Center for Research on Economic and Social Theory
and
Department of Economics
Working Paper Series

Estimating the Cyclicality of Marginal Costs Directly

Susanto Basu

January, 1993
Number 93-18

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109-1220

Revised January 1993

ESTIMATING THE CYCLICALITY OF MARGINAL COSTS DIRECTLY

Susanto Basu

Department of Economics
University of Michigan
Lorch Hall M106
Ann Arbor, MI 48109-1220

I would like to thank Brad DeLong, John Fernald, and Greg Mankiw for very helpful comments and advice. The usual disclaimer applies.

Abstract: This paper studies the cyclical behavior of marginal costs by examining changes in a constructed measure of total costs in response to changes in output. The main result is that marginal costs in U.S. manufacturing appear close to constant at annual frequency. This conclusion contrasts with previous work that has found marginal costs to be strongly procyclical. The finding of non-increasing marginal cost is robust to considerations of various types of bias, and issues of cyclical measurement error. The methodology of the paper is shown to be valid under different assumptions about economic behavior such as labor hoarding, increasing marginal prices of inputs, and the existence of fixed costs of production. The results do not support the hypothesis that marginal costs are constant over the business cycle because of large productive spillovers in manufacturing that are external at the two-digit level.

Whether marginal costs of production increase with the quantity of output produced is an issue whose resolution affects a number of questions of interest in modern macroeconomics. This paper attempts to shed light on these issues by taking a direct approach to estimating the cyclicality of marginal costs.

The slope of "the" marginal cost curve is a central issue in judging the plausibility of New Keynesian models of output fluctuations that are based on the idea that small "menu costs" of changing prices lead to nominal price rigidity. These models were based on the idea that small menu costs can explain large business cycles (Mankiw, 1985); Akerlof and Yellen, 1985). But it quickly became clear that if this result were to hold for non-negligible shocks firms must face marginal costs of production that are close to constant. If producers face sharply increasing marginal costs, either because of decreasing returns to scale or because of quickly increasing marginal disutility of labor, then the size of the output change that will be tolerated before producers change prices is likely to be quite small (Ball and Romer, 1989; Blanchard and Fischer, 1989, ch. 8). On the other hand, if marginal costs are close to constant, large output changes can plausibly be viewed as resulting from changes in aggregate demand — a vindication of the theoretical plausibility of traditional Keynesian views about output fluctuations.

A second issue whose resolution would be influenced by determining the degree to which marginal costs are procyclical is the behavior of inventories. As is well known, inventory investment is extremely volatile, often accounting for a high percentage of fluctuations in output. The volatility of inventories is, however, matched by economists' uncertainty about the determinants of inventory investment. The traditional model of inventories is that they are used to smooth production (Blinder, 1986) — a view that is sensible only if production costs are strictly convex. If, on the other hand, costs of production are either linear or nonconvex, the production smoothing model loses its *raison d'être*, and models of inventory behavior based on avoidance of stock-outs (Kahn, 1987) or the usefulness of inventories as factors of production (Ramey, 1989) must be invoked.

Finally, the cyclical behavior of marginal costs can shed light on the relevance of some supergame-theoretic models of oligopolistic competition for macroeconomics. Rotemberg and Saloner (1986) argue that supergame equilibria in oligopolistic industries may require firms to have

countercyclical markups.[1] Rotemberg and Woodford (1991) expand on this idea and claim to find evidence for countercyclical markups. However, Hall (1991) points out that their result depends on the assumption that the elasticity of labor demand is known a priori. With a different identifying assumption, Hall finds evidence of flat or negatively sloped marginal costs (and hence by implication acyclical or procyclical markups). I try to resolve this issue by arguing that one can directly measure profits, one of the variables that Rotemberg and Woodford treat as an unobservable. Given an estimate of profits, and hence of total costs, I can then try to address the cyclicality of marginal cost from the cost side rather than from the production side. Rotemberg and Woodford also assume that the production function has constant returns to scale after an initial fixed cost. I do not impose a similar assumption on the data. Clearly, if firms have increasing or decreasing returns in the variable parts of their production functions, their marginal costs may be very different than if they had constant returns.

In this paper, I try to resolve these important issues by estimating the cyclical behavior of marginal costs. The question of the cyclicality of marginal costs has, of course, been addressed previously. Two notable efforts are the papers by Bils (1987) and Ramey (1991).

Bils argues that the legally mandated overtime premium of 50 percent of normal wages represents a significant cost to employers of increasing labor hours, yet employers are observed to increase hours in booms. Since cost minimization implies that marginal costs can be inferred from optimizing behavior at any margin, Bils concludes that marginal costs increase sharply with output. However, Bils's procedure is identified only if there are convex costs of changing employment, as well as hours. This assumption is debatable: it is equally plausible to argue that processing employment applications requires large fixed costs (a new section of the personnel department), but the marginal cost of processing each application and each new employee is constant.

Ramey, estimating a structural model of output and inventory behavior, arrives at just the opposite conclusion. Her result can be viewed as a careful examination of the well-known stylized fact that the

1 The intuition is that if markups are too high during a period when demand is also high, it might be profitable for a firm to produce more output than the proposed equilibrium quantity. If rates of discount are sufficiently high, firms might cheat even if they would be punished (perhaps with a price war) in subsequent periods. So markups are countercyclical because of the requirement that firms' behavior be individually rational in every period of a supergame equilibrium.

variance of production is greater than the variance of sales. She allows for expected changes in costs of labor and materials, a stockout-avoidance motive, and costs of adjusting output. After controlling for cost shocks by using instruments, Ramey concludes that there is no production-smoothing incentive in the six industries she studies. In fact, there seems to be an incentive to bunch production, from which she concludes that marginal costs must be declining in output. Ramey's estimation procedure does not allow for fixed costs, however. It is likely that some of her findings of decreasing marginal costs stem from treating all costs as variable; essentially, she forces the cost function through the origin, so her finding of decreasing marginal cost may stem from a finding of decreasing average cost with constant marginal costs.

This paper presents a third, much more direct, approach to estimating the cyclical behavior of marginal costs. Essentially, I calculate an estimate of total costs and then examine the effect of changes in output on changes in costs. The strength of this method is that the simplicity of the procedure allows me to proceed with a minimum of auxiliary hypotheses. Also, since the methodology essentially involves constructing an estimate of marginal costs, I am then free to parametrize them as functions of any variables that might seem reasonable. Unlike Ramey's work, I allow for fixed costs in production, and model theproblems in estimating marginal costs that stem from the existence of fixed costs.

Viewed in this light, this paper is related to the literature in the industrial organization literature on estimating profit margins. In particular, it is related to the papers by Domowitz, Hubbard, and Petersen (1986a,1986b,1987) which study the behavior of price-cost margins. It is different from those papers, however, because it focuses on the cyclicality rather than the level of costs. Domowitz *et. al.* are interested in the level of profitability of firms in the industries they study. Their focus on the level rather than the change of costs leads them to make some unpalatable assumptions. First, they assume that marginal cost is equal to average cost. Bils (1987) demonstrated the potential dangers of this procedure. Second, they assume that all payments to capital represent fixed costs and all payments to labor and materials are variable costs. Obviously, it is unreasonable to consider all of capital as a fixed cost (which would imply that the quantity of capital is independent of the level of output). Their procedure also fails to

allow for the possibility of fixed costs from overhead labor or from advertising (which is counted as materials usage).

The paper is structured in five sections, the first of which is this introduction. Section I discusses the theory and empirical implementation of the model. Issues that might present difficulties for the econometric procedure employed are discussed in Section II. Section III examines the robustness of the estimation method to alternative hypotheses about the economic environment. Section IV discusses the data and estimation procedures employed. Section V presents the results. The last section offers some concluding comments.

I. The Basic Model

The theoretical framework for the estimation is very simple, involving few behavioral assumptions. It starts from the definition of economic profits for a firm i:

$$\pi_i \equiv P_i Q_i - VC_i - FC_i . \qquad (1)$$

Here P_i is the relative price of output in industry i, Q_i is the output produced, VC_i represents variable costs, and FC_i fixed costs. Clearly, if economic profits and fixed costs were observable, one could simply calculate variable costs for any observed level of output. Changes in variable costs attributable to changes in output would then yield an estimate of the marginal cost curve. As noted above, there is no presumption in equation (1) that fixed (or variable) costs are associated with certain inputs and not with others.

Noting that variable costs are just the sum of marginal costs of production, we can rewrite equation (1) as

$$P_i Q_i - \pi_i - FC_i = \int_0^{Q_i} MC(x,z)\,dx . \qquad (2)$$

Here marginal costs are assumed to be a function of output and of a shift factor, z, representing the level of technology.

Taking differences of (2) over time gives the basic idea of the paper:

$$d\,(P_i Q_i - \pi_i) - d\,FC_i = MC(Q_i, z)\,dQ_i + \left(\int_0^{Q_i} \frac{\partial MC(x,z)}{\partial z}\,dx \right) dz. \qquad (3)$$

Equation (3) just says that the change in total costs (total revenues minus profits) minus the change in fixed cost, equals the marginal cost of production times the change in the quantity produced, plus the change in variable costs resulting from changes in technology. This equation establishes the desired link between the observables, revenues and profits, and the unobservable quantity of interest, marginal cost. The terms representing changes in costs for reasons other than changes in quantity are nuisance parameters that must be dealt with in the estimation.

It is clear from equation (3) that this procedure allows us to estimate directly the level of marginal cost, rather than the slope of the marginal cost curve. But by parameterizing marginal cost as a function of quantity produced, it is easy to find how marginal cost of production varies with quantity.

It will be convenient to do the estimation in logs. Rearranging the terms in equation (1) and expressing them as log differences gives the following:

$$d \ln (P_i Q_i - \pi_i) \approx \left(1 - \frac{FC_i}{P_i Q_i - \pi_i}\right) d\left(\frac{FC_i}{P_i Q_i - \pi_i}\right) + d \ln (VC_i)\,. \qquad (4)$$

The approximation in this expression comes from neglecting the cubic and higher order terms in the Taylor expansion of ln(1+x). Abstracting from changes in technology, equation (4) shows that the growth of total costs equals the change in the fixed to total cost ratio (multiplied by one minus that ratio) plus the marginal cost of production multiplied by the change in quantity, now expressed as a fraction of total variable cost.

Equation (4) will be the basis of all the estimation done in the paper. In order to implement it econometrically, however, some economic and parametric assumptions are required.

First, a method must be devised to calculate economic profits. It is not difficult, following Hall and Jorgenson (1967), Hall (1986a, 1990) and Caballero and Lyons (1990, 1991), to compute a series for the user cost of capital, r. Total payments to capital are observable, as the residual of revenues after

payment of all other factors of production and taxes. Therefore, economic profits may be computed as the total payments to capital net of the user cost of capital, which is just rK.[2]

Unobservability of fixed costs is a greater problem. For the moment, we can assume that the ratio of fixed to total costs is well approximated by a constant. This assumption can be justified by assuming that the economy is on a long-run balanced growth path (hence fixed and total costs grow at the same rate) and that shocks to the growth rate of fixed and total costs are identical. In this case, of course, the change in the ratio of fixed to total costs is constant at zero. This assumption is obviously a very strong one; the consequences of relaxing it are examined below.

The other question is how one should parameterize the cost function. In this paper, I assume that the elasticity of marginal cost with respect to output is constant. This implies that

$$\frac{\partial^2 C}{\partial Q^2}\,\frac{Q}{\partial C/\partial Q} = \gamma^{-1} \qquad (5)$$

where C(Q) is the firm's cost function and γ is a parameter. Solving this equation for the implied cost function gives the conclusion that the cost function must be of the form

$$C(Q) = B + AQ^{\gamma} \qquad (6)$$

where A and B are constants with respect to Q. The parameter B will be interpreted as a fixed cost. AQ^{γ} is the variable cost function.

We still have to take into account the fact that variable costs are influenced by factors other than the quantity produced. The simplest parameterization of variable costs is that the cost of producing a given level of output depends on the relative price of inputs, the quantity of output produced, a term reflecting technological progress, and an error term which reflects shocks to technology:

$$VC_i = \exp(-\alpha_i\, t)\, \exp(\phi_i)\, P_{x,i}\, A_i\, Q_i^{\gamma_i} \qquad (7)$$

---

2 Using the Hall-Jorgenson procedure is not completely innocuous. Required payments to capital are computed by multiplying the required rate of return by the current-dollar value of the capital stock. Using the Hall-Jorgenson procedure requires one to assume that the current value of the capital stock is correctly calculated, i.e. that the current value is based on the price that the capital stock would fetch in a perfect market for (potentially used) capital goods. To the extent that this assumption is not satisfied, one faces the problem pointed out by Fisher and McGowan (1983) inherent in trying to estimate economic rates of profit from data on accounting profits and a book value of capital that is not necessarily related to the economic value.

Here α represents technological progress, which is presumed to lower variable costs for a given quantity of output produced, and ϕ_i represents technology shocks. A_i is an industry-specific constant. $P_{X,i}$ is a price index of input costs for industry i. The parameter of interest is γ_i. $\gamma_i = 1$ implies that marginal costs are constant in output in industry i. $\gamma_i > 1$ implies increasing marginal costs, and $\gamma_i < 1$ implies decreasing marginal costs. As noted above, $(\gamma$-1) is the elasticity of marginal cost with respect to output.

Taking differences of (7) and substituting into (4) gives the estimating equation:

$$d \ln (P_iQ_i - \pi_i) - d \ln (P_{X,i}) = (c_i - \alpha_i) + \gamma_i \, d \ln Q_i + \varepsilon_i \; . \qquad (8)$$

Here c_i is the constant reflecting changes in the ratio of fixed to variable costs in industry i, and ε is the error term that results from taking the difference of ϕ. For purposes of estimation, it is assumed that the stochastic process followed by ε is stationary, i.e., there is at most one unit root in output.

II. Econometric Issues

A. Aggregation

The theory that has been developed is based on the behavior of a firm. The estimation, however, will be done with industry data. A relevant question, therefore, is to ask how well this type of model aggregates, and discover what relation the estimated parameters bear to the concepts that are of economic interest.

The elementary theory of a perfectly competitive industry makes clear that the long-run industry supply curve need not bear a direct relation to firms' short-run marginal cost curves, where "long-run" and "short-run" are defined relative to the length of time that is needed for entry and exit to take place.

Suppose all firms have positive fixed costs and increasing marginal costs, so they have U-shaped average cost curves and produce, in equilibrium, at minimum average cost. If the time period under question is long enough to permit entry and exit and if the scale of the industry does not affect firms' costs, the industry will appear to have constant marginal costs — as indeed it will, since it produces more or less output by having a smaller or larger number of firms, each of which produces at the same marginal cost. But the fact that the industry has constant costs does not provide much information about

the slope of the firm-level marginal cost schedule — at the firm level, marginal costs may be monotonically increasing in output.

So while the short-run supply schedule of a competitive industry (or cost schedule, which is what is being estimated here) should just replicate firms' marginal cost curves, this will not be true over any horizon that allows for entry and exit.[3] Since the estimation here is done with annual data, it is undoubtedly the case that the industry will not have a fixed number of plants over the length of a year.

How relevant is this problem for issues of economic interest? The answer must depend upon the specific question, but one general issue that appears relevant is the seldom-made distinction between a firm and a plant. To make matters concrete, suppose that the slope of the marginal cost curve is at issue because one wants to judge the plausibility of menu-cost models of business cycles. In these models the slope of the marginal cost curve is relevant because, assuming a constant desired markup of price over marginal cost, the loss of profit is directly related to the (square of the) change in marginal cost from a change in output. What is relevant for this issue is the cost to a firm (an administrative unit) of producing more output, not the cost to a plant (a productive unit) of producing more output. If a firm faces an increased demand for its output because of sticky prices, it may well operate another plant and produce the desired output at constant marginal cost, even though plant-level marginal costs are sharply increasing in output. So if one allows for the possibility of multiplant firms, the industry-level marginal cost curve may be a more relevant concept for the question at issue than the plant-level marginal cost curve.

Nevertheless, the question of whether the estimates obtained using the procedure outlined above reflect a large amount of entry and exit is an interesting one. Without access to higher frequency or more disaggregated data, it is not possible to check whether entry and exit are relevant on time scales of less than a year. However, it is possible to redo the estimation at *lower* than annual frequencies and see whether the estimates change in a way that we would expect if changes in the number of plants have a

---

[3] Ramey's (1989, 1991) claim of successful aggregation from firm level parameters to industry level data implicitly assumes a constant number of productive units (or at least a number of firms that does not vary in a cyclical fashion).

large influence on marginal costs. From knowledge of behavior at frequencies of less than a year, it may be possible to make some educated guesses about the importance of entry over shorter time horizons.

B. The Treatment of Fixed Costs

So far, I have assumed that fixed costs happen to move in lockstep with variable costs. Obviously this happy state of affairs is unlikely. What are the consequences for the estimate of γ when this assumption is not satisfied?

There are two consequences. Abstracting for the moment from changes in relative input prices, the methodology of this paper is essentially to construct an estimate of total costs and then to construct an estimate of marginal costs by looking at the ratio of percent changes in total costs to percent changes in output. What would ideally be done, of course, is to obtain an estimate of *variable* cost, and then calculate marginal cost by examining the change in variable costs from a unit change in output. Substituting the first procedure for the second can result in either underestimating or overestimating the extent to which variable costs change in response to a change in output.

Suppose fixed costs were to increase at a time when output was increasing. Then the procedure of the paper would attribute the entire change in costs to the increase in output, thereby making it seem that marginal costs are sharply increasing. On the other hand, suppose output were to increase without an increase in fixed costs. Suppose that marginal costs are constant, so variable costs increase one for one with output. Then total costs will increase less than one for one, and the procedure here will report that marginal costs appear to be falling — since costs seem to increase less than one for one with output — when marginal costs are in fact constant.

We can attempt to get an idea of the biases by modeling explicitly the growth of fixed and variable costs. Suppose that the economy is on a balanced growth path where all variables — including fixed and variable costs — grow at the same rate.[4] Suppose also that there are independent shocks to the growth rate of fixed and variable costs. Then

[4] If the growth rates of fixed and variable costs differ, then the ratio of fixed to variable costs goes asymptotically to zero or infinity. In either case, as demonstrated below, the bias from unobservability of fixed costs goes to zero. So the assumption of a constant average ratio of fixed to variable costs is a worst-case scenario.

$$\frac{dFC}{FC} = \beta + \eta \tag{9a}$$

and

$$\frac{dVC}{VC} = \beta + \nu \tag{9b}$$

where η and ν are independent shocks with mean zero. It follows that

$$d\left(\frac{FC}{FC + VC}\right) = \eta - \nu. \tag{10}$$

The bias in the estimate of γ therefore is proportional to the covariance between percent changes in output and $(\eta - \nu)$, with a constant of proportionality depending on the ratio of fixed to variable costs. Recall from equation (4) that $(\eta - \nu)$ is also multiplied by a series of terms that depend on the ratio of fixed to variable costs.

From the model of variable costs presented in equation (5), the covariance of ν with percent changes in output is just γ times the variance of changes in output. The covariance of changes in output with shocks to the growth of fixed costs is harder to calculate. However, since the constant term of the regression is equal to the mean over the sample period of $(\eta - \nu)$, minus α, the rate of growth of technological progress, it is possible to use the constant as a lower bound on the rate of growth of fixed costs in excess of trend. In all the regressions the constant is found to be very small, but it is almost always positive. It is certainly possible for the excess growth of fixed costs to be positive but negatively covarying with output, but this sounds rather implausible. At worst, therefore, one would probably assume that the covariance of shocks to the growth rate of fixed costs was uncorrelated with output changes.[5,6]

It is clear, however, that shocks to the growth of variable costs create a negative bias in the estimate of γ. From the discussion above, it follows that

[5] If one thinks of fixed costs as being literally independent of the level of output at all times, then the covariance must be zero.

[6] Even if fixed costs at the firm level are uncorrelated with output changes, fixed costs at the industry level will be positively correlated with changes in output. The reason is that, with procyclical entry, the total of industry fixed costs will increase when output increases. Therefore, the formula for the bias derived below, which is done under the assumption that the covariance of fixed costs with output is zero, should be interpreted as giving an upper bound for the true bias.

$$\text{plim}\,\gamma^{e} = \gamma\left(1 - \left(1 - \frac{1}{1 + VC/FC}\right)\left(\frac{1}{2 + VC/FC + FC/VC}\right)\right) \qquad (11)$$

where γ^{e} is the estimated γ. From this formula, it appears that the maximum bias occurs when fixed costs are just equal to variable costs, and is approximately 12.5 percent. Extending the Taylor approximation of ln(1+x) to the fifth term gives an upper limit to the bias of about 17 percent.[7] The actual size of the bias expected depends on one's estimate of the average ratio of fixed to variable costs in manufacturing.

One way to examine this issue is to decide on plausible figures for the average markup and the average rate of pure profit (as a percentage of sales) in U.S. manufacturing. The average rate of pure profit is widely held to be low, perhaps even zero. The profit figures calculated here support that assertion. If profits are close to zero, but prices exceed marginal cost on average, then the revenue in excess of variable costs must be consumed by fixed costs. To put it another way, if the rate of profit is zero, then the ratio of fixed to average costs must be given by (μ^{*}-1), where μ^{*} is the markup on *average* variable costs. The high bias case corresponds to $\mu^{*} = 2$ with an assumption of zero profits.

Domowitz, Hubbard, and Petersen (1988) used highly disaggregated panel data on four-digit S.I.C. manufacturing industries and estimated the markup of price on *marginal* cost to be about 1.6.[8] Their figures are more appropriate than the ones obtained by Hall (1986b, 1988), since Domowitz *et. al.* use gross output data, as is done here, rather than following Hall and using value-added. Assuming that the rate of profit is zero, a markup of 1.6 implies that the estimated γ would be biased down by 17 percent (very close to the maximum bias). Assuming that the rate of profit is 10 percent, the bias is reduced to 16 percent. So changes in the assumed rate of profit, over a realistic range, do not greatly change the expected bias. On the other hand, changes in the average markup do have a large effect. Assuming an

7 Since the expansion of ln(1+x) is an alternating convergent series for x<1, it follows that taking the expansion to a last term that is positive term leaves a net negative remainder. The fifth term of the expansion is positive.
8 If there are no fixed costs, then using the markup on marginal cost in place of the markup on average cost overstates the bias if marginal costs are falling in output and understates it if marginal costs are rising in output.

average markup of 1.1 (what Rotemberg and Woodford, 1991, term their "extremely conservative" case), the expected bias is about 8 percent, even assuming zero profits.[9]

A low estimate of the markup (on the order of 1.1) may be the correct one, however. The procedure used by Hall and Domowitz *et. al.* assumed that firms have constant returns to scale. A methodology developed by Caballero and Lyons (1990,1991) estimates the degree of internal returns to scale, and finds it smaller than one. Using their estimate of the internal returns to scale and an estimate of the rate of profit, one can derive a figure for the average markup. Hall (1986b) calculates the average profit of U.S. manufacturing to be about 12 percent. Based on these figures, estimates of the markup are uniformly lower than 1.1, implying that fixed costs (and thus the bias from ignoring them) must be very small.[10,11]

C. Pure User Cost of Capital

The method I use to obtain an estimate of total costs is to decompose payments to capital into the rental cost of capital and pure profit, and then to subtract profits from total revenues. When calculating the rental cost of capital, I assume is that the pure user cost of capital is zero — that is, that capital depreciates strictly as a function of time, and not as a function of use. To the extent that this assumption is false, the model is not correctly specified. This misspecification does not pose as many problems as the unobservability of fixed costs, however. The reason is that the omitted variable is now changes in the ratio of user cost of capital to total cost. Since both the numerator and the denominator change in the same direction in response to changes in output, the change in their ratio is likely to be small. In fact, as this discussion makes clear, even the sign of the bias that results from ignoring pure user costs of capital is not theoretically determinate.

9 The bias figures are all based on taking the Taylor approximation to the fifth term. Thus, in each case, they represent an upper bound.
10 Basu and Fernald (1991) repeat Caballero and Lyons's tests with gross output data and find much the same results. In fact, the problem with these estimates of μ is that many of them are too low. For some of the results, both papers find that μ is less than one, which is incompatible with (monopolistic) profit maximization.
11 As will be seen below, the results of the estimation here imply that firms may have increasing returns to scale. But for plausible magnitudes of the degree of returns to scale, and the low profit figures obtained here, the markup will still be low. Bils (1991) finds returns to scale on the order of 1.15. With zero average profits, this implies that μ^{*} is also 1.15, not far from the figure of 1.1 suggested here.

For simple models of the pure user cost of capital, the expected bias is generally not very large. Suppose capital depreciates exponentially over time and in relation to intensity of use, so that at each instant of time the rate of depreciation is

$$\delta_0 + \delta_1 \frac{Q}{K} \qquad (12)$$

In that case, the bias from ignoring the depreciation in relation to use is

$$-\delta_1 T \left(1 - \gamma \frac{VC}{VC + FC}\right) \qquad (13)$$

where T is a term reflecting tax adjustments to the required rate of return on capital and is approximately equal to one.[12]

As discussed in the previous section, there are good reasons for believing that the ratio of variable to total costs is close to 1. Therefore, unless δ_1 is large, the bias is likely to be small, unless γ is very different from one. But it is unlikely that pure user cost is an important component of depreciation (that is, δ_1 seems to be small). Hulten and Wykoff (1981) examined the age-price profiles of a variety of types of capital goods and concluded that the exponential decay assumption very closely fit the data. So deviations from constant exponential depreciation, if they exist, are not very large. Bils (1991) estimates from an examination of used car prices that a maximum of one third of depreciation results from use. Given that total depreciation averages 0.13 per year, this suggests that δ_1 is quite small.

D. Expected Capital Gains and Adjustment Costs

As will be seen below, the computed rate of return to capital, which is subtracted from the total of payments to capital to yield an estimate of economic profits, does not include a term for expected capital gains. Any bias in the estimate of γ resulting from omitting capital gains should be positive, however. Capital gains are strongly procyclical, so it is reasonable to assume that expected capital gains are also procyclical. In that case, neglecting to subtract a term reflecting expected capital gains will lead one to

12 This expression assumes that the covariance between changes in the growth rate of output and shocks to the growth rate of fixed costs is zero. Under the assumption that the growth of fixed costs is perfectly positively correlated with output growth, the term in parentheses would be $(1-\gamma)$.

conclude that profits are less procyclical than they in fact are, and thus to conclude that total costs are more procyclical than is actually the case.

The Hall-Jorgenson formula for the required rate of return to capital is computed under the assumption that the capital stock can be adjusted costlessly. A series of papers deriving microeconomic foundations for Tobin's (1969) "q" theory of investment have assumed that adjustment costs are positive and convex. If the adjustment costs are modeled as proportional to the ratio of investment to the capital stock (as in Blanchard and Fischer, 1989, ch. 2), then the required rate of return to capital is less than that predicted by the Hall-Jorgenson formula. The difference comes from the fact that an investment in an additional unit of capital increases the size of the capital stock and, *ceteris paribus*, lowers future costs of investment.[13] Since the investment/capital ratio is strongly procyclical, neglecting the adjustment-cost-savings term leads to the same type of bias as ignoring capital gains. In both cases, assuming increasing marginal costs to be the null hypothesis, ignoring expected capital gains and costs of adjustment work against rejecting the null.

III. Different Assumptions about Variable Costs

The model presented in Section I may seem very special, but it is applicable under quite general circumstances. This section investigates the robustness of the model to several common assumptions about the behavior of firms over the business cycle. In some cases minor modifications to equation (8) have to be made, but these turn out to be straightforward to accommodate within the basic framework of this approach.

A. Labor Hoarding

Suppose that industry i is characterized by complete labor hoarding and (for simplicity) assume that labor is its only input in production. In this case, the industry has only fixed costs and a marginal cost of production of zero. Would estimation of equation (8) give these results? In particular, this might

13 This effect is not present in a model where the cost of adjustment is solely a function of the absolute size of investment, as in Abel (1981).

seem to be a problem since by taking wages as part of the input cost index, it may appear that labor is forced to be a variable factor.

Labor hoarding does not cause problems for the econometric procedure employed here, however. As long as there are no changes in wages, treating labor as a fixed input implies that changes in revenue translate one for one into increases in profits. But since the first term on the left hand side of (8) takes the difference of revenues minus profits, there would be no change in the dependent variable from an increase in output. If there are changes in wages, the first term on the left hand side will be different from zero — changes in revenues will not equal changes in profits. But the second term will adjust for the fact that all of the difference in revenues minus profits is due to the change in wages, rather than any increase in the use of inputs resulting from the higher level of output. So once again the dependent variable will take on the value of zero. Therefore, one would expect that any error-minimizing econometric procedure would set the estimate of γ to zero, which is the correct result for an industry that always has a marginal cost of zero.

B. Increasing Marginal Input Prices

As noted above, Bils (1987) based his inference of sharply increasing marginal costs on the fact that manufacturing industries are legally required to pay a 50 percent overtime premium to workers who work more than full time. He observed that, despite the premium, many industries still make use of considerable overtime labor during booms. Since using overtime is presumed to be very expensive, but such an action was taken by firms that, by assumption, are minimizing costs, Bils concluded that the marginal cost curve must be steep. Bils stressed that although straight time wages are only mildly procyclical, or even acyclical, the *marginal* wage facing manufacturers is sharply increasing due to the overtime premium.

To accommodate Bils's insight, it is necessary to allow for the fact that the cost of labor (or, more generally, all inputs) might be increasing faster than would be predicted simply by the observed input price index. Hence, the correct parametrization of variable costs is

$$VC_i = \exp(-\alpha_i t)\, \exp(\phi_{i,t})\, P_{X,i}\, [1 + Q_i^{\theta_i}]\, A_i\, Q_i^{\Psi_i} \qquad (14)$$

The term multiplying $P_{X,i}$,$1+ Q_i^{\theta}$, represents the fact that increasing output may increase the marginal cost of inputs to an extent greater than is captured by the input price index. If labor were the only input to production, then $P_{X,i}$ would be the straight wage and $1+Q_i^{\theta}$ the marginal wage schedule. The marginal wage schedule would be a function of the overtime premium and the percentage of workers who must take overtime in response to a given increase in output. Even allowing for the Bils correction, however, the estimating equation is essentially unchanged. Based on the cost function (7), the new equation is

$$d \ln (P_i Q_i - \pi_i) - d \ln (P_{X,i}) = (c_i - \alpha_i) + (\theta_i + \psi_i)\, d \ln Q_i + \epsilon_i \qquad (15)$$

Now the coefficient on $d\ln Q_i$ represents not only a technological fact — whether a one percent change in output demands more or less than a one percent change in inputs — but also reflects the fact that an increase in output requires an increase in the marginal cost of inputs. The sum of these two effects is, however, correctly regarded as being a true measure of marginal costs. The marginal cost facing a firm that contemplates an increase in output reflects both the percent change in inputs required to effect a one percent change in output and the fact that an increase in output will require paying higher input prices for the marginal inputs used.

C. Do Prices Reflect Costs?

A number of authors have argued that markets for inputs such as intermediate manufactured goods are characterized by long-term relationships between buyers and sellers. These markets may clear along several dimensions (quality, time to delivery, etc.), of which price is only one. Evidence for this hypothesis is provided by the fact that delivery lags in manufacturing are strongly procyclical (Carlton, 1987), and by the fact that industries that are more likely to be in long-term relationships with buyers have smaller price changes and larger quantity changes in response to demand shocks (Bradburd and Caves, 1987). Bradburd and Caves argue that in relationships characterized by long-term arrangements, the observed spot prices are not necessarily the implicit transactions prices specified by the long-term contract.

This point is, of course, a familiar one in the context of payments to labor and the debate over whether the observed real wage is allocative. Hall (1980) critiqued the Keynesian "sticky nominal wage" literature of the 1970s by arguing that reasonable employers and workers in a long-term relationship should avoid deadweight losses from unnecessary output fluctuations. With observable nominal shocks, or in a world where real shocks are not important, the employment contract should specify a constant level of hours worked, with spot wages an installment payment on the long-term wage bill.

The arguments of both the industrial organization literature and the long-term labor contracting literature present problems for estimating changes in marginal cost with respect to output.[14] The problem seems less acute insofar as it pertains to intermediate inputs. The evidence of Bradburd and Caves is consistent with the interpretation that suppliers compete with each other on the basis of factors other than price, and one of these factors is the ability to increase supply to the downstream industry when it experiences a change in demand. So while suppliers of inputs may be in long-term agreements with their customers, this does not imply that the customers perceive the real cost of their inputs to be anything other than the prices they pay.[15,16]

There is strong evidence that matters are different when it comes to labor. Recently Trejo (1991) examined the issue of the required overtime premium that Bils relies upon to show that the marginal wage is procyclical. Trejo asked whether the required overtime payments are allocative, i.e. whether they increase total wage payments or whether they simply reduce the straight wage while leaving the wage bill unaffected. He found using one method that approximately one half of the overtime premium is offset through changes in the straight wage. Another method implied that the offset is in fact complete. He concluded that the offset is substantial, though not necessarily complete.

The method used here is less vulnerable to Trejo's criticism than is Bils's. Bils's method is invalid if changes in the wage from overtime payments do not reflect changing labor costs. The evidence presented by Trejo is that the wage is adjusted downward to compensate for the required 50 percent premium on a week-to-week basis. But if the adjustment to the straight wage takes place over a period

[14] It should be noted that these problems are not unique to the present paper. The estimation done by both Bils (1987) and Ramey (1991) use prices of labor and other inputs as measures of true costs.

[15] It may be the case, of course, that one party or the other has to pay a premium for engaging in this stable relationship.

[16] An independent problem is that reported intermediate input prices are not always the true transactions prices. Stigler and Kindahl (1970) document the existence of differences between the indices of quoted prices as compiled by the BLS and actual transaction prices.

of time of less than one year, the annual wage bill should be substantially unaffected. This is all that is required for the methodology of this paper to be correct. But, in principle, one could have very long-term implicit contracts where the wages paid average to the true cost of labor over a period of years rather than weeks. If that were the case, long-term contracts would present a problem for measuring the true cost of labor input.

D. Productive Externalities from Aggregate Activity

A substantial volume of theoretical work has suggested that there may be external economies in production.[17] That is to say, an increase in the aggregate level of output may lead to an increase in productivity of individual firms or industries. Some evidence that this effect is empirically large in both U. S. and European manufacturing has been presented by Caballero and Lyons (1990, 1991). If it is the case that changes in aggregate output affect sectoral productivity, then a model that assumes marginal costs are only a function of own output is misspecified. Suppose the true model is

$$d \ln (P_i Q_i - \pi_i) - d \ln (P_{X,i}) = (c_i - \alpha_i) + \gamma_i \, d \ln Q_i + \kappa_i \, d \ln Q + \varepsilon_i , \qquad (16)$$

where $d \ln Q$ is the percent change in aggregate output. Suppose, however, that instead of estimating equation (16), we estimate equation (8). Given that previous work leads us to expect that κ is negative (an increase in aggregate output lowers sectoral costs) and the correlation between sectoral and aggregate output is certainly positive, standard omitted variable bias reasons would lead us to expect that the estimates of γ obtained by estimating equation (8) would be biased downward.

Actually, given that the true model is (16), the estimate of γ from (8) is a useful quantity. It measures the increase in costs from a one percent increase in sectoral output, conditional on the fact that aggregate output increases by the average correlation between aggregate and sectoral output. If one is interested in the cyclicality of marginal costs, this parameter may be the relevant one. For example, if a firm is debating whether it should smooth production by building up inventories in a recession, it will certainly take into account the existence of external economies, which might make production cheaper in a boom. But it would be desirable to disentangle the two effects econometrically.

[17] For a survey, see Cooper and John (1988).

There is, however, another well-known effect which would mitigate the cost savings of external economies. This is simply the effect of tight markets on input costs in ways that are not measured by observed prices (cf. the discussion in the previous section). These effects may operate by increasing the effective price of inputs — for example, the same wages paid to lower quality labor hired in a boom would represent an increase in the cost of labor services. In this case, the complete parametrization of variable costs would be

$$VC_i = \exp(-\alpha_i t) \exp(\phi_{i,t}) P_{X,i} [1+ Q_i^{\theta_i}] A_i Q_i^{\Psi_i} Q^{\lambda_i} Q^{\kappa_i} . \qquad (17)$$

where λ_i reflects the cost-raising effects of tight markets. Just as in the model without external effects but with overtime payments, the correct estimating equation is still given by (16). But again, it is impossible to disentangle the productive effects of aggregate output from the opposite effect produced by tight markets. One possibility is to try and capture the pecuniary externality coming from tight markets by using a variable that directly captures this effect. One candidate is labor market tightness, as represented by one minus the unemployment rate.[18] This suggests a third estimating equation:

$$d \ln (P_i Q_i - \pi_i) - d \ln (P_{X,i}) = (c_i - \alpha_i) + \gamma_i \, d \ln Q_i + \kappa_i \, d \ln Q + \lambda_i \, d \ln (1-u) + \epsilon_i \qquad (18)$$

IV. Data and Estimation

A. Data

The data for this paper consist of annual figures on U.S. 2-digit manufacturing industries from 1953 to 1985. The data were collected by Jorgenson, Gollop, and Fraumeni (1987) and are extensively described in their book. Data used include gross output, payments to the four factors of production (capital, labor, energy, and materials), the prices of the inputs, total revenues, and estimates of the current value of the capital stock by industry.

[18] The relevance of this variable is suggested by the work of Beaudry and DiNardo (1991). Actually, since they work with individual data, they suggest one minus the minimum unemployment rate that has been observed since the worker took his current job. This variable is suggested by an implicit contracting model with costless worker mobility.

Two price series had to be constructed to perform the estimation. The first is the user cost of capital, r_k, for each type of capital k. Following the references cited above, this was constructed so that

$$r_k = (\rho+\delta_k) \frac{1 - c - \tau d}{1 - \tau} . \qquad (19)$$

Here ρ is the real interest rate, δ_k the rate of depreciation of each type of capital k, c the investment tax credit, τ the corporate tax rate, and d the present value of depreciation allowances. Unlike the work of Hall (1986a,b; 1990) and Caballero and Lyons (1990,1991), the required rate of return is calculated for each type of capital separately. The Jorgenson *et. al.* data reports separate figures for 55 different types of capital, each with its own rate of depreciation. The user cost of capital in an industry is calculated as the product of the current value of each type of capital multiplied by its own required rate of return, $r_k P_k K_k$. Again following previous work, ρ was taken to be the dividend yield on the S&P 500, and data on c, τ, and d were obtained from Alan Auerbach. As noted above, expected capital gains and future adjustment cost savings were not incorporated into the constructed required rate..

The other price series to be constructed was the input price index for each sector, $P_{X,i}$. The first step in constructing the index was to construct cost shares for each of the four inputs. The cost shares differ from the usual revenue shares in that the cost of each type of capital is taken to be $r_k P_k K_k$, where P_k is the specific deflator for the type of capital k constructed by Jorgenson *et. al.* Having constructed the cost shares, the input price index was created as a Divisia index from the prices of the four inputs, where the weights were the cost shares. The price of capital is taken to be its rental cost.

Finally, since the equations had to be estimated by instrumental variables, different sets of instruments were obtained. The first set are those used by Ramey (1989, 1991) and Hall (1988, 1990). They are the rate of change of national defense spending, the percent change in the world price of crude oil, and a dummy variable that indicates the party of the president in office. A second instrument is simply the rate of growth of aggregate output. A third follows Barro (1977, 1978) in constructing a series of residuals from a money (M1) growth equation. These residuals are taken to represent unanticipated money growth and are then used as instruments.

B. Estimation

Equations (8), (16), and (18) are estimated with the three sets of instruments described above. Instrumental variables estimation is clearly required, since the quantity of output produced is expected to be (negatively) correlated with the error term, which represents shocks to the cost function. The identifying assumption for the second instrument (changes in total output) is that there are no aggregate shocks to technology, i.e. that there are no "real business cycles" at the economy-wide level. This instrument, and a defense of the assumptions implicit in its use, are found in Hall (1986b).

Estimations are done using panel data on two-digit manufacturing industries using Three-Stage Least Squares (3SLS). This estimation technique allows for a free contemporaneous covariance matrix of the disturbance terms from the different industry equations, which is then adjusted to produce heteroscedasticity-robust standard errors. Estimation of the system of equations is done both with and without imposing the assumption that some of the key parameters are equal across industries. (In the case that the slope parameters are free to vary, the estimation technique becomes SUR.) Tests of these overidentifying restrictions are reported in the next section. The sample period for all the regressions is 1953 to 1985.

V. Results

Equations (8), (16), and (18) were estimated as systems, both with and without the constraint that the parameters of interest (particularly γ) are equal across industries. Table 1 gives the 3SLS results for the constrained estimation, along with tests of the overidentifying restrictions. The point estimates and standard errors are reported for each of the three sets of instruments described above. For comparison purposes, SUR estimates (uninstrumented) are also reported. (All the regressions included industry-specific constants, which have not been reported in the results.)

The results are remarkably similar for all three sets of instrumental variable estimates, and those results to the SUR figures. The point estimates uniformly imply that a one percent increase in output increases variable costs by about 0.8 to 0.9 percent — strong evidence that the marginal cost curve for

aggregate manufacturing does not increase rapidly with output. The parameters are also very precisely estimated — the 95 percent confidence interval does not include large values of γ, which would have implied sharply increasing marginal costs.

Table 2 presents the same results with a correction for the bias that would result from different assumptions about the profit rate and the size of the markup, assuming that the covariance of changes in fixed costs with the growth rate of output is zero. All of the biases are upper bounds for the possible bias, obtained by taking a fifth-order Taylor expansion. This conservative procedure results in a maximum value for γ of 1.02. An estimate of the markup on the order of 1.1 seems most reasonable because it is derived from the results of Caballero and Lyons (1989a,b), who allowed for the possibility of non-constant returns to scale. A markup of 1.1 implies a γ of 0.94, which in turn implies that marginal costs curve fall somewhat as output increases.[19]

An estimate of γ very close to one is encouraging because of its congruence with other results derived from inventory data. Kahn (1991a) notes that an empirical examination of inventory data implies that stockout-avoidance is a large part of the motivation for holding inventories (Kahn, 1987). By simulating a model with both stockout-avoidance and production smoothing motives for holding inventories, he concludes that unless marginal costs are very close to constant, production-smoothing (or production-bunching) motives would swamp stockout-avoidance in the data. He concludes, therefore, that marginal costs must be very close to constant in order to mimic the stylized facts in the data: stockout-avoidance, and a variance of production that is larger than the variance of sales.[20] Kahn's results are also somewhat at odds with Ramey's (1991) findings that marginal costs are strongly negatively sloped.

Although the overidentifying restriction that γ is constant across industries is never rejected, it is instructive to look at individual industry estimates of γ. Accordingly, Table 3 presents results that relax

[19] Calculations indicate that the estimated standard errors are positively biased if the estimated γ is biased down. Therefore the confidence intervals around the bias-corrected estimates of γ are even smaller than the reported standard errors in Table 1 would indicate.
[20] One caveat is that a large variance of temporary cost shocks could overturn this result (cost shocks that are expected to be very persistent should not much change the variance of production relative to sales). Kahn (1991b) does not find support for the hypothesis that cost shocks are a major driving force in inventory investment.

the assumption that the slope coefficient is equal across industries.[21] The results indicate that a number of industries have very negatively sloped marginal costs — the lowest point estimate is uniformly that for food, while some have sharply increasing marginal costs, though there is no uniformity as to which industry has the fastest increasing costs. Point estimates from different sets of instruments are generally fairly close. The major exception is the tobacco industry, which according to the first set of instruments has very negatively sloped marginal cost, but according to the the second and third has sharply increasing costs. The coefficients that are estimated with small standard errors are uniformly quite reasonable in magnitude. The auto industry, which is often held to have strongly increasing returns, does not appear to have quickly falling marginal costs. But the intuition for increasing returns in the auto industry probably comes from large fixed costs rather than falling marginal costs. Nothing in these results contradicts that view.

It is interesting to compare the industry-specific estimates obtained in Table 3 with Ramey's results. For comparison purposes, it is best to use the first set of instruments (defense spending, etc.), since these are the same as Ramey's. Ramey estimates the slope of marginal costs for seven of the two-digit industries covered here. In four cases her estimates are significant at the five percent level. The ranking of these four industries, from most negatively sloped marginal cost to least, is: autos, tobacco, apparel, and rubber. A similar ranking can be constructed for the industries here. Ordering the industries by the same criterion gives: tobacco, autos, rubber, and apparel. In both Ramey's results and the ones here, the estimates for rubber and apparel are very close, so one can consider them virtually identical. The major difference between Ramey's rankings and the ones in Table 3 is the switching of the order of tobacco and autos (though in both estimations, the 95 percent confidence intervals for the slope coefficients overlap). Additionally, both sets of results indicate that food is the industry with the most negatively sloped marginal costs (though both estimates have large standard errors).

This general congruence of rankings from two very different estimation techniques is reassuring. However, as noted previously, the magnitudes of the two sets of estimates are quite different. Ramey finds strongly negatively sloped marginal costs for all seven industries. The estimates in Table 3 indicate

[21] The 21 equations are still estimated as a system, however, not equation by equation.

that two of those industries, rubber and apparel, have increasing marginal costs. Why it is that Ramey's results and the ones here broadly agree as to the ranking of industries by slope of marginal cost, but disagree substantially as to the magnitude of the slope is an open question. One possibility is that omitting any treatment of capital gains in the required rate of return series, which was shown above to lead to a positive bias in the estimated elasticity, is in fact very significant.

The importance of aggregation and the issue of entry is explored by using three-year differences. The intuition is that if entry is important for the results that have been reported so far, redoing the estimation using longer time differences to allow for more entry should yield smaller estimates of γ. This intuition receives some support from the results reported in Table 4, which repeats the estimation of Table 1 for three-year differences. The point estimates from the instrumented regressions are generally somewhat smaller, though not spectacularly so. Interestingly, the point estimate from the SUR regression is much smaller than before, perhaps indicating the greater importance of cost shocks for output movements over long horizons.

It is possible, of course, that entry and exit take place mostly over a period of less than one year. This hypothesis receives some support from the preliminary results of Bresnahan and Ramey (1991). They find that in the U.S. automobile industry, plant closings and openings are one of the primary means of varying productive capacity. These take place over a time horizon of weeks or months. On average, a plant is closed for an entire week every two months for reasons not related to holidays. If this time horizon for entry is general, it provides a means of reconciling the results here with those of Beaulieu, Mackie-Mason, and Miron (1991). They find costs to be quite convex at seasonal frequencies. If entry and exit (in the sense of changes in the number of operating plants) take place at intra-annual but inter-seasonal frequencies, then much of the disparity can be resolved. The correct slope of marginal costs then depends on the time horizon in question; over business cycle frequencies, the slope from annual data is probably more appropriate. The entry-exit hypothesis does not, however, resolve the difference between the results obtained here and those of Ramey (1991), who uses monthly data.

The regressions of equation (16), which estimate γ while controlling for changes in aggregate output., give results very similar to those derived without incorporating aggregate output effects. These estimates of γ are found in Table 5, and are very similar to the estimates in the first line of Table 1 which were estimated without controlling for aggregate output changes. With the first two sets of instruments, the aggregate output coefficient is negative, as expected, but very small and statistically insignificant. Only with the unanticipated money instrument is the coefficient economically and statistically significant. Again, the estimates are obtained with a high degree of precision.

Both results — constant or diminishing marginal costs and insignificant external effects — are somewhat at odds with the findings of Caballero and Lyons (1990,1991). They find that two-digit manufacturing industries have decreasing returns to scale to their own inputs, but have positive and very significant increases in productivity from increases in aggregate output. The results are much more consistent with those obtained by Basu and Fernald (1992). Their results also give a similar view of U.S. manufacturing: marginal costs and returns to scale appear to be more or less constant. External effects from aggregate output may or may not be statistically different from zero, but the point estimates are very small. All told, the results support a standard view of U.S. industry, at least at the two digit level, with the exception that marginal costs are found to be roughly constant.

When the externality parameter is allowed to vary by industry, many of the coefficients are negative and significant, as would be predicted by the previous work.[22] On the other hand, for some industries the estimates are positive and strongly significant, and for most industries with most instruments the results are insignificant. Examining the individual industry results does not shed much light on the puzzle posed by the pooled manufacturing estimates.

On the other hand, the results are quite consistent with the findings of Bils (1991) and the engineering studies he cites. They find returns to scale that are mildly increasing, as one would predict if these industries had relatively small fixed costs and constant or slightly declining marginal costs.

The last set of regressions to be estimated are those based on equation (18). This equation attempts to separate the productive effects of output externalities from the cost-raising effects of tight markets. The effect of tight markets is proxied for by one minus the unemployment rate. Results for this estimation are reported in Table 6. In all cases the output externality is estimated as large and

22 The restriction that the externality parameter is equal across industries is not rejected, so these results have not been reported here.

productive, while the effect of tight markets is to raise costs. These results must be regarded with caution, since the failure of unemployment to comove with output over the sample period may be due to long-term changes in the labor force that took place in the 70s rather than to cyclical phenomena. Nevertheless, when the estimation is repeated with pre-1970 and post-1970 dummies, the results do not change significantly. So even though the results are by no means conclusive, they suggest that the small net externality found in the estimations of Table 5 is composed of a large productive external effect that is almost completely offset by a large, cost-raising, tight-markets effect.

VI. Conclusion

This paper has examined the cyclical behavior of marginal costs by examining changes in a constructed measure of total costs in response to changes in output. Marginal costs in U.S. manufacturing appear close to constant at annual frequency. This conclusion is robust to considerations of various types of bias and different assumptions about economic behavior, such as labor hoarding and increasing marginal prices of inputs. The results, in conjunction with previous work, provide weak support for the proposition that changes in the number of plants helps keep marginal costs constant. The data do not support the hypothesis of large productive spillovers in manufacturing that are external at the two-digit level.

REFERENCES

Abel, Andrew B. (1981). "Dynamic Effects of Permanent and Temporary Tax Policies in a q Model of Investment." *Journal of Monetary Economics* 9 353-73.

Akerlof, George, and Janet Yellen (1985). "A Near-Rational Model of the Business Cycle, with Wage and Price Inertia." *Quarterly Journal of Economics* 100 (Aug.) 823-38.

Ball, Lawrence, and David Romer (1989). "Are Prices too Sticky?" *Quarterly Journal of Economics* 104 (Aug.) 507-524.

Barro, Robert J. (1977). "Unanticipated Money Growth and Unemployment in the United States." *American Economic Review* 67 (March) 101-15.

______ (1978). "Unanticipated Money , Output, and the Price Level in the United States." *Journal of Political Economy* 86 549-80.

Basu, Susanto and John G. Fernald (1992). "Productive Externalities in U.S. Manufacturing: Do They Exist, and Are They Important?" Mimeo, Harvard University.

Beaudry, Paul and John DiNardo. "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data." *Journal of Political Economy* 99 (Aug.) 665-688.

Beaulieu, J., J. Mackie-Mason, and J. Miron (1991). "Why Do Countries and Industries with Large Seasonal Cycles Also Have Large Business Cycles?" *Quarterly Journal of Economics* (forthcoming).

Bils, Mark (1987). The Cyclical Behavior of Marginal Cost and Price." *American Economic Review* 77 (Dec.) 838-855.

Blanchard, Olivier and Stanley Fischer (1989). *Lectures on Macroeconomics.* Cambridge: MIT Press.

Blinder, Alan S. "Can the Production Smoothing Model of Inventory Behavior be Saved?" 101 *Quarterly Journal of Economics* (Aug.) 431-53.

Bradburd, Ralph M. and Richard E. Caves (1987). "Transaction-Cost Influences on the Adjustment of Industries' Prices and Output." *Review of Economics and Statistics* (Nov.) 575-83.

Bresnahan, Timothy F. and Valerie A. Ramey (1991). "The Nature of Output Fluctuations: Evidence From Plant Level Data." Mimeo, National Bureau of Economic Research.

Caballero, Ricardo J., and Richard K. Lyons (1990). "Internal Versus External Economies in European Industry." *European Economic Review* 34, 805-30.

______, and ______ (1991). "External Effects in U.S. Procyclical Productivity." Mimeo, Columbia University. Forthcoming in *Journal of Monetary Economics.*

Carlton, Dennis W. (1987). "The Theory and the Facts of How Markets Clear: Is Industrial Organization Useful for Understanding Macroeconomics?" In R. Schmalensee and R. Willig (eds.) *Handbook of Industrial Organization.* Amsterdam: North:Holland.

Cooper, Russell, and Andrew John (1988). "Coordinating Coordination Failures in Keynesian Models." *Quarterly Journal of Economics* 103 (Aug.) 441-464.

Domowitz, Ian, R. Glenn Hubbard, and Bruce C. Petersen (1986a). "Business Cycles and the Relationship between Concentration and Price-Cost Margins." *Rand Journal of Economics* 17 1-17.

______, ______, and ______ (1986b). "The Intertemporal Stability of the Concentration-Margins

Relationship." *Journal of Industrial Economics* 35 13-34.

______ , ______ , and ______ (1987). "Oligopoly Supergames: Some Empirical Evidence on Prices and Margins." *Journal of Industrial Economics* 35 13-34.

______ , ______ , and ______ (1988). "Market Structure and Cyclical Fluctuations in U.S. Manufacturing." *Review of Economics and Statistics* 55-66..

Fisher, Franklin M. and John G. McGowan (1983). "On the Misuse of Accounting Rates of Return to Infer Monopoly Profits." *American Economic Review* 73 (March) 82-97.

Hall, Robert E. (1980). "Employment Fluctuations and Wage Rigidity." *Brookings Papers on Economic Activity* no. 1, 91-123..

______ (1986a). "The Relation Between Price and Marginal Cost in U.S. Industry." National Bureau of Economic Research Working Paper 1785.

______ (1986b). "Market Structure and Macroeconomic Fluctuations." *Brookings Papers on Economic Activity* no. 2 285-322.

______ (1988). "The Relation Between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy* 96 (Oct.) 921-947.

______ (1990). "Invariance Properties of Solow's Productivity Residual" in Peter Diamond (ed.) *Growth, Productivity, Employment* (Cambridge: MIT Press).

______ (1991). "Labor Demand, Labor Supply, and Employment Volatility" in Olivier J. Blanchard (ed.) *NBER Macroeconomics Annual* (Cambridge: MIT Press).

______ and Dale W. Jorgenson. "Tax Policy and Investment Behavior." *American Economic Review* 57 (June) 391-414.

Hulten, Charles R. and Frank C. Wykoff (1981). "The Estimation of Economic Depreciation Using Vintage Asset Prices: An Application of the Box-Cox Power Transformation." *Journal of Econometrics* 15 (Aug.) 367-96.

Jorgenson, Dale W., Frank Gollop, and Barbara Fraumeni (1987). *Productivity and US Economic Growth.* Cambridge: Harvard University Press.

Kahn, James A. (1987). "Inventories and the Volatility of Production." *American Economic Review* 77 (Sept.) 667-79.

______ (1991a). "The Slope of Marginal Cost: Evidence from Inventory Behavior." Mimeo, National Bureau of Economic Research .

______ (1991b). "Why is Production More Volatile than Sales? Theory and Evidence on the Stockout-Avoidance Motive for Inventory-Holding." *Quarterly Journal of Economics* (forthcoming).

Mankiw, N. Gregory (1985). "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly." *Quarterly Journal of Economics* 100 (May) 529-539.

Ramey, Valerie A. (1989). "Inventories as Factors of Production and Economic Fluctuations." *American Economic Review* 79 (June) 338-54.

______ (1991). "Nonconvex Costs and the Behavior of Inventories." *Journal of Political Economy* 99 (April.) 306-334.

Rotemberg, Julio J. and Garth Saloner (1986). "A Super Game Theoretic Model of Price Wars during

Booms." *American Economic Review* 77 (Dec.) 917-926.

______ and Michael Woodford. "Markups and the Business Cycle" in Olivier J. Blanchard (ed.) *NBER Macroeconomics Annual* (Cambridge: MIT Press.)

Stigler, George J. and James K. Kindahl (1970). *The Behavior of Industrial Prices.* New York: National Bureau of Economic Research .

Tobin, James (1969). "A General Equilibrium Approach to Monetary Theory." *Journal of Money, Credit, and Banking* 1 (Feb.) 15-29.

Trejo, Stephen J. (1991). "The Effect of Overtime Pay Regulation on Worker Compensation." *American Economic Review* 81 (Sept.) 710-40.

Table 1: Economy-Wide Estimate of Changes in Cost as a Function of Output

Estimated system: $d \ln (P_i Q_i - \pi_i) - d \ln (P_{x,i}) = (c_i - \alpha_i) + \gamma_i \, d \ln Q_i + \varepsilon_i$
Restriction: $\gamma_i = \gamma$ for all i.

Parameter	SUR	Inst. 1	Inst. 2	Inst. 3
γ	0.82	0.88	0.88	0.85
	(0.0087)	(0.012)	(0.013)	(0.015)
Wald statistic for $\gamma_i = \gamma_j$	22.66	21.39	23.86	19.69

(χ^2 (0.05,20) = 31.41)

Table 2: Estimates of Results found in Table 1 Corrected for Fixed Cost Bias

	SUR	Inst. 1	Inst. 2	Inst. 3
$\mu^* = 1.1$, $\pi = 0$	0.88	0.95	0.95	0.92
$\mu^* = 1.6$, $\pi = 0.1$	0.95	1.02	1.02	1.00
$\mu^* = 1.6$, $\pi = 0$	0.96	1.02	1.02	1.00
$\mu^* = 2$, $\pi = 0$	0.96	1.02	1.02	1.00

*Standard Errors in Parentheses

Inst. 1: Price of Oil, National Defense Spending, Party in Power
Inst. 2: Aggregate Output
Inst. 3: Unanticipated Money

Table 3: Individual Industry Estimates of the Changes in Cost as a Function of Output

Estimated system: $d \ln (P_i Q_i - \pi_i) - d \ln (P_{X,i}) = (c_i - \alpha_i) + \gamma_i \, d \ln Q_i + \varepsilon_i$

Estimates of γ_i

Industry	SUR	Inst. 1	Inst. 2	Inst. 3
Food and Kindred Products (7)	0.67 (0.063)	0.27 (0.216)	0.55 (0.268)	0.22 (0.209)
Tobacco Manufactures (8)	0.31 (0.060)	0.64 (0.182)	1.29 (0.648)	1.46 (0.333)
Textile Mill Products (9)	0.97 (0.036)	0.86 (0.076)	1.10 (0.096)	0.87 (0.080)
Apparel and Other Textiles (10)	0.91 (0.021)	1.02 (0.045)	0.95 (0.049)	0.93 (0.048)
Lumber and Wood Products (11)	0.80 (0.037)	1.03 (0.121)	1.42 (0.160)	1.41 (0.187)
Furniture and Fixtures (12)	0.92 (0.021)	0.87 (0.038)	0.82 (0.036)	0.78 (0.040)
Paper and Allied Products (13)	0.78 (0.028)	0.73 (0.044)	0.64 (0.066)	0.63 (0.067)
Printing and Publishing (14)	0.90 (0.048)	0.72 (0.094)	0.61 (0.090)	0.61 (0.091)
Chemicals and Allied Products (15)	0.79 (0.030)	0.61 (0.060)	0.57 (0.074)	0.54 (0.081)
Petroleum Refining (16)	0.87 (0.060)	0.81 (0.233)	0.94 (0.214)	0.79 (0.230)
Rubber and Plastic Products (17)	0.99 (0.031)	0.97 (0.054)	0.98 (0.051)	0.97 (0.051)
Leather and Leather Products (18)	0.38 (0.042)	1.11 (0.148)	1.28 (0.236)	1.33 (0.211)
Stone, Clay, and GlassProducts (19)	0.60 (0.020)	0.66 (0.035)	0.71 (0.039)	0.75 (0.043)
Primary Metals (20)	1.02 (0.014)	0.97 (0.026)	0.96 (0.029)	1.00 (0.037)
Fabricated Metal Products (21)	0.80 (0.018)	0.78 (0.034)	0.78 (0.031)	0.82 (0.042)
Machinery,except Electrical (22)	0.63 (0.011)	0.73 (0.026)	0.76 (0.034)	0.81 (0.041)
Electrical Machinery (23)	0.88 (0.021)	0.87 (0.032)	0.76 (0.039)	0.75 (0.046)
Motor Vehicles (24)	0.87 (0.014)	0.86 (0.026)	0.87 (0.025)	0.87 (0.033)
Other Transport Equipment (25)	1.19 (0.028)	1.09 (0.051)	0.94 (0.152)	0.86 (0.143)
Instruments (26)	1.02 (0.032)	1.06 (0.075)	0.71 (0.079)	0.62 (0.074)
Miscellaneous Manufacturing (27)	0.79 (0.040)	1.00 (0.124)	0.73 (0.095)	0.86 (0.108)

---

*Standard Errors in Parentheses

*Sectoral index numbers, from 7 to 27, refer to the 35-sector Jorgenson data set. These differ from standard SIC only in the separation of Motor Vehicles (24) from Other Transport Equipment (25).

Inst. 1: Price of Oil, National Defense Spending, Party in Power
Inst. 2: Aggregate Output
Inst. 3: Unanticipated Money

Table 4: Effects of Three-Year Changes in Output on Three-Year Changes in Costs

Estimated system: $(1-L^3)\ln(P_iQ_i - \pi_i) - (1-L^3)\ln(P_{x,i}) = (c_i - \alpha_i) + \gamma\ (1-L^3)\ln Q_i + \varepsilon_i$

Parameter	SUR	Inst. 1	Inst. 2	Inst. 3
γ	0.76	0.85	0.84	0.85
	(0.0048)	(0.0053)	(0.0066)	(0.00054)

Table 5: External Effects of Aggregate Output on Costs

Estimated System: $d\ln(P_iQ_i - \pi_i) - d\ln(P_{x,i}) = (c_i - \alpha_i) + \gamma\, d\ln Q_i + \kappa\, d\ln Q + \varepsilon_i$

Parameter	SUR	Inst. 1	Inst. 2	Inst. 3
γ	0.84	0.91	0.91	0.89
	(0.0091)	(0.014)	(0.017)	(0.021)
κ	0.18	-0.049	-0.047	-0.14
	(0.018)	(0.031)	(0.026)	(0.045)

*Standard Errors in Parentheses

Inst. 1: Price of Oil, National Defense Spending, Party in Power
Inst. 2: Aggregate Output
Inst. 3: Unanticipated Money

Table 6: Positive External Effects Distinguished from Tight Market Effects

Estimated System: $d\ln(P_iQ_i - \pi_i) - d\ln(P_{x,i}) = (c_i - \alpha_i) + \gamma\, d\ln Q_i + \kappa\, d\ln Q + \lambda\, d\ln(1-u) + \varepsilon_i$

Parameter	SUR	Inst. 1	Inst. 2	Inst. 3
γ	0.84	0.90	0.91	0.89
	(0.0091)	(0.014)	(0.017)	(0.021)
κ	0.055	-0.32	-0.45	-0.54
	(0.033)	(0.060)	(0.192)	(0.128)
λ	0.28	0.84	1.20	1.17
	(0.026)	(0.17)	(0.61)	(0.37)

*Standard Errors in Parentheses

Inst. 1: Price of Oil, National Defense Spending, Party in Power
Inst. 2: Aggregate Output
Inst. 3: Unanticipated Money

DATE DUE

NOV 09 1996	

DEMCO, INC. 38-2931

MCO