

MIT LIBRARIES DUPL



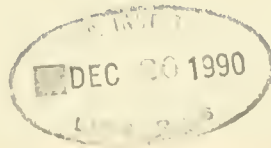
3 9080 00706965 8



HD28
.M414

no. 3216-

90



WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

CYCLICAL MARKUPS: THEORIES AND EVIDENCE

by

Julio J. Rotemberg
Massachusetts Institute of Technology

Michael Woodford
University of Chicago

WP#3216-90-EFA

November 1990

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

CYCLICAL MARKUPS: THEORIES AND EVIDENCE

by

Julio J. Rotemberg
Massachusetts Institute of Technology

Michael Woodford
University of Chicago

WP#3216-90-EFA

November 1990

DEAD 1990

**Cyclical Markups:
Theories and Evidence ***

Julio J. Rotemberg
Massachusetts Institute of Technology

Michael Woodford
University of Chicago

current version: November 6, 1990

Abstract

If changes in aggregate demand were an important source of macroeconomic fluctuations, real wages would be countercyclical unless markups of price over marginal cost were themselves countercyclical. We thus examine three theories of markup variation at cyclical frequencies. The first assumes only that the elasticity of demand is a function of the level of output. In the second, firms face a tradeoff between exploiting their existing customers and attracting new customers. Markups then depend also on rates of return and future sales expectations; a high rate of return or expectations of low sales growth lead firms to assign a lower value to future revenues from new customers. Firms thus raise prices and markups. In the third theory, markups are chosen to ensure that no one deviates from an (implicitly) collusive understanding. Increases in rates of return or pessimistic expectations then lead firms to be less concerned with future punishments so that markups fall. Aggregate post-war data from the U.S. are most consistent with the predictions of the implicit collusion model.

* We wish to thank Alan Blinder, Robert King, N. Gregory Mankiw and Hal Varian and participants at the NBER Summer Institute for suggestions, Changyong Rhee for providing us with his data on q and the NSF for research support.

Cyclical Markups: Theories and Evidence

Julio J. Rotemberg

Michael Woodford

Real wages are not strongly countercyclical. As pointed out already by Dunlop (1938) and Tarshis (1939), this presents difficulties for models in which technological possibilities do not vary at business cycle frequencies. For, why should firms be willing to pay more to workers when output is high and hence the marginal product of labor is low? One obvious possibility, suggested by Keynes (1939), is that the desired markup of price over marginal cost is low when output is high.

In this paper we consider three leading models of endogenous markup variation which might explain variations in real wages at business cycle frequencies. The first simply postulates that the elasticity of demand facing the representative firm varies over the business cycle. This is the underlying idea behind Robinson (1932) and Bilal (1989). The second is based on the customer market model of Phelps and Winter (1970), the macroeconomic consequences of which have been explored by Gottfries (1989), Greenwald and Stiglitz (1988) and Phelps (1989). Finally, our third model is based on the model of implicit collusion of Rotemberg and Saloner (1986) and Rotemberg and Woodford (1989).

The first model is essentially static; it does not depend on expectations firms hold about the future. The other two are dynamic. The price charged in the customer market model is the result of a tradeoff between exploiting the existing customer base and attracting new customers whose profitable purchases come in the future. Similarly, pricing in the implicit collusion model involves a tradeoff between current profits from undercutting one's competitors and the future profits from maintaining collusion in one's industry.

While expectations about the future matter in both of the latter models, their effect is rather different in the two cases. In the customer market model, large purchases in the future induce

firms to try to enlarge their market share, so that prices and markups tend to fall. In the implicit collusion model, by contrast, high expected future demand increases the costliness of a price war, making it possible for firms to raise prices and markups. The first model is thus one where low prices are like an investment; prices are cut when the future looks better than the present. The second is one where a rosy future leads to price increases.

This is the implication of the two models that we actually test. Therefore, our test is somewhat broader than a test of the specific models that we consider. To test this implication of the models, we need to measure markup variations. Like Bilal (1987), we do this by making assumptions on the aggregate production function and exploring the consequences of the equality between the markup and the ratio of the marginal product of labor to the real wage. Unlike him, we do not impose the assumption that the elasticity of substitution between hours and other inputs equals one. Also, our production function simultaneously allows for fixed costs (overhead labor), but does not imply that output can be increased by adding only production workers. In our general formulation, the marginal product of labor, as a function of output and factor inputs, depends upon the state of technology. Hence we use a generalization of Solow's (1957) method for measuring changes in technological possibilities and, armed with this measure, we obtain estimates of markup variations that depend only on observable variables.

We show that, for plausible parameter values, these measured variations in the markup tend to be quite countercyclical. One parameter that we do not measure directly and which influences our result is the average level of the markup. Given that measured profits over and above the required return to capital seem to be small, a higher level of the average markup means that there is a wider gap between average costs and marginal cost. If this gap is due to fixed costs, economies with high markups also have high fixed costs. But, in this case, a given percentage change in factor inputs corresponds to a much larger percentage change in the factors that are productive at the margin. Thus high estimates of the average markup imply that marginal cost rises substantially in booms, thus leading to countercyclical markups. This is particularly true of the estimates of the average markup derived by Hall (1988a) from observations on the response of total factor productivity to changes in aggregate demand.

While the assumption of a high average markup produces measures of the markup that, by

necessity, rise when employment falls, it does not have any direct implication for the relationship between markups and expectations about future profitability. We show that our constructed markups tend to rise when the rate at which firms discount future cash flows is low (and when the discounted value of future profits is high). We thus provide some direct evidence for the class of models where high prices are like an investment.

The paper proceeds as follows. In Section 1 we present the framework underlying all three models. Section 2 presents the static model of time varying elasticities of demand. Section 3 develops the customer market model, while Section 4 develops the model of implicit collusion. Section 5 gives the details of how we construct our measures of markup variations. Section 6 explains two methods for testing the implications of the three models. Section 7 discusses our data and section 8 gives our empirical results. Finally, section 9 puts our results in context by discussing the role of the various models of markup variation in explaining fluctuations in aggregate activity.

1. The Basic Setup

We consider economies with many symmetric firms whose total number is normalized to equal one. We will focus on symmetric equilibria, so that in equilibrium all firms charge the same price at time t , P_t . For simplicity we will treat the output of these symmetric firms as the numeraire so that, in units of the numeraire, P_t is one.

These symmetric firms have access to a technology of the form

$$y_t^i = F(K_t^i, z_t(H_t - \bar{H}_t)) \quad (1)$$

where y_t^i , H_t^i and K_t^i represent respectively firm i 's output, labor input and capital input at time t . The variable z_t represents the state of technology at time t , so that a higher z corresponds to a more productive period, while \bar{H}_t is the amount of labor devoted to fixed costs. The allowance for an overhead labor requirement is a way of introducing decreasing average costs, of the kind needed to reconcile an assumed markup of price over marginal cost with the apparent absence of significant pure profits in U.S. industry.¹

Each firm has access to competitive markets for labor and capital services. At time t , firm i

¹ For evidence on the existence of increasing returns, in the sense of average costs in excess of marginal cost on average, in U.S. industry, see Hall (1987).

must pay a wage w_t for each unit of labor and it must pay r_t for each unit of capital that it rents. Given the homogeneity of F and competitive factor markets, marginal cost at t is independent of the number of units that the firm produces and is equal to

$$\min_{h,k} w_t h + r_t k \quad \text{s.t.} \quad F(k, z_t h) = 1 \quad (2)$$

The assumption that F is homogeneous of degree one so that marginal cost is constant is not essential for the models to be presented below. However, it simplifies our analysis by allowing us to write the ratio of two firms' prices as the ratio of their respective markups. We denote the equilibrium markup by μ_t ; this is the equilibrium ratio of the price charged by all firms to marginal cost. Since both w_t and r_t are denominated in the units of the typical firm's output, marginal cost in (2) is simply equal to $1/\mu_t$. Letting firm i 's ratio of price to marginal cost be denoted by μ_t^i , firm i 's profits gross of fixed costs in units of the numeraire are equal to

$$\Pi_t^i = \left(\frac{\mu_t^i - 1}{\mu_t} \right) y_t^i \quad (3)$$

At a symmetric equilibrium all firms charge the same price and, given our normalization, the sales of each equal the aggregate level of sales Y_t . We denote by X_t each firm's expected present discounted value at t of the stream of individual profits from period $t + 1$ onward:

$$X_t = E_t \sum_{j=1}^{\infty} \frac{q_{t+j}}{q_t} \left(\frac{\mu_{t+j} - 1}{\mu_{t+j}} \right) Y_{t+j} \quad (4)$$

Here E_t takes expectations conditional on information available at t , and q_{t+j}/q_t is the stochastic asset pricing kernel, so that any random yield z_{t+j} (in units of period $t + j$ goods) has a present discounted value in period t of $E_t(q_{t+j} z_{t+j}/q_t)$.

We now distinguish between three models that differ in both the specification of demand and of market structure.

2. The Static Monopolistic Competition Model

In this model each firm behaves like a monopolistic competitor in that it takes as given the prices of all other firms, the level of marginal cost and the level of aggregate demand. As in the "symmetric" monopolistic competition model of Dixit and Stiglitz (1977), we assume that the

demand for firm i depends on the ratio of its price to the average price charged by all other firms. Equivalently, firm i 's demand at t depends on the ratio of its own markup μ_t^i to the markup charged by all other firms in the symmetric equilibrium we will consider, μ_t . Thus we write firm i 's demand as²

$$y_t^i = D\left(\frac{\mu_t^i}{\mu_t}, Y_t\right) \quad (5)$$

where the firm's demand depends on aggregate demand through the level of aggregate sales Y_t . To preserve symmetry we require that the demand for each firm be equal to Y if they all charge the same price. Thus we require that $D(1, Y) = Y$. A special case to which we will return has homothetic preferences so that demand is the product of a function of relative prices and aggregate demand Y_t . In this special case both D and the partial derivative of D with respect to relative prices, D_1 , are proportional to Y .

Since the firm's problem is static we can obtain its decision rule by substituting (5) into (3) and maximizing with respect to μ_t^i . This yields the familiar formula

$$D + \frac{\mu_t^i - 1}{\mu_t} D_1 = 0 \quad (6)$$

In a symmetric equilibrium all firms charge the same markup, so that the markup can rise if and only if $-D_1(1, Y)/D(1, Y) = -D_1/Y$, the elasticity of demand evaluated at the point where all prices are the same, falls. Thus the markup can rise with a change in Y_t if and only if preferences are not homothetic. There is little *a priori* reason to expect either direction of deviation from homotheticity, so that markups seem as likely to rise with increased sales as to fall.

3. The Customer Market Model

The customer market model continues to have each firm maximizing profits with respect to its markup taking the markup in all other firms as given. It differs in that demand has a dynamic pattern. A firm that lowers its current price not only sells more to its existing customers, but also expands its customer base. Having a larger customer base leads future sales to be higher at any given price. One simple formulation that captures this idea involves writing the demand for firm i at time t as

²In writing the demand function (5) we have avoided considering the effect of changes of the composition of demand on its overall price elasticity and hence on markups. We have done this because we are interested in the effects of changes in aggregate demand and no particular compositional shift seems plausibly associated with a large fraction of changes in aggregate demand.

$$y_t^i = \eta\left(\frac{\mu_t^i}{\mu_t}, Y_t\right) m_t^i \quad \eta_1 < 0, \quad \eta(1, Y) = Y \quad (7)$$

The variable m_t^i is the fraction of total demand Y_t that goes to firm i if it charges the same price as all other firms. The market share m_t^i depends on past pricing behavior according to the rule

$$m_{t+1}^i = g\left(\frac{\mu_t^i}{\mu_t}\right) m_t^i \quad g' < 0, \quad g(1) = 1 \quad (8)$$

so that a temporary reduction in price raises firm i 's market share permanently. Equations (6) and (7) capture the idea that customers have switching costs, in a manner analogous to the models of Gottfries (1986), Klemperer (1987), and Farrell and Shapiro (1988).³ A reduction in price attracts new customers who are then reluctant to change firms for fear of having to pay these switching costs. One obvious implication of (6) and (7) is that the long run elasticity of demand, i.e., the response of eventual demand to a permanent increase in price, is larger than the short run elasticity of demand. In our case, a firm that charges a higher price than its competitors eventually loses all its customers, though this is not essential for our analysis.

The firm's expected present discounted value of profits from period t onward is thus

$$E_t \sum_{j=0}^{\infty} \frac{q_{t+j}}{q_t} \left(\frac{\mu_{t+j}^i - 1}{\mu_{t+j}}\right) \eta\left(\frac{\mu_{t+j}^i}{\mu_{t+j}}, Y_{t+j}\right) m_t^i \prod_{z=0}^{j-1} g\left(\frac{\mu_{t+z}^i}{\mu_{t+z}}\right) \quad (9)$$

Firm i chooses μ_t^i to maximize (9), taking as given the stochastic processes $\{\mu_t\}$ and $\{Y_t\}$.

Therefore

$$\begin{aligned} & \eta\left(\frac{\mu_t^i}{\mu_t}, Y_t\right) + \eta_1\left(\frac{\mu_t^i}{\mu_t}, Y_t\right) \left[\frac{\mu_t^i - 1}{\mu_t}\right] + \\ & g'\left(\frac{\mu_t^i}{\mu_t}\right) E_t \sum_{j=1}^{\infty} \frac{q_{t+j}}{q_t} \left[\frac{\mu_{t+j}^i - 1}{\mu_{t+j}}\right] \eta\left(\frac{\mu_{t+j}^i}{\mu_{t+j}}, Y_{t+j}\right) \prod_{z=1}^{j-1} g\left(\frac{\mu_{t+z}^i}{\mu_{t+z}}\right) = 0. \end{aligned} \quad (10)$$

where subscripts denote partial derivatives. At a symmetric equilibrium where all firms charge the same price, each has an m_i equal to one and g equals one in all periods. So the expectation term in (10) is equal to the common present discounted value of profits given by (4). Therefore, (10) gives

³This idea has been applied to the analysis of international pricing issues by Gottfries (1988) and Froot and Klemperer (1989).

the markup μ_t as:

$$\mu_t = \mu(X_t, Y_t) \equiv \frac{\eta_1(1, Y_t)}{Y_t + \eta_1(1, Y_t) + g'(1)X_t} \quad (11)$$

The second order condition for a maximum of profits implies that the denominator of (11) is negative. Therefore, the derivative of μ with respect to X is negative. An increase in X means that profits from future customers are high so that each firm lowers its price in order to increase its market share. The effect of current sales Y_t on the markup is more ambiguous. In the homothetic case where η_1 is proportional to Y , (11) implies that the markup depends only on the ratio X_t/Y_t ; the elasticity of the markup with respect to Y is equal to the negative of the elasticity with respect to X . A high value of Y means that current customers are relatively profitable so that, in the homothetic case, raising prices and exploiting existing customers is relatively attractive. This intuition must be modified when the elasticity of demand facing an individual firm depends on the level of sales. Differentiating (11) and ignoring time subscripts, the derivative of μ with respect to Y is

$$\frac{-\mu + (1 - \mu)\eta_{12}}{Y + \eta_1(1, Y) + g'(1)X}$$

which is positive in the homothetic case where η_{12} , the second partial of η with respect to relative prices and Y , is zero. This derivative can be negative if η_{12} is sufficiently negative so that demand becomes much more elastic as output rises. However, because this term is multiplied by $(1 - \mu)$, the derivative is negative only if the magnitude of η_{12} is substantial, particularly if the typical markup is small.

Put broadly, equation (11) says that lower prices are a form of investment, an investment in market share. Such an investment is attractive when the present discounted value of the future returns from investment (X) are high relative to the payoff from current consumption, which in the homothetic case is represented by Y . While the story is logically distinct from the static model, they are closely related. An increase in X through, for instance, a fall in interest rates and discount rates does not make the demand curve more elastic. However, it raises the importance of the sales that go to customers with relatively elastic demand, thus promoting reductions in price.

4. The Implicit Collusion Model

The model in this section is a simplified presentation of Rotemberg and Woodford (1989). We consider an economy with many industries, each of which consists of n firms. The n firms in each industry collude implicitly in the sense that there is no enforceable cartel contract, but only an implicit agreement that firms that deviate from the collusive understanding will be punished. On the other hand, the firms in each industry, even when acting in concert, take other industries' prices, the level of aggregate demand, and the level of marginal cost as given. Abusing the language somewhat, we can view industries as monopolistic competitors in the usual sense, while the firms within each industry collude implicitly.

Keeping this distinction in mind, we write the demand for firm i in industry j as

$$y_t^{ij} = D^i\left(\frac{\mu_t^{1j}}{\mu_t}, \dots, \frac{\mu_t^{nj}}{\mu_t}, Y_t\right) \quad D^i(1, \dots, 1, Y) = Y \quad (12)$$

The function D^i is symmetric in its first n arguments except the i th, and the functions D^i (for $i = 1, \dots, n$) are all the same after appropriate permutation of the arguments. Using (3), profits for firm i in industry j when all other firms in industry j charge the markup μ_t^j , while firms in other industries all charge μ_t , equal

$$\Pi_t^{ij} = \frac{\mu_t^{ij} - 1}{\mu_t} D^i\left(\frac{\mu_t^j}{\mu_t}, \dots, \frac{\mu_t^{ij}}{\mu_t}, \dots, \frac{\mu_t^j}{\mu_t}, Y_t\right). \quad (13)$$

If each firm lived for only one period, it would maximize (13) with respect to its own markup treating the markups of all other firms as given. The resulting Bertrand equilibrium in the industry would have a markup equal to $\mu^B(\mu_t, Y_t)$. If the firms in an industry charged more than $\mu^B(\mu_t, Y_t)$, individual firms would benefit from undercutting the industry's price. Higher prices, with their attendant higher profits, can only be sustained as a subgame perfect equilibrium if deviators are punished after a deviation. If firms interact repeatedly and have an infinite horizon, there are many equilibria of this type and these differ in the price that is charged in equilibrium.

We assume that firms succeed in implementing that symmetric equilibrium that is jointly best for them. That is, their implicit agreement maximizes the present discounted value of expected equilibrium profits for each firm in industry j , taking as given the stochastic processes for $\{\mu_t\}$ and $\{Y_t\}$. As shown by Abreu (1982), the punishment for any deviation is as severe as possible in the

optimal symmetric equilibrium. Therefore, a deviating firm sets price to maximize current period profit Π_t^{ij} . The result is that the single period profits of a deviating firm equal:

$$\Pi_{dt}^j = \max_{\mu_t^{ij}/\mu_t} \frac{\mu_t^{ij} - 1}{\mu_t} D^i\left(\frac{\mu_t^j}{\mu_t}, \dots, \frac{\mu_t^{ij}}{\mu_t}, \dots, \frac{\mu_t^j}{\mu_t}, Y_t\right) \quad (14)$$

After any deviation, the firms in the industry punish the deviator to the maximum possible extent. Because of the possibility of exit, the voluntary participation of the firm that is being punished precludes it earning an expected present value lower than zero after a deviation. We give conditions that ensure that a deviator indeed earns a present discounted value of zero in Rotemberg and Woodford (1989).⁴

Let X_t^j denote, by analogy to (3), the expected present discounted value of the profits that firms in industry j can expect to earn in subsequent periods if there are no deviations. Then, if the expected present value of profits after a deviation equal zero, firms in industry j will not deviate as long as

$$\Pi_{dt}^j \leq \pi_t^j + X_t^j \quad (15)$$

where π_t^j is the value of Π_t^{ij} when firm i charges the same price as the other firms in its industry. We consider the case where the incentive compatibility constraint (15) is always binding.⁵

At a symmetric equilibrium, all industries have the same markup so that each firm sells Y_t and X_t^j equals X_t . Using $D(\rho, Y)$ to denote $D^i(1, \dots, \rho, \dots, 1, Y)$, we then have from (13)-(15)

$$\max_{\rho} \left[\rho - \frac{1}{\mu_t} \right] D(\rho, Y_t) = \left[1 - \frac{1}{\mu_t} \right] Y_t + X_t \quad (16)$$

where ρ represents the relative price chosen by the deviating firm. Equation (16) can be solved for μ_t yielding once again $\mu_t = \mu(X_t, Y_t)$. The relevant solution of (16) is the one where μ_t exceeds the Bertrand level, so that deviators undercut the equilibrium price and ρ is less than 1.

Differentiation of (16) yields

⁴The main condition requires that there exist a $\bar{\mu}$ smaller than one such that when all firms in industry j charge a markup of $\bar{\mu}$ while the firms in other industries charge a markup greater than or equal to one, a deviating firm cannot sell positive quantities by charging a price in excess of marginal cost. This assumption requires that the goods produced by firms in the industry be relatively good substitutes. It ensures that the deviating firm cannot make positive profits in the periods following a deviation by deviating from the behavior it is expected to follow after the deviation.

⁵In Rotemberg and Woodford (1989) we give conditions under which a deterministic steady state exists in which (15) is always binding. We also show that, for small enough stochastic shocks, there continues to exist a perturbed equilibrium in which (15) always binds.

$$\mu_X = \frac{\mu^2}{D(\rho, Y) - Y} \quad (17)$$

Since ρ is less than one, $D(\rho, Y) > D(1, Y) = Y$ and μ_X is positive. An increase in X , which raises the cost of deviating, raises the equilibrium markup. Such an increase in the markup is necessary to maintain the equality between the costs and the benefits of deviating.

We can also bound the response of the markup to changes in X from above. In particular

$$X = (\rho - 1/\mu)D(\rho, Y) - (1 - 1/\mu)Y < (1 - 1/\mu)[D(\rho, Y) - Y] = \frac{\mu(\mu - 1)}{\mu_X} \quad (18)$$

where the first equality follows from (16), the inequality from $\rho < 1$, and the last equality from (17). Therefore, the elasticity of μ with respect to X , while positive, is smaller than $\mu - 1$.

The effects of changes in Y are more ambiguous. In the homothetic case, where $D_Y = D/Y$ for all prices, (16) implies that μ depends only on the ratio X/Y . Thus an increase in Y raises the benefits to deviating now and the markup falls. More generally, μ_Y is negative as long as increases in Y raise the left hand side of (16) more than they raise the right hand side. This occurs as long as

$$\frac{\Pi_d(\mu, Y)D_2(\rho, Y)}{D(\rho, Y)} > \frac{\Pi(\mu, Y)}{Y}$$

While this must hold in the homothetic case where D_2/D equals $1/Y$, it could fail more generally if YD_2/D is sufficiently less than one for $\rho < 1$. This quantity is increasing in ρ only if the elasticity of demand faced by a deviating firm, $-\rho D_1(\rho, Y)/D(\rho, Y)$, is a decreasing function of Y . For goods that are close substitutes, the optimal deviating ρ is only slightly less than one, even though Π_d is much larger than Π . Since $YD_2(1, Y)/D(1, Y) = 1$, it seems likely that YD_2/D is not much smaller than one, so that $\mu_Y > 0$ is implausible in this model.

5. Construction of a Time Series for Markup Variations

Empirical estimation of a markup equation requires first that we construct a time series for cyclical variations in the markup over the postwar period for the U.S. Our method is quite simple. We assume (as in the theoretical models discussed above) an aggregate production function of the form (1).⁶ The markup of price over marginal cost is then

⁶Our results are little affected by the choice of the functional form (1) over a form such as

$$Y_t = F(K_t, z_t H_t) - \Phi_t.$$

$$\mu_t = \frac{F_H(K_t, z_t(H_t - \bar{H}_t))}{w_t} \quad (19)$$

We can thus construct a markup series from aggregate time series for output, factor inputs, and real wages, given a quantitative specification of the production function F (including a value for \bar{H}_t), and given a time series for the productivity shocks $\{z_t\}$. The productivity shocks present an obvious difficulty, since they are not directly observed. In our previous paper (Rotemberg and Woodford (1989)), we measured the effects of a particular type of aggregate demand shock on the markup by choosing a shock (innovations in real military purchases) that could be argued to be uncorrelated with variations in $\{z_t\}$. This will not, however, suffice if we wish to construct a time series for cyclical variations in the markup over the entire postwar period. Here we propose instead to reconstruct a series for $\{z_t\}$ from (1), using what is essentially the familiar Solow (1957) method, corrected for the presence of imperfect competition and increasing returns to scale.⁷

We consider a log-linear approximation to (1) around a steady-state growth path along which H_t grows at the same rate as \bar{H}_t , while K_t and Y_t grow at the same rate as $z_t\bar{H}_t$.⁸ This approximation yields

$$\hat{y}_t = \left(\frac{F_1 K}{Y}\right) \hat{k}_t + \frac{z F_2 (H - \bar{H})}{Y} \left[\hat{z}_t - \left(\frac{H}{H - \bar{H}}\right) \hat{h}_t\right] \quad (20)$$

where hatted lower case variables refer to log deviations from trend values, and where the other expressions represent constant coefficients evaluated at the steady-state growth path.

We assume that, for both factors, the marginal product equals μ^* times the factor price in the steady-state growth path, where μ^* is the steady-state markup. Therefore, $F_1 K/Y$ and $z F_2 H/Y$ are respectively equal to $\mu^* s_K$ and $\mu^* s_H$, where s_K and s_H are payments to capital and labor as a share of output's value. Because F is homogeneous of degree 1, Euler's equation implies that

By contrast, the assumed size of the fixed costs in relation to total costs (or more generally, of average cost in relation to marginal cost), represented here by the average size of H_t/H_t , is important to our conclusions.

⁷ Bills (1987) avoids the need to construct series for $\{z_t\}$ altogether by assuming a Cobb-Douglas production function with no overhead requirement (at least for production hours) so that F_H in (19) can be replaced by $\alpha Y_t/H_t$. We show that this restrictive functional form is not necessary and are able to consider the consequences of alternative assumptions regarding factor substitutability and the size of fixed costs.

⁸ The assumption that the overhead labor requirement grows at a constant rate allows us to obtain a stationary equilibrium with growth (in which, among other things, the ratio of fixed costs to total costs fluctuates around a constant value). Presumably this should be due to growth in the variety of goods produced as the economy grows, although we do not model that explicitly here. We could have assumed instead that the overhead labor requirement is constant in per capita terms. Because, per capita hours appear stationary, this too would have allowed us to apply our techniques.

$$\mu^* s_K + \mu^* s_H \frac{H - \bar{H}}{H} = 1 \quad (21)$$

Using (21), (20) can be written as

$$\hat{z}_t = \frac{\hat{y}_t - \mu^* s_K \hat{k}_t - \mu^* s_H \hat{h}_t}{1 - \mu^* s_K} \quad (22)$$

This allows us to construct a time series for \hat{z}_t from the variations in detrended output and factor inputs, given average factor shares, and given values for the single free parameter μ^* . This parameter is set to 1 in Solow's original method.⁹

Assuming that w_t and z_t have the same trend growth rates, the analogous log-linear approximation of (19) yields

$$\hat{\mu}_t = \hat{z}_t - \hat{w}_t + \frac{\mu^* s_K}{e} \left(\hat{k}_t - \hat{z}_t - \frac{\mu^* s_H}{1 - \mu^* s_K} \hat{h}_t \right)$$

where e represents the elasticity of substitution between the two factors in F , evaluated at the factor ratio associated with the steady-state growth path. Substituting (22) for \hat{z}_t this becomes

$$\hat{\mu}_t = \frac{e - \mu^* s_K}{e - e\mu^* s_K} \hat{y}_t + \frac{(1 - e)\mu^* s_K}{e - e\mu^* s_K} \hat{k}_t - \frac{\mu^* s_H}{1 - \mu^* s_K} \hat{h}_t - \hat{w}_t \quad (23)$$

Hence we need to specify only the parameters e and μ^* in addition to the observable factor shares to construct our markup series. Assigning numerical values to e and μ^* is admittedly somewhat problematic. Our basic strategy is to determine ranges of plausible values, and then to check the degree to which our results are sensitive to the exact values chosen for e and μ^* within those ranges. The parameter e is often “calibrated” in real business cycle studies on the basis of observed long run trends. The absence of a significant trend in factor shares, in the face of a significant trend in relative factor prices over the last century, is sometimes taken to indicate an elasticity of substitution near 1. But this is not a particularly persuasive justification. First, this fact might simply indicate that most technical progress is labor-augmenting, as in (1), rather than a long run elasticity of 1.

Second, there need not be much relationship between the long run elasticity and the short run elasticity (relevant for our purposes). On the one hand, if one assumes a “putty-clay” technology,

⁹ Technically, Solow's calculation also differs from (22) in allowing the factor shares to be time-varying. This amounts to preserving some higher-order terms in the Taylor series expansion of (1), but there is then little reason to drop other second-order terms. We thus stick here to a simple log-linear approximation.

the short run elasticity of substitution might be much less than that indicated by long run trends. But, on the other hand, cyclical variations in capital utilization might make the relevant short run elasticity even greater than the long run elasticity. Suppose that the current production function is not (1) but

$$Y_t = F(u_t K_t, z_t(H_t - \bar{H}_t))$$

where u_t represents the degree of utilization of the capital stock. Then if u_t varies positively with $z_t(H_t - \bar{H}_t)/K_t$ at cyclical frequencies, the relevant elasticity e in the above calculations is the one associated with the reduced-form production function

$$Y_t = \tilde{F}(K_t, z_t(H_t - \bar{H}_t)) = F\left(u\left(\frac{z_t(H_t - \bar{H}_t)}{K_t}\right) K_t, z_t(H_t - \bar{H}_t)\right)$$

But, if the long run utilization of capital is constant and thus independent of trends in factor prices, the elasticity one would infer from growth observations would be that associated with the true production function evaluated at constant u .¹⁰ In this case, the measured long run elasticity of substitution would be smaller than the relevant short run elasticity. We must thus admit that the relevant elasticity is not easily measured. We take as our baseline case the value $e = 1$ (Cobb-Douglas), the value most often used in real business cycle studies, but we also consider the possibilities $e = 0.5$ and $e = 2$.

We are similarly unable to directly observe μ^* . Hall (1988a) proposes to measure it on the basis that the \hat{z}_t series given by (22) should be orthogonal to changes in variables such as real military purchases or the party of the President. Hall uses value added as his measure of output and finds values above 1.8 for all seven of his 1-digit industries. Domowitz, Hubbard, and Petersen (1988) use gross output instead and obtain smaller estimates of μ^* for most industries; a value of around 1.6 is typical of their findings. However, since we study the behavior of value added, their estimate would have to be adjusted upward to be appropriate for our analysis. Nonetheless, we take 1.6 as our baseline case, but also consider the value 2. As some readers may be skeptical about the existence of markups even as high as 60%, we present some results for a markup variation series

¹⁰ The difference in response to cyclical as opposed to secular changes might be due to adjustment costs, so that changes in capital utilization would be used more in the case of transitory fluctuations. Properly taking into account such adjustment costs would, of course, require a more complicated specification of production possibilities.

constructed under the assumption $\mu^* = 1.1$, although we regard this as an extremely conservative choice.

Figures 1, 2, and 3 illustrate the constructed series for markup growth rates over the postwar period, under different assumptions regarding μ^* and e . These are constructed by ignoring the departures of capital from trend, \hat{k} , and using the data described in section 8 below. Because we make an assumption about the average level of the markup in order to construct the series, we present here only our constructed series for markup changes, to make it clear that we do not pretend to have directly measured the level. Figure 1 represents our baseline case, $\mu^* = 1.6$, $e = 1$. Figure 2 shows the consequences of assuming instead $e = 0.5$, while Figure 3 presents the case $\mu^* = 2$, $e = 1$. In each case, the growth rate of hours is shown as well; it is clear that for each of these sets of parameters the constructed series displays strongly countercyclical markup variations.

The effects of parameter variation are easily understood. Assuming a lower elasticity e implies a sharper decline in the marginal product of hours in booms, and so increases the amplitude of the countercyclical variation in the series constructed for $\hat{\mu}_t$. Assuming a higher μ^* implies a higher steady state \bar{H}/H because of (21), and hence a larger estimate of the percentage increase in $H_t - \bar{H}_t$ for any given observed increase in H_t . For any given e , this then implies a sharper decline in the marginal product of hours in booms, so that a higher μ^* results in a greater amplitude of countercyclical variation in $\hat{\mu}_t$. (Note the different scales for the markup series in Figures 1-3.)

Our results on the countercyclical pattern in the markup confirm the conclusion of Bilts (1987), although we obtain this result for a different reason. Focusing on the baseline case of $e = 1$, (23) becomes

$$\hat{\mu}_t = \hat{y}_t - \frac{\mu^* s_H}{1 - \mu^* s_K} \hat{h}_t - \hat{w}_t = -\hat{s}_{Ht} - \left(\frac{\mu^* s_H}{1 - \mu^* s_K} - 1 \right) \hat{h}_t \quad (24)$$

where \hat{s}_{Ht} denotes log deviations of the share of hours. If μ^* equals one, and given that $s_H + s_K = 1$ (so that there are then also no fixed costs), $\hat{\mu}_t$ is simply the negative of \hat{s}_{Ht} , which is not very strongly cyclical. But if we assume $\mu^* > 1$ (and hence increasing returns), then a countercyclical term is added to $\hat{\mu}_t$. Bilts assumes instead constant returns (and ignores the final term in (24)), but points out that the relevant wage \hat{w}_t is the marginal wage (the wage paid for marginal hours) rather than the average wage. These two can differ if the utilization of overtime labor is cyclical and if overtime

hours must be paid more than straight-time hours. With this correction, he obtains

$$\hat{\mu}_t = -\hat{s}_{Ht} - \hat{u}_t$$

where \hat{u}_t represents the log deviation of the ratio of the marginal wage to the average wage. In the Appendix we show how to compute this correction with our data. Bils' method for estimating \hat{u}_t depends crucially upon regarding the overtime premium as allocative. For a criticism, see Hall (1988b). Because we are uncertain of the extent to which Bils' treatment of the overtime premium is justified, we present most of our results without this correction.

Our specification of production possibilities is obviously overly simple in many respects, and many of its shortcomings deserve more careful attention in the future. We should, however, note that many of the most obvious corrections to our simple measure would tend to imply even stronger evidence for countercyclical markup variation. One might wish to consider adjustment costs for hours. In this case (19) becomes

$$\mu_t = \frac{F_H(K_t, z_t(H_t - \bar{H}_t))}{w_t + \lambda_t}$$

where λ_t represents the shadow cost of increasing hours in period t in addition to the wage. Assuming a convex adjustment cost function, λ_t will be positive when hours are increasing (due to current adjustment costs) or higher than they are expected to be in the future (due to expected future adjustment costs), and similarly negative when hours are decreasing, or lower than they are expected to be in the future. Hence λ_t should be a procyclical correction, and imply an even more countercyclical markup.

As another example, one might wish to consider composition biases due to the heterogeneity of different workers' hours. As many studies have shown (see, e.g., Kydland and Prescott (1988), Barsky and Solon (1989)), the most important such bias has to do with the greater cyclical variability of low wage (and presumably low-productivity) hours. The precise effect of such bias on our conclusions depends upon many aspects of the assumed true model of heterogeneous hours, but it seems likely that it reduces the extent to which measured markup variations are countercyclical. Suppose that low-wage and high-wage hours are two distinct factors of production, and assume a Cobb-Douglas production function. We can measure the markup as the ratio of the marginal

product of low-wage hours to the low wage.¹¹ Then, corresponding to (24), one obtains

$$\hat{\mu}_t = -\hat{s}_{HLt} - \left(\frac{\mu^* s_{HL}}{1 - \mu^* s_K} - 1 \right) \hat{h}_{Lt}$$

where \hat{h}_{Lt} represents the log deviation of low-wage hours from trend, s_{HL} represents the trend value of the share of payments to low-wage hours in output, and so on. Both \hat{s}_{HLt} and \hat{h}_{Lt} should be more procyclical than the corresponding \hat{s}_{Ht} and \hat{h}_t in (24). These considerations would both tend to make $\hat{\mu}_t$ more countercyclical when hours are disaggregated. On the other hand, s_{HL} is smaller than s_H , so the direction of the overall bias is not certain.

6. Method for evaluating the competing theories

Our three theories all yield relationships of the form $\mu_t = \mu(X_t, Y_t)$. For purposes of estimation, we adopt the log-linear approximation

$$\hat{\mu}_t = \epsilon_X \hat{x}_t - \epsilon_Y \hat{y}_t + \eta_t \tag{25}$$

where, again, hatted variables represent logarithmic deviations from trend values, while η_t represents a possible stochastic disturbance. Examples of stochastic disturbances of this type include changes in antitrust enforcement and changes in the degree of foreign competition. The static theory of section 2 implies that ϵ_X is zero. The customer market model implies that $\epsilon_X < 0$ while the implicit collusion model has $0 < \epsilon_X < (\mu^* - 1)$. If one imposes the additional requirement that the preferences are homothetic, all models imply that ϵ_Y is equal to ϵ_X . Even without imposing homotheticity, the dynamic models imply that ϵ_Y has the same sign as ϵ_X unless the elasticity of demand is extremely dependent on the level of demand.

The problem with estimating (25) is that we lack direct observations on \hat{x}_t . We have two methods for dealing with this issue. The first uses measurements of Tobin's q , the ratio of firms' market value to the value of their capital in place. The total market value of all firms is equal to $X_t + K_t - \Phi_t + N_t$ where K_t equals the current returns to capital plus the present value of the depreciated capital stock at the beginning of next period Φ_t is the present value of fixed costs and N_t captures any additional influences on market values such as the the present discounted value of

¹¹One can equivalently consider the marginal product of high wage hours but adjustment costs are more likely to distort the results in this case.

taxes levied from firms as well as random misvaluations of the stock market. Then the logarithmic deviation of Tobin's q should equal

$$\hat{q}_t = \frac{X}{X + K - \Phi + N} \hat{x}_t + \frac{\Phi - X - N}{X + K - \Phi + N} \hat{k}_t - \frac{\Phi}{X + K - \Phi + N} \hat{\phi}_t + \frac{1}{X + K - \Phi + N} \hat{n}_t$$

where the ratios with $(X + K - \Phi + N)$ in the denominator represent steady state values, and where \hat{n}_t represents a deviation (rather than a logarithmic deviation) from the steady state value N .

Tobin's q is one on average¹² so that $(X + N - \Phi)$ equals 0. Letting ν_t represent $\frac{K\epsilon_X}{X}$ times the last two terms in the previous equation, and using (25) we have

$$\hat{\mu}_t = \frac{K}{X} \epsilon_X \hat{q}_t - \epsilon_Y \hat{y}_t + \nu_t + \eta_t \quad (26)$$

Equation (26) can be estimated by ordinary least squares if one is willing to assume that the residual $\nu_t + \eta_t$ is uncorrelated with any two of the three variables in the equation. In particular, if $\nu_t + \eta_t$ is uncorrelated with \hat{q}_t and \hat{y}_t , we can recover its coefficients by regressing \hat{q}_t on the other variables. For the residual to have this property, ν_t would have to be unimportant and η_t would have to be correlated only with μ_t . However shocks to η_t such as changes in antitrust enforcement might well have a direct effect on Y . Moreover, such shocks are likely to be serially correlated so that positive realizations of η_t raise X_t and \hat{q}_t as well. As a result, our estimate of $\frac{K}{X} \epsilon_X$ is likely to be biased upwards. On the other hand, the existence of important variation in ν_t would tend to bias this coefficient toward zero if these variations affect only q_t .

If, instead, η_t is unimportant and ν_t is uncorrelated with $\hat{\mu}_t$ and \hat{y}_t , then the coefficients in (26) can be recovered by regressing \hat{q}_t on the other variables. Examples of shocks to ν_t with this property might include those that affect N_t and some of those that affect Φ_t . One immediate difficulty with this reverse regression is that increases in real discount rates lower Φ_t and ν_t so that, insofar these raise $\hat{\mu}_t$, the coefficient of $\hat{\mu}_t$ will be biased downwards. Another difficulty is that important variations in η_t would bias the coefficient of $\hat{\mu}_t$ in the reverse regression toward zero so that the estimated ϵ_X would be too large.

Our second procedure starts from the observation that

¹²This is consistent with the absence of equilibrium pure profits ($X = \Phi$) and with an average N of zero.

$$X_t = E_t \left\{ \frac{q_{t+1}}{q_t} [\Pi_{t+1} + X_{t+1}] \right\} \quad (27)$$

In the steady state where capital, output and profits grow at the rate g , the trend value of X_t equals the trend value of Π_t divided by $(r^* - g)$ where r^* is the trend value of the real rate at which profits are discounted. Therefore, the log-linearization of (27) gives

$$\hat{x}_t = E_t \left\{ \frac{(r^* - g)\epsilon_X \hat{\pi}_{t+1} + (1 + g)\hat{x}_{t+1} - \hat{r}_{t+1}}{1 + r^*} \right\}$$

where \hat{r}_t is the deviation from trend of the real rate of return between $t - 1$ and t . Moreover, (3) implies that $\hat{\pi}_t$ is equal to $\hat{y}_t + \hat{\mu}_t / (\mu^* - 1)$. Hence (25) and (27) together imply that

$$\begin{aligned} \hat{\mu}_t + \epsilon_Y \hat{y}_t = E_t \left\{ \left(1 - \frac{(r^* - g) \left[1 + \frac{\epsilon_X}{\mu^* - 1} \right]}{1 + r^*} \right) (\hat{\mu}_{t+1} - \eta_{t+1}) \right. \\ \left. \left(\epsilon_Y + \frac{(r^* - g)(\epsilon_X - \epsilon_Y)}{1 + r^*} \right) \hat{y}_{t+1} - \frac{\epsilon_X}{1 + r^*} \hat{r}_{t+1} \right\} + \eta_t \end{aligned} \quad (28)$$

If one eliminates the expected value operator from (28) and ignores the η terms one obtains an equation whose residual is supposed to be uncorrelated with information available at t . Following the suggestion of Hansen (1982) we estimate this equation by instrumental variables.

The presence of the *eta*'s might affect the results from this procedure. In this case, however, the estimate of ϵ_X is biased only if such shocks have effects on both the expected rate of change in the markup (since $\hat{\mu}_{t+1}$ enters with a coefficient almost equal to one) and on the expected rate of return on financial assets. While this is certainly a possibility it seems less likely than that such shocks affect the *levels* of q and the markup simultaneously.

7. Data

Our time series for Tobin's q comes from Blanchard, Rhee and Summers (1989). Our measure of the output (value added) of the private sector is obtained from the NIPA as the difference between GNP and the value added by the Federal, State and local governments. Our index of the prices of goods is the ratio of nominal to real private value added. Our measure of private hours is obtained from the establishment survey as the difference between total hours in nonagricultural payrolls and hours employed by the government. These hours do not have exactly the same coverage as our

output series. Thus, for our measures to be strictly accurate, the percentage changes in agricultural hours must equal the percentage changes in the hours of private nonagricultural establishments.

We employ two measures of wages. The principal one is a measure of hourly compensation. This measure equals private employee compensation from the NIPA (i.e. total compensation minus government compensation) over our measure of private hours. The second measure is average hourly earnings in manufacturing. One advantage of the compensation series is that it has a larger coverage both in terms of the sectors whose payments are recorded and in terms of the forms of compensation that are included.¹³

8. Empirical Determinants of Markup Variation

We present results both from estimating equation (26) via ordinary least squares and from estimating (28) with instrumental variables. We start with estimates of equation (26) with the markup on the left and q on the right hand side. This specification makes sense if the variations in ν_t can be neglected while those of η_t affect only μ_t .

Since the variables are supposed to be logarithmic deviations from trend we include the logarithms of q , y and the constructed markup as well as a constant and a linear trend. Our baseline markup variation series is constructed assuming an average markup μ^* equal to 1.6 and an elasticity of substitution of capital for labor e equal to 1.0, and ignoring the overtime premium. We also allow the residual η_t to have first order serial correlation, so that it equals $\rho\eta_{t-1}$ plus an i.i.d. disturbance. Under this specification, estimation of (26) for the period 1947.III to 1988.IV yields

$$\mu_t = -0.72 - 0.002 t - 0.42 y_t + 0.035 q_t$$

$$(0.6) \quad (0.0007) \quad (0.09) \quad (0.014)$$

$$\text{Period: 1952:II-1988:4; } \quad \rho = 0.934 \quad R^2 = 0.997; \quad \text{D.W.} = 1.54$$

Both coefficients are positive as is predicted by the implicit collusion model and thus of the opposite

¹³A second advantage is that there is reason to believe the compensation series has smaller measurement error, at least in the way we use it. We use the real wage only to construct our series on markups. Ignoring fluctuations in capital, equation (23) gives the detrended markup as a function of the detrended levels of output, \hat{y}_t , hours, \hat{h}_t and the real wage, \hat{w}_t . A simple transformation allows one to write the detrended markup as a function of the detrended labor share ($\hat{s}_{Ht} = \hat{w}_t + \hat{h}_t - \hat{y}_t$, detrended output and detrended hours. The use of the two different wage series is thus equivalent to the use of the corresponding two series for fluctuations in the labor share. To see which series has more classical measurement error we use US data from 1947.III to 1989.I to run regressions of the logarithm of one share on the other including a trend and a correction for first order serial correlation. When the share using hourly earnings is on the right hand side its coefficient equal 0.73 and is statistically different from one. When that using compensation is on the right hand side, its coefficient is 0.93 and is not statistically different from one. We thus cannot reject the hypothesis that the earnings share equals the compensation share plus noise.

sign than the coefficients predicted by the customer market model. Moreover, since both coefficients are significantly different from zero at conventional significance levels, the customer market model is statistically rejected. The fact that ϵ_X is statistically different from zero also leads us to reject static models of the markup where the only determinant of the markup is the current level of output.

According to (26), the coefficient on y_t is ϵ_Y while that on q_t is $\frac{K\epsilon_X}{X}$. To obtain an estimate of ϵ_X we thus must obtain a measure for $\frac{X}{K}$. According to our model, this equals $\frac{(1-1/\mu^*)Y}{(r^*-g)K}$ which equals $30Y/K$ for our base case. The coefficient on q_t must thus be multiplied by $30Y/K$ to obtain an estimate of ϵ_X . Since Y/K is roughly 10,¹⁴ the implied value for ϵ_X is approximately 0.1. This is consistent with the restriction that ϵ_X be smaller than $\mu - 1$.

We show in Table 1 how these coefficients vary as we vary μ^* and e . Increases in μ^* raise the variability of the markup. In particular, they amplify the reduction in $\hat{\mu}_t$ for a given increase in \hat{h}_t . As a result, a given increase in \hat{y}_t reduces the markup by more. This explains why the coefficient on y_t falls as μ^* rises. What is somewhat more unexpected is that increases in μ^* also raise the coefficient on q so that the implied value of ϵ_X rises as well.

For a given average markup, increases in e raise the coefficient on y_t while having no effect on the coefficient on q_t . The reason for this apparently anomalous result can be seen from the formula (23) giving our measure of markup variations. For given μ (and hence $\frac{H}{H-R}$) changes in e affect markup variations only by affecting the influence of private output on the markup. In particular increases in e raise the weight of changes in output on the measured markup. These increases therefore raise the estimated effect of y_t on μ_t .

We now turn to estimation of the same equation but with q_t on the left hand side. This produces less biased coefficients if there are important variations in ν_t and unimportant movements in η_t . We again, let the residual in the equation have first order serial correlation. For our baseline series on markup variations, the estimation of such an equation including both a constant and a trend yields:

¹⁴ See Rotemberg and Woodford (1989).

$$q_t = -4.66 - 0.006 t + 1.29 y_t + 1.20 \mu_t$$

$$(3.5) \quad (0.006) \quad (0.52) \quad (0.48)$$

$$\text{Period: 1952:II-1988:4; } \rho = 0.969 \quad R^2 = 0.952; \quad \text{D.W.} = 1.81$$

where the coefficient on the markup equals $\frac{X}{K\epsilon_X}$ and that on private value added equals $\frac{X\epsilon_Y}{K\epsilon_X}$. The estimates of both ϵ_Y and ϵ_X are positive. In addition, the ratio of the coefficient on \hat{y}_t over that on $\hat{\mu}_t$ gives ϵ_Y which is thus estimated to be near one.

To obtain an estimate of ϵ_X we must, again, obtain a measure for $30Y/K$. For plausible values of Y/K , the resulting estimate of ϵ_X is very large, too large to be consistent with any of our three models. One reason for this result may be that increases in expected returns lower X_t and Φ_t at the same time so that they also raise ν_t . Insofar as this increase in expected returns reduces markups, the coefficient on markups will be too small and our estimate of ϵ_X will be too large. Variations in η that are correlated only with μ_t have the same effect since they bias the coefficient on μ_t toward zero.

In Table 2 we show how the coefficients on $\hat{\mu}_t$ and \hat{y}_t vary as we vary μ^* and e . As we increase the average markup (and hence increase its variability) the correlation between the markup and stock prices falls so that the former falls. In contrast, the latter coefficient estimate rises as we increase the average markup.

For a given average markup, increases in e lower the estimated value of $\frac{X\epsilon_Y}{K\epsilon_X}$ while having no effect on the estimate of $\frac{X}{K\epsilon_X}$. The reason for this is, once again, that increases in e raise the influence of y_t on μ_t . Increases in e therefore reduce the regressions estimate of the independent effect of output on stock prices.

We now turn to the estimation of (28) via instrumental variables. There are several advantages to this procedure. First, the estimates are somewhat less subject to endogeneity bias. Second, the method does not require observations on the present discounted value of profits X . It does however require information on discount rates (or marginal rates of substitution). Given the inadequacies of various rates of return as discount rates we experiment with the return on the stock market, the return on Treasury Bills and the return on prime commercial paper. Third, it allows us to recover quantitative estimates for both ϵ_Y and ϵ_X more easily.

We include a constant and a trend as well as the logarithms of the markup, output, hours,

the real wage and the level of real returns in our estimation. As instruments we use a constant, a linear trend, the current and one lagged value of the logarithms of output, the labor input and the real wage as well as the *ex post* real return between $t - 1$ and t .

The results of estimating (28) for the period 1947.III to 1988.IV using our baseline markup series and the return on the stock market are presented in Table 1. We show estimates and summary statistics for both the case where $\epsilon_Y = \epsilon_X = \epsilon$, and for the case where ϵ_Y and ϵ_X are allowed to differ.

The summary statistics reported in table 1 concerning the fit of the two equations are encouraging. The Durbin-Watson statistic reveals that little serial correlation remains in the errors. Because we use more instruments than there are coefficients, the two equations are overidentified. The test statistic proposed by Hansen (1982) to test these overidentifying restrictions is reported in the row marked J and is distributed χ^2 with 5 and 6 degrees of freedom under the null hypothesis that the restrictions are valid. The actual values of this statistic are very small, which probably indicates that the instruments are quite collinear.

Turning to the estimates, consider first the case where ϵ_Y and ϵ_X are not constrained to be equal. A 1% increase in X is then estimated to raise the markup by about a fifth of a percentage point. A 1% increase in Y by contrast lowers the markup by about 1%. Both these coefficients are again and significant.

The estimates of ϵ_Y and ϵ_X are inconsistent with the homothetic versions of both dynamic models because they are statistically significantly different from each other. Once homotheticity is dropped, ϵ_Y can be larger than ϵ_X as long as the elasticity of demand is higher when Y is large. Then, increases in Y raise disproportionately the number of customers that a deviator gets for a given change in his markup. This disproportionate increase implies that deviations become much more attractive when Y increases. They thus require relatively large reductions in the markup.

Measurement difficulties provide an alternative explanation for the difference between the two coefficients. To gain some intuition into the source of this discrepancy suppose first that the average real discount rate r^* equals the average growth of private value added g . Then, (28) makes the expected change in the logarithm of the markup between t and $t+1$ a linear function of the expected change in logarithm of private value added (with coefficient ϵ_Y) and of the expected real return

rate between t and $t + 1$ (with coefficient ϵ_X).

Given that the difference between r and g is in fact quite small (it equals 0.0126), the finding that ϵ_Y exceeds ϵ_X is a finding that the expected change in private value added is more correlated with the change in the markup than the expected discount rate. This could well be due to the fact that the relevant discount rate for firms differs from the expected return on stocks, so that the measurement error in \hat{r}_t biases the estimate of ϵ_X downwards.

An additional prediction of the implicit collusion model is that ϵ_X should be less than $\mu - 1$. This restriction is satisfied whether ϵ_Y and ϵ_X are allowed to differ as in the first column or whether they are constrained to be equal as in the second column. In the latter column, the estimate the elasticity of the markup with respect to X/Y , ϵ , of 0.22 which is well below .6 while remaining significantly positive.

The difference between the J statistics reported in the two columns can be used to test whether the restriction that the two elasticities are the same is valid. This is the analogue of the likelihood ratio test proposed by Gallant and Jorgenson (1979) and it sometimes produces inferences which are at variance with those from Wald tests based on the standard errors of the coefficients. Indeed, in this case, the Wald test rejects the equality of the two coefficients but the difference between the two statistics is 1.35 which is well below the critical value for the χ^2 distribution with one degree of freedom.

In Tables 2, 3 and 4 we report variations on the model which are designed to gauge the robustness of our results. Tables 2 and 3 are devoted to obtaining estimates for different values of the average markup and for different values of the elasticity of substitution. We consider in particular elasticities of substitution equal to 0.5, 1 and 2 and average markups of 1.1 (which is much smaller than what is found by the methods of Hall), 1.6 (our base case) and 2. Table 2 is devoted to estimates when the two elasticities are equal while the estimates of Table 3 are obtained without imposing this restriction.

The two parameters μ^* and e affect the results. As explained in Section 5, increases in μ^* and reductions in e both increase the tendency of the markup to be countercyclical. It is thus not surprising that our estimates of ϵ in Table 2 and those of ϵ_Y in Table 3 rise with μ^* and fall with e . What is, once again, more surprising is that the estimates of ϵ_X in Table 3 which correspond

to estimates of the effect of expected rates of return on the markup also increases with μ^* and falls with ϵ . One notable feature of Table 2 is that, with the exception of the estimates for an elasticity of substitution of 0.5 and an average markup of 1.1, the estimates of ϵ are lower than the corresponding $\mu^* - 1$ as required by our theory. The estimates of ϵ_X in Table 3 are below $\mu^* - 1$ as well.

Table 4 presents other variations while holding the average markup and elasticity of substitution fixed at our base levels of 1.6 and 1. The first variation replaces one useful instrument of real returns on stocks, the lagged return, by another, namely the lagged dividend price ratio.¹⁵ This has no material effect on our results.

The next two lines present the estimates when we use either the Treasury Bill rate plus a constant or the commercial paper rate plus a constant as a discount rate. The reason we have to add constants to these rates is that the average real return on these instruments is lower than the economy's growth rate. Therefore a risk premium must be added to these rates to make the firm's problem well defined. Somewhat arbitrarily we choose risk premia which are equal to the difference between the economy's growth rate and the average return on the instrument that is being considered. We did this to ensure that the ϵ_X is estimated only from the variation in *ex ante* returns. However, small differences in this assumed risk premium have no effect on our conclusions.

What the second and third row of Table 4 show is that the use of alternative instruments does not affect the magnitudes of the coefficients although it does affect the standard errors. In particular, when we use the return on Treasury Bills, the estimate of ϵ_X is not statistically significantly different from zero. Because the commercial paper rate is probably a more accurate representation of firm's discount rates it is comforting that the results using this latter return are somewhat more significant. The last two rows show that the results are not sensitive to our use of hourly compensation instead of hourly earnings in manufacturing. Whether we use stock returns or return on Treasury Bills, this measure of wages produces essentially the same results as hourly compensation.

We finally consider the sensitivity of our results to the addition of the Bils correction for the difference between the average and marginal wage. We obtain this correction using the method spelled out in the Appendix. The resulting correction reasonably substantial. We estimate that

¹⁵ See Keim and Stambaugh (1986).

the increased use of overtime implies that, when hours rise by 1% the average wage rises by 0.056 of 1%, while the marginal wage rises by 0.417 of 1%. Using the resulting markup series estimation of (26) for our base case yields

$$\mu_t = -0.56 - 0.002 t - 0.66 y_t + 0.043 q_t$$

(0.7) (0.0009) (0.10) (0.017)

Period: 1952:II-1988:4; $\rho=0.944$ $R^2=0.998$; D.W.=1.54

The reverse equation with q on the left hand side yields instead

$$q_t = -4.92 - 0.006 t + 1.49 y_t + 1.06 \mu_t$$

(3.5) (0.006) (0.55) (0.41)

Period: 1952:II-1988:4; $\rho=0.969$ $R^2=0.952$; D.W.=1.82

In both cases, the estimate of ϵ_Y rises with the correction. This is not surprising since the correction makes marginal cost more procyclical. However, the estimates of ϵ_X are not very much affected by the correction. When we estimate (28) via instrumental variables using the corrected markup series, both estimates rise. The parameters ϵ_X and ϵ_Y are then estimated to equal 0.281 and 1.443 respectively, and their standard errors are 0.07 and 0.24. While both elasticities are now higher, ϵ_X remains much smaller than $\mu - 1$. Note also that the rejection of the two alternative models is even stronger in this case.

9. Conclusion

Our results provide further evidence that the markup of prices over marginal cost moves countercyclically over the business cycle. We have also found that the type of markup variations that occur are reasonably consistent with the predictions of the model of endogenous markup determination that we have previously discussed (Rotemberg and Woodford (1989)). Our results are quite inconsistent with the other leading "dynamic" model of markup determination, that proposed by Phelps and Winter (1970), and we are also able to reject a simple "static" specification, according to which the elasticity of demand varies with the level of aggregate demand. These conclusions are consistent with our previous analysis of the effects of military purchases on economic activity; in Rotemberg and Woodford (1989) we also found empirical regularities consistent with the implicit collusion model that are harder to reconcile with either of the alternatives.

These conclusions suggest that markup variations may well be an important mechanism by which changes in the demand for goods translate into changes in output. In the case of competitive product markets, firms' demand for hours H_t is a decreasing function of the current real wage w_t , given by the relation:

$$F_H(K_t, z_t H_t) = w_t$$

An increase in the demand for goods cannot shift this labor demand curve, because the capital stock is predetermined (in the short run) and the demand for goods should not affect the state of technology z_t . Hence an increase in the demand for goods, whether as a result of an increase in government purchases, a change in export or investment demand etc. can increase output (and hours) only insofar as the short run labor supply curve is shifted outward, and firms move down the labor demand curve in response to lower real wages. Such labor supply shifts can result from intertemporal substitution and, in the case of government purchases, be the consequence of wealth effects.

But, if this is the way demand shocks affect the economy, they cannot be very important since real wages fail to be countercyclical (and are furthermore procyclical when composition biases associated with aggregation of high-wage and low-wage hours are taken account of). Therefore, only shifts in the aggregate production function can be very important in explaining business cycles if one maintains perfect competition.

For shocks to the demand for goods to be important, they must increase firm's willingness to hire additional workers at a given real wage. As Kalecki (1938) and Keynes (1939) recognized, this requires that the markup fall in the modified demand for labor function:

$$F_H(K_t, z_t H_t) = \mu_t w_t \tag{29}$$

We have presented evidence that, as required, the markup is indeed countercyclical, and furthermore it moves in a way that is consistent with a coherent theory of endogenous markup determination. That theory, the oligopolistic collusion model, explains how transitory movements in aggregate demand for produced goods can result in fluctuations in labor demand of the same sign. The customer market model, by contrast, would be much less promising as a basis for such

a theory, for most of the sources of variation in aggregate demand that one might be interested in - for example, temporary increases in export demand, shifts of tastes toward present as opposed to future consumption, or investment booms due to discovery of especially productive investment opportunities - would all tend to raise current output relative to future profits, and to raise real interest rates. Therefore, markups would tend to increase shifting the labor demand curve inward so that an even greater labor supply shift and real wage decline is needed to increase output than in the competitive model.¹⁶

A simple "static" specification, $\mu_t = \mu(Y_t)$, would mean replacing (29) by:

$$F_H(K_t, z_t H_t) = \mu(F(K_t, z_t H_t))w_t \quad (30)$$

This still establishes a relation between w_t and H_t that cannot be shifted, in the short run, by anything other than a technology shock. Thus, in such a theory, aggregate demand variations can affect output and hours only by shifting the short run labor supply curve. If μ' is sufficiently negative, such variations along the fixed schedule (30) might involve an acyclical or even procyclical real wage. But it would seem to us undesirable to place the entire responsibility for the efficacy of aggregate demand variations upon their effects upon household labor supply.

If one assumes a representative consumer with preferences that are additively separable between periods, as is common in the real business cycle literature, the first order condition for optimal labor supply is:

$$\text{MRS}(C_t, H_t) = w_t$$

where C_t denotes aggregate consumption. It is usual to suppose that this marginal rate of substitution between consumption and leisure is increasing in consumption and decreasing in the quantity of leisure. Hence an outward shift in the labor supply curve must coincide with a decrease in consumption; but consumption is procyclical rather than countercyclical.¹⁷ Moreover, such a model would make it hard to understand the procyclical variation in vacancy rates (the relative constancy of the Beveridge curve) and quit rates (see, for instance, Parsons (1973)) which both suggest procyclical movements in firm's willingness to hire workers at a given real wage. Furthermore, such a

¹⁶In fact, increases in government purchases may well reduce output and employment in such a model. See Phelps (1989).

¹⁷In Rotemberg and Woodford (1989) we show that increases in military spending lead consumption to rise rather than to decline.

model would rely critically both upon labor-supplying households being able to quickly understand and respond to the future consequences of current aggregate shocks, and upon those households' effective integration into the capital markets, either of which might be doubted. Finally suppose that, as in many popular efficiency wage models, households are rationed in the amount of labor they can supply. Then, variations in desired household labor supply due to wealth effects or intertemporal substitution may have little effect upon which of the points consistent with (30) is realized in equilibrium.

Hence a "dynamic" model of markup determination, such as the oligopolistic collusion model, offers the greatest promise as a basis for understanding the role of aggregate demand variations in the generation of business cycles, quite apart from the evidence provided here in support of such a specification. Further quantitative investigations of the extent to which this model can account for the character of observed aggregate fluctuations, under various hypotheses about the ultimate driving shocks, would seem to be warranted.

Appendix

The Bils Correction

In this Appendix we consider the computation of \hat{u}_t , the variations in the ratio of the marginal wage to the average wage in the presence of varying use of overtime workers. Bils assumes that total wage payments can be written as:

$$w_t[H_t + pV(H_t)]$$

where w_t is the straight-time wage, p is the overtime premium which is assumed to equal 50%, and $V(H)$ indicates how many overtime hours firms employ as a function of total hours. Therefore, the marginal wage (the increased expenditure when hours rise by one unit) is:

$$w_t[1 + pV'(H_t)]$$

while the average wage is

$$w_t \left[1 + p \frac{V(H_t)}{H_t} \right]$$

Therefore the percent change in the marginal wage for a 1% increase in employment is

$$\gamma_M = \frac{pV''H^2}{H + pV'H}$$

while the corresponding percent change in the average wage is

$$\gamma_A = \frac{p(V'H - V)}{H + pV}$$

The logarithmic deviation of the ratio of marginal to average wage, \hat{u}_t is then equal to $(\gamma_M - \gamma_A)\hat{h}_t$.

To obtain estimates of γ_M and γ_A we use the available data on overtime which, unfortunately, cover only the manufacturing sector. We assume that the actual value of overtime hours V_t is given by $V(H_t)$ times a stationary residual. We thus run a regression of \hat{u}_t , the detrended logarithm of overtime hours, on \hat{h}_t and \hat{h}_t^2 . Allowing for an error with both first and second order serial correlation yields

$$\hat{v}_t = 7.01 \hat{h}_t + 2.69 \hat{h}_t^2 t$$

$$(0.59) \quad (8.11)$$

Period: 1956:III-1989:I; AR(1)=1.17 AR(2)=0.24 R²=0.97; D.W.=2.05

Assuming that the residual is uncorrelated with the right hand side variables, the coefficients in this regression are the coefficients in a second order logarithmic expansion of $V(H)$. Thus, the first coefficient equals $V'H/V$ while the second equals one half of $\frac{H^2 V''}{V} + \frac{V'H}{V} - (\frac{V'H}{V})^2$. Using these facts, together with knowledge that in our data V/H equals 0.0187, gives a value for γ_M of 0.417 and one for γ_A of 0.056. As in Bils' analysis, the former is about eight times larger than the latter. Bils's estimates are both somewhat larger because his index of total hours covers only production hours in manufacturing, so that his average V/H is higher.

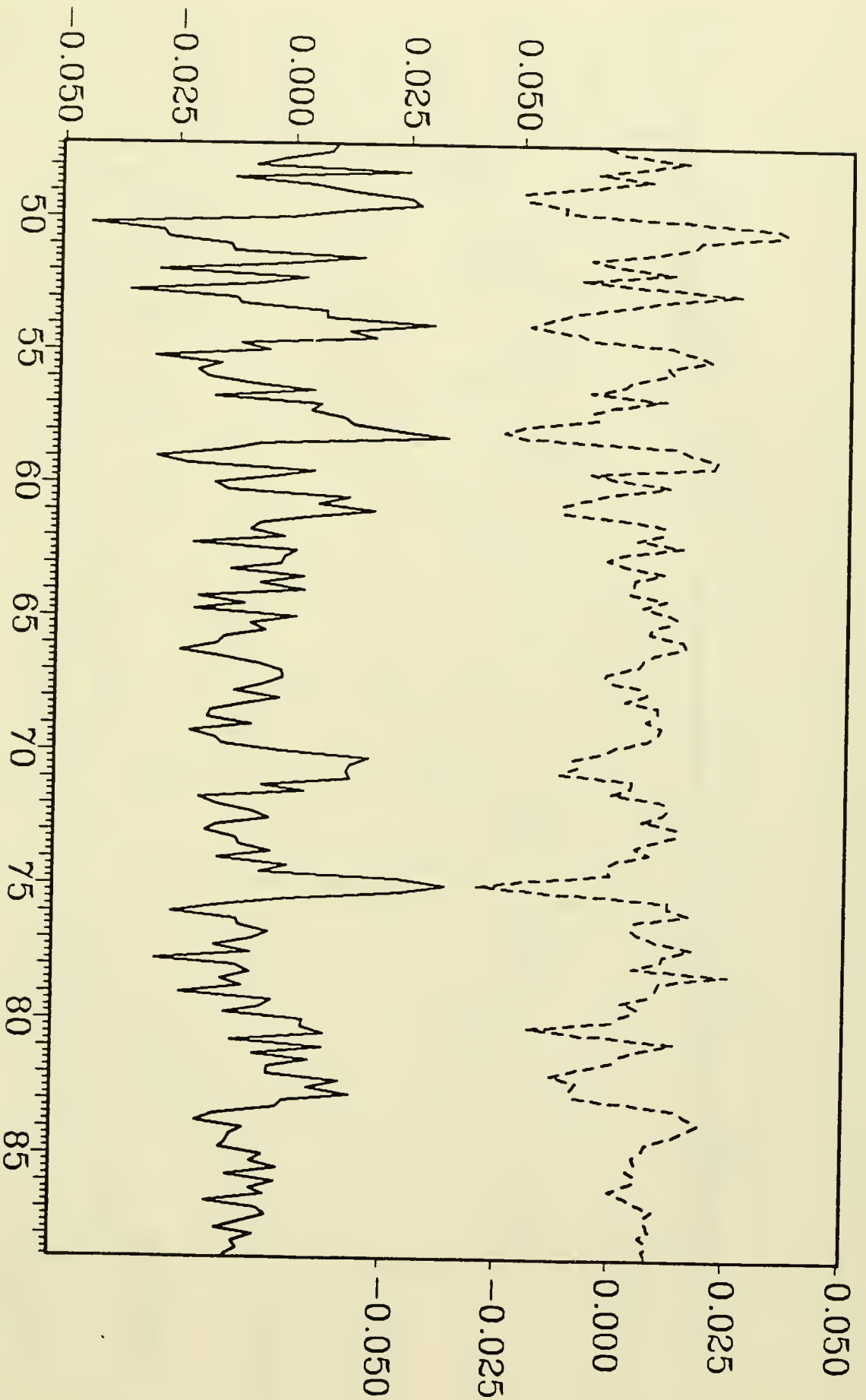
10. References

- Abreu, Dilip: "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, 39, 1986, 191-225.
- Barsky, Robert and Gary Solon: "Real Wages over the Business Cycle," NBER Working Paper 2888, March 1989.
- Bils, Mark: "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77, December 1987, 838-57.
- : "Pricing in a Customer Market," *Quarterly Journal of Economics*, November 1989.
- Blanchard, Olivier, Changyong Rhee and Lawrence Summers: "The Stick Market, Profit and Investment," National Bureau of Economic Research Working Paper 3370, May 1990.
- Dixit, Avinash and Joseph Stiglitz: "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 297-308, 1977.
- Domowitz, Ian, R. Glenn Hubbard, and Bruce C. Petersen, "Market Structure and Cyclical Fluctuations in U.S. Manufacturing," *Review of Economic Statistics*, February 1988.
- Dunlop, John T.: "The Movement in Real and Money Wage Rates" *Economic Journal*, 18, 1938, 413-34.
- Farrell, Joseph and Carl Shapiro: "Dynamic Competition with Switching Costs," *Rand Journal of Economics*, 19, Spring 1988, 123-37.

- Froot, Kenneth A. and Paul D. Klemperer: "Exchange Rate Pass-Through when Market Share Matters," *American Economic Review*, 79, September 1989, 637-54.
- Gallant, Ronald A. and Dale W. Jorgenson: "Statistical Inference for a System of Simultaneous, Non-Linear, Implicit Equations in the Context of Instrumental variables Estimation," *Journal of Econometrics*, 11, October/December 1979, 275-302.
- Gottfries, Nils: "Price Dynamics of Exporting and Import-Competing Firms," *Scandinavian Journal of Economics*, 88, 417-436.
- : "Market Shares and Pricing Behavior in the Tradeables Industry: An Examination of Swedish Manufacturing", Institute for International Economic Studies Paper No. 412.
- : "Customer Markets, Credit Market Imperfections and Real Price Rigidity" *Economica*, forthcoming, 1989 mimeo.
- Greenwald, Bruce and Joseph Stiglitz: "Financial Market Imperfections and Business Cycles," *NBER Working Paper 2494*, 1988.
- Hall, Robert E.: "The Relation Between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96, October 1988, 921-48.
- "Investment under Uncertainty: Theory and Test with Industry Data," *NBER Working Paper 2264*, May 1987.
- "Substitution over Time in Work and Consumption," National Bureau of Economic Research Working Paper 2789, December 1988b.
- Hansen, Lars Peter: "Large Sample Properties of Method of Moments Estimators," *Econometrica*, 50, July 1982, 1029-54.
- Kalecki, Michael: "The Determinants of the Distribution of National Income," *Econometrica*, 6, 97-112, 1938.
- Keim, D., and Robert Stambaugh: "Predicting Returns in the Bond and Stock Markets," *Journal of Financial Economics*, 17, 1986, 357-90.
- Keynes, John Maynard: "Relative Movements of Real Wages and Output," *Economic Journal*, 49, 34-51, 1939.
- Klemperer, Paul D.: "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, 102, May 1987, 375-94.
- Kydland, Finn E. and Edward C. Prescott: "Cyclical Movements of the Labor Input and its Real Wage," Federal Reserve bank of Minneapolis Working Paper 413, November 1988.

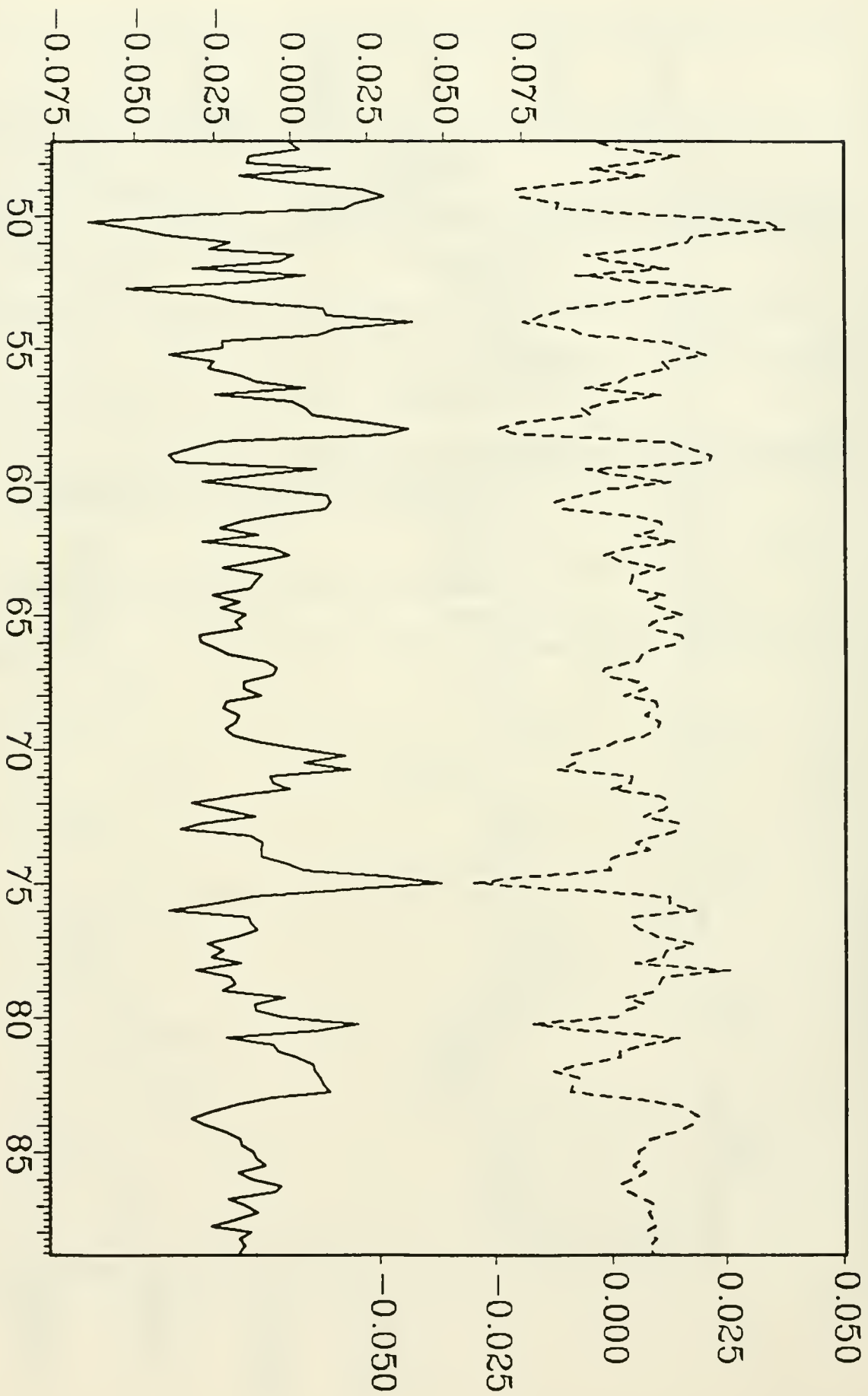
- Parsons, Donald O.: "Quit Rates over Time, A Search and Information Approach," *American Economic Review*, 63, June 1973, 390-401.
- Phelps, Edmund S.: "Aggregate Demand in a Non-Monetary Model of Labor-Market and Product-Market Equilibrium," mimeo, Columbia University, June 1989.
- Phelps, Edmund S. and Sidney G. Winter: "Optimal Price Policy under Atomistic Competition" in E. Phelps ed. *Microeconomic Foundations of Employment and Inflation Theory*, W.W.Norton and Co, New York, 1970.
- Robinson, Joan: "The Economics of Imperfect Competition," Macmillan, London, 1932.
- Rotemberg, Julio J. and Garth Saloner: "A Supergame-Theoretic Model of Price Wars during Booms," *American Economic Review*, 76, June 1986, 390-407.
- Rotemberg, Julio J. and Michael Woodford, "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," National Bureau of Economic Research, Working Paper No. 3206, Cambridge, Massachusetts, December 1989.
- Solow, Robert M.: "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39, August 1957, 312-20.
- Tarshis, L.: "Changes in Real and Money Wage Rates," *Economic Journal*, 19, 1939, 150-4.

Figure 1



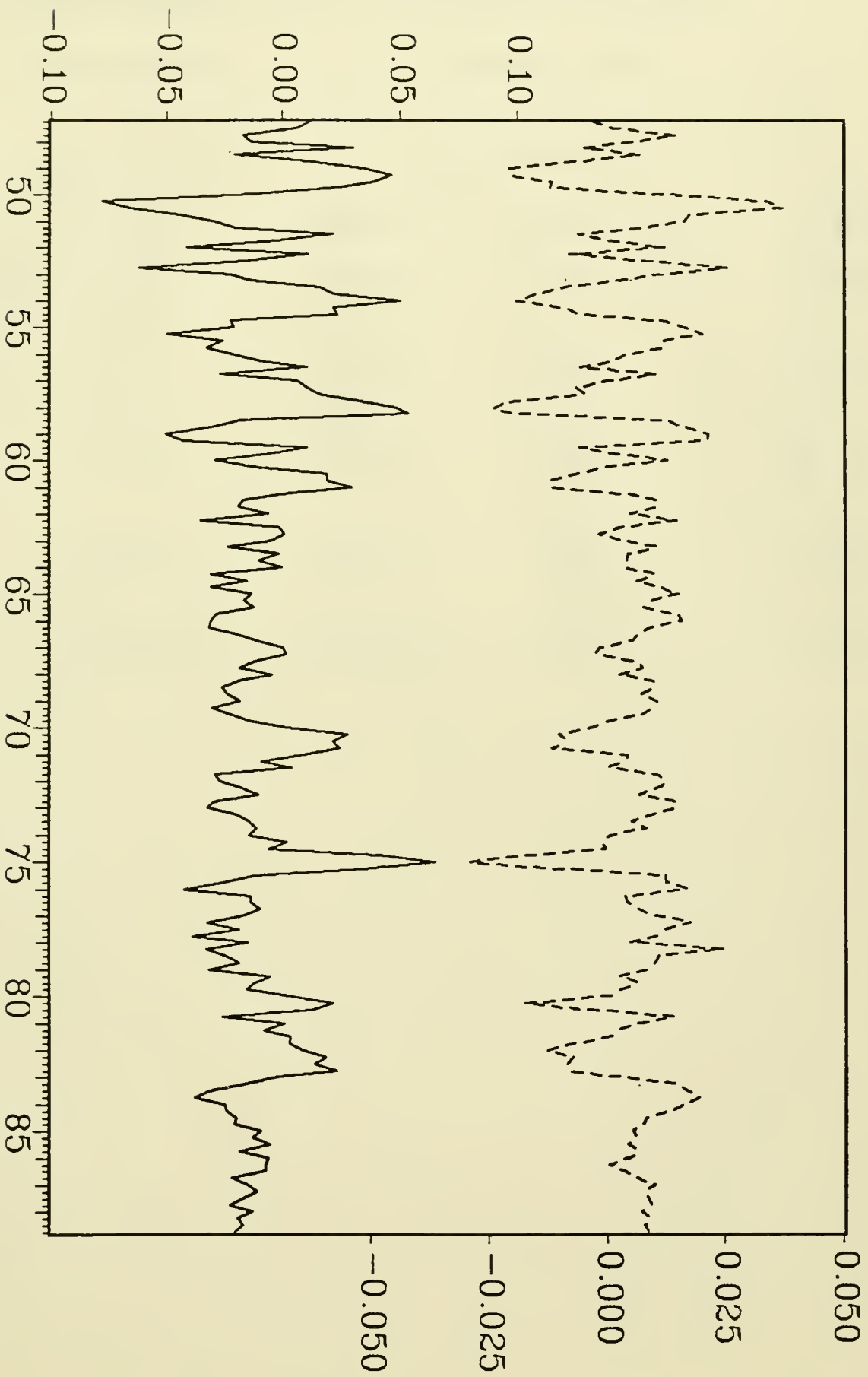
— DM16U1 - - - - DLH

Figure 2



DM16U05 - - - - - DLH

Figure 3



— DM2U1 - - - - DLH

TABLE 1

Estimation of equation (26) for different specifications

E
L
S
A
S
T
I
C
I
T
Y
O
F

		AVERAGE MARKUP		
		1.1	1.6	2
Coefficient on				
0.5	q _t	0.016 (0.009)	0.035 (0.014)	0.058 (0.022)
	y _t	-0.211 (0.05)	-1.083 (0.69)	-2.099 (0.132)
1	q _t	0.016 (0.009)	0.035 (0.014)	0.058 (0.022)
	y _t	0.169 (0.05)	-0.416 (0.50)	-1.099 (0.132)
2	q _t	0.016 (0.009)	0.035 (0.014)	0.058 (0.022)
	y _t	0.359 (0.05)	-0.083 (0.49)	-0.599 (0.132)

TABLE 2

The levels equation with q as the dependent variable

		AVERAGE MARKUP				
		1.1	1.6	2		
Coefficient on						
E L A S T I C I T Y O F	S U B S T I T U T I O N	0.5	μ_t	1.54 (0.82)	1.20 (0.48)	0.86 (0.32)
			y_t	1.12 (0.52)	2.09 (0.69)	2.59 (0.86)
	1	μ_t	1.54 (0.82)	1.20 (0.48)	0.86 (0.32)	
		y_t	0.53 (0.53)	1.21 (0.50)	1.73 (0.58)	
	2	μ_t	1.54 (0.82)	1.20 (0.48)	0.86 (0.32)	
		y_t	0.24 (0.60)	0.89 (0.49)	1.31 (0.52)	

TABLE 3The basic instrumental variables specifications
US data 1947:III-1988-IV

Parameter	Separate Coeffs.	Constrained Coeffs.
Constant	0.223 (0.06)	0.195 (0.05)
Coeff. on trend	0.66e-4 (0.4e-4)	-0.85e-5 (0.3e-4)
ϵ_Y	1.062 (0.20)	
ϵ_X	0.184 (0.06)	
ϵ		0.195 (0.05)
R^2	0.995	0.980
D.W.	2.36	2.00
J	1.09	1.46

Table 4

Instrumental Variables Method
Elasticity of the markup with respect to X/Y

		AVERAGE MARKUP			
		1.1	1.6	2	
E L A S T I C I T Y O F	S U B S T I T U T I O N	0.5	0.118 (0.04)	0.224 (0.07)	0.366 (0.12)
		1	0.099 (0.03)	0.204 (0.06)	0.335 (0.10)
		2	0.092 (0.03)	0.195 (0.05)	0.322 (0.09)

TABLE 5

Instrumental Variables Method
 Separate Elasticities with respect to Y and X

E
L
A
S
T
I
C
I
T
Y
O
F

S
U
B
S
T
I
T
U
T
I
O
N

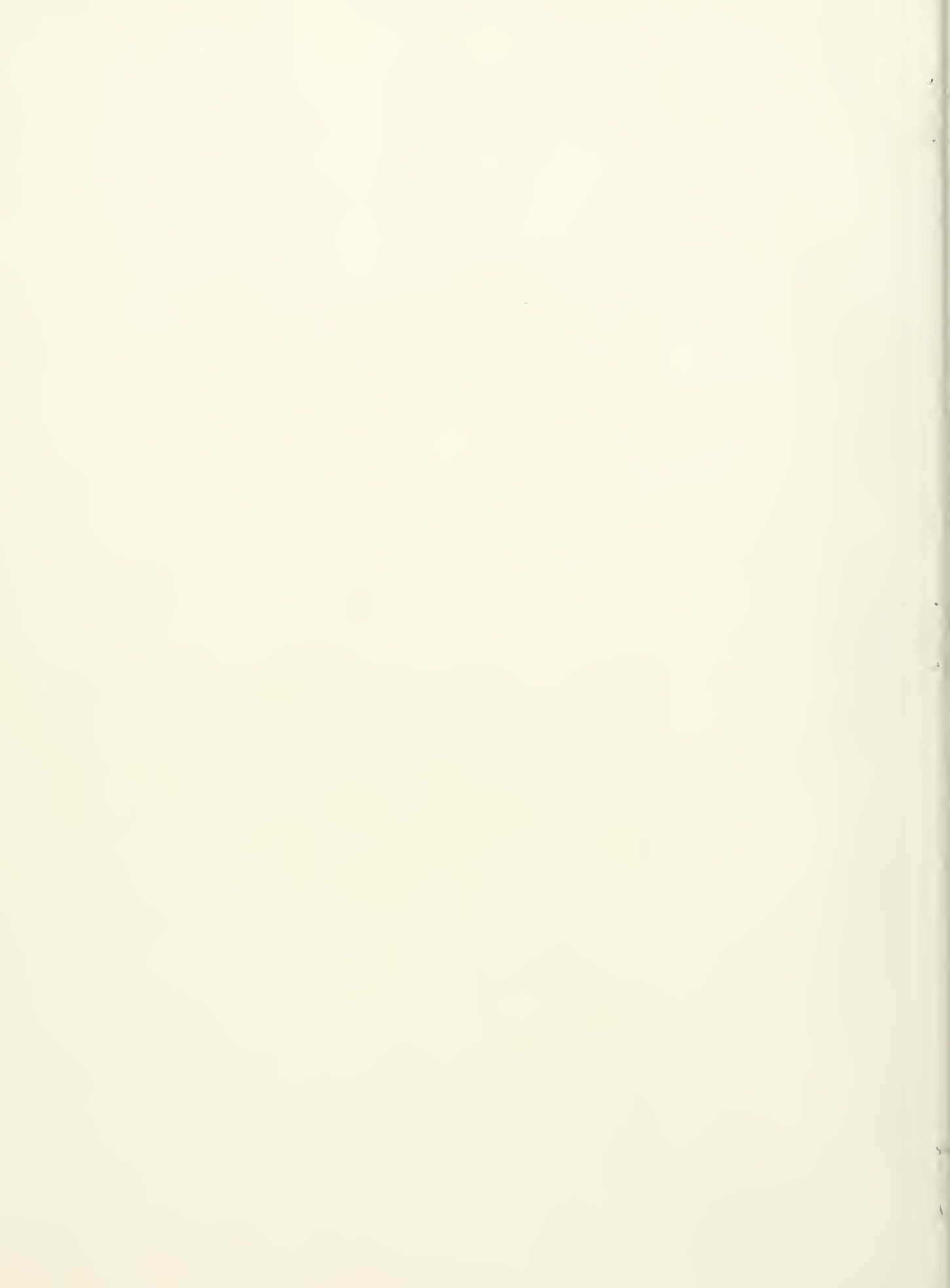
		AVERAGE MARKUP			
		1.1	1.6	2	
0.5	ϵ_Y	0.515 (0.11)	1.726 (0.20)	3.124 (0.32)	
	ϵ_X	0.099 (0.035)	0.185 (0.06)	0.396 (0.11)	
	ϵ_Y	0.139 (0.11)	1.062 (0.20)	2.127 (0.32)	
	1	ϵ_X	0.097 (0.03)	0.184 (0.06)	0.293 (0.09)
		ϵ_Y	-0.049 (0.11)	0.729 (0.20)	1.628 (0.32)
	2	ϵ_X	0.096 (0.03)	0.183 (0.06)	0.292 (0.09)

TABLE 6

Instrumental Variables Method
Variations with average markup equal to 1.6 and
elasticity of substitution equal to 1.

	ϵ_Y	ϵ_X
Use of lagged dividend/price ratio instead of lagged return as an instrument	1.184 (0.22)	0.165 (0.05)
Use of return on Treasury Bills instead of stock return	0.933 (0.17)	0.365 (0.25)
Use of return on commercial paper instead of stock return	0.916 (0.19)	0.455 (0.24)
Use of hourly earnings in manufacturing instead of hourly private compensation	1.274 (0.24)	0.255 (0.07)
Use of hourly earnings and return on Treasury Bills	1.055 (0.19)	0.655 (0.29)
Use of hourly earnings and return on commercial paper	0.878 (0.19)	0.953 (0.28)

2966 064



Date Due

1991

MAY 26 1993

MIT LIBRARIES



3 9080 00706965 8

