relationship between concentration and efficiency. A firm may stumble upon a cost-reducing process and then expand its share of the market. The two events yield distinguishable efficiency gains. The former is not caused by the increase in concentration, but both will be statistically related to it. More commonly, perhaps, efficiency does not come free, thus creating an immediate complication. Investment in search for efficiency will be induced by low costs of expansion, so in this sense the increase in concentration and the initial discovery are causally related.[49]

If the literature is incomplete, so is the rationale it provides for legal hostility to concentration. The possibility that an anticoncentration policy can retain most of the efficiency gains associated with concentration and yield a net improvement in resource allocation cannot be ruled out. But if the magnitudes of the effects we have measured here are close to correct, the odds are against that possibility. It is not clear that U.S. antitrust policy restricts concentration very much.[50] However, if it does, it is more likely to reduce efficiency, raise prices, and reduce owner wealth.

To get at the magnitude of the risks facing an anticoncentration policy, we can focus on industries which have a four-firm concentration ratio

[49] This is at least one way to interpret the importance of the interaction between growth and concentration in explaining efficiency. The growth can both lower expansion costs and increase the payoff to a cost-saving discovery.

[50] See B. Peter Pashigian, Market Concentration in the United States and Great Britain, 11 J. Law & Econ. 299 (1968), and George J. Stigler, The Economic Effects of the Antitrust Laws, 9 J. Law & Econ. 225 (1966).

greater than .5. The average concentration ratio in this sector is around .7, and the typical member spent something over 70 cents per dollar of output for payroll and raw materials. Now imagine that through a divestiture action the concentration ratio for such an industry is reduced to .5. Given our empirical results, this action could raise unit costs on the order of 20 per cent, which in turn would raise price by 10 to 15 per cent. Assuming unit elastic demand, the lower figure would impose a cost on consumers of around 9.6 cents per dollar's worth of output, of which 9.1 cents would be a transfer to producers. Resource costs would increase by around 12.7 cents per dollar of output, so producers would lose 3.6 cents per dollar, and the total loss would be just over 13 cents. Since this concentrated sector currently accounts for around one-fourth or 250 billion dollars of manufacturing sales, any extensive deconcentration program would risk imposing losses which are many times greater than the typical estimates of the benefits such a policy might have been thought to produce.[51]

[51] See, for example, Arnold C. Harberger, Monopoly and Resource Allocation, 44 Am. Econ. Rev., pt. 2, at 77 (Papers & Proceedings, May 1954).