ROBERT E. HALL
Stanford University and National Bureau of Economic Research

Market Structure and Macroeconomic Fluctuations

MARKET STRUCTURE and macroeconomic fluctuations are related to each other in two different ways. First, macroeconomic fluctuations reveal a good deal about market structure. Students of industrial organization have not generally exploited cyclical movements in their research; they have concentrated almost entirely on cross-sectional analysis. One of my goals in this paper is to look at some standard issues in industrial organization through time series variation in individual industries as it is associated with the aggregate business cycle. Second, market structure has an important role in the propagation of macroeconomic shocks. In competitive industries, there are strong forces pushing toward equilibrium. Hence, competitive market structure seems to require an equilibrium interpretation of fluctuations. Where sellers have market power, on the other hand, there is no presumption of full, efficient resource utilization. Fluctuations may be the perverse consequence of noncompetitive conditions.

The first part of the paper looks at the experience of some fifty industries at the two-digit standard industrial classification (SIC) code level, covering all sectors of the U.S. economy. It reaches two basic conclusions about the market structure of American industry. The first is that the majority of the industries are noncompetitive in an important way. Specifically, they choose to operate at a point where marginal cost is well below price, a policy that makes sense only if firms influence

This research is part of the National Bureau of Economic Research's program on economic fluctuations. I am grateful to Michael Knetter and David Bizer for outstanding assistance. The National Science Foundation supported this work. Data for this study are available from the author.

prices through their volumes of production, that is, if they are noncompetitive. I measure marginal cost in a straightforward way. Each boom or recession in the national economy causes increases or decreases in output and cost in an industry. Of course, changes in the prices of inputs also cause changes in cost, but such changes can be measured and eliminated. The ratio of the change in cost, adjusted for input price changes, to the change in output is marginal cost. Comparison of marginal cost to price, observed directly, completes the analysis. I estimate the markup ratio, the ratio of price to marginal cost. In competition, the markup ratio is 1, whereas with market power it exceeds 1. In most industries in my sample, the hypothesis of unit markup is rejected in favor of higher values. In many industries the markup ratio is above 1.5 and in a few it exceeds 3.

My second conclusion about market structure turns on the question of whether market power necessarily translates into excess profit. At a minimum, this investigation is needed in order to make my findings of significant market power credible in view of the fact that the total profitability of U.S. business is not far above the level that represents a reasonable return to capital. Some explanation is needed for market power's failure to bring much profit. My work asks whether firms minimize cost with respect to a constant-returns technology, or whether, on the contrary, they incur fixed costs or other types of costs in excess of that benchmark. The conclusion favors the second case: many industries have costs above the level implied by minimizing cost with respect to a constant-returns technology. The typical firm in these industries is operating on a decreasing portion of its average cost curve. Again, fluctuations in the overall economy are used to measure marginal cost. A firm that minimizes cost with constant returns will earn the market return on its capital when the return is calculated as profit, using marginal cost in place of price to value output. In most industries in the sample, the return to capital calculated in this way is negative, indicating that they cannot be minimizing cost with respect to a constant-returns technology. They hold chronic excess capacity because of a minimum practical scale of operation or they have true fixed costs.

These findings support a view of the typical industry originally proposed by Edward Chamberlin.[1] Through product differentiation or

1. Edward Hastings Chamberlin, *The Theory of Monopolistic Competition* (Harvard University Press, 1933).

geographical separation, firms have power in their own markets. However, there are no barriers to market entry, so firms enter each niche until profit is driven to zero. Because of a minimum practical scale of operation, the typical production unit has excess capacity in the zero-profit equilibrium.

The second part of the paper examines the implications of these findings for macroeconomics. The most straightforward implication is a simple explanation for the well-documented phenomenon of procyclical productivity. In the type of equilibrium consistent with my empirical findings, marginal cost falls considerably short of price. Hence, the calculation of total factor productivity through the method developed by Robert Solow, which assumes the equality of price and marginal cost, involves a bias. I show that this bias has the right sign and magnitude to explain the observed procyclical behavior of productivity.

The findings about industry structure also have important implications for macroeconomic fluctuations. It is now well understood that a noncompetitive economy does not have the automatic full-employment tendency of the competitive economy. Recent authors have built theoretical models in which market power implies that the equilibrium of the economy occurs at a point with unused labor. Some of these models have multiple equilibria. However, there is still a large gap between the theoretical models and empirical work.

The ultimate goal of research in this area is to build and estimate a model in which the economy moves from one equilibrium to another, each involving different levels of resource utilization. A recession and succeeding recovery would be explained as an episode during which output and employment as determined by the equilibrium of the model first shrank and later expanded. However, work has not yet reached this point. Therefore, I will limit my own consideration to the question of how market power and excess capacity diminish the strength of the economy's drive to full employment.

Consider a competitive firm with a well-defined level of capacity (capacity is the level of output where the marginal cost curve turns upward and becomes nearly vertical). Such a firm is unlikely to be satisfied with producing less than its capacity output. As the empirical work in the first part of the paper shows, marginal cost is low when output is below capacity. Unless price falls all the way to the low level of marginal cost along the flat part of the marginal cost schedule, there is substantial incremental profit to be made by putting more output on the market. The competitor never fears that added output will spoil the market, for the absence of that concern is the definition of competition. Hence, output rises to capacity. The only other possibility is for price to fall to a low level.

In the world described by my empirical findings, the incentive to keep output at capacity is nowhere near as strong. A business faced with disappointing sales in a recession hesitates to push more output on the market, because the market will absorb it only at a lower price. Profit will hardly rise. Indeed, profit may not rise at all—the decline in price may exactly offset the increase in sales volume.

For a business in a Chamberlinian equilibrium, the trade-off between sales volume and product price is a matter of low priority because it has only small implications for profit. Product design, cost control, and marketing are the important business decisions. It is true, of course, that recessions bring large reductions in profit for most businesses. However, they cannot recover profit by cutting price and raising volume. A minimum conclusion from my research, then, is that the incentives are weak for those business actions that would restore full employment.

Macroeconomic Fluctuations and Market Structure

Macroeconomic fluctuations continuously bring about natural experiments that reveal marginal cost. When a boom causes a firm to raise its output, the firm incurs extra cost to produce that output. The ratio of the cost increase to the output increase is marginal cost. The empirical work described in this section is no more than a refinement of this simple idea. A much more complete exposition of the technique is available in an earlier paper of mine.[2]

Some economists make a distinction between short-run and long-run marginal cost. For my purposes, that distinction is somewhat off the point. I define marginal cost as the derivative of the cost function with respect to output, holding the capital stock constant. In the out years of

2. Robert Hall, "The Relation between Price and Marginal Cost in U.S. Industry," Working Paper E-86-24 (Hoover Institution, Stanford University, June 1986). An earlier version appeared as Working Paper 1785 (National Bureau of Economic Research, January 1986).

a firm's plan, my measure of marginal cost will equal long-run marginal cost, because the firm will plan a cost-minimizing capital stock. In competition, the firm will equate its marginal cost to the market price of its product, where marginal cost is defined as I have indicated. The equality will hold whether or not the firm is capable of adjusting its capital stock to the current market price. If the firm cannot adjust its capital stock in the short run, then my definition corresponds to the usual concept of short-run marginal cost.

COMPARING MARGINAL COST AND PRICE

The simplest version of my method applies when output and employment change from one year to the next, but the capital stock remains constant. I will also assume that the firm does not use any materials as inputs; labor is the only variable input. Then I measure marginal cost, x, as

$$x = w \frac{\Delta N}{\Delta Q - \theta Q},$$

where w is the hourly wage, N is hours of work, Q is the quantity of output, and θ is the rate of technical progress. Note that the change in output, ΔQ, must be adjusted for the amount by which output would have risen, θQ, had there been no increase in labor input.

All the variables in the marginal cost formula are observed directly except for marginal cost and the rate of technical progress. Robert Solow exploited that fact in his famous paper on productivity measurement in which he assumed that marginal cost was equal to price and solved the equation for the rate of technical progress.[3] Not surprisingly, all of the calculations in Solow's paper are closely related to productivity measurement, and the results are intimately related to the cyclical behavior of productivity.

I will proceed in a somewhat different way from Solow. Instead of making the assumption that marginal cost equals price, I will make assumptions about technical progress and derive conclusions about the relation between marginal cost and price. The assumption is that tech-

3. Robert M. Solow, "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, vol. 39 (August 1957), pp. 312–20.

nical progress can be viewed as random deviations from an underlying constant rate:

$$\theta_t = \theta + u_t.$$

The randomness of the deviation is expressed in a particular way, which is absolutely central to all of the empirical work in this paper: u_t is uncorrelated with the business cycle. That assumption is based on two hypotheses of this work. First, the ups and downs of the economy, from whatever source, do not cause year-to-year changes in productivity. On-the-job learning by doing or research and development stimulated by a vibrant economy does not yield immediate improvements in productivity. The effects are spread over sufficiently many years that the correlation of u_t with the business cycle is negligible. Second, fluctuations in productivity growth do not themselves cause the business cycle. That is, I assume that recessions are not the result of a sudden reduced effectiveness of technology, nor are booms episodes in which output rises more than usual because production functions have shifted favorably. In this respect, my assumption conflicts squarely with the "real business cycle" school, which views variations in the rate of technical progress as one of the main driving forces of the business cycle.

Another assumption I make is that a firm's markup ratio—the ratio of price to marginal cost—can reasonably be approximated as a constant over time. The assumption does not commit me to a "markup" theory of pricing. Rather, it says only that the outcome of the decision process by which a firm chooses its marginal cost and, possibly, its price is such that the ratio of the two is approximately a constant. The assumption is completely compatible with competition, where the markup ratio must be 1. I denote the markup ratio as μ.

Inserting the assumptions about productivity growth and the markup ratio into the formula for marginal cost gives

$$\frac{p}{\mu} = \frac{w\,\Delta N}{\Delta Q - (\theta + u)Q}.$$

Solving for the change in output yields

$$\Delta Q = \mu\frac{w}{p}\Delta N + (\theta + u)\,Q.$$

In rates of change, this is

$$\frac{\Delta Q}{Q} = \mu \frac{wN}{pQ} \frac{\Delta N}{N} + \theta + u.$$

Letting Δq and Δn be the rates of change, and letting α be the revenue share of labor, wN/pQ, I get, finally,

$$\Delta q = \mu \, \alpha \, \Delta \, n + \theta + u.$$

To see what this formula means, consider first the case of competition, where μ is 1. In this case, the formula says that the rate of change of output is equal to the rate of change of labor input weighted by labor's share in revenue, α, plus the constant and random elements of productivity growth. Note that, under competition, the revenue share is a measure of the elasticity of the production function with respect to labor input.

Solow's method for measuring productivity growth is simply to move θ to the left-hand side of the equation and output growth to the right-hand side. The "Solow residual" is just $\Delta q - \alpha \, \Delta \, n$, the part of output growth not explained by growth in labor input.

As stated, my assumption is that u is uncorrelated with the business cycle. As is well known and amply confirmed by the results of this paper, the Solow residual is quite procyclical. Recall that Solow's approach assumes competition, that is, $\mu = 1$. The finding of a procyclical Solow residual leads to one of two conclusions: either my assumption of zero correlation is incorrect, or the firm is not competitive. My work follows the path of the second conclusion. A value of μ in excess of 1 will lower the correlation of the residual, u, and the business cycle. My approach, stripped to its absolute basics, is to choose as the estimate of μ the value that is just high enough to leave the residual uncorrelated with the business cycle. Plainly, the truth lies somewhere between the polar cases. Not every industry is perfectly competitive even in the most optimistic view, and some degree of correlation between productivity shifts and the business cycle would be conceded by any reasonable observer. However, my work proceeds on the assumption that the correlation is small enough to be ignored.

But, a chorus of readers will object, there are numerous sound economic reasons why productivity should be procyclical. All of those sound economic reasons, however, turn out to involve noncompetitive behavior. Consider the explanation based on labor hoarding. Productivity declines in a temporary slump because idle workers are kept at the firm in anticipation that their future employment will be profitable. However, in a competitive industry one of two things will happen in a slump. If workers are idle, marginal cost is at a low level because additional hours of labor are available for free. The industry price must fall to this low level of marginal cost, or it must fall far enough to stimulate demand to the point of eliminating all hoarded workers. In the latter case, neither output nor employment falls in the slump, so nothing happens to productivity. In the former case, the decline in the price has to be considered in the productivity calculation. The price decline makes the revenue share of labor, $\alpha = wN/pQ$, rise dramatically. The residual gives a much higher than normal weight to whatever employment decline occurs, enough so that measured productivity is unchanged.[4]

Other explanations of procyclical productivity either are also founded on assumptions of noncompetitive firms or fail to explain why Solow's method for measuring productivity has a procyclical bias. For example, it is true that productivity will be found to be procyclical in an industry where all firms operate chronically on the declining portion of their average cost curves. However, such industries cannot be competitive, because when price and marginal cost are equal and marginal cost is below average cost, firms would have losses at all times. Or, if firms have overhead labor but are competitive, then Solow's calculation will not give procyclical productivity. The reason is that Solow uses the observed real wage to adjust for the impact of those changes in employment that actually occur over the cycle. Under competition, where the real wage correctly measures the marginal product, his adjustment will operate exactly to offset the changes in output in the productivity calculation. It is true that overhead labor makes output per employee-hour procyclical, but it does not make Solow's productivity measure procyclical.

My argument that competition rules out procyclical productivity applies only to Solow's type of calculation, not to measures such as output per employee-hour. Other measures could easily be procyclical. But my work rests on Solow's measure, in which the response of price rules out procyclical productivity under competition.

4. See Hall, "Relation between Price and Marginal Cost," pp. 7–9, for the details.

Of course, in the real world, firms hoarding labor do not cut their prices to the level of marginal cost with free labor. Price remains high even when marginal cost falls. But this only confirms my point that procyclical productivity involves noncompetitive behavior.

Changes in the Capital Stock. My derivation so far has assumed that the capital stock does not change from one observation to the next. The computations are easily modified to handle the case where capital does change. In fact, all that is involved is redefinition of Δq as the proportional change in the output-capital ratio and of Δn as the proportional change in the labor-capital ratio.

Errors in measuring capital are a potential source of bias in my method for estimating the markup ratio. A bias in the estimate of μ will occur if the measurement error is correlated with the business cycle. The most likely source of bias is that the capital stock will be measured as the amount of capital available to the firm, whereas the calculations should use the amount of capital actually in use. Such an error would certainly be correlated with the business cycle, since capital utilization falls in a slump. However, the bias from this source depends on the pure user cost of capital. If the firm perceives the cost of higher capital utilization to be zero, because there is no pure user cost of capital, then the bias is zero. If the pure user cost is positive, the bias in the estimate of the markup ratio, μ, is positive.

Materials. Generalization of Solow's method for productivity measurement to include inputs of materials is straightforward. Each input appears on the right-hand side of the equation as the product of its revenue share and its rate of growth. Practitioners of total factor productivity have made calculations with dozens of different factors treated in this way. However, the data available for my work do not include explicit measures of materials. Instead, materials have been subtracted from output in order to arrive at value added. I have carried out a full analysis of the implications of applying my version of Solow's method to data on value added.[5] In general, the estimate of the markup coefficient, μ, obtained from the relation between labor input and real value added is an overstatement of the markup of the full price of output over marginal cost. The magnitude of the overstatement depends on the correlation of materials and output. In the unlikely case where the growth

5. Ibid., pp. 10–13.

of materials input is uncorrelated with the growth of output, there would be no overstatement of μ. At the opposite pole, which is more realistic, when materials grow in strict proportion to output, the overstatement is governed by a simple formula. The estimated markup coefficient, μ, is then interpreted as the ratio of the price deflator to marginal labor cost. Because the price deflator is the price less unit materials cost, and marginal labor cost is marginal cost less the same unit materials cost, the value added markup, μ, necessarily overstates the gross output markup, say μ^*. The formula governing their relation is

$$\mu^* = \frac{\mu}{1 + (\mu - 1)m},$$

where m is the ratio of materials cost to total revenue. The ratio m is available only in those years when the Commerce Department has compiled input-output tables, not on an annual basis.

The gross output markup coefficient, μ^*, measures the markup of the actual price of a product over its marginal cost of production, under the assumption that the firm is a price-taker in its materials markets. But μ^* understates the departure of any given price from its competitive level, because the materials suppliers are unlikely to be competitive. In principle, in order to find the degree of departure of a given price from its competitive level, one would have to carry out a full analysis of the upstream suppliers, using input-output data. I have not yet tried to do that. However, there is one simple case where the answer is obvious. Suppose the upstream suppliers are similar to the industry under examination—specifically, they have the same markups and the same materials shares. Then the value added markup for the industry is also the markup of price over full marginal cost, counting the upstream markups. The exercise just considered is made even more relevant by the fact that many of the firms studied here are vertically integrated into their upstream supply industries. Working with input-output data would involve the arbitrary transfer prices used by such firms for their reports to the Commerce Department. In a firm that used its upstream unit's actual marginal cost as the transfer price, all of its market power would be assigned to its downstream unit with my method. On the other hand, if the transfer price includes some monopoly profit, then the analysis of the downstream unit would understate the monopoly power of the integrated operation.

Econometric Method. The equation to be estimated is

$$\Delta q = \mu \alpha \Delta n + \theta + u.$$

The slope, μ, is the markup ratio and the constant, θ, is the average rate of technical progress. The general principle of estimation is to find the value of μ such that the residual, u, is uncorrelated with the business cycle. More specifically, I use an instrumental variable estimator with the rate of growth of real gross national product, Δy, as the instrument.

As indicated earlier, the assumption that u is uncorrelated with the change in real GNP derives from two basic hypotheses of this work. First, fluctuations in productivity growth in any given industry are not causes of fluctuations in total real GNP. That is, genuine productivity growth is not a driving force in the business cycle. Shifts in production functions do not occur quickly enough and do not have sufficiently widespread effects to make an important contribution to year-to-year changes in real GNP. Second, the aggregate business cycle does not itself cause fluctuations in productivity. The production function of a given industry does not shift when national output rises or falls. The actual fluctuations in productivity observed over the cycle are either the result of using a method of productivity measurement other than Solow's total factor productivity or the consequence of market power.

Many other instrumental variables could be considered in addition to the change in real GNP. I have experimented with real military spending, but it is inadequate by itself and has little incremental power when the change in real GNP is already used. In future research, I plan to explore the use of industry-specific instruments such as federal purchases of the output of the industry. Better instruments could improve the results in two ways. First, under the hypotheses that justify the use of real GNP as an instrument, additional instruments could reduce the standard error of the estimate of the markup ratio. In industries whose output is hardly correlated with GNP, the improvement could be substantial. Second, alternative instruments might enable me to test the assumption that the productivity disturbance in each industry is uncorrelated with the national business cycle.

Data. I have assembled data for forty-eight industries covering all sectors of the U.S. economy and for durables and nondurables within manufacturing.[6] From the national income and product accounts, I have taken real and nominal value added, indirect business taxes, hours of work of all employees, total compensation for each industry, and aggregate real GNP. In addition, I have used Bureau of Economic Analysis data on the net real capital stock for each industry. From the data, I have computed the price net of indirect business taxes as the ratio of nominal value added less indirect business taxes to real value added. Because compensation includes social security contributions and other fringe benefits, and the concept of price excludes sales taxes and other indirect taxes, the price and labor cost data are on a comparable basis. That is, a competitive industry would equate marginal cost based on this concept of labor cost to this concept of the price of output.

Because the national income and product accounts discontinued the compilation of the comprehensive measure of labor input after 1978, the sample period is 1949 through 1978. The data are annual.

I have used data for all of the two-digit industries included in the national income and product accounts except for the following, where problems in measuring output are so severe as to make the results questionable no matter how they come out: petroleum refining, banking, insurance carriers, real estate, holding companies, health services, and educational services. For petroleum refining, the calculation of value added seems to be severely distorted by the treatment of foreign income taxes. For banking, insurance, and holding companies, there are severe problems in adding back to purchases of services the value of the financial return paid to customers for their financial investments. For health and educational services, many transactions are outside the market. There remain forty-eight two-digit industries after these deletions.

Results. Table 1 shows the results of estimation for the forty-eight industries, which are divided into two groups: those in which cyclical fluctuations have enough impact on employment and output to shed

6. Most of the industries are at the two-digit SIC level; some are groups of two-digit industries and some are three-digit or groups of three-digit industries. The grouped industries are coal mining (code 11 in table, SIC codes 11 and 12); other transportation equipment (372 in table; SIC codes 372–379); farms (1 in table; SIC codes 1 and 2); agricultural services (7 in table; SIC codes 7, 8, and 9); construction (15 in table; SIC codes 15, 16, and 17); wholesale trade (50 in table; SIC codes 50 and 51); and retail trade (52 in table; SIC codes 52–57).

some light on the value of the markup ratio and those in which cyclical fluctuations are weak or absent and are thus uninformative. The criterion for choosing the informative cases is that the standard error of the estimate of μ be 1.0 or smaller. The criterion is loose and merely excludes the cases of completely useless results.

Only three of the forty-eight industries had inadmissible estimates of μ, below 1: other transportation equipment, security and commodity brokers, and agricultural services. All of them are within 1 standard error of the competitive value of 1. In thirteen industries, the hypothesis of competition is decisively rejected in favor of market power; the estimate of $\mu - 1$ is more than double its standard error. In most of these instances, the estimated value of μ indicates economically substantial market power. In six industries—paper, chemicals, primary metals, motor vehicles, railroad transportation, and trucking and warehousing—the value of μ exceeded 2 and departed from competition by at least 2 standard errors as well. Two of these industries—railroads and trucking—were regulated throughout the sample period. In a decade or so, it should be possible to determine whether deregulation has made them more competitive.

A few industries—textiles, lumber, and other transportation equipment—are shown to be reasonably close to competitive, in that the estimate of the markup ratio is at least 1 standard error below 1.4. In these industries, the data say that the chances are at least five out of six that the markup is 40 percent or less.

Table 1 also shows results for two aggregates within manufacturing—nondurables and durables. The markup ratios are estimated to be 1.61 and 1.62, respectively, with very small standard errors. The hypothesis of competition is decisively rejected for the aggregates as well.

In summary, most two-digit industries show signs of market power, and in a significant part of the economy, market power is substantial. The evidence is based on the finding that increases in output are achieved with increases in labor input that cost relatively little in comparison with the price charged for the output.

EXCESS CAPACITY

With all this market power, shouldn't American industry be inordinately profitable? But if it were, then a new puzzle would result: why

Table 1. Estimates of Markup Ratios, by Industry, 1949–78[a]

Standard industrial classification (SIC) code	Industry	Markup ratio	Summary statistic: Standard error	Summary statistic: Durbin-Watson
	Meaningful estimates[b]			
Substantial market power				
28	Chemicals	3.39	0.78	1.99
26	Paper	2.68	0.33	1.45
40	Railroad transportation	2.38	0.35	1.64
44	Water transportation	2.16	0.65	1.44
371	Motor vehicles	2.07	0.22	2.42
33	Primary metals	2.06	0.15	2.36
42	Trucking and warehousing	2.06	0.48	2.28
Some market power				
32	Stone, clay, and glass	1.81	0.22	2.21
11	Coal mining	1.68	0.51	0.71
27	Printing and publishing	1.61	0.66	1.74
76	Repair	1.60	0.23	2.39
31	Leather	1.59	0.33	2.66
70	Hotels and lodging	1.59	0.88	2.76
39	Miscellaneous	1.52	0.55	2.70
36	Electrical machinery	1.43	0.15	2.35
48	Communications	1.43	0.64	1.92
30	Rubber	1.41	0.20	2.41
35	Nonelectrical machinery	1.39	0.10	2.23
34	Fabricated metals	1.39	0.13	1.42
25	Furniture	1.38	0.17	2.19
23	Apparel	1.30	0.24	2.04
38	Instruments	1.29	0.15	2.38
95	Total nondurables	1.61	0.19	1.81
96	Total durables	1.62	0.09	1.87
Little market power				
15	Construction	1.11	0.34	1.43
22	Textiles	1.05	0.27	1.88
24	Lumber	1.00	0.21	1.87
7	Agricultural services	0.92	0.74	2.29
372	Other transportation equipment	0.91	0.18	1.65
62	Security and commodity brokers	0.56	0.92	2.02
	Unreliable estimates[c]			
10	Metal mining	2.80	1.23	2.16
45	Air transportation	3.28	1.33	1.40
483	Radio and TV broadcasting	2.00	1.40	2.08
78	Motion pictures	2.87	1.63	2.46
20	Food and beverages	3.09	1.64	1.55

Table 1. *(continued)*

Standard industrial classification (SIC) code	*Industry*	*Markup ratio*	*Summary statistic*	
			Standard error	*Durbin-Watson*
21	Tobacco	1.28	2.14	2.26
52	Retail trade	3.63	2.19	2.04
50	Wholesale trade	3.67	2.67	1.35
81	Legal services	4.09	2.75	1.78
75	Auto repair	−1.46	4.74	0.37
41	Local and interurban transit	−1.61	7.00	1.83
79	Amusement	0.35	7.97	1.78
61	Credit agencies	−0.81	8.10	0.93
49	Utilities	10.18	9.09	0.42
13	Oil and gas extraction	11.30	13.20	0.62
64	Insurance agents	−4.14	28.10	2.32
1	Farms	17.20	28.90	1.13
14	Nonmetallic minerals	20.30	104.00	1.61
46	Pipelines	50.50	182.00	1.94
73	Business services	−10.40	432.00	0.85

Sources: Author's estimates as described in text. The data used in the calculations are from the national income and product accounts.

a. The markup ratio, μ, is estimated from the equation: $\Delta q = \mu \alpha \Delta n + \theta + u$, using the rate of growth of GNP, Δy, as an instrumental variable. The dependent variable, Δq, is the change in output; α is the revenue share of labor, Δn is the change in hours of labor, and θ is a constant measuring the mean rate of technical progress.

b. Standard error of the estimate of the markup ratio, μ, is 1.0 or smaller.

c. Industries with too little cyclical variation to measure the markup ratio.

wouldn't new firms enter the market and compete away the profit? One powerful body of thought holds that competition is the only possible outcome in the long run in an industry without barriers to entry. According to that view, market power creates profit opportunities, so entry will occur up to the point that market power is fully dissipated by the multiplicity of sellers. Even more optimistically, the "contestable markets" school argues that the mere possibility of entry will enforce competition in a market with few sellers.[7]

The model of the coexistence of market power and free entry, first articulated by Edward Chamberlin and put on a more formal footing by Michael Spence, Avinash Dixit, and Joseph Stiglitz, has two essential ingredients.[8] First, there must be some separation between the markets

7. And the school is based far to the east of Chicago. See William J. Baumol, John C. Panzar, and Robert D. Willig, *Contestable Markets and the Theory of Industry Structure* (Harcourt, Brace, Jovanovich, 1982).

8. Chamberlin, *The Theory of Monopolistic Competition*; Michael Spence, "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies,*

of rivals. The formal treatments usually consider product differentiation, but geographic dispersion of markets will have the same effect. Second, there must be fixed costs of some kind associated with each distinct market. Absent market differentiation, a single seller could supply multiple markets from a single production unit. As many sellers came to do this, competition would be the result. Absent fixed costs, each market could be served by a great many sellers operating at a small scale, and competition would again be the result.

With differentiated markets and fixed costs, a zero-profit equilibrium with market power will emerge. The smallest markets will be served by a single seller. Although that seller may make some profit, each potential entrant foresees that it cannot cover its fixed costs at the price that would result from the competition between the two were the new seller to enter the market. Hence the market is in equilibrium with monopoly. Larger markets may be able to sustain more than one seller, but still there will be some market power in equilibrium. The details of the equilibrium with more than one seller depend on what theory of oligopoly governs their interaction. If the sellers reach the monopoly price by collusion or otherwise, then market power will remain strong. If the equilibrium is the type described by Cournot, the price will approach the competitive level as the number of sellers grows. In the version of the story with product differentiation, it generally pays for the entrant to adopt a differentiated product, so there is always just one seller in each market.

The empirical work in this paper does not attempt to test Chamberlin's model specifically, but instead examines the profit earned by various industries and compares it with the profit that would be earned by an industry with the degree of monopoly power found in the results of the previous section. The calculation of latent monopoly profit assumes that the technology has constant returns to scale. In particular, the calculation excludes the possibility of increasing returns in general or fixed costs of any kind; that is, it excludes the possibility that the firm operates most of the time on a decreasing portion of its average cost curve.

The basic finding is that profit is nowhere near as high as it would be under full exploitation of market power with constant returns. My interpretation is that firms face setup costs, advertising costs, or fixed

vol. 43 (June 1976), pp. 217–35; and Avinash K. Dixit and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, vol. 67 (June 1977), pp. 297–308.

costs that absorb a good part of the latent monopoly profit. In this interpretation, firms frequently operate on the decreasing portions of their average cost curves. Marginal cost is consequently well below average cost, and zero or low levels of actual pure profit are the result.

I retain the approach to the measurement of market power set forth in the previous section. It has the convenient property that market power is expressed as the ratio of price to marginal cost. Once marginal cost is known, then the profit-maximizing price is known directly; my analysis does not need to go through the steps of profit maximization. Similarly, the optimal level of employment is already implicit in the analysis. However, the third dimension of optimization, the choice of the capital stock, now has to be considered explicitly.

Let π be the actual rate of pure profit relative to sales:

$$\pi = \frac{pQ - wN - rK}{pQ},$$

where rK is the annual service cost of the capital stock. Under the hypothesis of zero expected pure profit, π would be a purely random element, sometimes positive and sometimes negative. On the other hand, if a firm could exercise its full monopoly power and choose its capital stock subject to a constant-returns-to-scale technology, then π would be substantially positive, on the average. To see how big it would be, we must consider the firm's optimal choice of capital.

The characterization of the cost-minimizing choice of capital under constant returns is remarkably simple. Think of the firm as divided into a marketing department and a production department. Marketing takes no inputs. Production sells to marketing at a transfer price equal to marginal cost. If production has chosen its capital stock optimally, the pure profit of the production department will be zero on the average. That is, the quantity

$$\pi_P = \frac{xQ - wN - rK}{pQ}$$

should be a purely random element with zero mean. I call π_P the firm's "production profit." The other part of the firm's total profit is

$$\begin{aligned}\pi_M &= \pi - \pi_P \\ &= (p - x)/p \\ &= 1 - 1/\mu,\end{aligned}$$

which I call "marketing profit." A properly run firm producing with constant returns to scale has production profit that averages zero over the cycle—production just covers the annual carrying cost of its capital stock. Its marketing department generates all of its pure profit, as measured by π_M. By contrast, a firm in the situation described by Chamberlin, with zero pure profit, will generate a substantial amount of marketing profit but will lose the corresponding amount in its production department. That is, π_P will be sufficiently negative to offset the profit generated by marketing. A competitive firm will have no pure profit in either the marketing or the production departments.

In this section, I will simply compute total profit and its two components, based on the estimates of market power from the previous section. That is, I will compute total profit directly from the data, impute marketing profit by inserting the earlier estimates of the markup ratio, μ, into the formula for π_M just derived, and then calculate the profit of the production department as the difference between total profit and marketing profit. This procedure does not try to deal fully with the statistical reliability of the decomposition, but other work of mine shows that the estimates of production profit are quite reliable.[9]

Because the imputation of marketing profit is entirely dependent on the earlier estimation of the markup coefficient, μ, there is no point in trying to make the calculations for this section except for those industries in which there is enough cyclical movement in employment and output to identify the markup. Hence, the results in this section are confined to those industries in the top part of table 1 for which the standard error of the estimate of μ was 1.0 or smaller.

Data. All the data for these calculations are the same as in the previous section except for one added series, the rental price of capital (r in the formulas above). Briefly, I computed the rental price according to the Hall-Jorgenson formula, using the dividend yield of the Standard and Poor's 500 as the real interest rate. I obtained values for the depreciation rate, the effective investment tax credit rate, and present discounted value of depreciation deductions from Jorgenson and Sullivan, and the value for the deflator for business fixed investment from the national income and product accounts.[10]

9. Robert E. Hall, "Chronic Excess Capacity in U.S. Industry," Working Paper 1973 (National Bureau of Economic Research, July 1986).

10. Robert E. Hall and Dale W. Jorgenson, "Tax Policy and Investment Behavior," *American Economic Review*, vol. 57 (June 1967), pp. 391–414; Dale W. Jorgenson and

Results. Table 2 shows the decomposition of pure profit per dollar of sales for the twenty-eight relevant industries. The first column shows total profit on the average over the period, together with the standard error of the estimate of the average. In all but the regulated industries, total pure profit per dollar of sales is positive and usually exceeds zero by many standard errors. A number of industries earn 20 cents or more in pure profit for each dollar of sales. The hypothesis of strict zero profit is rejected by the data. However, this rejection is subject to a number of qualifications. All earnings of the firms not paid out as compensation are treated as profit or the return to capital. The returns to the firm's investments in human capital, research and development, and advertising are included in profit. However, another accounting convention goes in the opposite direction—the costs of these investments are deducted from profit in the year the investments are made. Profit is overstated for slowly growing firms whose current investment falls short of the return earned from past investment and understated for quickly growing ones.

The profit calculations also overstate profit slightly because of the omission of inventories from the capital stock. I have been able to calculate the current market value of inventories for about half of the industries covered in table 2. Profit per dollar of value added is generally about 4 cents less than the numbers in table 2 when the service cost of inventories is subtracted from revenue.

The third column shows the marketing profit per dollar of sales. Marketing profit is a simple increasing function of the markup coefficient, $\pi_M = 1 - 1/\mu$. The fourth column then computes production profit as the residual. Production profit is invariably negative for firms with market power, sometimes substantially negative. Not surprisingly, the biggest production losses occur in regulated industries. But chemicals, paper, primary metals, trucking, and stone, clay, and glass all have production losses in excess of 30 cents per dollar of value added.

Production profits are negative because firms are unable to minimize costs by making a free choice of the scale of their productive units. Instead, many of their units are "too big" because they are at the minimum practical scale. Together with their associated marketing departments, they cover their costs, so they are reasonable investments.

Martin A. Sullivan, "Inflation and Corporate Capital Recovery," in Charles R. Hulten, ed., *Depreciation, Inflation, and the Taxation of Income from Capital* (Washington, D.C.: Urban Institute, 1981), pp. 171–237.

Table 2. Decomposition of Profit per Dollar of Value Added, 1949–78[a]

		Profit per dollar				
SIC code	*Industry*	*Total*	*Standard error*	*Marketing[b]*	*Production[c]*	*Markup ratio*
Substantial market power						
28	Chemicals	0.22	0.013	0.71	−0.49	3.39
26	Paper	0.14	0.009	0.63	−0.49	2.68
40	Railroad transportation	−0.30	0.011	0.58	−0.88	2.38
44	Water transportation	−0.36	0.018	0.54	−0.90	2.16
371	Motor vehicles	0.33	0.023	0.52	−0.19	2.07
33	Primary metals	0.07	0.016	0.51	−0.45	2.06
42	Trucking and warehousing	0.20	0.005	0.51	−0.31	2.06
Some market power						
32	Stone, clay, and glass	0.12	0.011	0.45	−0.32	1.81
11	Coal mining	0.17	0.018	0.40	−0.24	1.68
27	Printing and publishing	0.15	0.004	0.38	−0.23	1.61
76	Repair	0.31	0.230	0.38	−0.07	1.60
31	Leather	0.09	0.006	0.37	−0.28	1.59
70	Hotels and lodging	0.09	0.013	0.37	−0.28	1.59
39	Miscellaneous	0.16	0.006	0.34	−0.19	1.52
36	Electrical machinery	0.15	0.010	0.30	−0.15	1.43
48	Communications	−0.01	0.015	0.30	−0.31	1.43
30	Rubber	0.16	0.010	0.29	−0.13	1.41
35	Nonelectrical machinery	0.15	0.007	0.28	−0.13	1.39
34	Fabricated metals	0.10	0.006	0.28	−0.18	1.39
25	Furniture	0.12	0.007	0.28	−0.15	1.38
23	Apparel	0.11	0.003	0.23	−0.13	1.30
38	Instruments	0.15	0.010	0.22	−0.07	1.29
95	Total nondurables	0.21	0.006	0.38	−0.17	1.61
96	Total durables	0.15	0.009	0.38	−0.23	1.62
Little market power						
15	Construction	0.24	0.003	0.10	0.14	1.11
22	Textiles	0.07	0.009	0.05	0.02	1.05
24	Lumber	0.21	0.007	0.00	0.21	1.00
7	Agricultural services	−2.90	0.198	−0.09	−2.81	0.92
372	Other transportation equipment	0.01	0.015	−0.10	0.11	0.91
62	Security and commodity brokers	0.30	0.019	−0.79	1.09	0.56

Source: Author's estimates as described in text.
a. For those industries in the top part of table 1, where the standard error of the estimated markup was 1.0 or smaller.
b. Calculated as an increasing function of the markup coefficient, $\pi_M = 1 - 1/\mu$.
c. Residual of total profit per dollar and marketing profit per dollar. Figures are rounded.

However, they do not typically operate anywhere near their physical capacities.

QUALIFICATIONS AND POTENTIAL SOURCES OF BIAS

The fact that drives all of my results is well known and uncontroversial: for many industries and all broad aggregates, output can rise substantially with only a modest increase in measured labor input. All measures of productivity, from the simplest measure of output per employee-hour to the most sophisticated computation of total factor productivity, show a pronounced cycle that tracks the movements in employment and output. My work amounts to a new interpretation of this established fact; it attributes procyclical productivity to the existence of market power. As I have already noted, existing explanations of the procyclicality of productivity, such as the hypotheses of labor hoarding and overhead labor, also presuppose noncompetitive behavior and so are harmonious with my explanation.

There is not much doubt, as a matter of economic analysis, that market power distorts the total factor productivity calculations recommended by Solow. Solow's basic idea was to subtract from the growth of output the part that could be explained by the growth of labor input. He used a market measure of the marginal product of labor, the real wage, to provide the coefficient to put in front of labor growth in that calculation. Under competition, the real wage is a proper measure of the marginal product. But with market power, the real wage understates the marginal product of labor. Hence, Solow's calculation makes too small an adjustment for changing labor input in the presence of market power. In an expansion, output rises by more than can be explained by the increase in labor input. Measured productivity rises in the expansion.

The strong assumption that I make is that all of the cyclical behavior of total factor productivity is the result of the understatement of the marginal product of labor on account of market power. I exclude any other factor that does not operate through market power. Here I list and discuss other explanations that I reject by assumption.

—Productivity fluctuations as a driving force in the business cycle

Fluctuations in productivity have been central to the effort of the real business cycle school to find an explanation for aggregate fluctuations that does not rest on price-wage rigidity, market imperfections, or misperceptions about the state of the economy. The real business cycle school tries to use the same basic microeconomic principles that an economist would normally invoke to explain the ups and downs of, say, onion production. Aggregate output is set by the intersection of a supply function and a demand function; the prices mediating the two are the real interest rate and the real wage.[11] A favorable productivity shift

11. See Martin Eichenbaum and Kenneth J. Singleton, "Do Equilibrium Real Business Cycle Theories Explain Postwar U.S. Business Cycles?" in Stanley Fischer, ed., *NBER Macroeconomics Annual 1986* (MIT, 1986), pp. 91–135.

makes output rise; the public perceives a boom. In some other year when productivity was unusually low, the opposite would happen and people would complain of a recession.

In an economy with numerous industries, the productivity shocks would have to be economywide to create meaningful aggregate fluctuations. Were that not so and each of fifty industries had its own independent shock, the law of large numbers would make aggregate GNP almost immune from fluctuations driven by those shocks. In considering what types of productivity shocks might provide a competitive explanation for the findings of this paper, it is essential to restrict consideration to those shocks that operate in tandem across industries and to rule out innovations whose contributions are limited to particular industries.

One obvious common influence is the weather. However, the industry most affected by weather, agriculture, is almost unique in having output fluctuations that are completely uncorrelated with total GNP (see table 3 and the discussion that follows). The hypothesis that weather is an important driving force for total GNP surely faces an uphill battle with the data.[12]

I am also skeptical that process innovation has an important role in aggregate fluctuations. Even if an episode of rapid growth could be traced to the sudden adoption of improved technology, what about contractions in output? Are they periods when businesses throughout the economy choose simultaneously to abandon the most efficient methods?

The only problem with dismissing technology shocks as a driving force for aggregate output is that there must be some driving force—the economy does have important fluctuations. If the driving force is not technology, it must be shifts in preferences, government policies, terms of trade, and other determinants of economic activity. One could be just as scornful about the idea that there are spontaneous shifts in consumption or that investment is driven in part by animal spirits. But some of these forces must drive the cycle, or there would be no recessions and no booms.

—The wage does not control the allocation of labor

A fundamental hypothesis of my work is that the reported wage governs the firm's choice of labor input. The firm is seen as a price-taker

12. Research in progress by Jeffrey Miron has shown directly that fluctuations in temperature and precipitation have almost no relation to output by state and industry.

in the labor market. One alternative interpretation of my results is that firms have extensive monopsony power; they hire workers up to the point that the marginal compensation cost equals the marginal revenue product of labor. Under monopsony, marginal compensation cost will exceed the wage. Solow's productivity calculation will go off track because it uses the reported wage, not the higher level of marginal compensation cost, to infer the marginal product of labor. On this interpretation, my results show that the labor market is imperfect, not that the product market is imperfect.

In an economy where the majority of people work in labor markets with thousands of employers, and few workers are highly specialized in the type of work uniquely available from their employer, it seems implausible that monopsony power in its standard sense has much to do with my findings. Of more concern is the monopsony power that arises in the dealings of a firm with its established, long-term employees. That topic has been studied at length in the literature on employment contracts.

Under a long-term contract, it cannot be taken for granted that the wage set by the contract has anything to do with the cost of increasing labor input. The majority of American workers are paid by salary, and it is virtually the definition of a salary that compensation is the same amount each pay period independent of the actual amount of work. Salaried workers are expected to work harder and longer when there is more work to do. Of course, there must be some implicit cost of asking the existing work force to put in more hours, or management would ask them to work harder all the time. The typical salaried job involves an implicit or explicit arrangement whereby weeks with extra hours are balanced by short weeks or time off. In addition, those who put in extraordinary hours are more likely to earn raises.

As it happens, uncompensated fluctuations in work effort are not an important problem for my calculations, provided that the average amount of compensation correctly measures the implicit wage. Contract theory suggests that, on the average, the two should be equal. The firm should be indifferent whether to ask for more effort from its existing staff or to add new staff. The latter cost is just the average amount of compensation.

Average hourly compensation probably differs from the true implicit cost of labor over the cycle, understating the true cost in good years and overstating it in poor years. However, a cyclical bias has no impact on

Table 3. Statistics on the Rate of Change of Output by Industry, 1949–78

SIC code	*Industry*	*Standard deviation*	*Serial correlation*	*Correlation with cyclical industries*
	Meaningful estimates[a]			
Substantial market power				
28	Chemicals	0.053	−0.318	0.795
26	Paper	0.082	−0.177	0.701
40	Railroad transportation	0.074	−0.074	0.788
44	Water transportation	0.110	−0.193	0.459
371	Motor vehicles	0.195	−0.381	0.743
33	Primary metals	0.126	−0.284	0.871
42	Trucking and warehousing	0.057	−0.143	0.764
Some market power				
32	Stone, clay, and glass	0.067	−0.243	0.847
11	Coal mining	0.095	−0.166	0.463
27	Printing and publishing	0.034	0.067	0.533
76	Repair	0.039	0.122	0.216
31	Leather	0.077	−0.425	0.686
70	Hotels and lodging	0.038	−0.072	0.365
39	Miscellaneous	0.052	−0.400	0.765
36	Electrical machinery	0.089	0.119	0.881
48	Communications	0.020	−0.250	0.426
30	Rubber	0.092	−0.044	0.861
35	Nonelectrical machinery	0.092	0.057	0.766
34	Fabricated metals	0.077	−0.090	0.945
25	Furniture	0.090	−0.222	0.799
23	Apparel	0.053	0.061	0.759
38	Instruments	0.073	0.127	0.804
95	Total nondurables	0.038	−0.111	0.888
96	Total durables	0.083	−0.102	0.987
Little market power				
15	Construction	0.051	0.211	0.709
22	Textiles	0.074	0.126	0.640
24	Lumber	0.085	0.103	0.622
7	Agricultural services	0.052	−0.281	0.313
372	Other transportation equipment	0.125	0.590	0.417
62	Security and commodity brokers	0.067	0.334	0.273
	Unreliable estimates[b]			
10	Metal mining	0.108	−0.370	0.443
45	Air transportation	0.065	0.445	0.613
483	Radio and TV broadcasting	0.065	0.116	0.127
78	Motion pictures	0.068	0.232	0.270
20	Food and beverages	0.033	−0.212	0.554

Table 3. (*continued*)

SIC code	*Industry*	*Standard deviation*	*Serial correlation*	*Correlation with cyclical industries*
21	Tobacco	0.051	−0.087	−0.028
52	Retail trade	0.028	−0.140	0.734
50	Wholesale trade	0.030	−0.145	0.726
81	Legal services	0.042	0.113	−0.033
75	Auto repair	0.048	0.122	0.547
41	Local and interurban transit	0.048	0.213	0.246
79	Amusement	0.031	0.301	0.160
61	Credit agencies	0.030	0.487	0.496
49	Utilities	0.032	0.145	0.169
13	Oil and gas extraction	0.039	0.090	0.661
64	Insurance agents	0.029	0.041	0.559
1	Farms	0.034	−0.503	−0.380
14	Nonmetallic minerals	0.051	−0.126	0.777
46	Pipelines	0.051	0.210	0.651
73	Business services	0.035	0.059	0.519

Source: Author's estimates as described in text.
a. Standard error of the estimate of the markup ratio, μ, is 1.0 or smaller.
b. Industries with too little cyclical variation to measure the markup ratio.

my calculations.[13] Although the first differences of output and labor input are the essential input to the calculations, it is only the level of the wage, as it appears in labor's share, α, that matters. A cyclical error in α is unimportant. Suppose that the error is procyclical, as suggested above. In strong years, employment growth is positive, the growth of real GNP is positive, and the error is positive. Their product is positive. In weak years, all three components are negative, and the product is negative. The net contribution of the error to my calculations is zero, because the weak years offset the strong years.

—Other cyclical errors

The same argument applies to any error whose influence on my calculations is only to introduce a cyclical error in labor's share. Adjustment costs for labor are a good example. With adjustment costs, half the time the firm sees the marginal cost of increasing labor input as above the wage (when growth is high) and the other half of the time it sees the marginal cost of increasing labor input as less than the wage,

13. See Hall, "Relation between Price and Marginal Cost," pp. 36–41, for a formal demonstration of this point.

because of the saving in downward adjustment costs. Another example of a benign cyclical error is price rigidity that is not associated with market power. If the price is less flexible than the competitive price, but the two are equal on the average, then the only result is a cyclical error in labor's share, and that has no impact on my conclusions about market power.

—Errors in measuring labor input

The same argument that shows the irrelevance of cyclical errors in labor's share also demonstrates the sensitivity of my calculations to cyclical errors in measuring labor input. Suppose that the error in measuring hours is negative in strong years and positive in weak years. Then its product with the growth of GNP will be negative in strong years (when the error is negative and the change in GNP is positive) and negative in weak years as well (when the error is positive and the change in GNP is negative). The strength of the association of the change in labor input with the change in GNP will be understated. My estimate of the markup ratio, μ, is the ratio of the covariance of output and GNP changes to the covariance of labor input and GNP changes. That ratio will be overstated in the presence of cyclical errors in measuring labor input.

Two types of cyclical errors in measuring labor input are possible; they both create the same bias. First, fluctuations in reported hours of work may understate actual fluctuations in hours, because firms and workers report a standard forty-hour week and not their actual, more variable work week. My data on hours of work use all available sources to measure actual hours. In particular, the national income and product accounts use the Bureau of Labor Statistics' household survey to measure the hours of nonproduction workers. However, it is likely that there is some element of cyclical understatement of fluctuations in hours.

The second type of cyclical error escapes measurement altogether—fluctuations in the intensity of work effort. One dimension of the proposition that people work harder when there is more work to do is that they get more done per hour of work in the peak than in the trough. It should not be taken for granted, however, that this phenomenon is quantitatively large. It is less persuasive on the downside: in a slump, why would people want to keep coming to work for their usual hours and accomplish less per hour, when they could enjoy more time at home by working as hard as usual, but spending fewer hours at work? Even in normal times, they face the same opportunity, to work closer to capacity and spend fewer hours at work, or, for that matter, to work normal hours and earn more.

The only study I know that has examined work effort over the business cycle finds a small increase, not a decrease, during a slump. Jon Fay and James Medoff surveyed almost 200 managers of manufacturing plants and asked whether the work effort of blue-collar workers changed during a large cyclical contraction. A slight majority said effort increased.[14]

Lessons for Macroeconomics

The results developed here have implications for several important issues in macroeconomics. They add to our understanding of why measured productivity varies cyclically; they demonstrate that economic supply or capacity can be highly elastic; and they explain why market forces provide no strong tendency to move the economy to high-employment levels of operation.

CYCLICAL PRODUCTIVITY

At a minimum, macroeconomists should be aware that market power may have an important role in explaining cyclical fluctuations in measured productivity. None of my work tests the alternative hypothesis that cyclical fluctuations in productivity are an exogenous driving force in a competitive model. Rather, I assume that there is no important pattern of true productivity shifts common across industries that create recessions and booms. Those macroeconomists who believe, as I do, that productivity changes do not drive the business cycle should be at least partly convinced that noncompetitive conditions explain cyclical fluctuations in measured productivity as a response to changes in the forces that cause recessions and booms.

Labor hoarding is an important ingredient in the explanation of why small fluctuations in employment accompany large fluctuations in output. A competitive firm is unlikely to let its work force remain

14. Jon A. Fay and James L. Medoff, "Labor and Output Over the Business Cycle: Some Direct Evidence," *American Economic Review*, vol. 75 (September 1985), pp. 638–55.

idle. It can sell added output without depressing the price. Unless the price is so low that it cannot cover the cost of materials, the firm can make added profit by putting all of its workers to work. But a firm with market power may well hoard workers during a temporary downturn, because the alternative of dumping output on the market is unattractive on account of its depressing effect on price.

CAPACITY CONSTRAINTS

Another important implication of my findings is that it is physically possible for aggregate supply to be highly elastic. In the equilibrium I described in the first part of the paper, numerous firms inhabit market niches with surplus capacity because the constraint of minimum scale is binding. Each is capable of increasing output above its normal level by hiring only a little new labor. Because price far exceeds marginal cost, the increment to GNP from the added output will be worth more than the added wage cost. The output of the economy is constrained by demand in this type of equilibrium. An episode such as a major war or a dramatic, prolonged monetary stimulus can draw forth huge increases in GNP.

If some stimulus—fiscal, monetary, or other—raises demand in a way that is expected to be long-lasting, even more capacity will be created. Higher demand will raise profit in existing niches, stimulating the entry of new capacity in them, and will also make new niches sufficiently profitable for exploitation. In the new equilibrium, expected profit will be zero once again, but at a higher level of total capacity.

INCENTIVES TO EXPAND TO FULL EMPLOYMENT

Perhaps the most important implication of excess capacity and market power in many industries is that businesses have little or no incentive to expand to full capacity. A number of theoretical models have made this point recently. Oliver Hart's model of general equilibrium with market power posits market power in both product markets and labor markets and a single equilibrium in which output and employment are below their competitive levels.[15] In Hart's model, economic activity is sensitive to

15. Oliver Hart, "A Model of Imperfect Competition with Keynesian Features," *Quarterly Journal of Economics*, vol. 97 (February 1982), pp. 109–38. Another important

government interventions that would have little impact in a competitive economy. Procompetitive government policies that would increase the number of sellers in each product market would increase output and welfare. Hart's work offsets earlier partial equilibrium analysis that concluded that the welfare costs of market power were only small Harberger triangles summing to a fraction of a percent of GNP.

Walter P. Heller's more recent work considers a related model with a multiplicity of equilibria.[16] One of the equilibria is similar to Hart's. Others involve even lower output. There is no obvious economic force that will take the economy from its poorer equilibria to the best one. And even the best one has lower output than does competition.

There is a tremendous gap between the theoretical models just described and the actual U.S. economy. Rather than discuss any more elaborate general equilibrium models, I want to consider some features of partial equilibrium with excess capacity and market power at a somewhat more practical level. In particular, I will examine the issue of the incentives that a firm perceives to expand output when it is below its equilibrium output. I will enlarge upon an idea first advanced in the "small menu costs" literature, which has argued that prices are rigid in response to small changes in market conditions.[17] When a firm with market power sets a price to maximize profit, it picks the price where profit is locally unaffected by small changes in the price—the curve showing profit as a function of price is flat at its maximum. Consequently, within some region, the firm cannot improve its profit by enough to justify even small costs of changing its price. It keeps its price at its previous level even though new conditions would justify a different price if the change were costless.

---

paper is Martin L. Weitzman, "Increasing Returns and the Foundations of Unemployment Theory," *Economic Journal,* vol. 92 (December 1982), pp. 787–804.

16. Walter P. Heller, "Coordination Failure under Complete Markets with Applications to Effective Demand" (University of California, San Diego, August 1985).

17. George A. Akerlof and Janet L. Yellen, "A Near-Rational Model of the Business Cycle, with Wage and Price Inertia," *Quarterly Journal of Economics,* vol. 100 (supplement, 1985), pp. 823–38; Olivier J. Blanchard and Nobuhiro Kiyotaki, "Monopolistic Competition, Aggregate Demand Externalities and Real Effects of Nominal Money," Working Paper 1770 (National Bureau of Economic Research, December 1985); N. Gregory Mankiw, "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *Quarterly Journal of Economics,* vol. 100 (May 1985), pp. 529–39; and Julio J. Rotemberg and Garth Saloner, "The Relative Rigidity of Monopoly Pricing," Working Paper 1943 (National Bureau of Economic Research, June 1986).

The literature on small menu costs has considered the relation between price and profit, because the firm with market power is normally considered as setting a price and meeting the demand forthcoming at that price. However, the principal task of macroeconomics, in my view, is to explain the behavior of output, not prices. Price rigidity is significant to the extent that it brings about excessive fluctuations in output, not because prices are intrinsically important. Hence, it is important to look at the relation between profit and output. In doing so, I am not suggesting that firms consider output to be their control variable. Rather, I continue to assume that firms set prices and let their customers choose the quantity sold, but I look at the implications in terms of the resulting relation between output and profit. It turns out that for a broad class of circumstances, that relation is extremely flat.

When a firm finds itself out of equilibrium, with a level of output different from the profit-maximizing one, the incentive to make an adjustment depends on the flatness of the output-profit curve. The flatness depends, in turn, on the degree of market power—that is, on the elasticity of demand facing the seller, on the way that the elasticity changes with output, and on the shape of the marginal cost curve.

Constancy of Marginal Cost. Here I will demonstrate a proposition that is central to the view put forth in this paper: an industry that achieves its equilibrium along a flat portion of the marginal cost curves of its firms is more likely to have a nearly indeterminate equilibrium than is an industry at equilibrium along a rising portion of the marginal cost curve. Consider a firm facing given factor prices and stable behavior on the part of its rivals. The firm is thinking about alternative levels of its own output, achieved by setting different prices. If the firm's marginal cost schedule is steep, an increase in output moves the firm into a region where cost rises more steeply with output and hence profit falls rapidly. The maximum of profit is well defined. On the other hand, with flat marginal cost, only the decline in marginal revenue makes profit begin to decline as output rises above the point where profit is maximized.

What type of industry achieves equilibrium with its firms operating along flat parts of their marginal cost schedules? I will argue that this outcome is much more likely in Chamberlinian equilibrium than in competition, though it is not inevitably a feature of the Chamberlinian equilibrium. A competitive industry generates an expected return high enough to attract capital by having a level of capacity small enough, in equilibrium, so that demand occasionally presses against capacity and high prices are the result. Absent these periods of scarcity pricing, revenue would cover only variable costs, and capital would earn an inadequate return. On the other hand, when firms have market power in equilibrium, the profit derived from that power is itself an attractor of capital. Periods of scarcity pricing only add to the attraction of investment in the industry. It is perfectly possible for equilibrium to occur with sufficient underutilized capital that output never enters a region of rising marginal cost. To summarize, constant marginal cost is an impossibility under competition, because it cannot generate the revenue to pay for the capital stock, but constant marginal cost is completely consistent with an equilibrium with market power. Hence a finding of market power points in the direction of constant marginal cost.

In my findings in the first part of this paper, competition is ruled out for industries with markup ratios substantially above 1. While the results are consistent with the explanation I have just given for flat marginal cost, the argument is not conclusive. There are alternative explanations for the findings of market power with little profit, and not all of them require that marginal cost be flat. For example, suppose that an advertising campaign is needed to establish brand-name recognition in order to enter the industry. The technology has constant returns to scale and there is no minimum practical scale. The equilibrium will not involve excess capacity and a level of output on a flat part of the marginal cost schedule. Instead, all of the latent profit from market power will be dissipated by advertising a sufficiently large number of products.

My results to date support the hypothesis that the marginal cost curve is flat but are also consistent with noncompetitive alternatives. Unfortunately, a direct empirical attack on the problem is difficult because of the cyclical measurement errors that are likely to pervade the data. These errors have a benign effect on my measures of market power and profitability, but stand in the way of measuring the slope of the marginal cost schedule.

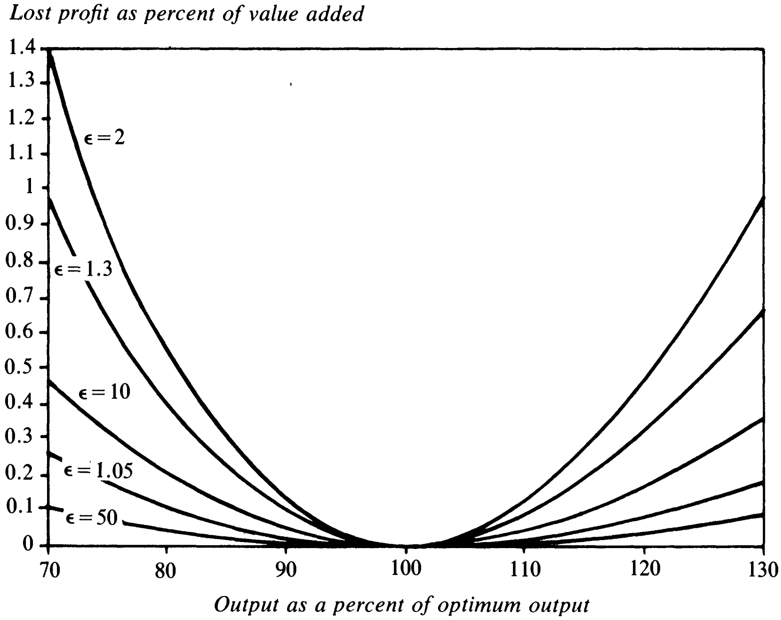
Implications of Constant Marginal Cost. The findings of the first part of this paper are consistent with an industry equilibrium along a flat part of the marginal cost schedule of each firm. In order to draw out the implications of constant marginal cost, I will make the additional assumption that the demand schedule perceived by each firm has constant elasticity. Figure 1 shows various output-profit curves for different elasticities. Each curve is labeled with its elasticity, ϵ; the corresponding markup, μ, is related by $\mu = 1/(1-1/\epsilon)$. Profit is most sensitive to output when the elasticity of demand is around 2. The curve for $\epsilon = 2$ shows that profit falls short of its maximum by about 0.6 percent of value added when output is 20 percent below its optimum and by about 0.5 percent when output is 20 percent too high. Even in this worst case, profit is hardly sensitive to output deviations of 20 percent. And when market power is either greater ($\epsilon = 1.3$ or 1.05) or smaller ($\epsilon = 10$ or 50), profit falls short of its maximum by only one or two tenths of a percent of value added for output deviations of 20 percent.

Figure 1 measures lost profit in relation to value added. Other normalizations might generate larger percentages, but it is important to understand how the normalizations differ. In particular, normalization by the value of profit itself is problematical. Profit is zero on the average in competitive industries and I have argued that it is fairly close to zero even with market power because of the process of entry. Hence a normalization by pure profit would not make sense. Normalization by the total earnings of capital, which are between a quarter and a third of value added, would triple or quadruple the percentages shown in figure 1, but they would still be very small.

The case of constant elasticity of demand is no more than illustrative. The elasticity of demand can decrease with output or it can increase. The linear demand curve is an example in which the elasticity of demand is higher at points of high price and low quantity. And any demand curve that intersects the vertical axis has at least a region where elasticity decreases with output. When elasticity increases mildly with output, profit is even less sensitive to output than it is in the case of a constant elasticity. There are good reasons to think that elasticity may increase with output for many products. Suppose that a product is sold to a number of groups of customers and the groups have different, but constant, elasticities. Then the total demand for the product must exhibit increasing elasticity. As price falls, the demands of the more elastic groups increase as a fraction of total demand. The elasticity of total demand is the weighted average of the elasticities of the groups, so it must rise when the more elastic groups are a larger part of the total. When elasticity increases with output at just the right rate, profit will be perfectly flat for a range of levels of output. That is, it is possible for marginal revenue to be a constant corresponding with marginal cost.

The upshot of this investigation of the implications of constant

Figure 1. Profit and Output with Constant Marginal Cost[a]



Source: Author's calculations as described in text.
a. Assumes that firms face constant marginal cost and constant elasticity of demand. Each profit-output curve is labeled with its demand elasticity, ϵ.

marginal cost is that the incentive to set exactly the profit-maximizing price and produce exactly the corresponding quantity is weak and may even be absent when the elasticity of demand is constant or increases with output. Output can be nearly indeterminate over a wide range with the right curvature of the demand schedule.

With indeterminacy, a firm perceives itself as capturing a fixed amount of profit no matter what its price and output are. Even if indeterminacy does not hold, the incremental profit from adjusting output by, say, 10 percent, is tiny. In businesses with the flat marginal cost curves suggested by my results, fine-tuning output and price is not a matter of priority for management. Managers perceive that lowering a price will raise volume, but they also know that the volume and price effects will cancel each other to a first approximation. Other areas of management, such as better products, more effective promotion, and reduction in overhead and production costs, receive higher priority because there is no automatic offset to their benefits.

Everyday economic life is full of examples of near-indeterminacy at work. In most communities today, the prices of essentially identical gasoline at neighboring stations can differ by several cents a gallon. Dispersion in prices persists for months, but the pattern is not permanent. Chevron was once at the top of the distribution, but is now near the bottom. Stations with low prices do vastly more business, but their profit per gallon is enough less that they are no more profitable than the high-priced ones. Indeterminacy of quantity seems the only reasonable explanation. A condominium in Maui has raised its daily rate 10 dollars each year for the past seven years, quite heedless of what has happened to other rates or to the general state of the market. The occupancy rate is down considerably, but the owners are satisfied that they could not earn more (or less) by bringing their rate back to its historical relation to other rates. During the Great Depression, International Nickel, then a monopoly, did not bother to change the dollar price of nickel for eight full years, in spite of large reductions in costs and output.

Time Series Implications. This line of thought implies that output in sectors with constant marginal cost is close to indeterminate. If a shock depresses the output of a firm governed by the constant marginal cost hypothesis, there is no strong force tending to restore output to its previous level. The shock may depress profit, but the firm cannot raise profit by adjusting its price so as to raise its output. However, indeterminacy does not imply any particular time series behavior for output. A firm could choose to stabilize its output and let price absorb all shocks. Or, if it had market power, it could stabilize price and let quantity track shifts in demand. Both strategies would yield the same stream of profits. It is tempting to jump to the conclusion that price stabilization is the prevailing mode in industries with market power. The industries shown in table 1 to have market power seem to be ones where management sets a price and customers choose the quantity. However, the data on the time series properties of output by industry do not give strong support to that proposition. Table 3 shows the standard deviation and the serial correlation of the rate of growth of output of the industries in table 1.

The standard deviation of the rate of change of output is a summary measure of the variability of output in an industry. Under competition, the equilibrium profit-maximizing level of output should rise along with capacity. This proposition would remain true even if there were shifts in demand, because the competitive industry operates at capacity and absorbs demand shifts through price variations. Only shifts in the capacity level of output could explain significant random variation in output growth in the competitive industry.[18] Table 3 does show a slight positive relation between market power and output instability, but plainly market power is not the prime determinant of instability. For example, output growth in durables has a standard deviation of 8.3 percent as against 3.8 percent in nondurables, even though measured market power is identical and substantial in both. Textiles is found to be a competitive industry, but its standard deviation is 7.4 percent. In fact, none of the industries found to be approximately competitive has smooth output growth. But those industries with the most market power do tend to have the most unstable output.

Another dimension of the time series behavior of output is the duration of departures from equilibrium. The second column of table 3 shows a simple measure of persistence, the serial correlation of the rate of change of output. A serial correlation of zero means that output is a trended random walk. Shocks are infinitely persistent; there is no tendency for an increase in one year to be followed by a decrease in the following year as output returns to its equilibrium level. Negative serial correlation means that shocks are temporary. The recent literature on random-walk components of macroeconomic variables has called attention to the importance of the serial correlation of the first difference of output.[19] For total real GNP, the serial correlation is roughly zero. The market power–excess capacity hypothesis is consistent with random-walk behavior of output but does not mandate it. Table 3 shows quite clearly that output is more like a random walk in competitive industries than in

18. If capacity changes abruptly from year to year, a competitive industry could show considerable fluctuations in the rate of growth of output. To check this point, I calculated the standard deviation of the rate of growth of the output-capital ratio for each of the industries. In competition, this ratio would remain nearly stable over time. In fact, the standard deviations of the rates of growth of output and the output-capital ratio are almost the same in each industry. Irregular capacity growth is not the explanation of irregular output growth.

19. Charles R. Nelson and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics,* vol. 10 (September 1982), pp. 139–62; John Y. Campbell and N. Gregory Mankiw, "Are Output Fluctuations Transitory?" Working Paper 1916 (National Bureau of Economic Research, May 1986).

those with market power. All of the industries found to have substantial market power have negative serial correlations of output changes. Only one of the competitive industries shows a negative serial correlation. A glance at the data for the industries with substantial market power shows why the serial correlation is negative. Each recession brings a large reduction in output. In the succeeding year, part, but not all, of that decline is usually reversed. These observations dominate the calculation of the serial correlation.

The fundamental identifying hypothesis of this line of research holds that productivity fluctuations are not a driving force in the business cycle. This hypothesis enables me to interpret the empirical findings of the paper as revealing that market power is extensive. Under the alternative hypothesis—that productivity shifts are the driving force of macroeconomic fluctuations—the results have just the opposite interpretation. To the extent that it is hypothesized that productivity is procyclical, there is less room for market power to explain the same facts. At the polar extreme, if all observed productivity fluctuations are taken to be exogenous driving forces, then there is an implicit assumption of pure competition.

The third column of table 3 presents a rough test of the hypothesis that productivity shifts are not a driving force in aggregate fluctuations, based on the following logic: suppose the hypothesis is wrong. The origin of what we observe as the cycle is basically durables and construction. The cycle must be driven by productivity shifts in those industries. In years when productivity is low and the economy is in recession, labor should move out of those industries and into unrelated industries. Output in those industries should rise as a result. The competitive real business cycle model seems to require that at least some industries should be countercyclical. Column 3 shows that essentially every industry is procylical. The only industry with a meaningfully negative correlation with the cyclical industries is farming.

The only way to save the real business cycle view is to appeal to a systematic tendency for productivity shifts to occur simultaneously in the same direction in most sectors. The data are inconsistent with the notion that the cycle has its origin in productivity shifts in the most cyclical industries and that other sectors merely respond to those shifts.

The negative correlation for farming has another implication, unfavorable to the real business cycle view as well: the one common influence on productivity on which we can all agree is the weather. Probably the bulk of the fluctuations in real output in farming are the result of changes in the weather. Hence, farming serves as a proxy for the influence of weather on all industries. But the negative cyclical correlation of farm output casts doubt on the one reasonable influence that operates across all industries.

Conclusions

The findings of this paper support a view about the operation of product markets in the U.S. economy that is consistent with the observed pattern of large, persistent movements in aggregate real output. However, the view is anything but firmly established. In essence, the view is that many industries are in equilibrium along a flat part of each firm's marginal cost schedule. The rents associated with efficient use of resources, with equilibrium on a steeper part of marginal cost, would attract additional entry. A firm with constant marginal cost is virtually indifferent to alternative levels of output. When output is 10 or 20 percent below the profit-maximizing point, profit is only a few tenths of a percent below its maximum, as a proportion of sales.

With extremely weak incentives to restore previous levels of output, it is no mystery that industries and the entire economy can undergo large and persistent fluctuations in output. This insight does not lead to accurate predictions about movements in output and the corresponding movements in prices. Rather, it supplies the answer to the question that has acutely troubled disequilibrium business cycle theorists for the past two decades: how are sluggish price adjustment and large output fluctuations consistent with rational economic behavior?

All of the conclusions of this paper follow from the fact that total factor productivity is procyclical. The measures of the markup coefficient are no more than an interpretation of that fact. And the conclusions about excess costs are based entirely on the interpretation that procyclical productivity reveals market power.

The competing explanation for procyclical productivity appears in the active and growing literature on real business cycles. According to this view, exogenous productivity shifts, positively correlated across industries, are a prime moving force in the business cycle. Consequently, productivity growth and output growth are positively correlated in each industry. Real business cycle theorists tend to assume competition, and this assumption is consistent with my results: if cyclical shifts in productivity are an important reason for the procyclical behavior of productivity, market power must be correspondingly less important.

As the evidence now stands, one has a choice between these two very different views, both consistent with the principal evidence. Prior beliefs about the plausibility of large exogenous shocks in productivity are the primary basis upon which the choice has to be made. My own view is that productivity shocks, in the narrow sense of shifts of production functions, are not an important source of aggregate fluctuations. Hence, I believe that the observed procyclical behavior of measured productivity is in some considerable part the result of market power. Moreover, I think that the finding of market power in many industries opens up avenues of explanation of the vulnerability of total output to many other types of shocks, including shifts in the terms of trade, spontaneous shifts in consumption and investment, and changes in government policy.

Comments and Discussion

Olivier Jean Blanchard: This is an exciting paper on a potentially important topic for macroeconomics. Much of our attention in explaining business cycles has focused on the structure of the labor market, with the structure of the goods market usually being given low billing. In studies of wage-price dynamics, for example, most of the emphasis is on wage dynamics; prices are simply assumed to reflect standard average costs, up to a markup that is largely independent of the level of output. But, as was pointed out by William Nordhaus in his survey of empirical price equations in 1972, if such a characterization of price behavior is accurate, it points to imperfect competition in the goods market.[1] If imperfect competition is indeed prevalent, it is clearly something that should be taken into account explicitly: it may help us understand not only pricing behavior but investment demand, labor demand, and so on. Hall's paper is part of a research effort aimed at understanding the implications of imperfect competition for macroeconomics. The paper has two distinct parts, the first aimed at documenting the existence of imperfect competition, the second drawing implications for macroeconomics. I shall deal with them in turn.

The first part of the paper is based on three facts.

The "productivity fact." In the regression

$$(1) \qquad y = an + e,$$

where y and n denote, respectively, the logarithms of value added and man-hours, and e is the logarithm of total factor productivity, the

1. William D. Nordhaus, "Recent Developments in Price Dynamics," in Otto Eckstein, ed., *The Econometrics of Price Determination* (Board of Governors of the Federal Reserve System, 1972), pp. 16–50.

coefficient a is estimated to be around 1.0, at both the aggregate and the sectoral levels. Hall also shows that it remains around 1.0 at the sectoral level when aggregate value added is used as an instrument in the regression.

The "share fact." While the share of labor in value added varies in the business cycle, its average value is equal to 0.75.

The "profit fact." The rate of pure profit, that is, profit in excess of the normal return to capital, is small. While Hall looks directly at profit, an alternative approach is to compute average q ratios, which have the advantage of avoiding some of the issues associated with the differences between accounting and economic profits. Average q ratios rarely exceed 1.0 for long periods of time.

All three facts have long been well known. The insight of the paper is to ask how they can be consistent. Its conclusion is that they are consistent only if goods markets are imperfect. Let me play devil's advocate and return to an alternative explanation based on perfect competition. Hall mentions but excludes the alternative on grounds of implausibility, but it is worth examining further.

Under the alternative I offer, there is perfect competition and decreasing returns to labor in the relevant range. And, contrary to Hall's maintained assumption, productivity shocks and GNP fluctuations are correlated.

In this explanation, perfect competition explains the profit fact, and decreasing returns explain the share fact. The correlation between GNP and productivity explains why the estimate of the coefficient in equation 1, which has a true value equal to the labor share, is biased upwards and is equal to 1. The bias comes from the fact that movements up the upward sloping marginal cost curve are associated with shifts to the right of the curve: the estimated marginal cost curve is much flatter than the true one.

Two conditions are necessary to get such a positive and substantial correlation. The first is that productivity shocks be correlated across sectors, so that the aggregate productivity shock is highly correlated with sectoral shocks. The second is that aggregate productivity shocks and GNP be correlated.

I find rather convincing Hall's arguments that the correlation between productivity shocks across sectors is likely to be small. While one can think of innovations, such as the increased use of computers, that affect productivity in all sectors, they appear to account for a small proportion of productivity innovations. But Hall could have supplemented his arguments with a formal test. Given that the assumption that productivity shocks are uncorrelated across sectors is not used directly in estimation, it is an assumption that can be tested under the maintained assumption that aggregate productivity shocks are uncorrelated with GNP. The way to test it is to compute the set of correlations of the estimated productivity shocks obtained by Instrumental Variables (IV) estimation across sectors. A table giving these correlations would be of interest; I would not be surprised to see positive correlations of estimated productivity shocks across most sectors. If this were the case, it would certainly weaken Hall's argument.

I find less convincing Hall's arguments that aggregate productivity shocks and GNP are likely to be uncorrelated. One does not need to be a believer in real business cycles to expect a positive correlation. In the most old-fashioned Keynesian model with fixed nominal wages, positive productivity shocks will decrease nominal prices at any level of output, increase real money, and increase demand and output. It is also plausible that, even if the relation is not causal, investment and productivity shocks are correlated. If they are, aggregate demand will move with productivity shocks. This suggests the use of truly exogenous aggregate demand variables instead of GNP in equation 1. While the search for such instruments has proved elusive, Hall has in the past argued that defense spending was such an instrument; it could be used here as well.[2]

To get a feel for how large the correlation between GNP and productivity shocks must be in order to reconcile the productivity and share facts—that is, to imply a bias of about 0.25 in the estimated coefficient when equation 1 is estimated by the IV method of Hall—consider the following example. Assume that productivity shocks are perfectly correlated across sectors, that each sector has a production function given by equation 1, with y and n in rates of change, and that e, the innovation in productivity growth, is white noise. Assume that movements in GNP growth are given by

$$y = v + be, \tag{2}$$

2. See Robert Hall, "The Role of Consumption in Economic Fluctuations," in Robert J. Gordon, ed., *The American Business Cycle: Continuity and Change* (University of Chicago Press, 1986).

where v are white noise movements in GNP growth uncorrelated with the productivity shock e, and b measures the effect of e on y. The variables v and e have variances respectively equal to σ^2_v and σ^2_e. The bias in the estimated coefficient a_{iv} obtained by using y as an instrument in equation 1 is a function of b and x, the ratio of the variance of v to the variance of e. Then $a_{iv} = a(1 + b^2x)/[1 + (b^2 - b)x]$. A value of $a_{iv} = 1.0$ while $a = 0.75$ is consistent, for example, with $x = 1.0$ and $b = 0.26$, thus a relatively small value of b. If the productivity shocks are not perfectly correlated across sectors, the value of b required to explain a bias of 0.25 increases but, as long as the correlation between productivity shocks is high, the values of b required to explain the bias are not implausibly high.

To summarize, there are good grounds to doubt that Hall's assumption of no correlation between productivity shocks and GNP fluctuations is correct. Small deviations from this assumption may explain a substantial part of the discrepancy between the share of labor and the estimated coefficient in equation 1. Because the perfect competition explanation is a plausible way of reconciling the three facts, Hall's paper will probably not change many minds.

On the other hand, because in my own view imperfect competition is an important characteristic of the economy, I like Hall's explanation. It rejects perfect competition and points to monopolistic competition with fixed costs and constant returns to labor in the relevant range.

Constant returns to labor explain the productivity fact. Imperfect competition implies that price exceeds marginal cost, which under constant returns is equal to the wage. Under constant returns, the labor share is proportional to the product wage and thus less than 1.0: that explains the share fact. Finally, free entry explains the profit fact by implying that profits must be dissipated, either by large fixed costs compared with the market or by overaccumulation of capital or other quasi-fixed factors.

One way of assessing the plausibility of the results is to look at the cross-section implications. I have computed, for manufacturing, the rank correlation between the index of market power derived in Hall's table 1 and the four-firm concentration ratio, taken from Rotemberg and Saloner.[3] It is approximately equal to zero. Whether this is bad news for

3. Julio J. Rotemberg and Garth Saloner, "A Supergame-Theoretic Model of Price Wars During Booms," *American Economic Review*, vol. 76 (June 1986), pp. 390–407.

Hall or bad news for the concentration ratio can be assessed only on a case-by-case basis, and I know far too little about individual sectors to do so.

The second part of Hall's paper, which assesses the macroeconomic implications of imperfect competition, is less convincing. In it, Hall chooses to emphasize the implications of imperfect competition when marginal cost is flat. But there is nothing in the analysis of the first part that implies that marginal cost is flat. Indeed, one of the advantages of the method used to estimate market power in the first part is that it is robust, for example, to the presence of adjustment costs; put another way, it gives no information as to the size of these adjustment costs and thus to the slope of the short-run marginal cost curve.

Does the logic of imperfect competition imply that marginal costs are flatter than they are under perfect competition at normal levels of output? The answer is at best maybe. What is true is that excess profits from market power must, in the long run, be dissipated to prevent entry. They can be dissipated through excess capacity, as Hall argues. If that is the case, then, other things equal, marginal cost may indeed be relatively flat. But that is one of many ways to reduce profits: they can also be dissipated by introducing new products until fixed costs associated with producing each product absorb excess profit. In that case, firms need not have excess capacity, and marginal cost may be very steep.

What do we know about the slope of the marginal cost curve? Unless we believe that there is a large bias in the estimated coefficient, a, in equation 1, we know that there are approximately constant returns to labor; the textbook rationale for why marginal cost is increasing in output is thus ruled out. But we have other facts that strongly point to upward sloping marginal cost. As pointed out by Mark Bils, firms use overtime, which is costly.[4] If there were other and cheaper ways of increasing production, they would presumably use them; that suggests that there are adjustment costs in adjusting the labor force, that the short-run marginal cost curve is steep. Evidence on labor demand suggests slow adjustment of labor demand to target, pointing again to adjustment costs. Evidence on inventory behavior, while being inconsistent with the simple production-smoothing model, also suggests costs of adjustment in production. Taken as a whole, the evidence leads me to conclude that

4. Mark Bils, "The Cyclical Behavior of Marginal Cost and Price" (Ph.D. dissertation, Massachusetts Institute of Technology, 1985).

marginal cost is not flat. Thus, the challenge remains one of explaining why prices do not respond to changes in marginal cost.

Even if imperfect competition does not deliver a flat marginal cost curve and thus a straightforward explanation of why firms have little incentive to change prices and can easily accommodate variations in demand, all is not lost.

First, under imperfect competition, there are other explanations for why prices will not necessarily move with marginal costs. Among them are variable demand elasticities, customer attachment, and bilateral relationships. While I do not yet find any of them fully convincing, I remain hopeful. Also, because price decisions are taken by firms rather than by an auctioneer and because the price decisions of one firm depend, both through the prices of its inputs and the prices of competing outputs, on the price decisions of other firms, small lags in adjusting prices at the individual firm level can cumulate into substantial price inertia at the aggregate level. Thus, imperfect competition can explain both a small adjustment of prices to movements in demand and a slow adjustment of prices to changes in wages, two important ingredients of the Keynesian model.

Second, whether or not imperfect competition can explain why aggregate demand movements affect output, it clearly can explain many of the features of the actual response of the economy to aggregate demand shocks. To take just two examples, fixed costs provide a simple explanation for procyclical movements in productivity—if they are partly in the form of overhead labor—and for procyclical movements in profitability. Procyclical movements in profitability imply in turn that the threat of entry is low or nonexistent in recessions, and may in turn explain why workers and firms may not be eager to decrease nominal wages and prices.[5]

R. Glenn Hubbard: The merger of topics and modeling strategies in macroeconomics and industrial organization, implicit in economic research for decades, is—deservedly—receiving renewed attention. Imperfections in product markets, labor markets, and capital markets provide a foundation for reconciling the predictions of formal microec-

5. These arguments are presented more formally in Olivier Blanchard and N. Kiyotaki, "Monopolistic Competition and Aggregate Demand" (MIT, September 1986).

onomic models with observed movements of aggregate prices and quantities of interest to macroeconomists and policymakers. Examples of this new line of research include models of contracting and staggered price setting, "menu costs" and sticky prices, and imperfect competition and equilibrium output and employment levels. Hall's paper breaks new ground in this area, linking cyclical movements in productivity—a phenomenon long studied in empirical macroeconomics—to market power and discussing implications for output movements over the cycle.

There is much to recommend this paper. It is both bold and simple, centered on three basic ideas. First, price substantially exceeds marginal cost in most U.S. industries—a manifestation of product market power. Second, market power coexists with low average profit rates, a relationship that Hall attributes to substantial excess capacity in industries with a high markup of price over cost. Third, U.S. industries face roughly constant marginal costs, and thus have only a weak incentive to maintain output at capacity or to restore full employment following a decline in aggregate demand; this last finding does not appear to depend much on there being substantial market power. Hall's work is careful and direct; the key assumptions are highlighted and defended.

Refocusing attention on the Solow residual as conveying information about the relationship between price and marginal cost is very useful.[1] My principal comments have to do with interpretations of the spread—both quantitatively, as a reasonable definition of margins, and in terms of representing substantial product market power. Hall's basic message about the importance of market structure considerations for macroeconomics is an important one, however, and I want to present some supplementary empirical evidence using more disaggregated data in support of some of his conclusions.

1. Micha Gisser also finds, but does not emphasize, a positive relationship between the growth rate of output and the Solow residual in a Cobb-Douglas production function. He also finds that increases in concentration in originally unconcentrated industries are associated with larger measures of total factor productivity, arguing that a small group of efficient firms increases the productivity of an industry and concentrates it. Micha Gisser, "Price Leadership and Dynamic Aspects of Oligopoly in U.S. Manufacturing," *Journal of Political Economy,* vol. 92 (December 1984), pp. 1035–48.

Hall is quite right to point out that labor hoarding is consistent with his argument. With labor hoarding in a recession, the marginal cost of labor is small, so that large spreads between price and marginal cost are associated with positive Solow residuals, even though the markup of price over average variable cost may be small. Assuming that the markup is constant may not be innocuous, however.

One of the goals of the paper is to reconcile high price-cost markups with the observation that most industries do not appear to have supra-competitive average profit rates. With a high "marketing profit," as measured by Hall, a loss in "production profit" is required to match the roughly zero average profit in the data. One issue here is ascertaining the markup relevant for market power–market structure considerations. Hall's definition of marginal cost does not include the cost of materials, which have a more significant share in the value of output than labor in many industries. Including materials costs in the expression for marginal cost, x, (assuming that industries are input price-takers and that the materials-output ratio is constant) implies that

$$x = w\frac{\Delta N}{\Delta Q - \theta Q} + p_M\frac{M}{Q},$$

where p_M and M denote the price and quantity of materials used, respectively.

The point is more than definitional. The concept of marginal cost captured in the standard Lerner-index measure relates to output and is logically inclusive of materials, whose use is presumably almost entirely marginal in most industries. For a given industry, a firm's price-cost margin can be expressed as

$$\frac{P - MC_i}{P} = \frac{s_i(1 + \lambda_i)}{\eta},$$

where s_i is the ith firm's market share, λ_i is its conjectural variation (the ith firm's guess about the output response of all other firms), and η is the industry demand elasticity. Some reference points of interest include the monopoly outcome, $PCM = 1/\eta$, and the Cournot outcome, $PCM = s_i/\eta$.

Hall makes use of industry data; we can derive industry expressions by aggregating across firms. If MC is assumed to equal AVC for each firm, then

$$\frac{P - AVC}{P} = \frac{\Sigma s_i^2(1 + \lambda_i)}{\eta},$$

where AVC is the industry weighted average variable cost. Again, we can consider the monopoly outcome,

$$\frac{P - AVC}{P} = \frac{1}{\eta},$$

and the Cournot-Nash outcome,

$$\frac{P - AVC}{P} = \frac{\Sigma s_i^2}{\eta} = \frac{H}{\eta},$$

where H is the Herfindahl index of concentration.

It is difficult to consider issues of relative markups at the two-digit level of aggregation, which obscures substantial variations in markups within a two-digit grouping. In a study of SIC four-digit manufacturing industries with Ian Domowitz and Bruce Petersen, I found that price-cost margins as defined above were much closer to the Cournot predictions than to the monopoly predictions (given reasonable assumptions about Herfindahl indexes and demand elasticities).[2] Price-cost margins never approximate those predicted by collusion even in very highly concentrated industries.

Table 1 presents some summary information about industry markups under alternative definitions. Calculations are based on a panel data base of 312 four-digit manufacturing industries from the Census of Manufactures from 1958 to 1981 constructed by Domowitz, Petersen, and myself. Tabulations are averages over that period. To preserve complementarity with Hall's results, I report calculations at the two-digit level of aggregation. The first two columns of results report the labor and materials shares in the value of output; there is considerable variation in the data, but in all cases materials shares are large relative to labor shares. The next two columns contrast Hall's estimates of the ratio of price to marginal cost with direct calculations from the data including materials in marginal cost. The fifth column reports estimates of the price-cost ratio (again including materials) using Hall's modeling approach. Comparing the rankings in the third and fourth columns points up some differences. For example, while industry 28 (chemicals and allied products) has a high price-cost ratio in either case, the same does not hold for paper, primary metals, and transportation equipment, whose price-

2. Ian Domowitz, R. Glenn Hubbard, and Bruce C. Petersen, "Oligopoly Supergames: Some Empirical Evidence on Prices and Margins," *Journal of Industrial Economics* (forthcoming).

Table 1. Price-Cost Markups under Alternative Definitions, 1958–81

Standard industrial classification code	*Industry*	*Labor share in output (percent)*[a]	*Materials share in output (percent)*[a]	*Hall's price-cost ratio*	*Alternative price-cost ratio*[a]	*Estimated price-cost ratio*[b]
20	Food and kindred products	5.7	46.6	3.09	1.31	1.44
21	Tobacco products	8.6	56.6	1.28	1.48	1.72
22	Textile mill products	16.9	58.2	1.05	1.25	1.35
23	Apparel	18.3	51.0	1.30	1.28	1.48
24	Lumber and wood products	17.4	58.2	1.00	1.26	1.39
25	Furniture and fixtures	16.3	41.1	1.38	1.34	1.41
26	Paper and allied products	16.5	52.2	2.68	1.33	1.48
27	Printing and publishing	20.7	31.4	1.61	1.52	1.96
28	Chemicals and allied products	10.0	45.5	3.39	1.60	1.59
29	Petroleum and coal products	6.9	67.1	n.a.	1.28	1.37
30	Rubber and miscellaneous plastic products	20.6	46.2	1.41	1.36	1.56
31	Leather and leather products	19.7	51.9	1.59	1.27	1.39
32	Stone, clay, and glass products	20.1	38.8	1.81	1.44	1.76
33	Primary metals	15.3	59.6	2.06	1.25	1.36
34	Fabricated metals	17.0	45.1	1.39	1.35	1.55
35	Machinery, except electrical	17.0	40.1	1.39	1.37	1.51
36	Electric machinery, electronic equipment	15.4	42.7	1.43	1.44	1.61
37	Transportation equipment	17.5	54.4	2.07	1.24	1.35
38	Instruments and related products	17.1	35.7	1.29	1.54	1.87
39	Miscellaneous manufacturing industries	17.8	46.4	1.52	1.38	1.57

Sources: Author's calculations based on a panel data base of 312 four-digit manufacturing industries. Results are averages over the period 1958–81. Results are averages over that period at the two-digit level of aggregation. The data base is discussed in Ian Domowitz, R. Glenn Hubbard, and Bruce C. Petersen, "Business Cycles and the Relationship between Concentration and Price-Cost Margins," *Rand Journal of Economics*, vol. 17 (Spring 1986), pp. 1–17.

n.a. Not available.

a. The Census price-cost margin (PCM) is used as the basis for the calculation. Most studies in industrial organization construct the price-cost margin only with respect to the value of sales. The value of sales may differ considerably from the value of output because of inventory changes. The *PCM* here is calculated as:

$$PCM = \frac{Value\ of\ sales + \Delta\ Inventories - Payroll - Cost\ of\ materials}{Value\ of\ sales + \Delta\ Inventories},$$

which is identical to (*Value added* − *Payroll*)/(*Value added* + *Cost of materials*), given the Census's definition of value added.

b. Output, not value added, is the quantity unit. Four-digit data were pooled in each two-digit category. Results are based on the standard fixed-effects within-group estimator. See Ian Domowitz, R. Glenn Hubbard, and Bruce C. Petersen, "Market Structure and Cyclical Fluctuations in U.S. Manufacturing" (Northwestern University, 1986).

cost margins are less than that for all industries on average. I have no information on the regulated nonmanufacturing industries comprising the remainder of Hall's industries with "substantial market power."

One contribution of Hall's suggested approach is that it can be applied to consider influences of measures of market power on industry markups.

Table 2. Industry Concentration, Unionization, and Price-Cost Markups, 1958–81[a]

	All industries	*Industry concentration*[b]		*Level of unionization*[c]	
		High	*Low*	*High*	*Low*
Estimated margin	0.352	0.369	0.343	0.279	0.371
	(0.004)	(0.012)	(0.005)	(0.007)	(0.004)
Implied markup ratio	1.54	1.58	1.52	1.39	1.59

Source: Author's calculations.

a. Estimates are based on four-digit Census manufacturing data. Data covering 288 industries (excluding "miscellaneous" and "not elsewhere classified" industries) were used. Fixed-effects estimates are reported above. The equation estimated is

$$\Delta q_{it} - \alpha L \Delta L_{it} - \alpha M \Delta M_{it} = \beta \Delta q_{it},$$

where β is the estimated margin. β is estimated for all industries, for high and low levels of industry concentration, and for unionization. An instrumental-variables procedure is used, with current and lagged values of real GNP growth as instruments. See Domowitz, Hubbard, and Petersen, "Market Structure."

b. "High" refers to industries in which the average four-firm concentration ratio is greater than 0.50; other industries are classified as having "low" measures of concentration.

c. "High" refers to rates of unionization greater than the average for all industries; "low" refers to rates of unionization less than the average for all industries. Union data were taken from Richard B. Freeman and James L. Medoff, "New Estimates of Private Sector Unionism in the United States," *Industrial and Labor Relations Review*, vol. 32 (January 1979), pp. 143–74.

Returning to the four-digit manufacturing data described before, we can consider, for example, effects of industry concentration and unionization on Hall's margins. Table 2 reports revised estimates of the price-cost markup, again relative to labor and materials costs, for "high" and "low" levels of industry concentration, as measured by the four-firm concentration ratio, and of unionization, as measured by the fraction of the work force covered by union bargaining agreements. Markups are higher in concentrated industries, but the effect is not large.[3] Differences by high and low levels of unionization reveal that unions depress margins, a result consistent with previous studies by Richard Freeman, Michael Salinger, and Domowitz, Petersen, and myself.[4] These results are only illustrative, but they suggest that Hall's methodology could be applied to a range of structural models of the influence of product- and labor-market characteristics on industry margins and the cyclical behavior of industry prices and costs.

3. This result is not particularly sensitive to the chosen point of division.

4. Richard B. Freeman, "Unionism, Price-Cost Margins, and the Return to Capital," Working Paper 1164 (National Bureau of Economic Research, July 1983); Michael A. Salinger, "Tobin's q, Unionization, and the Concentration-Profits Relationship," *Rand Journal of Economics*, vol. 15 (Spring 1984), pp. 159–70; Ian Domowitz, R. Glenn Hubbard, and Bruce C. Petersen, "The Intertemporal Stability of the Concentration-Margins Relationship," *Journal of Industrial Economics*, vol. 35 (September 1986), pp. 13–34.

Such findings also point up the possibility that the assumption of a constant markup of price over cost may be inaccurate. For example, the markup might vary over the business cycle, which could explain in part Hall's results, or in response to changes in import competition.[5]

Finally, a crude way of examining Hall's assumption that aggregate productivity shocks are not responsible for the observed pattern in the Solow residual would be to consider data before and after the oil shocks of the 1970s. Though not reported here, if one estimates Hall's basic model (again using panel data on four-digit manufacturing industries) over the 1958–73 and 1974–81 periods, Hall's results still hold for the 1958–73 period. Indeed, the estimated markup ratio is larger.

As noted, a particularly interesting feature of Hall's paper is the proposed reconciliation of high price-cost markups and low average profit rates. Hall's assertion of Chamberlinian competition may well be true for industries engaged in the production of "consumer goods," where product differentiation is important. One can imagine that advertising and investment in building "brand loyalty" are the important fixed costs. What then about the bulk of manufacturing industries engaged in the production of homogeneous "producer goods"? Hall argues that excess capacity is an important feature of the industries studied and that "production profits" are negative. The situation is consistent with equilibrium in an industry because firms must operate at (large) minimum efficient scales.

My reservation about the "excess capacity" argument is that minimum efficient scales in manufacturing are typically quite small, so that it is difficult to imagine an industry equilibrium with substantial excess capacity for this reason alone.[6] Moreover, engineering and economic studies have largely concluded that long-run cost curves at the plant level are much less steep at suboptimal plant scales than suggested by

5. In fact, Census price-cost margins are procyclical in concentrated industries, and spreads of price-cost margins across concentration levels narrowed dramatically in the 1970s. See Ian Domowitz, R. Glenn Hubbard, and Bruce C. Petersen, "Business Cycles and the Relationship between Concentration and Price-Cost Margins," *Rand Journal of Economics*, vol. 17 (Spring 1986), pp. 1–17.

6. F. M. Scherer and others, *The Economics of Multi-Plant Operation: An International Comparisons Study* (Harvard University Press, 1975). See also Joe S. Bain, *Barriers to New Competition* (Harvard University Press, 1956); and the review in F. M. Scherer, *Industrial Market Structure and Economic Performance* (Rand McNally, 1970).

many textbook diagrams.[7] For example, F. M. Scherer and others calculate the percentage increase in unit costs in the long run as a consequence of operating at only one-third of the size of the minimum efficient scale, and find them generally to be not very significant.[8]

Outside of manufacturing, in such industries as trucking, it would be difficult to imagine high minimum efficient scales. For regulated industries in general, the phenomenon of "excess capital" traceable to the Averch-Johnson effect is well known.

A more promising extension would be to consider the possibility of "excess labor" (the labor hoarding described by Hall), based on, say, specific human capital considerations in manufacturing industries or on "excess employment" in regulated industries. If labor were perceived incorrectly as being entirely variable cost, then measured average variable cost would exceed marginal cost, reconciling the issues raised by Hall and explaining the finding I mentioned previously of procyclical Census price-cost margins, which are defined with respect to average variable cost.

The paper concludes with three "lessons for macroeconomics," addressing issues of cyclical productivity, capacity constraints and the shape of the aggregate supply schedule, and incentives to expand to full employment. With respect to the first point, I agree wholeheartedly that more research by macroeconomists on links between market structure and cyclical fluctuations is needed. While agreeing with the finding that price exceeds marginal cost, I would only caution that this spread need not be indicative of substantial realized product market power. I am less sanguine about the policy implications coming from the second point; I am not persuaded by the claim that "numerous firms inhabit market niches with surplus capacity because the constraint of minimum scale is binding." The third lesson is perhaps most interesting in suggesting promising directions for future research.

Hall's calculations of the relatively small impact of sales reductions on profits are very interesting. With constant marginal cost and a constant price elasticity of demand, there seems to be little sensitivity of profit to

7. See Scherer and others, *The Economics of Multi-Plant Operation;* and Leonard W. Weiss, "Optimal Plant Size and the Extent of Suboptimal Capacity," in Robert I. Masson and P. David Qualls, eds., *Essays on Industrial Organization in Honor of Joe S. Bain* (Ballinger Press, 1975), pp. 123–42.

8. Scherer and others, *The Economics of Multi-Plant Operation.*

output movements. The finding is not particularly sensitive to the existence of large markups; for elasticities near unity and for large elasticities, Hall's calculated reduction in profits accompanying a decline in output is small relative to sales. I have two reservations about the interpretation of this finding. First, I am not convinced that the case of increasing price elasticity with increases in output is any more obvious than one of decreasing price elasticities as output increases. Losses would be substantially larger under a linear demand curve, for example. Second, despite Hall's reservations, some measure of the level of initial profit rather than sales would seem to be a better indicator of "large" and "small" effects. The reductions are much larger as a fraction of initial profit (quite large in the linear demand case), a fact presumably of concern to management. The "near-indeterminacy" finding is nonetheless important and suggestive.[9] Industry-level estimates of demand elasticities and marginal cost schedules are needed to sort these issues out.

These reservations notwithstanding, Hall's paper provides an important step in linking methodologies and agendas in macroeconomics and industrial economics.

General Discussion

Christopher Sims applauded Hall's focus on the macroeconomics of imperfectly competitive markets but noted that his quantitative measures of market power are highly unreliable. Although Hall considers but dismisses potential sources of bias by arguing that none is likely to be large, Sims stressed that even small biases may add up to significant estimation problems. Sims also cautioned that period-by-period changes in the wage bill may understate the true marginal costs of production. For example, if a firm expands output but postpones needed maintenance in doing so, the true marginal cost of output exceeds the current period

9. Long-term contracts can also explain price rigidity of the sort noted in Hall's International Nickel example, with associated implications for "output indeterminacy." Robert Weiner and I have considered the role of constant marginal cost in reconciling contracting-based price-rigidity outcomes under competition and monopoly. No unambiguous outcome can be delineated, but, in general, results depend on the slopes of the demand and marginal cost curves. Price stickiness will be relatively greater under monopoly the flatter is the marginal cost curve or the steeper is the demand curve. See R. Glenn Hubbard and Robert J. Weiner, "Nominal Contracting and Price Flexibility in Product Markets," Working Paper 1738 (National Bureau of Economic Research, October 1985).

increase in costs because the true cost also includes the deferred maintenance expenses. Relatedly, Robert Gordon commented that Hall's estimates are likely to be quite sensitive to the time period over which changes in output and labor input are computed. While labor may not be fully adjusted in response to changes in output over a single year, it is likely to be adjusted over two to three years. Taking Hall's measures literally, Gordon continued, one would conclude that the market power of individual Japanese firms, which do not vary employment much in response to short-run output fluctuations, was greater by a factor of two to three than the market power of individual U.S. firms. Gordon found such a conclusion implausible.

Gregory Mankiw noted that Hall's intuitively appealing argument that true productivity shocks could not be so closely correlated across industries as to account for observed fluctuations in aggregate output should apply equally well to long-term productivity growth. Yet there has been an aggregate productivity slowdown since the mid-1970s, which suggests that productivity developments may be correlated across industries. There could, therefore, be something to the notion that productivity shocks are a source of business cycles; if so, it would mean Hall's marginal cost estimates are wrong. William Nordhaus suggested that, in attempting to estimate the response of output and labor to demand changes, it would be better to use variables, such as defense spending, that are unambiguously demand shifts, rather than GNP. Major macroeconomic shocks, such as OPEC price changes, apparently affected productivity trends across many industries during the 1970s and early 1980s. If so, the GNP variable is, at least in part, capturing supply rather than demand shocks.

Lawrence Summers found it easy to believe that most firms would be delighted to sell more output at current prices, implying that most markets are characterized by monopolistic competition, as Hall suggests. However, Summers doubted that the excess capacity associated with monopolistic competition reflects minimum efficient scales of operation. Summers's preferred explanation is that physical plants that are large enough to meet peak demand are less than fully utilized at nonpeak periods.

Robert Pindyck was surprised at some of the industries Hall classifies as having substantial market power. The Chamberlinian monopolistic competition model, which relies on product differentiation, does not seem applicable to coal, paper, primary metals, and stone, clay, and glass; yet all are categorized by Hall as industries in which firms have substantial market power. The paper would be more convincing, Pindyck continued, if there were more discussion of the source of market power for the industries with high price-cost margins. Summers noted that the absence of a strong correlation between Hall's price-cost margin estimates and conventional concentration estimates does not necessarily undermine Hall's estimates. Previous work by Michael Salinger has shown that Tobin's q and concentration have no systematic relationship until unionism is taken into account, revealing that there are large rents in concentrated sectors but that unions capture them.

Kenneth Judd cautioned against jumping too quickly from positive to normative conclusions. The recent theoretical research cited by Hall has indeed shown that Chamberlin's imperfect competition theory could be modeled formally, but it has also shown that what Chamberlin labeled excess capacity was in fact not inconsistent with efficiency.