NBER TECHNICAL PAPER SERIES

# SOME EVIDENCE ON FINITE SAMPLE BEHAVIOR OF AN INSTRUMENTAL VARIABLES ESTIMATOR OF THE LINEAR QUADRATIC INVENTORY MODEL

Kenneth D. West

David W. Wilcox

Technical Paper No. 139

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1993

We thank John Hulbert for research assistance and Ricardo Fiorito, Glenn Rudebusch, and participants in seminars at the Federal Reserve Board of Governors, the Johns Hopkins University, the ISIR session at the 1993 AEA meetings, the 1993 meeting of the NBER Working Group in Common Elements in Trends and Fluctuations, the Research Triangle Econometrics Workshop, and the University of Sienna for helpful comments. West thanks the National Science Foundation, the Sloan Foundation and the University of Wisconsin Graduate School for financial support. The views expressed here are those of the authors and not necessarily those of the Board of Governors of the Federal Reserve System or of other members of its staff. This paper is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Technical Paper #139
July 1993

# SOME EVIDENCE ON FINITE SAMPLE BEHAVIOR OF AN INSTRUMENTAL VARIABLES ESTIMATOR OF THE LINEAR QUADRATIC INVENTORY MODEL

## ABSTRACT

We evaluate some aspects of the finite sample distribution of an instrumental variables estimator of a first order condition of the Holt et al. (1960) linear quadratic inventory model. We find that for some but not all empirically relevant data generating processes and sample sizes, asymptotic theory predicts a wide dispersion of parameter estimates, with a substantial finite sample probability of estimates with incorrect signs. For such data generating processes, simulation evidence suggests that different choices of left hand side variables often produce parameter estimates of an opposite sign. More generally, while the asymptotic theory often provides a good approximation to the finite sample distribution, sometimes it does not.

Kenneth D. West
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706, U.S.A.
and NBER

David W. Wilcox
Division of Monetary Affairs
Board of Governors of the
Federal Reserve System
Washington, DC 20551, U.S.A.

## I. Introduction

The linear quadratic inventory model has been one of the mainstays of empirical work in inventories since its formulation by Holt et al. in 1960. It has recently been applied to inventory movements in much U.S. data, including those for two digit manufacturing (West (1986), Eichenbaum (1989), Ramey (1991)), for a number of industries with physical product data (Krane and Braun (1991)) and for the automobile industry both pre- and post-World War II (Blanchard (1983), Kashyap and Wilcox (1993)).

Instrumental variables estimates of a first order condition from the model are, however, rather sensitive to what seem to be minor changes in specification or sample period. One illustration of this is the dispersion of the parameter estimates produced by Eichenbaum (1989), Ramey (1991) and West (1986), all of whom applied the model to the same set of two digit manufacturing industries (but using somewhat different sample periods, instruments, methods for treatment of unobserved serial correlation, etc.). Among inventory experts, it is well known that a key parameter ($a_1$, in the notation of the model introduced in the next section) was found to be negatively signed by Ramey, positively signed by Eichenbaum and West; as emphasized by Ramey, the sign of this coefficient is economically important since it influences whether firms bunch or smooth production.

One possible explanation of the current lack of consensus is that some of the differences in specification are important. The model allows for an unobservable cost shock, and it may be important whether or not one allows this shock to be serially correlated (as do Eichenbaum and Ramey but not West). In addition, while all three authors use instrumental variables estimators, the instruments vary from author to author; Ramey argues that the instruments used by others are not valid, and that consistent estimates can be obtained only with "truly exogenous" instruments of the sort that she uses. That such differences in technique are not the entire story is suggested by a second sense in which instrumental variables estimates seem to be sensitive to specification: estimates sometimes change dramatically when one does nothing more than change the left hand side variable. This sensitivity is noted by Ramey (1991), Krane and Braun (1991) and Kashyap and Wilcox (1993), with the latter two finding that the sign of the key parameter mentioned in the previous paragraph ($a_1$) tends to be negative when Ramey's normalization is used, positive when another normalization is used. (Similar sensitivity to choice of left hand side variables has been noted in estimation of the consumption-CAPM [e.g., Hansen and Singleton (1990)].)

Such sensitivity might well reflect model misspecification, in all these papers. While we recognize the need to consider such a possibility, in the present paper we focus on examining finite sample performance assuming a correctly specified model. In line with the papers cited above, we work with a simple linear quadratic model in which costs are quadratic functions of production, of changing production, and of the deviation of inventories from a target proportion of sales. We assume that the first order condition, or Euler equation, of the model is estimated by instrumental variables using lags of inventories and sales as instruments. We generate data in accord with the model, under the simplifying assumption that sales are exogenous.

Given a data generating process, we use conventional asymptotic theory to solve analytically for an approximate finite sample variance-covariance matrix of the parameter estimates. We find that for plausible cost parameters and sales processes, the implied dispersion of parameter estimates sometimes is large, with substantial areas of the probability distribution falling on both sides of zero; this suggests substantial probability of obtaining a wrong-signed estimate of a parameter from any given realization of the data. In one extreme case, we conclude that if the asymptotic approximation accurately describes the finite sample distribution, roughly 30,000 observations on monthly data (i.e., about 2500 years) would be required before a certain parameter estimate would have a 95 percent probability of having the correct sign.

The large dispersion of parameter estimates raises the possibility that sampling error accounts for the above-noted sensitivity to specification, including in particular sensitivity to choice of left hand side variable. Conventional asymptotic theory does not, however, appear to be particularly helpful on this score. For this reason, and to establish more generally the applicability of the conventional asymptotic approximation, we conduct a set of Monte Carlo experiments.

For each of several data generating processes, we generate 1000 datasets, each with 300 observations, 300 being approximately the number of monthly observations on real inventories and sales available at the two digit SIC code level in the United States. Using three different left hand side variables, we estimate the Euler equation by instrumental variables and tabulate the distribution of the resulting point estimates.

We find that, in many respects, the asymptotic approximation works well. In general, confidence intervals constructed from the simulated data are narrow when the asymptotic confidence interval is narrow, large when the asymptotic one is large. And usually there is little bias, in that the median of most parameter estimates is within a fraction of an asymptotic standard error of the population value.

But for all normalizations, the estimators tend to be somewhat more disperse than is predicted by the asymptotic theory, and in a few cases they have substantial bias as well. (Similar results have been obtained in studies of finite sample properties of instrumental variables estimators of asset pricing models (Tauchen (1986), Kocherlakota (1990), Ferson and Foster (1991)).) We also find that different choices of left hand side variable have a nontrivial tendency to produce estimates of different sign. Moreover, consistent with the possibility raised above, this tendency is most apparent in those DGP's in which our asymptotic approximation suggests a relatively large probability of deriving a wrong-signed estimate. Interestingly, often but not always the normalization that in empirical work has tended to produce a negative estimate of the parameter we denote $a_1$ tends to do so in our simulations as well.

Neither our simulations nor our asymptotic theory are rich enough to enable us to conclude that one normalization is better than another in this model, still less to produce guidelines useful for practitioners using other models. Instead, we take the message of the asymptotic calculations and Monte Carlo simulations to be as follows. At least for some data, it will be difficult to obtain sharp estimates of the parameters of this model, and one should not be surprised if minor changes in specification, estimation technique, or even choice of left hand side variable cause parameter estimates to change sign or otherwise shift dramatically.

Two warnings are appropriate before we turn to the details of the study. First, we consider in detail only point estimates but not test statistics, the latter not being central to the question we wish to study. Second, we do not claim to be comprehensive in our choice of data generating processes. In particular, we recognize that whatever results we establish under our simplifying assumption that sales are exogenous might not hold under a more sophisticated, and, in our view, more plausible, setup in which sales are endogenous. Nor, of course, is it assured that our results will obtain if, in contrast to the present study, the model is inconsistent with the data.

The paper is structured as follows. The second section presents the model and solves for the reduced form. The third section presents our data generating processes. The fourth section describes our instrumental variables estimators. The fifth section considers an asymptotic approximation to the distribution of our parameters. The sixth section presents simulation evidence on this distribution. The seventh section concludes. An Appendix contains some algebra, as well as some results omitted from the body of the paper that are likely to be of interest mainly to a specialist interested in conducting a closely related study.

## II. The Model

The model follows Holt et al. (1960). A representative firm maximizes the expected present discounted value of future cash flows, with a cost function that includes linear and quadratic costs of production and of changing production and of holding inventories. Let $p_t$ be real price, $S_t$ real sales, $Q_t$ real production, $H_t$ real end of period inventories, $C_t$ real costs, b a discount factor, $0 \le b < 1$, $E_t$ mathematical expectations conditional on information known at time t, assumed equivalent to linear projections, and $u_t$ a cost shock that is observable to the firm but unobservable to the econometrician. The objective function is

$$\max \lim_{T\to\infty} E_t \Sigma_{j=0}^{T} b^j (p_{t+j}S_{t+j} - C_{t+j}) \qquad (2.1)$$

$$\text{s.t. } Q_{t+j} = S_{t+j} + H_{t+j} - H_{t+j-1},$$

$$C_{t+j} = .5a_0\Delta Q_{t+j}^2 + .5a_1 Q_{t+j}^2 + .5a_2(H_{t+j-1} - a_3 S_{t+j})^2 + H_{t+j}u_{t+j}$$

$$+ \text{ linear terms } + (\text{linear x trend}) \text{ terms}$$

For the moment, the $a_i$'s are all assumed to be positive. Our omission of shocks that shift the marginal cost of production or of changing production (i.e., terms of the form shock x $Q_{t+j}$ or shock x $\Delta Q_{t+j}$) is for notational economy and without economic substance.

As in West (1986), Eichenbaum (1989), Ramey (1991), Krane and Braun (1991) and Kashyap and Wilcox (1993), the instrumental variables technique that we consider works off a first order condition, or Euler equation. An optimizing firm will not be able to cut costs by increasing production by one unit this period, storing the unit in inventory, and producing one less unit next period, holding revenue unchanged throughout. Formally, differentiating (2.1) with respect to $H_t$ gives

$$E_t\{ a_0(\Delta Q_t - 2b\Delta Q_{t+1} + b^2\Delta Q_{t+2}) + a_1(Q_t - bQ_{t+1}) + ba_2(H_t - a_3 S_{t+1}) \quad (2.2)$$
$$+ \text{deterministic terms} + u_t \} = 0,$$

where the deterministic terms result from the linear and (linear x trend) terms in the cost function (2.1).

We aim to evaluate instrumental variables estimators of the parameters of (2.2). Our estimators are described in the next section of the paper. The remainder of this section describes how we generate the artificial data necessary to evaluate the estimators.

For simplicity, we generate data assuming that sales are exogenous to the firm. The equilibrium decision rule implied by the Euler equation (2.2) is (West (1992))

$$H_t = (\lambda_1+\lambda_2)H_{t-1} - \lambda_1\lambda_2 H_{t-2} + \quad (2.3)$$
$$b^{-1}\lambda_1\lambda_2(\lambda_1-\lambda_2)^{-1}\Sigma_{j=0}^{\infty}[(b\lambda_1)^{j+1}-(b\lambda_2)^{j+1}]E_t D_{t+j} + \text{deterministic terms},$$
$$D_t = -(\Delta S_t - 2b\Delta S_{t+1} + b^2\Delta S_{t+2}) - a_0^{-1}a_1(S_t - bS_{t+1}) + a_0^{-1}ba_2a_3S_{t+1}$$
$$- a_0^{-1}u_t,$$

$\lambda_1$, $\lambda_2$ the two smallest (in modulus) roots of:

$$\lambda^4 - b^{-2}a_0^{-1}[ba_1+2a_0b(1+b)]\lambda^3 + b^{-2}a_0^{-1}[a_0(1+4b+b^2)+a_1(1+b)+ba_2]\lambda^2$$
$$- b^{-2}a_0^{-1}[a_1+2a_0(1+b)]\lambda + b^{-2} = 0.$$

The above assumes for simplicity that $\lambda_1 \neq \lambda_2$. If $\lambda_1$ and $\lambda_2$ are complex, they are complex conjugates, so that $\lambda_1+\lambda_2$ and $\lambda_1\lambda_2$ are real.

We assume that $S_t$ is forecast from a trend-stationary AR(2) and that the cost shock is white noise:

$$S_t = \phi_1 S_{t-1} + \phi_2 S_{t-2} + \text{constant} + \text{trend} + \epsilon_{St}, \qquad (2.4)$$
$$(u_t, \epsilon_{St}) \sim \text{i.i.d. } N(0,\Sigma),\ \Sigma \text{ positive definite.}$$

In closed form, (2.3) is then

$$H_t = (\lambda_1+\lambda_2)H_{t-1} - \lambda_1\lambda_2 H_{t-2} + \delta_1 S_t + \delta_2 S_{t-1} \qquad (2.5)$$
$$+ \text{constant} + \text{trend} + \psi u_t$$

for certain $\delta_i$'s and a certain $\psi$ that depend on b, $\lambda_1$, $\lambda_2$ and the parameters in (2.4). Exact formulas are given in the appendix. Equations (2.4) and (2.5) can then be combined to obtain a reduced form data generating process

$$H_t = (\lambda_1+\lambda_2)H_{t-1} - \lambda_1\lambda_2 H_{t-2} + \pi_1 S_{t-1} + \pi_2 S_{t-2} \qquad (2.6a)$$
$$\text{constant} + \text{trend} + \epsilon_{Ht},$$
$$\pi_1 = \delta_1\phi_1+\delta_2,\ \pi_2 = \delta_1\phi_2,\ \epsilon_{Ht} = \psi u_t+\delta_1\epsilon_{St}.$$
$$S_t = \phi_1 S_{t-1} + \phi_2 S_{t-2} + \text{constant} + \text{trend} + \epsilon_{St}, \qquad (2.6b)$$

where (2.6b) simply repeats (2.4).

## III. Generating the Synthetic Data

In all data generating processes, the discount factor b was set to 0.995 (appropriate if the data are assumed to be monthly). We experiment with four sets of cost parameters, given in Table IA. All are based on studies using U.S. data of one sort or another. Parameter set A is roughly consistent with the estimates for post-war aggregate data in West (1990) and those for automobile data in Blanchard and Melino (1985), parameter sets B and C with those for post-war two-digit manufacturing in Ramey (1991) and West (1986) respectively, parameter set D with those for auto data from the 1920's and 1930's in Kashyap and Wilcox (1993). See Ramey (1991) for an argument for the reasonableness of the negative values for $a_1$ in parameter sets B and D.

Table IB reports parameters for exogenous processes. The autoregressive coefficients of 0.7 and 0.25 were chosen to match roughly the estimates of an AR(2) around trend fit to real sales of nondurable goods manufacturing industries, monthly, 1967-1990. The sales innovation variance of 0.120833 was chosen so that the implied unconditional variance of sales is 1 (a harmless normalization). The variance of the cost shock $u_t$ and its correlation with the sales shock $\epsilon_{St}$ were chosen so that, in conjunction with the cost parameters of parameter set A (Table IA), the implied ratio $var(H_t)/var(S_t)$ and the implied correlation $\rho(H_t, S_t)$ approximately matched that of monthly nondurables manufacturing industries, 1967-1990, with $H_t$ total inventories. Coefficients on trend terms were chosen so that the implied coefficients of variation of $\Delta S_t$ and $\Delta H_t$ approximately match those of monthly nondurables manufacturing, 1967-1990; because different choices of the cost parameters imply different autoregressive coefficients in (2.6a), the coefficient on the trend term in (2.6a) varies from data generating process to data generating process.

A complete data generating process (DGP) is specified by combining a given set of cost parameters (A, B, C or D) with the sales and cost shock processes. Given a DGP, we generate data as follows. As indicated in (2.4), the vector of shocks $(u_t, \epsilon_{St})$ is assumed to be iid normal. This implies that $H_t$ and $S_t$ are normally distributed. We first draw a vector of initial values from the unconditional normal distribution of the 4x1 vector $(H_0, H_{-1}, S_0, S_{-1})'$. We then use (2.6) to generate 10,004 observations. Our experiments employ a sample size of 300, so we use observations 1 and 2 for lags, observations 303 and 304 for leads, and discard the final 10,004-304 = 9700 observations. These 9700 additional observations were reserved for some additional experiments that have yet to be concluded. 1000 samples were generated for each data generating process.

Table IC displays the implied values of the parameters of the inventory equation (2.6a) for each of our DGP's. The values of $\lambda_1+\lambda_2$ and $-\lambda_1\lambda_2$, the coefficients on inventories lagged once and twice, are similar for A, C and D, and suggest slow adjustment of inventories to shocks; the values for B suggest quick adjustment, which may be counterfactual for much inventory data. (If the cost shock is serially correlated, as is assumed by Ramey (1990) and by us in a specification presented Appendix Table A2, adjustment will be slow.)

Table ID displays the second moments of inventories and sales that are implied by the various DGP's. As noted above, the values of $var(H_t)/var(S_t)$ and of $\rho(H_t,S_t)$ for DGP A are approximately those for monthly nondurables in manufacturing, 1967-1990. The values of $var(H_t)/var(S_t)$ and of $\rho(H_t,S_t)$ for DGP's B, C and D are rather different, but no doubt are representative of some other inventory data! Across DGP's, the values of first order autocorrelation coefficients are similar for inventories and are of course identical for sales.

## IV. Estimating the Parameters

### A. Choice of Left Hand Side Variable

Given the deterministic terms present in our data generating processes, (2.2) becomes

$$E_t\{ a_0(\Delta Q_t-2b\Delta Q_{t+1}+b^2\Delta Q_{t+2}) + a_1(Q_t-bQ_{t+1}) + ba_2(H_t-a_3S_{t+1}) + d + \delta t + u_t\} = 0. \tag{4.1}$$

We include $d+\delta t$ only to make clear exactly what we did; our interest is in the $a_i$'s, and we will not investigate the sampling distribution of estimates of the coefficients on the constant and trend terms.

In (4.1), note that the parameters $a_0$, $a_1$, $a_2$, d and $\delta$ are identified only up to scale: if $\{a_0,a_1,a_2,a_3,d,\delta\}$ set $u_t$ orthogonal to the instrument set, then so does $\{\alpha a_0,\alpha a_1,\alpha a_2,a_3,\alpha d,\alpha\delta\}$ for any nonzero $\alpha$. Thus by estimating (4.1) alone, one cannot recover absolute magnitudes of the parameters but only their magnitudes relative to some linear combination of themselves. Given a choice of "denominator" (a choice of linear combination), values of any two of $a_0$, $a_1$ and $a_2$ relative to this denominator determine the value of the third relative to the chosen denominator. Our aim, then, is to analyze three parameter estimates: (i)two of $a_0$, $a_1$, and $a_2$ relative to some "denominator," and (ii)$a_3$.

In reporting parameter estimates, we follow much empirical work and (1)let choice of left hand side variable dictate which parameter estimates to report, with the coefficient on the variable moved to the left hand side being the "denominator" used in reporting, and (2)report $a_3$ regardless of left hand side variable. To illustrate our approach, focus for the moment on the normalization that puts $ba_2H_t$ on the left hand side and then divides both sides of the equation by $ba_2$. This normalization was used by Ramey

(1991) and, in part, by Krane and Braun (1991) and Kashyap and Wilcox (1993). In the tables below this is called the HH normalization:

$$H_t = (a_0/a_2)X_{0t+2} + (a_1/a_2)X_{1t+1} + a_3S_{t+1} + (d/ba_2) + (\delta/ba_2)t + v_{t+2}, \tag{4.2}$$

$$= X_t'\beta + v_{t+2},$$

$$X_{0t+2} = -b^{-1}(\Delta Q_t - 2b\Delta Q_{t+1} + b^2\Delta Q_{t+2}),$$

$$X_{1t+1} = -b^{-1}(Q_t - bQ_{t+1}),$$

$$v_{t+2} = -(ba_2)^{-1}(u_t + ba_0(X_{0t+2} - E_tX_{0t+2}) + ba_1(X_{1t+1} - E_tX_{1t+1})) - a_3(S_{t+1} - E_tS_{t+1}),$$

$$X_t = (X_{0t+2}, X_{1t+1}, S_{t+1}, 1, t)',$$

$$\beta = (a_0/a_2, a_1/a_2, a_3, d/ba_2, \delta/ba_2)'.$$

As is typical in empirical work, we impose a value of b; the value chosen was that used in generating the data, b=.995. With a value of b imposed, we can construct $X_{0t}$ and $X_{1t}$, and estimate $\beta$ linearly with a conventional instrumental variables technique described in detail in section B below. For this normalization, $a_2$ is the "denominator" referenced above, and in our tables below we report the small-sample distribution of estimates of $a_0/a_2$, $a_1/a_2$ and $a_3$.

In this context, choice of left hand side variable is irrelevant asymptotically, provided the "denominator" is nonzero in the population. But as was noted in the introduction, Ramey (1991), Krane and Braun (1991) and Kashyap and Wilcox (1993), using various datasets, found that estimated parameters sometimes varied widely for different choices of left hand side variable. We therefore consider two alternative choices of left hand side variable, in order to evaluate the possibility that such variation is likely even when the model is correctly specified.

The first of these alternatives is the Legendre-Clebsch or LC normalization used in Kashyap and Wilcox (1993) and experimented with in Ramey (1991). Define $c_t$ as the present value of future costs, $c_t = E_t \Sigma_{j=0}^{\infty} C_{t+j}$. This normalization puts $(\partial^2 c_t/\partial H_t^2)H_t = [a_0(1+4b+b^2)+a_1(1+b)+ba_2]H_t = c_1 H_t$ on the left hand side and then divides both sides of the equation by $c_1$. Then (2.2) may be rewritten

$$H_t = (a_0/c_1)X_{2t+2} + (a_1/c_1)X_{3t+1} + (a_2a_3/c_1)(bS_{t+1}) \qquad (4.3)$$

$$(d/c_1) + (\delta/c_1)t + v_{2t+2},$$

$$X_{2t+2} = bX_{0t+2} + (1+4b+b^2)H_t$$

$$X_{3t+1} = bX_{1t+1} + (1+b)H_t$$

$$v_{2t+2} = -c_1^{-1}\{u_t + a_0(X_{2t+2}-E_tX_{2t+2}) + a_1(X_{3t+1}-E_tX_{3t+1})$$

$$+ ba_2a_3(S_{t+1}-E_tS_{t+1})\},$$

$$c_1 = a_0(1+4b+b^2)+a_1(1+b)+ba_2.$$

Here, the "denominator" is $c_1$, and in our tables below we report estimates of $a_0/c_1$, $a_1/c_1$ and $a_3$. We obtain $\hat{a}_3$ using $a_3 = b(a_2a_3/c_1)$ / $[1-(1+4b+b^2)(a_0/c_2)-(1+b)(a_1/c_2)]$.

The third and final normalization is that used in West (1986) and Krane and Braun (1991), which puts $[(1+b)a_0+a_1](bQ_{t+1}-Q_t) = c_2(bX_{1t+1})$ on the left hand side and divides both sides of the equation by $bc_2$. We call this the QC normalization since $c_2$ is the slope of the marginal cost of production $Q_t$. In this case the regression equation is

$$X_{1t+1} = (a_0/c_2)X_{4t+2} + (a_2/c_2)(bH_t) + (a_2a_3/c_2)(-bS_{t+1}) + \qquad (4.4)$$

$$(d/c_2) + (\delta/c_2)t + v_{3t+2},$$

$$X_{4t+2} = (b^2X_{1t+2}+bX_{1t}),$$

$$v_{3t+2} = bX_{1t+1}-E_tbX_{1t+1} +$$

$$c_2^{-1}( u_t - a_0(X_{4t+2}-E_tX_{4t+2}) + ba_2a_3(S_{t+1}-E_tS_{t+1}) ),$$

$$c_2 = (1+b)a_0+a_1.$$

Here, the "denominator" is $c_2$, and in our tables below we evaluate estimates of $a_0/c_2$, $a_2/c_2$ and $a_3 = (a_2a_3/c_2)/(a_2/c_2)$. Table II lists the three sets of coefficients.

B. Estimation Technique

We use (4.2) to illustrate the estimation technique. Let $Z_t$ be a 6x1 vector of instruments consisting of the variables that appear in the reduced form (2.6),

$$Z_t = (H_{t-1},H_{t-2},S_{t-1},S_{t-2},1,t)'. \tag{4.5}$$

(Because cost shocks are present, period t values of $H_t$ and $S_t$ are not legitimate instruments; see (2.4) and (2.5).) Let T be the sample size, where T=300 in our experiments. Let Z be a Tx6 matrix whose t'th row is $Z_t'$, $X = [X_t']$ be the Tx5 matrix of right hand side variables, $Y = [H_t]$ be the Tx1 vector of the left hand side variable. In the Monte Carlo experiments, we follow much recent empirical work and use the instrumental variables estimator that has the smallest possible asymptotic variance-covariance matrix given the set of instruments used,

$$\hat{\beta} = (X'Z\hat{W}Z'X)^{-1}X'Z\hat{W}Z'Y, \tag{4.6}$$

where $\hat{W}$ is a qxq matrix that is an estimate of the inverse of the spectral density at frequency zero of the 6x1 vector $Z_tv_{t+2}$, i.e., the inverse of $\Sigma_{j=-\infty}^{\infty}EZ_tZ_{t-j}'v_{t+2}v_{t+2-j}$. Since the cost shock $u_t$ is iid in our data generating processes, $v_{t+2}$ and $Z_t v_{t+2}$ are MA(2) and this infinite sum collapses to

$$W = (\Sigma_{j=-2}^{2} EZ_t Z_{t-j}' v_{t+2} v_{t+2-j})^{-1}. \qquad (4.7)$$

Two technical notes: First, given trend stationarity of $H_t$ and $S_t$ (as opposed to stationarity around a constant mean), the expectation $EZ_t Z_{t-j}' v_{t+2} v_{t+2-j}$ depends on t, and so W as defined in (4.7) varies with t. Technically, the asymptotic theory requires scaling the elements of $Z_t$ (and $X_t$) by certain diagonal matrices whose elements are functions of T, after which the relevant probability limits do not vary with t (West (1988)). For the sake of simplicity, we slur over such complications in our discussion here and in the definitions of $\hat{\Gamma}_j$ and V (equations (4.8) and (5.1) below).

Second, since $Z_t v_{t+2}$ is not white noise, more efficient estimates would be obtained if additional lags of $H_t$ and $S_t$ were used, even though such lags do not appear in the reduced form. See Hansen (1985) for a general statement, West and Wilcox (1993) for discussion in the context of the linear quadratic inventory model.

To construct $\hat{W}$ given our choice of $Z_t$, let $\dot{v}_{t+2}$ be the two stage least squares residual, and let

$$\hat{\Gamma}_j = T^{-1}\Sigma_{t=j+1}^{T} Z_t Z_{t-j}' \dot{v}_{t+2} \dot{v}_{t+2-j} \qquad (4.8)$$

for $j \geq 0$. Let

$$m = \min(10, [\hat{\gamma} T^{1/3}])$$

where

$$\hat{\gamma} = 1.1447(\hat{s}^{(1)}/\hat{s}^{(0)})^{2/3}, \ \hat{s}^{(1)} = 2\hat{\sigma}_1 + 4\hat{\sigma}_2, \ \hat{s}^{(0)} = \hat{\sigma}_0 + 2\hat{\sigma}_1 + 2\hat{\sigma}_2,$$

$\hat{\sigma}_j = w'\hat{\Gamma}_j w$, $w = (1,1,1,1,1,1)'$.

We set

$$\hat{W} = \{\hat{\Gamma}_0 + \Sigma_{j=1}^{m}[1-j/(m+1)](\hat{\Gamma}_j + \hat{\Gamma}_j')\}^{-1}. \quad (4.9)$$

The weights $1-j/(m+1)$ guarantee that $\hat{W}$ is positive definite. Newey and West (1992) provide analytical and simulation evidence on this technique for estimating W (although that paper did not consider truncating m at 10 or at any bound less than the sample size; we do that here to speed computation).

Equations (4.3) and (4.4) were estimated in analogous fashion, with appropriate changes in left and right hand side variables, but with the same instruments and estimation technique.

## V. Asymptotic Approximation to Distribution of Parameter Estimates

In this section, we use conventional asymptotic theory to approximate the sampling distribution of the parameter estimates. To explain this asymptotic theory, we focus on the HH normalization (equation (4.2)). As in (4.2), let $\beta$ be defined as the vector of coefficients on the right hand side under this normalization, $X_t$ be the vector of right hand side variables, $Z_t$ be the vector of instruments. Also let W be defined as in (4.9), and let $\hat{\beta}$ be a sample estimate computed as in (4.6). For a sufficiently large sample size T, we interpret Hansen (1982) as showing that

$$\hat{\beta} \approx N(\beta, V/T), \quad (5.1)$$

$$V = (EX_t Z_t' W E Z_t X_t')^{-1}.$$

(In the definition of V, we once again slur over the complications induced by trend stationarity.) Let $V_{ii}$ be the i'th diagonal element of the (5x5) matrix V. The square root of $V_{ii}/T$ is the asymptotic standard error for the i'th element of $\beta$, for a sample size of T. In normalizations LC and HH, in which $\hat{a}_3$ is obtained as a nonlinear function of the regression coefficients (see section IV), we obtain an asymptotic standard error using the conventional delta method.

Table III presents asymptotic standard errors for each DGP for the HH normalization. Results for the other two normalizations are similar; see the Appendix. (Exception: for QC, DGP B, standard errors for $a_0/c_2$ and $a_2/c_2$ are far larger [and implied t-statistics far smaller] than the comparable figures for HH, apparently because the "denominator" in this case is very small relative to the coefficients on the right hand side variables [for DGP B, $c_2 = (1+b)a_0+a_1 = -.005$].)

According to the asymptotic approximation, the distribution of some of the parameter estimates is very diffuse. Consider the case of $a_1/a_2$ in DGP A. Taken literally, (5.1) implies that there is only a 57 percent chance that $a_1/a_2$ will have the correct sign (since the probability is .57 that a $N(1,(5.7)^2)$ random variable is positive). In order to have a 95 percent chance of estimating $a_1/a_2$ with the correct sign, an investigator would need roughly 30,000 monthly observations (since a sample over 100 times bigger than that assumed in Table III is required to get the standard error to fall to about 0.50). Similar but less extreme statements apply to $a_3$ in DGP's A and D and to all three parameters in DGP C.

A small amount of experimentation suggests that the distribution of parameter estimates often is diffuse even when one departs dramatically from the parameter settings assumed so far. In particular, such dispersion still obtains if $S_t$ and $u_t$ are difference-stationary, or if one makes large changes in the variance-covariance matrix of $(u_t, \epsilon_{St})$. Details are in the Appendix.

Moreover, even if one uses not instrumental variables but full information maximum likelihood and estimates (2.6a) and (2.6b) jointly, imposing the cross equation restrictions (e.g., Blanchard (1983)), some of the parameter estimates remain diffuse (although less so, of course). In DGP A, for example, the asymptotic standard error for the full information maximum likelihood estimator of $a_1/a_2$ is about 2.2 for a sample size of 300. While the 2.2 figure is much less than the 5.7 in Table III, it is still big enough that, according to (5.1), more than 5000 monthly observations would be required before an investigator would have a 95 percent probability of calculating a positive estimate.

We therefore interpret the dispersion exhibited in Table III as reflecting two factors. First, sample realizations from some plausible DGP's may not be very informative about the values of the underlying cost parameters, as evidenced by the inability of even the FIML estimator to deliver precise estimates of those parameters in some cases. Second, the instrumental variables estimator that we study sometimes is not very efficient relative to maximum likelihood at extracting such information about the cost parameters as is embedded in the data.

That the cost parameters may sometimes be ill-determined statistically does not necessarily imply that they are ill determined economically; it is conceivable that the economic implications of a given parameter might not be very sensitive to a one- or two-standard deviation perturbation in the assumed value of that parameter. The following example indicates that such perturbations in parameter values do, however, have economically important implications in at least some contexts.

Suppose, for example, that one is interested in using the cost parameters to derive the implications of the model for the ratio of the variance of production to that of sales. Blinder and Maccini (1991) point out that estimates of this ratio typically are greater than one, maintain (as do many others) that it is central that an inventory model explain this stylized fact, and argue (as do some but not all others) that a plausible explanation should hold even in the absence of cost shocks. With no cost shocks, the conditions that allow such an explanation include: $a_1/a_2$ negative and sufficiently large in absolute value, and/or $a_3$ positive and sufficiently large. (See West (1992).) Would plausible perturbations in the cost parameters give rise to economically meaningful variations in the implied variance ratio? We investigate this question for $a_1/a_2$, for DGP A.

The first line of Table IV indicates that with $a_1/a_2$ set at its population value, DGP A implies a ratio equal to 1.00 (by coincidence, not by design). It is the positive value of $a_3$ that explains why this ratio is not below 1. When, however, the value of $a_1/a_2$ instead is -6 (about 1.2 asymptotic standard errors below $a_1/a_2$-1, according to Table III), the implied ratio is 1.19 (line (2) of Table IV), roughly the median value reported for two-digit data by Blinder (1986). Thus, given an estimate of -6, as well as estimates of $a_0/a_2$ and $a_3$ that are at or (by continuity) close to their population values, one might well interpret the model as adequately explaining the stylized fact; as indicated in line (5), given a comparable overestimate, one probably would not. Lines (3) and (4) indicate that 0.8 standard error under- or overestimates also imply variance ratios that some observers might consider qualitatively different. We read Table IV as suggesting that sampling variation in point estimates might lead to qualitatively different interpretations of the underlying economic environment.

## VI. Simulation Evidence on Distribution of Parameter Estimates

The previous section indicates that the estimates of the instrumental variables estimator are quite diffuse for some plausible specifications. This diffuseness suggests that sampling error might account for the sensitivity to choice of left hand side variable that was noted in the introduction (along with sensitivity to some other seemingly minor changes in specification). To investigate this possibility, and, more generally, to investigate the accuracy of the asymptotic approximation, we conduct a set of Monte Carlo experiments.

Table V presents some Monte Carlo results on the distribution of the parameter estimates. The top three panels summarize results of estimating the cost parameters using each of our three normalizations and four DGP's. The panel labelled "asymptotic" gives the corresponding asymptotic quantities. These are independent of DGP by virtue of the way we report results: we standardize each estimated parameter by subtracting the population parameter value and then dividing by the population asymptotic standard error. (We used a population rather than estimated standard error because our interest is in the distribution of parameter estimates and not test statistics.) According to the asymptotic theory, the resulting quantity should be approximately N(0,1).

For each of the three parameters, the column labelled "50% CI" gives a 50 percent confidence interval constructed by dropping the largest 250 and smallest 250 of the 1000 parameter estimates, or, for "asymptotic", the values appropriate for a N(0,1) variable. The difference between the upper and lower bounds of these confidence intervals is the interquartile range. "Median" gives the median of the 1000 estimates, "Trimmed MSE" a mean squared error computed by (1)dropping all entries greater than 3.0 in absolute value, (2)calculating the average squared value of the remaining observations, and (3)dividing by 0.9735, which is the variance of a N(0,1) variable doubly truncated at -3 and +3 (Johnson and Kotz (1970, p83)). We trimmed before computing the MSE because the simultaneous equations literature suggests that second moments of our estimator may not exist, since our equation has only one more instrument than right hand side variable (e.g., Phillips (1983)). The decision to truncate at 3.0 was arbitrary, and similar conclusions follow if one instead trims at 2.0. The number of observations excluded typically was fewer than 25 for $a_0/()$, $a_1/()$ and $a_2/()$, and between 100 and 200 for $a_3$.

We read Table V as indicating that in some respects the asymptotic theory works reasonably well. The upper and lower bounds of the 50 percent CI's generally are close to the theoretical values, and usually have roughly the predicted width (median width is 1.5, as compared to the asymptotic width of 1.4). The median of the "trimmed MSE" column is 1.16, indicating that in half the entries the MSE is at most 16 percent bigger than the asymptotic theory would predict. Asymptotic theory thus often provides good guidance to how disperse parameter estimates will be. Also, the median of the absolute value of the "median" column is 0.20, indicating that for half of the 36 parameters, the median of the 1000 estimates is within 0.20 asymptotic standard errors of the true parameter value.

On the other hand, we also read Table V as indicating some departures from the asymptotic approximation. While the 50% CIs generally look reasonable, 7 of them do lie entirely on one or the other side of zero. This means that at least three-fourths of the parameter estimates in these 7 cases fell on one side of the true value. All such instances occurred with normalization HH. Another indication of bias is that 5 of the median estimates are between .5 and 1 asymptotic standard errors away from the true parameter value ($a_0/a_2$: C; $a_1/a_2$: A, C; $a_3$: C/HH, B/QC), 3 more than one asymptotic standard error away ($a_0/a_2$: A, D; $a_1/a_2$: D).

In addition, the MSE's do tend to be greater than 1.00, the value predicted by the asymptotic approximation. While, as noted above, the median value is 1.16, only 8 of the 36 are below 1.00, and 2 of these 8 occur in QC, DGP B (which, recall, had huge asymptotic standard errors). The MSE is between 1.5 and 2.0 in 4 cases ($a_0/a_2$: A; $a_3$: B/HH, B/LC, D/QC), and is greater than 2 in three cases ($a_0/a_2$: D, $a_1/a_2$: D; $a_3$: B/QC).

Another respect in which asymptotic theory does not hold is that different normalizations perform differently. Some evidence to this effect has just been noted, in that HH's confidence intervals are more poorly centered than are LC's or QC's. In addition, its variability (as measured by the trimmed MSE) is more erratic.

Additional evidence on differences across normalizations is given in Tables VI and VII. These are contingency tables giving the probability that the signs of the estimates of $a_1/()$ agree for each pair of normalizations, where $() = a_2$ for HH, $() = c_1$ for LC, $() = c_2$ for QC.

We focus on $a_1$ because a number of authors have noted that in empirical work the sign of this parameter changes with normalization, tending to be negative for HH (Krane and Braun (1991), Ramey (1991), Kashyap and Wilcox (1993)). And, as discussed above, the sign of this parameter has key economic implications, since a negative value tends to induce production bunching, a positive value production smoothing.

Table VI presents the contingency table for the differences between estimated and true parameters, Table VII for the raw estimates themselves. Each panel in these two tables compares the estimates of $a_1/()$ from two normalizations. We consider three normalizations, so there are $3!/2! = 3$ panels. Within each panel, the 2x2 blocks present the results for the different DGP's. To see how the 2x2 blocks are calculated, consider the

2x2 block in the upper left hand corner of Table VI. The "0.564" indicates that in 564 of the 1000 simulations, the point estimates of both $a_1/a_2$ and $a_1/c_1$ were less than the true values (-1 and .16), the "0.299" that in 299 simulations, the point estimate of $a_1/a_2$ was less than the true value, the point estimate of $a_1/c_1$ greater than the true value, the "0.137" that in the remaining 137 simulation the point estimates of both were greater than the true values. Thus, HH's estimate of $a_1/a_2$ was less than the true value in 863 = 564+299 of the simulations. If both normalizations were median unbiased, the sum of the entries in each row and column would be 0.500. If, further, an overestimate from one normalization were invariably accompanied by an overestimate from the other, and similarly for underestimates, the diagonal elements would each be 0.500, the off-diagonals 0.000; if, on the other hand, both normalizations were median unbiased but an overestimate from one were accompanied by an overestimate from the other exactly half the time, and similarly for underestimates, each of the four elements would be 0.250.

Consistent with what one might have guessed from Table V, Table VI indicates some tendencies of the different normalizations to produce parameter estimates that are biased in different directions. Panels A and B suggest that, as compared to either LC or QC, HH produces more estimates that are (1)negatively biased for DGP's A and C, (2)positively biased for DGP D. (Panels B and C indicate an even more substantial conflict for QC, DGP B, which presumably is an artifact of the numerically small value of $c_2$ for that DGP.)

Table VII suggests that substantively different economic implications might be drawn from different normalizations. As discussed above, in DGP A, the asymptotic theory indicates that $a_1/()$ is likely to be estimated imprecisely, in the sense that there is likely to be a substantial probability of an incorrectly signed (negative) estimate. It may be seen that while this happens for LC in 48 percent of the samples, and for QC in 41.4 percent of the samples, it happens for HH in 80.6 (= 48.0 + 32.6) percent of the samples. In about a third (≈ .326 or .392) of the replications, HH yielded a negative estimate while LC or QC yielded a positive one. A similar pattern obtains in DGP C. In DGP D, there is little difference across the normalizations, as one might expect, given the asymptotic standard errors presented in Table III. (In DGP B, QC tends to spuriously yield positive estimates, which we once again consider uninteresting.)

## VII. Conclusions

Asymptotic and Monte Carlo results indicate considerable dispersion in estimates of the parameters of the Holt et al. (1960) linear quadratic inventory model, when the estimates are obtained by applying instrumental variables to a first order condition of the model. Alternative normalizations have substantial probability of delivering differently signed estimates of the parameters of the model. A priority for future work is investigation of alternative estimators, such as ones that pool data from various industries.

References

Blanchard, Olivier J., 1983, "The Production and Inventory Behavior of the American Automobile Industry," Journal of Political Economy 91, 365-400.

Blanchard, Olivier J., and Angelo Melino, 1986, "The Cyclical Behavior of Prices and Quantities: The Case of the Automobile Market," Journal of Monetary Economics 17, 379-408.

Blinder, Alan S., 1986, "Can the Production Smoothing Model of Inventories Be Saved?", Quarterly Journal of Economics CI, 431-54.

Blinder, Alan S. and Louis J. Maccini, 1991, "Taking Stock: A Critical Assessment of Recent Research on Inventories," Journal of Economic Perspectives 5, 73-96.

Eichenbaum, Martin S., 1989, "Some Empirical Evidence on the Production Level and Production Cost Smoothing Models of Inventory Investment," American Economic Review 79, 853-864.

Ferson, Wayen E. and Stephen E. Foerster, 1991, "Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," manuscript, University of Chicago.

Granger, C.W.J., and Tae Hwy Lee, 1989, "Investigation of Production, Sales and Inventory Relationships Using Multicointegration and Non-Symmetric Error Correction Models," Journal of Applied Econometrics 4, S145-S159.

Gregory, Allan W. and Michael R. Veale, 1985, "Formulating Wald Tests of Nonlinear Restrictions," Econometrica 53, 1465-1468.

Hansen, Lars Peter, 1982, "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica 50, 1029-54.

Hansen, Lars Peter, 1985, "A Method for Calculating Bounds on the Asymptotic Variance-Covariance Matrices of Generalized Method of Moments Estimators," Journal of Econometrics 30, 203-228.

Hansen, Lars Peter, and Kenneth J. Singleton, 1988, "Computing Semi-Parametric Efficiency Bounds for Linear Time Series Models," manuscript, University of Chicago.

Hansen, Lars Peter, and Kenneth J. Singleton, 1990, "Efficient Estimation of Linear Asset Pricing Models with Moving-Average Errors," NBER Technical Working Paper No. 86.

Holt, Charles C., Modigliani, Franco, Muth, John F., and Simon, Herbert A., 1960, Planning Production, Inventories and Work Force, Englewood Cliffs, NJ: Prentice Hall.

Johnson, Norman L., and Samuel Kotz, 1970, Continuous Univariate Distributions-1, Boston: Houghton Mifflin Company.

Kashyap, Anil K and David W. Wilcox, 1993, "Production and Inventory Control at the General Motors Corporation During the 1920s and 1930s," American Economic Review 83, 383-401.

Kocherlakota, Narayana, 1990, "On Tests of Representative Consumer Asset Pricing Models," Journal of Monetary Economics 26, 285-304.

Krane, Spencer D. and Steven N. Braun, 1991, "Production Smoothing Evidence from Physical Product Data," Journal of Political Economy 99, 558-581.

Newey, Whitney K., and Kenneth D. West, 1987, "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica 55, 703-708.

Newey, Whitney K., and Kenneth D. West, 1992, "Automatic Lag Selection in Covariance Matrix Estimation," manuscript, University of Wisconsin.

Phillips, Peter C. B., 1983, "Exact Small Sample Theory in the Simultaneous Equations Model," 449-516 in Z. Griliches and M.D. Intriligator (eds.) Handbook of Econometrics Volume 1, Amsterdam: North Holland.

Phillips, Peter C. B., and Joon Y. Park, 1988, "On the Formulation of Wald Tests of Nonlinear Restrictions," Econometrica 56, 1065-1084.

Ramey, Valerie A., 1991, "Nonconvex Costs and the Behavior of Inventories," Journal of Political Economy 99, 306-334.

Tauchen, George, 1986, "Statistical Properties of Generalized Method of Moments Estimators of Structural Parameters Obtained from Financial Markets Data," Journal of Business and Economic Statistics 4, 397-425.

Theil, Henri, 1972, Principles of Econometrics, New York: John Wiley and Sons.

West, Kenneth D., 1986, "A Variance Bounds Test of the Linear Quadratic Inventory Model," Journal of Political Economy 94, 374-401.

West, Kenneth D., 1988, "Asymptotic Normality, When Regressors Have a Unit Root," Econometrica 56, 1397-1418.

West, Kenneth D., 1990, "The Sources of Fluctuations in Aggregate Inventories and Sales," Quarterly Journal of Economics CV, 939-971.

West, Kenneth D., 1992, "Inventory Models," manuscript.

West, Kenneth D., and David W. Wilcox, 1993, "Finite Sample Behavior of Alternative Instrumental Variables Estimators of a Dynamic Linear Model," manuscript in preparation.

## Appendix

This appendix presents:

1. The parameters in (2.6).
2. Asymptotic t-statistics for all three normalizations, T=300 (Table A1).
3. Asymptotic standard errors alternative parameters for exogenous processes (Tables A1 and A2).

### 1. The parameters in (2.6).

Define the scalars $\rho_1$, $\rho_2$, $w_1$, $w_2$, $w_3$, and $w_4$, the (1x2) vector e' and the (2x2) matrices $\Phi$ and D as

$$\rho_1 = \lambda_1+\lambda_2,\ \rho_2 = -\lambda_1\lambda_2,$$

$$w_1 = b^2\rho_2,\ w_2 = -\rho_2[b^2+2b+b(a_1/a_0)+(ba_2a_3/a_0)],$$

$$w_3 = \rho_2[2b+1+(a_1/a_0),\ w_4 = -\rho_2,\ e' = (1\ 0),$$

$$\Phi = \begin{pmatrix}\phi_1 & \phi_2\\ 1 & 0\end{pmatrix},$$

$$D = [I-b\rho_1\Phi-b\rho_2\Phi^2]^{-1}.$$

Then

$$\psi = (\rho_2/a_0),$$

$$(\pi_1,\pi_2)' = e'D(w_1\Phi^3+w_2\Phi^2+w_3\Phi+w_4I),$$

$$\delta_1 = \pi_2/\phi_2,\ \delta_2=\pi_1-\delta_1\phi_1.$$

### 2. Asymptotic t-statistics for all three normalizations, T=300.

| Norm. | DGP | $a_0/a_2$ | $a_1/a_2$ | $a_3$ |
|---|---|---|---|---|
| HH | A | 2.15 | 0.20 | 0.26 |
| | B | 2.69 | -1.46 | 25.36 |
| | C | 1.17 | 0.63 | 0.66 |
| | D | 2.63 | -6.61 | 1.26 |

| Norm. | DGP | $a_0/c_1$ | $a_1/c_1$ | $a_3$ |
|---|---|---|---|---|
| LC | A | 6.13 | 0.20 | 0.26 |
| | B | 2.36 | -1.23 | 25.36 |
| | C | 2.44 | 1.58 | 0.66 |
| | D | 13.25 | -2.02 | 1.26 |

| Norm. | DGP | $a_0/c_2$ | $a_2/c_2$ | $a_3$ |
|---|---|---|---|---|
| QC | A | 3.95 | 1.48 | 0.26 |
| | B | -0.006 | -0.006 | 25.36 |
| | C | 1.92 | 0.86 | 0.66 |
| | D | 7.14 | 1.95 | 1.26 |

## Table A1

### Asymptotic Standard Errors, T=300, Alternative Trend Stationary Specifications

### A. Parameters of Exogenous Processes

| Mnemonic | $\phi_1$ | $\phi_2$ | $var(\epsilon_S)$ | var(u) | $\rho(\epsilon_S, u)$ | cv(ΔH) | cv(ΔS) |
|---|---|---|---|---|---|---|---|
| none | .75 | .20 | .120833 | 3.5 | -.5 | .2 | .2 |
| 3 | .75 | .20 | .120833 | 1.0 | -.2 | .2 | .2 |
| 4 | .75 | .20 | .120833 | 7.0 | -.8 | .2 | .2 |
| 5 | .75 | .20 | .120833 | 1.0 | -.8 | .2 | .2 |
| 6 | .75 | .20 | .120833 | 7.0 | -.2 | .2 | .2 |
| 7 | 1.70 | -.72 | .120833 | 3.5 | -.5 | .2 | .2 |
| 8 | .10 | .05 | .120833 | 3.5 | -.5 | .2 | .2 |

### B. Asymptotic Standard Errors, T=300

| DGP | Parameter $a_1/a_0$ | $a_2/a_0$ | $a_3$ |
|---|---|---|---|
| | .10 | .10 | .10 |
| A | (.53) | (.05) | (.38) |
| A3 | (.51) | (.05) | (.43) |
| A4 | (.51) | (.05) | (.43) |
| A5 | (.65) | (.06) | (.32) |
| A6 | (.49) | (.04) | (.52) |
| A7 | (.31) | (.03) | (.13) |
| A8 | (.54) | (.05) | (69.) |

Notes:

This presents information analogous to Table III, when the parameters of the exogenous processes are varied as indicated in the table. DGP A is the one studied in the text.

## Table A2

### Asymptotic Standard Errors, T=300, Difference Stationary Specifications

#### A. Parameters of Exogenous Processes

| Mnemonic | $\phi_1$ | $\phi_2$ | $var(\epsilon_S)$ | $var(\bar{u})$ | $\rho(\epsilon_S,\bar{u})$ | $cv(\Delta H)$ | $cv(\Delta S)$ |
|---|---|---|---|---|---|---|---|
| 2 | -.2 | .1 | .941111 | 7.0 | 0.4 | .2 | .2 |

#### B. Asymptotic Standard Errors, T=300

| | Parameter | | | | Parameter | | |
|---|---|---|---|---|---|---|---|
| DGP | $a_1/a_0$ | $a_2/a_0$ | $a_3$ | DGP | $a_1/a_0$ | $a_2/a_0$ | $a_3$ |
| A2 | .10<br>(.53) | .10<br>(.05) | .1<br>(35.5) | A | .10<br>(.53) | .10<br>(.05) | .10<br>(.38) |
| B2 | -2.0<br>(1.0) | 6.0<br>(2.7) | .5<br>(.7) | B | -2.0<br>(0.9) | 6.0<br>(2.2) | .50<br>(.02) |
| C2 | 2.0<br>(2.0) | .10<br>(.06) | 1.0<br>(86.3) | C | 2.0<br>(2.1) | .10<br>(.08) | 1.0<br>(1.5) |
| D2 | -.50<br>(.20) | .10<br>(.04) | .5<br>(20.2) | D | -.50<br>(.21) | .10<br>(.04) | .50<br>(.40) |

Notes:

1. The left hand half of panel B presents information analogous to that in Table III, when (1)(2.4) is replaced by: $\Delta S_t = \text{constant} + \phi_1\Delta S_{t-1} + \phi_2\Delta S_{t-2} + \epsilon_{St}$, $u_t = u_{t-1} + \bar{u}_t$, $(\bar{u}_t, \epsilon_{St}) \sim$ i.i.d. $N(0,\Sigma)$, $\Sigma$ positive definite, with the panel A values for $\phi_1$, $\phi_2$, $\Sigma$ roughly calibrated to estimates for monthly nondurables manufacturing, 1967-1990; (2)the regression equation is in differences rather than levels; (3)the instrument vector is $Z_t=(\Delta H_{t-1},\Delta H_{t-2},\Delta S_{t-1},\Delta S_{t-2},1)'$.

2. The right hand half of panel B presents information analogous to that in Table III, and is included for comparison.

Table I

Data Generating Processes

A. Parameters of Cost Function

| Mnemonic | $a_0$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|---|
| A | 1. | .1 | .1 | .1 |
| B | 1. | -2.0 | 6.0 | .5 |
| C | 1. | 2.0 | .1 | 1.0 |
| D | 1. | -.5 | .1 | .5 |

B. Parameters of Exogenous Processes

| $\phi_1$ | $\phi_2$ | $var(\epsilon_S)$ | var(u) | $\rho(\epsilon_S,u)$ | cv(ΔH) | cv(ΔS) |
|---|---|---|---|---|---|---|
| .75 | .20 | .120833 | 3.5 | -.5 | .2 | .2 |

C. Implied Coefficients of Inventory Equation

| DGP | $\lambda_1+\lambda_2$ | $-\lambda_1\lambda_2$ | $\pi_1$ | $\pi_2$ |
|---|---|---|---|---|
| A | 1.22 | -0.42 | 0.14 | -0.12 |
| B | 0.24 | -0.14 | 0.38 | 0.05 |
| C | 1.07 | -0.22 | 0.10 | -0.09 |
| D | 1.43 | -0.69 | 0.33 | -0.15 |

D. Implied Second Moments

| DGP | $var(H_t)/var(S_t)$ | $\rho(H_t,S_t)$ | $\rho(H_t,H_{t-1})$ | $\rho(S_t,S_{t-1})$ |
|---|---|---|---|---|
| A | 2.5 | 0.23 | 0.86 | 0.93 |
| B | 0.3 | 0.91 | 0.81 | 0.93 |
| C | 0.6 | 0.27 | 0.88 | 0.93 |
| D | 10.7 | 0.35 | 0.86 | 0.93 |

Notes:

1. The cost function (2.1) includes $.5a_0\Delta Q_t^2 + .5a_1Q_t^2 + .5a_2(H_{t-1}-a_3S_t)^2$; $\phi_1$ and $\phi_2$ are the autoregressive parameters of the sales process defined in (2.4); $\epsilon_S$ is the sales shock defined in (2.4); u is the cost shock defined in (2.1); $\lambda_1+\lambda_2$, $-\lambda_1\lambda_2$, $\pi_1$, and $\pi_2$ are the coefficients of the reduced form inventory equation (2.6a).

2. "var" denotes variance, "$\rho$" correlation, "cv" coefficient of variation.

Table II

Parameters to be Estimated, Alternative Normalizations

| Normalization | Parameters to be Estimated | | |
|---|---|---|---|
| (1) HH | $a_0/a_2$ | $a_1/a_2$ | $a_3$ |
| (2) LC | $a_0/c_1$ | $a_1/c_1$ | $a_3$ |
| (3) QC | $a_0/c_2$ | $a_2/c_2$ | $a_3$ |

Notes:

1. The cost function (2.1) includes $.5a_0\Delta Q_t^2 + .5a_1Q_t^2 + .5a_2(H_{t-1}-a_3S_t)^2$.

2. In row (2), $c_1 = [a_0(1+4b+b^2)+a_1(1+b)+ba_2]$. In row (3), $c_2 = (1+b)a_0+a_1$.

3. The corresponding equations in the text are: HH: (4.2); LC: (4.3); QC: (4.4).

Table III

Asymptotic Standard Errors, HH Normalization, T=300

| DGP | $a_0/a_2$ | $a_1/a_2$ | $a_3$ |
|---|---|---|---|
| A | 10.0 | 1.0 | .10 |
| | (4.7) | (5.7) | (.38) |
| B | .16 | -.33 | .50 |
| | (.06) | (.23) | (.02) |
| C | 10.0 | 20.0 | 1.0 |
| | (8.6) | (31.9) | (1.5) |
| D | 10.0 | -5.0 | .50 |
| | (3.8) | (0.8) | (.40) |

Notes:

1. The cost function (2.1) includes $.5a_0\Delta Q_t^2 + .5a_1Q_t^2 + .5a_2(H_{t-1}-a_3S_t)^2$.

2. The parameter values are repeated from Table II, for convenience.

2. In parentheses are the standard errors implied by the asymptotic theory, for a sample of size 300, assuming instrumental variables estimation as described in the text.

Table IV

Implied Ratios of Variance of Q to Variance of S, No Cost Shocks

| | $a_0/a_2$ | $a_1/a_2$ | $a_3$ | var(Q)/var(S) |
|---|---|---|---|---|
| (1) | 10. | 1. | .1 | 1.00 |
| (2) | 10. | -6. | .1 | 1.19 |
| (3) | 10. | -4. | .1 | 1.05 |
| (4) | 10. | 6. | .1 | 0.98 |
| (5) | 10. | 8. | .1 | 0.97 |

Notes:

1. The cost function (2.1) includes $.5a_0\Delta Q_t^2 + .5a_1Q_t^2 + .5a_2(H_{t-1}-a_3S_t)^2$.

2. Q is production; S is sales.

3. The implied variances of Q and S are solved for under the assumption that there are no cost shocks ($u_t=0$), that $a_2=.1$ and that the cost parameters $a_0/a_2$, $a_1/a_2$ and $a_3$ are as indicated. Line (1) presents results for DGP A. Lines (2) through (5) consider values of $a_1/a_2$ that are approximately -1.2, -0.8, 0.8, and 1.2 asymptotic standard errors away from 1.; as indicated in Table III, this standard error is 5.7.

Table V

Distributions of Standardized Parameter Estimates, From Simulations

A. Normalization HH

| DGP | $\widehat{(a_0/a_2)}-(a_0/a_2)$ 50% CI | Median | Trimmed MSE | $\widehat{(a_1/a_2)}-(a_1/a_2)$ 50% CI | Median | Trimmed MSE | $\hat{a}_3-a_3$ 50% CI | Median | Trimmed MSE |
|---|---|---|---|---|---|---|---|---|---|
| A | (-1.8,-0.6) | -1.11 | 1.87 | (-0.7,-0.3) | -0.54 | 0.46 | (-0.8, 0.4) | -0.20 | 1.03 |
| B | (-0.6, 0.7) | 0.04 | 0.88 | (-0.9, 0.3) | -0.42 | 1.05 | (-1.1, 0.7) | -0.26 | 1.63 |
| C | (-1.1,-0.6) | -0.95 | 0.89 | (-0.7,-0.5) | -0.64 | 0.49 | (-0.7,-0.4) | -0.57 | 0.54 |
| D | (-2.0,-0.4) | -1.10 | 2.28 | ( 0.2, 2.2) | 1.14 | 2.62 | (-0.9, 0.7) | -0.02 | 1.30 |

B. Normalization LC

| DGP | $\widehat{(a_0/c_1)}-(a_0/c_1)$ 50% CI | Median | Trimmed MSE | $\widehat{(a_1/c_1)}-(a_1/c_1)$ 50% CI | Median | Trimmed MSE | $\hat{a}_3-a_3$ 50% CI | Median | Trimmed MSE |
|---|---|---|---|---|---|---|---|---|---|
| A | (-0.7, 0.8) | 0.14 | 1.14 | (-0.8, 0.6) | -0.15 | 1.14 | (-1.2, 0.7) | -0.31 | 1.46 |
| B | (-0.6, 0.9) | 0.21 | 1.10 | (-0.9, 0.5) | -0.30 | 1.16 | (-1.1, 0.8) | -0.12 | 1.73 |
| C | (-0.7, 0.8) | 0.10 | 1.13 | (-0.8, 0.7) | -0.11 | 1.14 | (-0.7, 0.2) | -0.33 | 0.93 |
| D | (-0.6, 0.8) | 0.15 | 1.14 | (-0.8, 0.6) | -0.15 | 1.16 | (-1.5, 0.4) | -0.43 | 1.48 |

C. Normalization QC

| DGP | $\widehat{(a_0/c_2)}-(a_0/c_2)$ 50% CI | Median | Trimmed MSE | $\widehat{(a_2/c_2)}-(a_2/c_2)$ 50% CI | Median | Trimmed MSE | $\hat{a}_3-a_3$ 50% CI | Median | Trimmed MSE |
|---|---|---|---|---|---|---|---|---|---|
| A | (-0.8, 0.7) | -0.00 | 1.15 | (-0.7, 0.8) | 0.00 | 1.31 | (-1.3, 0.8) | -0.29 | 1.47 |
| B | ( 0.0, 0.0) | 0.01 | 0.00 | ( 0.0, 0.0) | 0.01 | 0.00 | (-2.0, 0.6) | -0.70 | 2.07 |
| C | (-0.8, 0.7) | -0.08 | 1.12 | (-0.7, 0.9) | 0.15 | 1.40 | (-0.8, 0.3) | -0.31 | 1.03 |
| D | (-0.8, 0.7) | -0.03 | 1.21 | (-0.8, 0.7) | -0.06 | 1.31 | (-1.6, 0.5) | -0.45 | 1.55 |

D. Asymptotic

| 50% CI | Median | Trimmed MSE | 50% CI | Median | Trimmed MSE | 50% CI | Median | Trimmed MSE |
|---|---|---|---|---|---|---|---|---|
| (-0.7, 0.7) | 0.00 | 1.00 | (-0.7, 0.7) | 0.00 | 1.00 | (-0.7, 0.7) | 0.00 | 1.00 |

Notes:

1. The cost function (2.1) includes $.5a_0\Delta Q_t^2 + .5a_1Q_t^2 + .5a_2(H_{t-1}-a_3S_t)^2$. The regression equations are: HH: (4.2); LC: (4.3); QC: (4.4).

2. The differences between estimated and population parameters are standardized by dividing by asymptotic standard errors.

2. The "50% CI" is a 50 percent confidence interval constructed using the 250'th and 750'th largest of the 1000 estimates; "Median" is the 500'th largest such entry; "Trimmed MSE" is a mean squared error computed after dropping observations greater than 3.0 in absolute value, and is expressed relative to the MSE for a standard normal similarly trimmed.

Table VI

Frequency Distribution of Signs of $a_1/()-a_1/()$, From Simulations

A. LG vs. HH

| LC: $(\widehat{a_1/c_1})-(a_1/c_1)$ | | HH: $(\widehat{a_1/a_2})-(a_1/a_2)$ DGP A ≤0 | DGP A >0 | DGP B ≤0 | DGP B >0 | DGP C ≤0 | DGP C >0 | DGP D ≤0 | DGP D >0 |
|---|---|---|---|---|---|---|---|---|---|
| | ≤0 | 0.564 | 0.000 | 0.563 | 0.021 | 0.540 | 0.004 | 0.162 | 0.387 |
| | >0 | 0.299 | 0.137 | 0.081 | 0.335 | 0.418 | 0.038 | 0.055 | 0.396 |

B. QC vs. HH

| QC: $(\widehat{a_1/c_2})-(a_1/c_2)$ | | HH: $(\widehat{a_1/a_2})-(a_1/a_2)$ DGP A ≤0 | DGP A >0 | DGP B ≤0 | DGP B >0 | DGP C ≤0 | DGP C >0 | DGP D ≤0 | DGP D >0 |
|---|---|---|---|---|---|---|---|---|---|
| | ≤0 | 0.500 | 0.000 | 0.644 | 0.356 | 0.465 | 0.004 | 0.159 | 0.330 |
| | >0 | 0.363 | 0.137 | 0.000 | 0.000 | 0.493 | 0.038 | 0.058 | 0.453 |

C. QC vs. LC

| QC: $(\widehat{a_1/c_2})-(a_1/c_2)$ | | LC: $(\widehat{a_1/c_1})-(a_1/c_1)$ DGP A ≤0 | DGP A >0 | DGP B ≤0 | DGP B >0 | DGP C ≤0 | DGP C >0 | DGP D ≤0 | DGP D >0 |
|---|---|---|---|---|---|---|---|---|---|
| | ≤0 | 0.500 | 0.000 | 0.584 | 0.416 | 0.469 | 0.000 | 0.489 | 0.000 |
| | >0 | 0.064 | 0.436 | 0.000 | 0.000 | 0.075 | 0.456 | 0.060 | 0.451 |

Notes:

1. The cost function (2.1) includes $.5a_0\Delta Q_t^2 + .5a_1Q_t^2 + .5a_2(H_{t-1}-a_3S_t)^2$. The regression equations are: HH: (4.2); LC: (4.3); QC: (4.4).

2. Each entry in a given 2x2 matrix gives the fraction of the replications for which indicated sign pattern occurred. For example, the ".564" in the first 2x2 matrix in panel A indicates that in 564 of the 1000 replications, the estimate of $a_1/a_2$ from normalization HH was less than the population value of $a_1/a_2$ <u>and</u> the estimate of $a_1/c_1$ from normalization LC was less than the population value of $a_1/c_1$. In a given 2x2 table, the four entries sum to 1.

3. For QC, the estimate of $a_1/c_2$ was computed as $1-(1+b)\widehat{a_0/c_2}$.

## Table VII

Frequency Distribution of Signs of $a_1/()$, From Simulations

A. LC vs. HH

| LC: $\widehat{a_1/c_1}$ | HH: $\widehat{a_1/a_2}$ DGP A ≤0 | DGP A >0 | DGP B ≤0 | DGP B >0 | DGP C ≤0 | DGP C >0 | DGP D ≤0 | DGP D >0 |
|---|---|---|---|---|---|---|---|---|
| ≤ 0 | 0.480 | 0.000 | <u>0.889</u> | 0.000 | 0.060 | 0.000 | <u>0.967</u> | 0.000 |
| > 0 | 0.326 | <u>0.194</u> | 0.034 | 0.077 | 0.476 | <u>0.464</u> | 0.032 | 0.001 |

B. QC vs. HH

| QC: $\widehat{a_1/c_2}$ | HH: $\widehat{a_1/a_2}$ DGP A ≤0 | DGP A >0 | DGP B ≤0 | DGP B >0 | DGP C ≤0 | DGP C >0 | DGP D ≤0 | DGP D >0 |
|---|---|---|---|---|---|---|---|---|
| ≤ 0 | 0.414 | 0.000 | <u>0.098</u> | 0.000 | 0.047 | 0.000 | <u>0.944</u> | 0.000 |
| > 0 | 0.392 | <u>0.194</u> | 0.825 | 0.077 | 0.489 | <u>0.464</u> | 0.055 | 0.001 |

C. QC vs. LC

| QC: $\widehat{a_1/c_2}$ | LC: $\widehat{a_1/c_1}$ DGP A ≤0 | DGP A >0 | DGP B ≤0 | DGP B >0 | DGP C ≤0 | DGP C >0 | DGP D ≤0 | DGP D >0 |
|---|---|---|---|---|---|---|---|---|
| ≤ 0 | 0.414 | 0.000 | <u>0.098</u> | 0.000 | 0.047 | 0.000 | <u>0.944</u> | 0.000 |
| > 0 | 0.066 | <u>0.520</u> | 0.791 | 0.111 | 0.013 | <u>0.940</u> | 0.023 | 0.033 |

Notes:

1. See notes to Table VI. This table differs from that table only in that it considers the sign of the estimated parameters, rather than the sign of the difference between the estimated and actual.

2. In each 2x2 matrix, the entry that is underlined is the one that would be 1.00 if both normalizations happened to yield the correct sign in all 1000 simulations.

3. Population values (asymptotic standard error) for $a_1/a_2$: DGP A: 1.0 (5.7); DGP B: -.33 (.23); DGP C: 20.0 (31.9); DGP D: -5.0 (0.8).