INTERNATIONAL ECONOMIC REVIEW
Vol. 32, No. 2, May 1991

AGGREGATION AND TESTING OF THE PRODUCTION SMOOTHING HYPOTHESIS*

BY KON S. LAI[1]

This paper examines the aggregate implications of the production smoothing model. The analysis indicates that aggregation can be a source of bias distorting tests of production smoothing based on the relative variance of production and sales. It is shown that, depending upon the relative variability of different types of market shocks firms face, the aggregation bias can be so severe as to render the test of the production smoothing hypothesis invalid.

1. INTRODUCTION

An important hypothesis regarding the dynamic behavior of production and inventory is based on the production smoothing model. The model suggests that a firm facing convex cost functions will adjust only partially its production to variations in sales, with changes in inventories absorbing some of the variations in sales. Accordingly, one would expect to observe that production is less variable than sales. Blinder (1986) observes, however, that the variance of production exceeds that of sales in most two-digit manufacturing industries. The evidence presented by Blinder calls into doubt the empirical validity of the production smoothing model.

In this paper the aggregate implications of the production smoothing model for the relative volatility of production and sales are examined. Specifically, the potential bias arising from the use of aggregate data in testing for production smoothing behavior is investigated.

Various approaches have been suggested to modify the production smoothing model to make it more consistent with the empirical findings. For example, if the "desired" inventory level is an increasing function of expected next-period sales, then the variance of production can exceed the variance of sales. Apart from the apparently ad hoc nature of this approach and possible conceptual difficulties that arise in defining "desired" inventory, as noted in Blinder (1982), the modified model is not supported by data based on variance bound tests (see West 1986).

An alternative approach is to suggest that the production smoothing model can be consistent with the data by allowing for persistent cost shocks, as shown in Glick and Wihlborg (1985) and Blinder (1986). The relevance of the proposition of cost shocks has been examined in Eichenbaum (1988) and Miron and Zeldes (1988b), and the empirical evidence in general appears mixed.

In a recent paper Kahn (1987) examines the role of the stockout-avoidance

* Manuscript received October 1988; revised December 1989.

[1] I am grateful to two anonymous referees for valuable comments and suggestions. An earlier version of this paper was presented at the 64th Annual Western Economic Association International Conference.

motive in understanding the "excess" volatility of production. Kahn shows that even in the absence of productivity or cost shocks, the firm's optimal behavior can be consistent with the "excess" volatility of production, if demand is serially correlated and the firm can backlog excess demand. Unfortunately, this approach lacks direct evidence of the importance of stockouts; one may question the quantitative significance of the stockout-avoidance motive in explaining the "excess" volatility of production.

Still there is another approach which explores the role of nonconvex production costs in explaining the "excess" volatility of production. Such approach, mentioned briefly in Blinder (1986), suggests that if the firm frequently operates in a region of decreasing marginal costs, then small shifts in demand can cause, at times, production to jump substantially; consequently, production can vary more than sales. Blinder's suggestion is formally examined in Ramey (1987). Ramey shows that declining marginal costs will always cause the variance of production to be higher than the variance of sales. The empirical evidence presented by Ramey appears supportive of the claim that firms behave as if they face not increasing, but decreasing marginal costs. If Ramey's analysis is correct, it suggests a role for the nonconvex cost model in understanding the volatility of production.

Ramey's result should be interpreted with caution, nonetheless. While nonconvex costs imply imperfect competition, Ramey's model ignores price as a decision variable and considers sales exogenous to the firm. Clearly changing price represents a potentially important channel of response to demand shocks. As long as the firm can exercise control over its sales by varying its selling price, the changes in price can have intertemporal substitution effects, which tend to smooth the production path. Furthermore, it is not at all obvious that nonconvex costs typify the manufacturing technology. To explain the observed volatility of production without altering the basic technological assumptions that underlie almost all neoclassical theory certainly remains an important question.

In contrast to the various attempts at modifying the production smoothing model to make it more consistent with Blinder's (1986) empirical findings, an interesting study by Ghali (1987) examines the validity of tests of the production smoothing hypothesis when aggregate seasonally adjusted data are used, as in Blinder. Ghali conjectures that the relative size of the variances of the seasonally adjusted production and sales may not provide valid tests of the production smoothing model. In addition, aggregating over firms where the seasonal patterns differ may also distort tests of production smoothing. Based on data of the portland cement industry, the empirical evidence presented by Ghali appears supportive of Ghali's conjecture. While aggregate seasonally adjusted data generally dismiss production smoothing as an empirically relevant hypothesis, disaggregate seasonally unadjusted data strongly support the production smoothing hypothesis.[2]

The import of Ghali's findings is that there may indeed not exist any tension

[2] Recent studies by Fair (1989) and Miron and Zeldes (1988a) have also suggested that negative results concerning the production smoothing hypothesis can be due to the use of poor data. Fair observes that data problems may be particularly important in testing the production smoothing hypothesis, where we are looking for differences in the paths of two series that are possibly small relative to the average levels of the paths.

between theory and fact to be resolved. The conventional production smoothing model can be actually consistent with the data. Of course, further empirical verification of Ghali's proposition is desirable when unadjusted disaggregated data for many other industries are available. Nonetheless, the empirical findings reported in Ghali appear significant enough to require theoretical explanations.

While Ghali's empirical study indicates potentially substantial biases in testing the production smoothing hypothesis when aggregate seasonally adjusted data are used, the theoretical basis for the biases appears not well discussed in the literature. Specifically, little effort has been devoted to exploring the aggregate implications of inventory policies of manufacturing firms. In a related context of retail inventories Caplin (1985), motivated by Blinder (1981), formally examines the aggregate implications of (S, s) inventory policies of retail firms for the volatility of orders relative to sales. According to the (S, s) policy, inventories are allowed to dwindle to some minimum level, s, whereupon they are restored to some maximum level, S. This inventory policy, which can be shown to be optimal in the presence of nonconvex costs in placing orders, implies that orders are more volatile than sales at the microeconomic level. When inventory policies are aggregated over retailers, Caplin's analysis allows for interdependence in sales processes between retail firms. For example, a positive correlation between retailers' sales can result from common factors in sales such as changes in aggregate demand, while a negative correlation between their sales can result from fights among retailers for shares of a fixed market, other things the same.

A major result of Caplin's analysis is that even allowing for some correlations in sales between firms, the aggregation in the (S, s) economy is "neutral" with respect to variances, neither increasing nor diminishing the impact of individual retailers on their own sales processes. It follows that the magnification of variance in the aggregate is simply the sum of individual components at the microeconomic level. This result is important in that the implications of (S, s) inventory policies for microeconomic variables can then fully apply to the corresponding aggregate variables.

Caplin's neutrality result on aggregation appears rather strong. The result depends on two major assumptions, namely retailers operate identical inventory policies and the levels of inventories follow jointly a Markov process. When these assumptions are relaxed, it is far from clear that the neutrality result can still hold.

In this paper we examine the source, direction, and potential degree of bias when aggregate data are used for testing the production smoothing hypothesis based on the volatility of production relative to sales. In contrast to Caplin's analysis of (S, s) inventory policies, we consider production smoothing policies of manufacturing firms facing convex costs. We focus on examining the possible bias arising from aggregation of inventory policies over firms, and the effects of seasonal adjustment, as stressed in Ghali, will be ignored in this analysis. It is shown that the direction and the size of the aggregation bias depend upon the relative variability of different types of market shocks. Specifically, when negatively correlated shocks between firms are relatively significant, in a sense to be defined, the bias is guaranteed to be upward. The resulted upward bias can be substantial, especially when the cost structures differ significantly between firms. Importantly, the aggregation bias can be so severe as to render tests of the production smoothing hypothesis misleading and invalid.

The paper is organized as follows. Section 2 outlines the basic model of production smoothing. Section 3 examines the implications of the model for the volatility of production at the aggregate level. Section 4 discusses the effects of relative cost shocks on aggregate behavior. Section 5 concludes.

2. THE MODEL

We examine a simple linear-quadratic model of production smoothing based on convex costs. For simplicity, we consider an industry of only two firms, indexed by $k = 1, 2$. While this specification provides a simple and tractable framework to study the effects of aggregation, it is not so restrictive as it may seem since aggregation can be done by aggregating one more firm at a time.

The production cost function of firm k is assumed to be approximated by

$$C_k(Y_{kt}) = c_k Y_{kt}^2/2, \qquad c_k > 0, \tag{1}$$

where Y_{kt} denotes the production level for firm k in period t. The condition $c_k > 0$ indicates that the firm operates in a region of rising marginal costs. In this model only finished goods inventories are considered and backlogging is allowed. The inventory holding cost function of firm k is assumed to be represented by

$$H_k(N_{kt}) = h_k N_{kt}^2/2, \qquad h_k > 0, \tag{2}$$

where N_{kt} denotes the inventory level for firm k at the end of period t. Since linear terms will show up as constants in the model solution, they can be omitted in the cost functions without affecting the main analysis.

It is noteworthy that while linear cost terms will not affect our later analysis in terms of variances, they can alter the optimal levels of individual variables. A difficulty with our model, similar to Blinder's (1982) model, is that it does not prevent output and inventories from being negative, as noted in Schutte (1983). In this regard, considering linear cost terms in the model can be useful in avoiding the problem. This point appears to be not well understood in the literature and is demonstrated in the Appendix, without digressing from the main discussion.

The familiar accounting identity relating output, sales, and inventories is given by

$$Y_{kt} = X_{kt} + N_{kt} - N_{kt-1}, \tag{3}$$

where X_{kt} is the amount of sales for firm k in period t.

Firm k makes plans at the end of period t given sales and initial inventory holdings, and its objective is to minimize the sum of discounted expected cost

$$E_t \sum_{j=t}^{\infty} (1+r)^{t-j}\{C_k(Y_{kj}) + H_k(N_{kj-1})\} \tag{4}$$

subject to equation (3), where E_t is the expectation operator conditional on the information set available at the end of period t, and r is the interest rate. The stochastic Euler equations for the problem can straightforwardly be derived and are given by

$$(5) \qquad E_j\{N_{kj+2} - (2 + r + h_k/c_k)N_{kj+1} - (1 + r)N_{kj} + X_{kj+1} - (1 + r)X_{kj}\} = 0,$$

$j = t, t + 1, \dots$. The transversality condition is

$$(6) \qquad \lim_{T\to\infty} E_t(1 + r)^{-T}\{c_k(N_{kT} - N_{kT-1} + X_{kT}) + h_k N_{kT}/(1 + r)\} = 0.$$

The sufficient condition for minimization is satisfied under the model assumptions of convex inventory holding cost function, $H_k(N_{kt})$, and convex production cost function, $C_k(Y_{kt})$. The firm's problem is to find a stochastic process $\{N_{kj}, j = t, t + 1, \dots\}$ that satisfies (5) and (6).

Using the lag operator, denoted by L, (5) can be written as

$$(7) \qquad (1 - \phi_k L)(1 - \tilde{\phi}_k L)E_{t-1}N_{kt+1} = E_{t-1}\{(1 + r)X_{kt} - X_{kt-1}\},$$

where ϕ_k and $\tilde{\phi}_k$ are the roots of the characteristic equation,

$$(8) \qquad f(z) = z^2 - \{2 + r + h_k/c_k\}z + (1 + r) = 0.$$

We observe that this equation has two real positive roots: one is smaller than unity and one greater than unity. Let ϕ_k be the smaller root and $\tilde{\phi}_k$ be the larger root. By solving the unstable root forward, the solution of the stochastic difference equation (7) can be found to be

$$(9) \qquad N_{kt} = \phi_k N_{kt-1} - E_t \sum_{j=1}^{\infty} (1 + r)^{-j}\phi_k^j\{(1 + r)X_{kt+j-1} - X_{kt+j}\}.$$

This describes the optimal inventory policy for firm k as a function of its last-period inventory level and expected future sales.

To close the model, we need to specify the sales process. As in Caplin, we consider that sales processes of individual firms can be correlated, and the nature of the correlation depends upon the type of shocks buffeting the market. Caplin observes that the interdependence can arise from fluctuations in aggregate demand or fights for shares given a fixed market, for example. Other things being the same, the former gives rise to a strong positive correlation in sales, and the latter results in a strong negative correlation in sales. To capture such interdependence in sales between firms, we assume that the sales of firm k ($k = 1, 2$) follow a simple stationary process with three mutually independent and serially uncorrelated components, given by

$$(10) \qquad X_{kt} = \bar{X}_k + w_{kt} + u_t + e_{kt},$$

where $\bar{X}_k$ is some stationary level of sales for firm k; w_{kt} is firm-specific shocks that are uncorrelated between firms; u_t is common shocks, originated from changes in aggregate demand; and e_{kt} is the shocks in relative market share such that $\Sigma e_{kt} =$

0. The assumption of serially uncorrelated processes is a convenient specification that will serve our analytical purpose. Consideration of serially correlated processes such as an autoregressive process will raise the algebraic burden of solving the aggregation problem substantially, but adds little insights. For convenience we also assume that var (w_{kt}) = var (w) for different firms.

In view of (9) and (10), the optimal inventory adjustment rule for firm k can be given in a simple form as follows (see also the Appendix):

$$N_{kt} = \phi_k N_{kt-1} - \phi_k X_{kt} + \phi_k(1 - \phi_k)\bar{X}_k/(1 + r - \phi_k). \tag{11}$$

This implies an optimal inventory investment policy as a distributed lag function of past changes in sales

$$\nabla N_{kt} = -\phi_k \sum_{j=0}^{\infty} \phi_k^j \nabla X_{kt-j}, \tag{12}$$

where ∇ denotes the difference operator (e.g., $\nabla N_{kt} = N_{kt} - N_{kt-1}$). It follows that the (unconditional) variance of inventory investment is given by

$$\text{var}\,(\nabla N_k) = 2\phi_k^2 \text{ var}\,(X_k)/(1 + \phi_k). \tag{13}$$

Furthermore, the covariance between inventory investment and sales is given by

$$\text{cov}\,(X_k, \nabla N_k) = -\phi_k \text{ var}\,(X_k). \tag{14}$$

It indicates that sales and inventory investment are negatively correlated, a standard result in the literature on the theory of production smoothing.

The accounting identity (3) leads to the following decomposition of the (unconditional) variance of production for an individual firm:

$$\text{var}\,(Y_k) = \text{var}\,(X_k) + \text{var}\,(\nabla N_k) + 2 \text{ cov}\,(X_k, \nabla N_k). \tag{15}$$

It follows from (13) and (14) that

$$\text{var}\,(Y_k) = (1 - \phi_k) \text{ var}\,(X_k)/(1 + \phi_k) < \text{var}\,(X_k). \tag{16}$$

This will hold as an equality only in the limiting case where $\phi_k = 0$, and that happens when the production cost function is linear. Hence, if the production smoothing model is correct, we should expect to observe that production does not vary more than sales.[3] While the prediction presumably applies to the microeconomic level, aggregate data are generally used for testing the production smoothing hypothesis because of a problem in the availability of disaggregate data. Unfortunately, it appears not at all straightforward that the prediction can apply correspondingly to the variables in the aggregate.

[3] The consideration of unconditional variances conforms to the literature. Empirical analyses necessarily rely on estimates of variances based on a finite sample period. Under the usual assumption of stationarity and ergodicity, the empirical variances may provide *point* estimates of the corresponding theoretical variances. One certainly cannot claim that the assumption of stationarity and ergodicity is generally realistic.

3. THE VOLATILITY OF AGGREGATE PRODUCTION

In this section we address the problem of the aggregation of decisions across firms. An important question is whether or not the implications of the production smoothing model can still stand at the aggregate level. Specifically, can aggregation result in the variance of aggregate production exceeding that of aggregate sales?

A simple decomposition of variances yields

$$\text{var}(Y) - \text{var}(X) = \text{var}(Y_1) + \text{var}(Y_2) - \text{var}(X_1) - \text{var}(X_2) + 2\{\text{cov}(Y_1, Y_2) - \text{cov}(X_1, X_2)\}, \tag{17}$$

where $Y = Y_1 + Y_2$, the aggregate production; $X = X_1 + X_2$, the aggregate sales. The decomposition of variances indicates that the difference between cov (Y_1, Y_2) and cov (X_1, X_2) is crucial in determining the direction and the size of the aggregation bias. When cov (Y_1, Y_2) is greater than cov (X_1, X_2), aggregation tends to raise the variability of production relative to sales. When cov (Y_1, Y_2) is less than cov (X_1, X_2), however, aggregation tends to lower the variability of production relative to sales.

The difference between cov (Y_1, Y_2) and cov (X_1, X_2) can be expressed in terms of the structural variances in our model. Such a relationship can provide insights into understanding the source and the direction of the bias due to aggregation. To establish the relevant relationship, we need to find out the dynamics of production over time. Combining (3) and (12) yields

$$Y_{kt} = X_{kt} - \phi_k \sum_{j=0}^{\infty} \phi_k^j \nabla X_{kt-j}, \qquad k = 1, 2. \tag{18}$$

Using (18), we can establish the following result:

$$2\{\text{cov}(Y_1, Y_2) - \text{cov}(X_1, X_2)\} = \theta\{\text{var}(e) - \text{var}(u)\}, \tag{19}$$

where

$$\theta = 2(\phi_1 + \phi_2 - 2\phi_1\phi_2)/(1 - \phi_1\phi_2) > 0 \tag{20}$$

for $0 < \phi_1, \phi_2 < 1$. Note that (19) does not depend upon var (w). This suggests that aggregation is neutral with respect to variances as far as firm-specific shocks are concerned. It follows that when firms are subject to only uncorrelated shocks between them, aggregation of decisions over firms induces no bias with respect to variances.

In general, the nature of interdependence in sales between firms is important in determining the direction of the aggregation bias with respect to the variance of production relative to sales in the aggregate. Negatively correlated shocks tend to add to the variance of aggregate production relative to sales, while positively correlated shocks tend to reduce it. When var (e) is greater than var (u), aggregation tends to raise the variance of aggregate production as compared to sales. On the other hand, when var (e) is less than var (u), aggregation tends to lower the relative variance of aggregate production to sales.

Moreover, the size of the aggregation bias depends upon the values of the parameters $\phi_k's$, which in turn depend upon the parameters of the cost functions of individual firms. It is straightforward to establish that

$$(21) \qquad d\theta/d\phi_1 > 0, \qquad d\theta/d\phi_2 > 0$$

for $0 < \phi_1, \phi_2 < 1$. In the limiting case where ϕ_1 and ϕ_2 both tend to zero, the value of θ also approaches zero. In this limiting case, aggregation is neutral with respect to variances, neither raising nor diminishing the variances of individual firms. Combining this with our earlier result, we observe that unless the production cost functions of individual firms are all linear, the aggregation of variances over firms will display an upward bias when relative shocks (e) are more variable than common shocks (u).

Analyses similar to that of the variance can apply also to the covariance between sales and inventory investment, which is an important component determining the volatility of production via relationship (15). A simple decomposition of the covariance between aggregate sales and aggregate inventory investment yields

$$(22) \qquad \text{cov}(X, \nabla N) = \text{cov}(X_1, \nabla N_1) + \text{cov}(X_2, \nabla N_2) + \text{cov}(X_1, \nabla N_2) + \text{cov}(X_2, \nabla N_1),$$

where $\nabla N = \nabla N_1 + \nabla N_2$, the aggregate inventory investment. This indicates that the sum of the cross covariance terms, cov $(X_1, \nabla N_2)$ and cov $(X_2, \nabla N_1)$, is crucial in determining the direction and the size of the aggregation bias with respect to the covariance. In view of (10) and (12), the sum of cov $(X_1, \nabla N_2)$ and cov $(X_2, \nabla N_1)$ can be expressed in terms of the structural variances:

$$(23) \qquad \text{cov}(X_2, \nabla N_1) + \text{cov}(X_1, \nabla N_2) = (\phi_1 + \phi_2)\{\text{var}(e) - \text{var}(u)\}.$$

From (22) and (23) we observe that while positively correlated shocks tend to diminish the covariance of sales and inventory investment in the aggregate, negatively correlated shocks tend to add to the covariance in the aggregate. More specifically, when var (e) is greater than var (u), aggregation tends to raise the covariance in the aggregate, making it less negative. On the other hand, when var (e) is less than var (u), aggregation tends to reduce the covariance in the aggregate, making it more negative. As in the case of the variance, aggregation is neutral with respect to the covariance when only firm-specific shocks are considered.

The analysis thus far indicates that aggregation can induce bias with respect to variances (and covariances) in testing the production smoothing hypothesis. It has not provided, nevertheless, a complete answer to the question of whether or not aggregation can lead to the variance of production rising above that of sales. To answer the question, we need to explicitly derive the value of var (Y) − var (X) in terms of the structural variances. Combining (16), (17), and (19) gives

$$(24) \qquad \text{var}(Y) - \text{var}(X) = (\theta - \beta)\,\text{var}(e) - (\theta + \beta)\,\text{var}(u) - \beta\,\text{var}(w),$$

where θ is defined in (20) and β is given by

$$\beta = 2\{\phi_1/(1 + \phi_1) + \phi_2/(1 + \phi_2)\}. \tag{25}$$

A necessary condition for the aggregation bias to be significant enough to lead to the "excess" volatility of production is that $\theta - \beta > 0$. From (20) and (25) we obtain

$$\theta - \beta = 2(\phi_1 - \phi_2)^2/\{(1 - \phi_1\phi_2)(1 + \phi_1)(1 + \phi_2)\}. \tag{26}$$

Hence, an equivalent condition for $\theta - \beta > 0$ is that $\phi_1 \neq \phi_2$. Note that ϕ_1 and ϕ_2 are decision parameters that are dependent ultimately upon parameters of the inventory cost and production cost functions of the corresponding firms. Specifically, ϕ_k is a function of h_k/c_k for firm $k = 1, 2$ (see equation (8)). A necessary condition for the "excess" volatility of production in the aggregate is therefore that the relative cost structures of inventory-holding and production, as summarized by the values of h_k/c_k, are different between firms.

Equation (24) also indicates that a necessary and sufficient condition for the "excess" volatility of aggregate production in our model is given by

$$\text{var}(e) > (\theta + \beta)\, \text{var}(u)/(\theta - \beta) + \beta\, \text{var}(w)/(\theta - \beta). \tag{27}$$

It follows that for a sufficiently large value of var (e) relative to var (u) and var (w), var (Y) can be greater than var (X) in the aggregate even when var (Y_k) is less than var (X_k) for individual firms, provided that the relative cost structures are not the same across firms (i.e., $\phi_1 \neq \phi_2$).

Condition (27) indicates that relative (negatively correlated) shocks play a key role in generating a relatively high volatility of aggregate production as compared to sales. The point is that a change in the relative demand between firms, other things the same, will cause some firms to increase output and some others to decrease output. Since firms with different cost structures will react in a different degree to the same demand shock, the increase in output is not likely to be matched exactly by the decrease in output. As a result, the relative sales shock can lead to a change in aggregate production even with aggregate sales being unchanged.

A simple example may be useful for illustration. Consider the following hypothetical episode: $X_t = 10{,}000$, $X_{1t} = 3{,}000$ and $X_{2t} = 7{,}000$; $X_{t+1} = 10{,}000$, $X_{1t+1} = 2{,}999$ and $X_{2t+1} = 7{,}001$. Note that the aggregate sales remain stable over time, while there is a small change in relative sales between the two firms. In period $t + 1$ the sales of Firm 1 decrease by one unit, but the sales of Firm 2 increase by the same amount. This change in relative sales will induce Firm 1 to lower output and Firm 2 to raise output. In the special case where the firms have exactly the same cost structure, a decrease in Firm 1's output will be matched by an equal increase in Firm 2's output, resulting in no variation in aggregation production. However, in the general case where the firms have different cost structures, the decrease in Firm 1's output will not be equal in magnitude to the increase in Firm 2's output. We thus have a change in aggregate production but no change in aggregate sales. It follows that aggregate production fluctuates more than aggregate sales over the episode. Indeed, in this example the variance ratio of var (Y)/ var (X) in the aggregate is infinitely large.[4]

The aggregation problem can be relatively more important for some series in some periods of time. For example, the aggregation problem would be more important in periods in which aggregate shocks are less significant, other things the same.[5] It should be noted, however, that aggregate shocks can often induce relative shocks, since different products have different income elasticities of demand and so they respond differently to changes in aggregate demand. Purely common shocks can hardly be found, therefore. The relevance of the aggregation problem can also depend on the underlying market structure. In general, we may expect that the aggregation problem is likely to be more significant for more competitive industries. Competing advertising campaigns, strategic price changes, and rises and declines of firms, for example, can lead to significant changes in their relative sales but with little change in the aggregate sales. One may also conjecture that the aggregation problem is more relevant for more highly aggregated data (e.g., all manufacturing industries as compared to an individual manufacturing industry), possibly because regional and sectoral shifts in demand would be important in generating relative shocks and because more heterogeneous goods are involved which, as noted earlier, can react rather differently to aggregate shocks and so are susceptible to changes in relative sales.

The results in this paper can be contrasted with that derived by Caplin in a related context of retail inventories, as noted earlier in Section 1. Based on the assumptions of identical inventory policies on the part of retail firms and the Markovian inventory dynamics process, Caplin is able to demonstrate that the aggregation in the (S, s) economy is neutral with respect to variances.

These two assumptions appear rather restrictive, however. Specifically, the consideration of identical retailers has in effect assumed away any distributional effects of random shocks in the aggregate. In terms of our model, this analytically corresponds to the assumption that firms adopt the same ϕ_k. In addition, following from the Markovian process assumption, Caplin allows for correlations in sales between retail firms in a stringent way such that the stationary probability density over inventories does not depend on the correlations in sales. As a result, the levels of inventories in individual firms are uncorrelated, at least in the stationary state.

Clearly, if we assume that firms adopt identical inventory policies and they are subject to only uncorrelated shocks in sales in our model, the aggregation in the production smoothing economy based on convex costs will also be neutral with respect to variances, as in the (S, s) economy Caplin considered. Since these assumptions are likely to be a special case rather than the general case, the aggregation of decisions across firms will likely introduce biases in hypothesis

[4] Blinder (1986) reports that the ratio of var (Y)/ var (X) ranges from a high of 2.40 to a low of 0.95 for twenty two-digit manufacturing industries, and is 1.14 for manufacturing as a whole. These empirical estimates are evidently not too large to be possibly explained by the aggregation problem.

[5] A referee has suggested that a possible way to measure the relevance of the aggregation problem is to compare periods of high unanticipated money variance with periods of low unanticipated money variance, and the aggregation problem would be relatively important in the latter period.

testing, though the biases can be upward or downward, depending on the actual variability of various market shocks individual firms face.

4. CONSIDERATION OF COST SHOCKS

The foregoing analysis has considered mainly demand shocks. It is shown that relative demand shocks tend to raise the variance of aggregate output relative to aggregate sales. Since the relative importance of demand and cost shocks remains an empirical question, and aggregate demand and cost shocks can be shown to have opposite effects on the relative variance of output and sales, an analysis of the effects of relative cost shocks is of interest.

The effects of relative cost shocks can be illustrated using a simple example similar to that considered earlier without going into additional modelling. Let s_{kt} be the cost shock to Firm k's marginal cost of production in period t.[6] Consider the following hypothetical episode: $X_t = 10{,}000$, $X_{1t} = 3{,}000$, $X_{2t} = 7{,}000$, $s_{1t} = 0.000$, and $s_{2t} = 0.000$; $X_{t+1} = 10{,}000$, $X_{1t+1} = 3{,}000$, $X_{2t+1} = 7{,}000$, $s_{1t+1} = 0.001$, and $s_{2t+1} = -0.001$. In this case sales remain unchanged over time, but there is a relative cost shock in period $t + 1$. This relative cost shock will induce Firm 1 to reduce output and Firm 2 to raise output. In the general case where the firms have different cost structures, the decrease in Firm 1's output will not be matched exactly by an equal increase in Firm 2's output. We thus have a change in aggregate output without any change in aggregate sales. It follows that aggregate output will fluctuate more than aggregate sales over the episode.

5. CONCLUDING REMARKS

In this paper the aggregate implications of the production smoothing model for the production and inventory behavior have been examined. The analysis indicates that aggregation can be a source of bias distorting tests of the production smoothing hypothesis. It is shown that the direction and the size of the aggregation bias depend on the relative variance of various components of market shocks, and that the bias can be so significant as to render tests of production smoothing invalid. Of course, the actual direction and degree of significance of the aggregation bias remain to be examined in empirical work. Theoretical discussions concerning the aggregate bias may be of little practical interest unless empirical evidence suggests that there is such a bias and the bias is qualitatively significant. In this regard, the empirical findings presented by Ghali are particularly interesting. According to Ghali's study, aggregation together with seasonal adjustment can induce serious distortions, leading to wrong inferences regarding the firm's behavior. Further empirical work using disaggregate data at the firm level is surely of interest. The use of disaggregate data can also minimize the data problem of measurement error, noted in Fair (1989) and Miron and Zeldes (1988a).

Finally, the notion of aggregation bias should be of general importance in

[6] The cost shock s_{kt} can be introduced formally into the production cost function as: $C_k(Y_{kt}) = s_{kt}Y_{kt} + c_k Y_{kt}^2/2$.

economic analyses and is by no means specific to the production smoothing analysis. The present analysis can be viewed as providing an example illustrating that great caution should be taken in generalizing the implications of microeconomic models for microeconomic variables to the corresponding macroeconomic variables.

California State University, Los Angeles, U.S.A.

APPENDIX

In this appendix the problem concerning negative inventories and output, noted in Schutte (1983), is examined. It will be shown that when linear terms in the cost functions are considered, the problem is not unavoidable.

Consider that cost functions (1) and (2) are given instead by

$$C_k(Y_{kt}) = c_{k0} + c_{k1}Y_{kt} + c_k Y_{kt}^2/2, \qquad c_{k1} < 0,\ c_k > 0, \tag{A1}$$

and

$$H_k(N_{kt}) = h_{k0} + h_{k1}N_{kt} + h_k N_{kt}^2/2, \qquad h_{k1} < 0,\ h_k > 0. \tag{A2}$$

The condition $c_{k1} < 0$ implies that the U-shaped cost curve has a minimum at positive output, which is inconsistent with the proposition of eventually diminishing marginal returns. The condition $h_{k1} < 0$ reflects that backlogging is more costly than storage—see Zabel (1986) on this point. Combining (A1), (A2), (3), and (4), the stochastic Euler equations for the problem are given by

$$\begin{aligned} E_j\{N_{kj+2} - (2 + r + h_k/c_k)N_{kj+1} \\ - (1+r)N_{kj} + X_{kj+1} - (1+r)X_{kj}\} = (h_{k1} + rc_{k1})/2c_k, \end{aligned} \tag{A3}$$

where $j \geq t$. It follows from (A3) that the stationary level of inventories (or the unconditional mean of N_k), denoted by $\bar{N}_k$, is given by

$$\bar{N}_k = (-h_{k1} - rc_{k1})/2h_k - rc_k\bar{X}_k/h_k. \tag{A4}$$

For $h_{k1} < 0$ and $c_{k1} < 0$, the stationary level of inventories can be positive, depending upon the actual values of different model parameters.

Solving (A3) yields, in contrast to (11), the following optimal rule:

$$N_{kt} = \phi_k N_{kt-1} - \phi_k X_{kt} - b(h_{k1} + rc_{k1})/2c_k + b(1 - \phi_k)\bar{X}_k, \tag{A5}$$

where $b = \phi_k/(1 + r - \phi_k) > 0$. Combining (3) and (A5) we have

$$Y_{kt} = (1 - \phi_k)(X_{kt} - N_{kt-1}) - b(h_{k1} + rc_{k1})/2c_k + b(1 - \phi_k)\bar{X}_k. \tag{A6}$$

Since $X_{kt} \geq 0$, Y_{kt} is positive if initial inventory N_{kt-1} is less than sales X_{kt}, or if h_{k1} or c_{k1} is sufficiently negative. Specifically, for a given path of sales X_{kt} over time, optimal output Y_{kt} will always be positive provided that h_{k1} or c_{k1} is sufficiently negative, the same condition required for the nonnegativity of the stationary level of inventories.

Note that the stationary level of inventories implied by (A5) is given by

$$\bar{N}_k = b\{-(h_{k1} + rc_{k1})/2c_k - r\bar{X}_k\}/(1 - \phi_k). \tag{A7}$$

Since ϕ_k is a root of equation (8), we have

$$(1 - \phi_k)(1 + r - \phi_k) = h_k\phi_k/c_k. \tag{A8}$$

Using (A8) we can then show that (A7) (and so (A5)) is consistent with (A4).

REFERENCES

BLINDER, A. S., "Retail Inventory Investment and Business Fluctuations," *Brookings Papers on Economic Activity* (1981), 443–505.

———, "Inventories and Sticky Prices: More on the Microfoundation of Macroeconomics," *American Economic Review* 72 (1982), 334–348.

———, "Can the Production Smoothing Model of Inventory Behavior be Saved?" *Quarterly Journal of Economics* 101 (1986), 431–453.

CAPLIN, A. S., "The Variability of Aggregate Demand with (S, s) Inventory Policies," *Econometrica* 53 (1985), 1395–1409.

EICHENBAUM, M., "Some Empirical Evidence on the Production Level and Production Cost Smoothing Models of Inventory Investment," Working Paper No. 2523, National Bureau of Economic Research, February 1988.

FAIR, R. C., "The Production-Smoothing Model Is Alive and Well," Working Paper No. 2877, National Bureau of Economic Research, March 1989.

GHALI, M., "Seasonality, Aggregation, and Testing of the Production Smoothing Hypothesis," *American Economic Review* 77 (1987), 464–469.

GLICK, R. AND C. WIHLBORG, "Price and Output Adjustment, Inventory Flexibility, and Cost and Demand Disturbances," *Canadian Journal of Economics* 18 (1985), 566–573.

KAHN, J. A., "Inventories and the Volatility of Production," *American Economic Review* 77 (1987), 667–679.

MIRON, J. A. AND S. A. ZELDES, "Production, Sales, and the Change in Inventories: An Identity That Doesn't Add Up," Working Paper No. 2765, National Bureau of Economic Research, November 1988a.

——— AND ———, "Seasonality, Cost Shocks, and the Production Smoothing Model of Inventories," *Econometrica* 56 (1988b), 877–908.

RAMEY, V. A., "Non-Convex Costs and the Behavior of Inventories," Discussion Paper No. 87-34, University of California, San Diego, December 1987.

SCHUTTE, D. P., "Inventories and Sticky Prices: Note," *American Economic Review* 73 (1983), 815–816.

WEST, K. D., "A Variance Bounds Test of the Linear Quadratic Inventory Model," *Journal of Political Economy* 94 (1986), 374–401.

ZABEL, E., "Price Smoothing and Equilibrium in a Monopolistic Market," *International Economic Review* 27 (1986), 349–356.