

ABSTRACT

DOORN, DAVID JOHN. Three Essays on Trend Analysis and Misspecification in Structural Econometric Models. (Under the direction of Alastair Hall)

The purpose of this research has been to look into several econometric issues of concern to researchers doing applied work in macroeconomics. The first essay looks at Bureau of Economic Analysis data on inventories and sales of finished goods often used in studies of inventory behavior. Applying recently developed methods, the series are rigorously tested to determine their stationarity properties. Results indicate that neither first differencing nor linearly detrending the data is appropriate. For most series a trend function with one or more breaks offers a better fit and also generates stationarity. The results are used to determine the impact on estimation in a simple production-smoothing model of inventory behavior. The impact of different trend specifications on univariate forecasting of inventories is also considered.

The second essay considers an alternative method of detrending time series data – the Hodrick-Prescott (HP) filter. Previous research has shown that HP filtering can have serious adverse consequences when used to analyze co-movements between data series at business cycle frequencies. Despite this, the filter has also been used to induce stationarity in a data series prior to estimation of structural econometric models. Little work has been done in analyzing the possible effects this may have on parameter estimates from such models. A simulation study is conducted to assess the impact of HP filtering on parameter estimation and a comparison is made to other detrending methods. It is shown that the HP filter induces bias in the parameter estimates and also increases the root mean squared error of the estimates from the simulations. In addition, there is some adverse impact on the size of certain test statistics.

The final essay looks at the impact of misspecification on estimation results from a structural econometric model when using a Generalized Method of Moments estimator. Simulated data consistent with a particular specification of the model is used to estimate two misspecified versions. It is shown that misspecification causes the probability limit of the estimator to differ from the true value. It is further shown that a popular specification test performs poorly in detecting the misspecification. An alternative method of model selection is shown to perform far better. Finally, because the use of conventional asymptotic theory is not appropriate in misspecified models, a recently proposed alternative asymptotic theory is tested to determine whether there is improvement in the ability to perform inference on the parameters from misspecified models.

**THREE ESSAYS ON
TREND ANALYSIS AND MISSPECIFICATION
IN STRUCTURAL ECONOMETRIC MODELS**

by
DAVID JOHN DOORN

A dissertation submitted to the graduate faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

ECONOMICS

Raleigh

2003

APPROVED BY:

Chair of Advisory Committee

Biography

David John Doorn was born January 11, 1965, to Keith and Carolyn Doorn in Cudahy, Wisconsin. He graduated from Waterford Union High School, Waterford, Wisconsin, in 1983. After many years as a machinist, David received his Bachelor of Arts in Economics from the University of Wisconsin at Milwaukee in 1995 and entered the Economics graduate program at North Carolina State University that same year. After teaching at the University of North Carolina at Greensboro from January 2000 to May 2002, David began work for the U.S. Department of Labor's Bureau of Labor Statistics in Kansas City, Missouri, in July 2002. David now lives in Kansas City with his wife, Kristine, and two sons, Alex and Devin.

Contents

List of Tables	v
List of Figures	vii
1 Trend Breaks in Finished Goods Inventories and Sales of Non-Durables	1
1.1 Introduction	1
1.2 The Data	3
1.3 Unit Root Testing	4
1.3.1 Augmented Dickey-Fuller Tests	6
1.3.2 Point-Optimal Tests	6
1.3.3 Tests Allowing For Breaks In The Trend Function	8
1.4 Choice of Detrending Models	10
1.5 Impact on Parameter Estimates from a Linear-Quadratic Model of Inventory Behavior	15
1.5.1 The L-Q Model	15
1.5.2 Impact on Parameter Estimates	18
1.5.3 Other Implications	19
1.6 Forecasting	20
1.7 Summary	23

2	Consequences of Hodrick-Prescott Filtering for Parameter Estimation in a Structural Model of Inventory Behavior	40
2.1	Introduction	40
2.2	The Hodrick-Prescott Filter	42
2.3	The Model	48
2.4	Estimation Technique	56
2.5	Empirical Example	60
2.6	Simulation Results	61
2.7	Conclusions	63
3	Misspecification Issues in Structural Econometric Models	79
3.1	Introduction	79
3.2	GMM Background and Definitions	82
3.3	The Model	84
3.3.1	The Full Model	84
3.3.2	Alternative Specifications	86
3.4	A Note on Moment Restrictions	89
3.5	Simulating the Data	92
3.6	Probability Limits of the Parameter Estimates	99
3.7	Specification Testing	101
3.7.1	Hansen's J-Test	102
3.7.2	Noise Ratios	104
3.8	Impact of HI Asymptotics	107
3.8.1	Estimation Results	109
3.9	Application	120
3.10	Summary and Conclusions	124
	Bibliography	141

List of Tables

1.1	Summary Statistics	24
1.2	Adjustment Factors	24
1.3	Prior Research	25
1.4	ADF Test: Optimal Lags and Results	25
1.5	Elliot et al. Critical Values	26
1.6	Test Results	26
1.7	Zivot and Andrews' Critical Values	26
1.8	Lumsdaine and Papell Critical Values	27
1.9	Structural Breakpoints and Test Results for Single-Break Models	27
1.10	Breakpoints and Test Results for Two Breaks in Trend	28
1.11	Selected Models For Detrending	28
1.12	Estimation Results From Deterministic Detrending Procedures	29
1.13	Choice of p from SIC	29
2.1	Cost Parameters for Simulation Study	66
2.2	Specification of Sales Process and Cost Shock Parameters	66
2.3	Implied Parameters of Inventory Process	66
2.4	Implied Values of Regression Coefficients	66
2.5	Actual data results	67
2.6	RMSE and Distribution Characteristics of Parameter Estimates	68
2.7	Simulation Results	69

3.1	Examples of Papers Performing Inference Based on Misspecified Models	127
3.2	Calibrated Parameter Values By Industry	128
3.3	Implied Population Values of Model Parameters	129
3.4	Large Sample Probability Limits By Specification & Weight Matrix	130
3.5	Rejection Rates of \mathbf{J}_T -test (%), Average Noise Ratios ($\bar{\mathbf{N}}\mathbf{R}$), and Noise Ratio Model Selection by Specification and Weight Matrix.	131
3.6	Limiting Distributions for Alternative GMM Estimators and the J_T Test Statistic	132
3.7	Estimation Results for the Full Model	133
3.8	Estimation Results for the Pure Model	135
3.9	Estimation Results for the Stockout Avoidance Model	136
3.10	t-statistic Rejection Rates: Pure Model	138
3.11	t-statistic Rejection Rates: Stockout Avoidance Model	139
3.12	Estimation with real data	140

List of Figures

1.1	Time Series of BEA Inventory and Sales Data	30
1.2	Time Series of BEA Inventory and Sales Data	31
1.3	Sequential t-tests: Selected Models — Inventories	32
1.4	Sequential t-tests: Selected Models — Sales	33
1.5	Detrending Models and Residuals	34
1.6	Detrending Models and Residuals	35
1.7	Detrending Models and Residuals	36
1.8	Forecasts	37
1.9	Forecasts	38
1.10	Forecasts	39
2.1	Frequency Response of HP Cyclical Filter	70
2.2	Spectral Density of Total Aggregate Inventories of Non-Durables	71
2.3	HP and Linear Trends	72
2.4	Residuals	73
2.5	ACF	74
2.6	Density of β_1 Estimates.	75
2.7	Density of β_2 Estimates.	76
2.8	Density of β_3 Estimates.	77
2.9	Density of β_4 Estimates.	78

Chapter 1

Trend Breaks in Finished Goods

Inventories and Sales of

Non-Durables

1.1 Introduction

Much research in recent years has focused on the role of inventories in explaining business cycles. While inventory investment is generally a very small component of GDP, it has been noted in the literature that most recessions tend to be periods of high inventory dis-investment, with changes in inventories accounting for 87% of the fall in GDP, on average, for postwar recessions in the U.S. (see, for example, Blinder and Maccini (1991) or Ramey and West (1997)). This indicates the importance of looking at inventory behavior when considering the possible causes of business cycles as well as their persistence. Unfortunately, most models of inventory behavior perform poorly with respect to the empirical facts and there has not yet been a widely accepted solution to this problem.

One possible cause of the poor performance of inventory models in empirical tests may be the inappropriate treatment of the data used to estimate it. Most estimation procedures require the data to be

stationary for the asymptotic econometric theory to hold. The widely applied Generalized Method of Moments (GMM) in particular requires stationarity of the variables contained in the moment condition on which estimation is based. Prior research has tended to choose data transformations that are based upon little or no formal testing prior to estimation, often choosing a linear or quadratic de-trending procedure. This study takes advantage of recent advances in the unit root literature and applies a number of previously unavailable testing procedures to data on inventory and sales for six industries and aggregate non-durables. This allows for a more informed determination of the stationarity properties of the data prior to estimation, rather than just choosing a convenient transformation based upon a small amount of evidence or common practice.

Initially, the commonly used Augmented Dickey-Fuller test for unit roots is applied to the data and the results presented. Then a more powerful class of tests is applied in an attempt to get stronger results. Lack of convincing evidence from these tests indicates the need for further testing, which initially includes the presence of a single break in the trend function for the stationary alternative in the unit root tests. These tests are then expanded to include the possibility of a second break in trend for these series. Most of the series tested are found to contain either one or two breaks in trend and this result is then applied in detrending the data. The resulting transformations do result in stationary series.

Given these results, a comparison is made to see if there is a significant impact on estimation results from a commonly used inventory model. The model is estimated using three different detrending procedures to attain stationarity in the data. These include detrending based on the presence of a linear trend, a quadratic trend, and a piecewise linear trend. Surprisingly, there is not a lot of difference in results obtained from the three deterministic detrending methods. We also find that first-differencing does give quite different results, but these tend to be implausible given both the underlying theory and the empirical facts.

In addition, the different detrending procedures imply quite different time paths for the series. A look at the implications of these results for forecasting of inventories is also undertaken. Not surprisingly, the time paths do diverge rather dramatically in many instances, implying quite different levels of inventories as the forecast horizon lengthens. This indicates the importance of being careful when

choosing a forecasting model and also illustrates the impact that the use of unit root tests may have on this selection.

The paper will proceed as follows: Section 1 describes the data series used in the analysis and includes an update of the adjustment suggested in West (1983) to put the inventory and sales series on a comparable measurement basis. Section 2 describes the unit root tests applied to the series and presents the results. Section 3 discusses the choice of detrending model for each series based on the tests in Section 2. Section 4 describes a simple model of inventory behavior and compares estimates using different detrending procedures. Section 5 discusses the implications of the test results from section 2 for forecasting of inventories. Section 6 concludes with a summary of the results.

1.2 The Data

The data series used in this study is one that is commonly employed in the Inventory literature. This is the Bureau of Economic Analysis (BEA) data on end-of-month finished goods inventories and sales. The BEA series are gleaned from the Census Bureau's *Manufacturer's Shipments, Inventories, and Orders* (M3) survey and are used in estimation of the National Income and Product Accounts (NIPA) for the calculation of GDP. The data are adjusted to constant chained 1992 dollars from the book values reported by firms in the raw M3 data. The procedures used are described in the *Survey of Current Business*, published monthly by the BEA. The data employed here are consistent with the procedures described in the August 1998 issue of the Survey. The monthly series are also seasonally adjusted, using the X-12 method, and cover the periods January 1959 through May 1998. This yields 473 observations.

To be consistent with prior research using the linear-quadratic model of inventory behavior, we concentrate on the six two-digit SIC industries commonly considered as being production to stock. These include Food, Tobacco, Apparel, Chemicals, Petroleum, and Rubber. Also considered is the series for overall non-durables manufacturing, which adds Textiles, Paper, Printing, and Leather to the above. Plots of these series are given in Figures 1.1 and 1.2 and summary statistics in Table 1.1.

West (1983) points out that in the BEA data sales and inventories are not valued in the same manner, with sales being reported by manufacturers at market value while inventories are generally valued at

cost. To make the data consistent across series it is necessary to adjust either inventories or sales by the ratio of market price to unit cost, evaluated in the base year. West suggests an approximation to this ratio that can be constructed from IRS data. This approximation takes the form: (business receipts) \div (cost of sales and operations + rent + repairs + depreciation + taxes). Table 1.2 gives the resulting adjustment factor using a 1992 base for each of the six two-digit industries and for total non-durables. In incorporating the adjustment the sales series were divided by the adjustment factor. The adjusted sales series should then be used in estimation. For much of the single series descriptive analysis that follows, however, this adjustment is inconsequential and so the unadjusted series are used.

A precursory look at the time series plots of the data suggests the possibility of a deterministic trend for all of the series, except perhaps for Tobacco sales and Tobacco inventories. It is of course also possible that the data follow a unit root with drift process. Most of the literature using this data tends to assume some kind of polynomial time trend and so de-trend the data using standard methods to attain stationarity. Very few of the studies actually discuss any unit root testing of the data to determine its stationarity properties, and those that do tend to report somewhat ambiguous results. Nearly all of the studies end up using deterministically detrended data, with some also repeating their work with differenced data, e.g. Eichenbaum (1989); Durlauf and Maccini (1995). Table 1.3 lists some of the prior research on inventories and what, if any, testing was reported on the data used, as well as the transformations applied prior to its use. Because there seems to be no compelling evidence in the literature, this study rigourously investigates the stationarity properties of the data using recently developed tests from the unit root literature.

1.3 Unit Root Testing

In analyzing the data several different tests for the presence of unit roots were applied. These include the standard Augmented Dickey-Fuller test, the point optimal tests of Elliot et al. (1996), and extensions of the ADF test that include the possibility of breaks in the trend function.¹ All of these tests are based

¹Phillips-Perron tests were also performed on the data. For the most part these tests gave the same results as the ADF tests. Given the potential size problems of the Phillips-Perron tests noted in the econometric literature(see Schwert (1989)), the actual results are not reported here.

on some form of equation 1.1 with $d(t)$ indicating a deterministic function of time. The specification of $d(t)$ is the primary difference between the tests applied.

$$y_t = \hat{\alpha}y_{t-1} + d(t) + \sum_{j=1}^k \hat{c}_j \Delta y_{t-j} + \hat{e}_t \quad (1.1)$$

The k autoregressive Δy_{t-j} terms are included to allow for the possibility of auto-correlation in the disturbances, which may cause problems in the limiting distributions of the test statistics. It has been noted in the literature on unit-roots that the number of extra regressors, k , included in Equation 1.1 can impact the power and the size of unit root tests. Hall (1994) shows that using pretest data-based model selection procedures to choose k can greatly increase the power of the ADF test. We employ two of these procedures here to choose the number of extra regressors to include for each data series in each test, the Akaike information criteria (AIC) and the Schwartz information criteria (SIC).²

Both the AIC and the SIC are designed to choose the k that minimizes a criterion function over all possible lag lengths considered, with the difference being in the specified criterion function. The criterion to be minimized takes the form:

$$T \ln(T^{-1} SSR_k) + CR, \quad (1.2)$$

where SSR_k is the sum of squared residuals from 1.3 and CR is $2(k+2)$ for AIC and $(k+2)\ln(T)$ for SIC. Under certain conditions it can be shown that the probability that \hat{k}_{SIC} is the true k is one in the limit, while there is a positive probability that \hat{k}_{AIC} is greater than the true k .³ This means that it is possible for AIC to overfit the model. In finite samples the nature of the penalty functions, CR, requires that $\hat{k}_{SIC} \leq \hat{k}_{AIC}$ once T attains a certain size. Of the following testing procedures, the three ADF tests make use of both AIC and SIC, while the rest only apply SIC to determine the number of lags to include. Unless otherwise noted, the maximum lag-length allowed was 25 for all series.

²We also used the Hannan and Quinn procedure reported in Hall (1994), but found the results to be virtually identical to either the SIC or AIC for every series and so chose not to report it.

³See Schwartz (1978) and Shibata (1976) for proofs of this.

1.3.1 Augmented Dickey-Fuller Tests

The first tests undertaken were standard augmented Dickey-Fuller (ADF) tests. Three versions of these tests were applied to the data, with the first being a test of the unit root with drift hypothesis against a single-mean stationarity alternative and the second being the same null against stationarity about a linear trend. In addition, some of the authors in Table 1.3 specified a quadratic trend function when detrending the data, so a third test of the unit root null against quadratic-trend stationarity was also applied. The trend function for these tests is of the form

$$d(t) = \mu + \beta t + \gamma t^2, \quad (1.3)$$

with β and γ restricted to zero for tests against the single-mean stationarity alternative and γ restricted to zero for the linear-trend specification.

The null hypothesis based on equation 1.1 is $H_0 : \alpha = 1$ and a t-statistic is constructed from the regression to test this assertion. This t-statistic is then compared to the tables formulated in Dickey and Fuller (1979, 1981) for the single-mean and linear-trend alternatives and H_0 is rejected for small values of t. For the quadratic-trend alternative p -values were calculated using the response-surface methods outlined in MacKinnon (1994). The number of k extra regressors chosen and the results of these tests are given in Table 1.4.

1.3.2 Point-Optimal Tests

In addition to the ADF tests, two tests proposed by Elliot et al. (1996) were used, both of which were shown to have better size and power properties than the above tests and several others that have been suggested in the literature. The first of these is a modified version of the ADF test used above, which Elliot et al. term the DF-GLS test, since it uses a type of generalized least-squares procedure in generating the test statistic. The second test is a simple scaled likelihood-ratio constructed by the authors which can be shown to encompass a number of other test statistics, including the ADF t -statistics used above, under certain assumptions. This they call the P_T test. Both of these are tests of the unit root hypothesis against a point alternative, that is, $H_0 : \alpha = 1$ versus $H_A : \alpha = \bar{\alpha}$, where α is that in

Equation 1.1. $\bar{\alpha}$ is chosen to be the alternative for which maximal power of the tests are approximately .5, which Elliot et al. show will cause the statistic to lie close to the Gaussian power envelope over a large range. They suggest using $\bar{\alpha} = 1 + \frac{\bar{c}}{T}$, where $\bar{c} = -7$ for tests against a constant mean and $\bar{c} = -13.5$ for tests against a linear trend specification. The tests used here are only those against the alternative of linear trend stationarity.

The DF-GLS test uses a procedure to demean and detrend the series of interest, y_t , to obtain y_t^D , which is then used in place of y_t in equation 1.1. In this case $d(t)$ becomes zero, so the test using the transformed series is against zero-mean stationarity.

The detrending procedure involves a number of steps. To begin with, the data series is transformed to obtain $y_{\bar{\alpha}} = (y_1, y_2 - \bar{\alpha}y_1, \dots, y_T - \bar{\alpha}y_{T-1})'$. Then define the trend specification in a $T \times q$ matrix, Z , where $z_t = (1)$ for the constant mean case and $z_t = (1, t)$ for the linear trend case. Transform Z to obtain $Z_{\bar{\alpha}} = (z_1, z_2 - \bar{\alpha}z_1, \dots, z_T - \bar{\alpha}z_{T-1})'$. We then run a regression of $y_{\bar{\alpha}}$ on $Z_{\bar{\alpha}}$ to obtain an estimate for the coefficient vector on the trend specification, $\hat{\beta}$. This estimate is then used to construct $y_t^D = y_t - \hat{\beta}'z_t$, the detrended and demeaned series used in the test.

The P_T -statistic is constructed as

$$P_T = \frac{[S(\bar{\alpha}) - \bar{\alpha}S(1)]}{\hat{\omega}^2} \quad (1.4)$$

where $S(x)$ is the sum of squared residuals from the regression of y_x on Z_x , with $y_x = (y_1, y_2 - xy_1, \dots, y_T - xy_{T-1})'$ and $Z_x = (z_1, z_2 - xz_1, \dots, z_T - xz_{T-1})'$ and the z_t are as above. The denominator, $\hat{\omega}^2$, is an estimator of $\omega^2 = \sum_{k=-\infty}^{\infty} \gamma(k)$, where $\gamma(k) = Ev_t v_{t-k}$ are the autocovariances of the error. The estimator used here is the autoregressive estimator described in Elliot et al.. This is

$$\hat{\omega}_{AR}^2 = \frac{\hat{\sigma}_{\eta}^2}{(1 - \sum_{i=1}^p \hat{a}_i)^2} \quad (1.5)$$

with $\hat{\sigma}_{\eta}^2$ and \hat{a}_i being the OLS estimates from the regression

$$\Delta y_t = a_0 y_{t-1} + \sum_{i=1}^p a_i \Delta y_{t-i} + a_{p+1} + a_{p+2}t + \eta_t \quad (1.6)$$

The choice of lag length was again made using the Schwartz information criterion mentioned above and are the same as were found in the initial ADF tests above.

These test statistics do not follow any of the usual tables of critical values, so Elliot et al. conduct Monte Carlo simulations to generate appropriate critical values. The critical values for each of these tests against the alternative of linear trend stationarity are reproduced in Table 1.5. The results for the inventory and sales series are given in Table 1.6, using the asymptotic critical values from Elliot et al..

1.3.3 Tests Allowing For Breaks In The Trend Function

One conclusion from the literature on unit roots has been that a mis-specification of the trend function, $d(t)$, can result in the appearance of a unit root when in fact one does not exist. Because of this we also conduct tests that allow for the possibility of shocks to the series at certain dates. These shocks, if they exist, may cause a break, or a shift, in the trend of an otherwise trend-stationary series. One way of testing this in the context of unit roots in general was derived by Perron (1989), in which several macroeconomic data series were tested. Perron performed a visual inspection of the series and concluded that there were two possible trend breaks, the Great Depression and the 1973 oil crisis. The series were then tested for unit roots using a specification that allowed for the possibility of a break in trend at one of these points.

For many of the inventory and sales series in the current study a visual inspection of the data is an unsatisfactory means of determining the possible existence of a structural breakpoint, and so other ways of making this determination are desirable. Zivot and Andrews (1992) and Banerjee et al. (1992) also take issue with Perron's method of determining breakpoint and devise procedures to choose the most likely date for a break in trend for each series. Their procedures assume that the structural change at these points is endogenous, and so data dependent, rather than exogenous. This leads to tests in which the null hypothesis of a unit root with drift, which implies there is no structural change, is set against the alternative of stationarity about a deterministic trend, with possible changes in the trend function at some point in time. This null can be formulated as

$$y_t = \mu + y_{t-1} + e_t$$

Support for the alternative can thus be gained by estimating breakpoints which give the greatest evidence for rejecting the null.

While these authors only consider the possibility of a single break in trend, Ben-David et al. (1996) and Lumsdaine and Papell (1997) derive a general extension to include two breaks in the trend function. The general form for the trend function in all of these models becomes:

$$d(t) = \hat{\mu} + \hat{\beta}t + \hat{\theta}DU1_t(\hat{\lambda}) + \hat{\gamma}DT1_t(\hat{\lambda}) + \hat{\omega}DU2_t(\hat{\lambda}) + \hat{\psi}DT2_t(\hat{\lambda}) \quad (1.7)$$

where $\lambda_i = T_{Bi}/T$, with T_{Bi} being the date at which the break occurs for $i = 1, 2$; $DUi_t(\lambda) = 1$ if $t > T\lambda$, 0 otherwise; $DTi_t(\lambda) = t - T\lambda$ if $t > T\lambda$, 0 otherwise.

We conduct tests based on equation 1.1 with six variants of this trend function, three that allow for the presence of a single break, and three that allow for two breaks. Model A imposes the restriction $\hat{\gamma} = \hat{\omega} = \hat{\psi} = 0$, and so only allows for the possibility of a one-time shift in the overall level of the series, or its mean. Model B includes $d(t)$ with $\hat{\theta} = \hat{\omega} = \hat{\psi} = 0$, which only allows for a one-time change in the slope of the deterministic trend. Model C includes $d(t)$ with $\hat{\omega} = \hat{\psi} = 0$, which allows for a change in both mean and level at a single date. These three models are identical to those considered in Zivot and Andrews (1992). Banerjee et al. also consider models A and B,⁴ but in a more complex notation, and attain similar results to those of Zivot and Andrews.

In addition we consider the two-break models of Ben-David et al. (1996) and Lumsdaine and Papell (1997). Model AA includes $d(t)$ with $\hat{\gamma} = \hat{\psi} = 0$, and so allows for two shifts in the mean of the series. Model CA imposes only $\hat{\psi} = 0$, and so allows for one break in which the mean shifts and another in which both the mean and the slope of the series change. Finally, Model CC makes no restrictions on $d(t)$ in equation 1.7, which allows for both a shift in mean and a change in slope at two possible breakpoints.

The testing strategy is to construct a t -statistic testing $\alpha = 1$ at each possible trend breakpoint, or combination of breakpoints, and then choosing those breakpoints that produce the minimum t -statistic as being the ones of interest.⁵ This results in choosing breakpoints which give the most weight to the alternative of trend stationarity. In practice this requires estimation of each model at each possible breakpoint, or set of breakpoints, using ordinary least squares. This process includes determining the number of k extra regressors to include for each possible breakpoint or set of breakpoints. In doing

⁴Their Model II, Cases A and B, page 275.

⁵For technical reasons described in Zivot and Andrews (1992) the possible breakpoints tested for here are $t = 2$ through $T - 1$.

this, the SIC procedure described above was employed for each possible λ , with the maximum possible lag length of 25 for the single-break models and, due to the extra computational burden, 12 for the two-break versions.⁶

As in the ADF testing above, the t-statistics here will not have the same distribution as the standard t-statistic, due to the inclusion of lagged first-differences of the regressor. Nor are they comparable to the ADF tables. For the single-break tests, Zivot and Andrews have compiled the appropriate tables for each of the three models mentioned above using Monte Carlo methods, both for the asymptotic distribution and for the finite-sample distribution of the test statistic. Due to the relatively large sample size employed here, we consider the asymptotic results to be the more appropriate tables to use for comparison. The findings of Banerjee et al. with regard to sample size indicate there is little variation in the critical values for sample sizes greater than 250, again lending support to the use of the asymptotic critical values. The critical values generated by Zivot and Andrews are reproduced in Table 1.7.⁷ Similarly, the appropriate critical values for the two-break tests have been calculated by Lumsdaine and Papell (1997) and are given in Table 1.8.

Table 1.9 gives the minimum t-statistic for each series and for each model in the single breakpoint case, also indicating the date at which the corresponding structural break is determined to have taken place. Similar results for the two-break tests are reported in Table 1.10. The significance of the test for each series and each model is also given. For the single-break tests it will also be useful to look at selected results graphically. Figures 1.3 and 1.4 plot a few of the time series along with the t-statistic generated using each date as a breakpoint. To save space only selected plots are given. The 1%, 5%, and 10% asymptotic critical values are indicated as well in each diagram.

1.4 Choice of Detrending Models

We now have plenty of test results with which to determine the fate of our data. By looking at each series separately we can combine the results from the standard unit root tests with those found using

⁶It should be noted that Zivot and Andrews use a different procedure to determine k . However, we chose the SIC method in order to be consistent with the methodology used in the ADF testing. Since the number of extra regressors should not affect the limiting distribution of the test statistic, this should not affect the critical values to be used in testing. (see Banerjee et al. (1992))

⁷Banerjee et al. produce asymptotic critical values for Models A and B that differ only slightly from those reported here.

other methods from the literature, along with visual inspection of the time series, to decide upon the most likely trend specification. In doing this we are ignoring the possible implications of the results for any type of modeling that might be done using these data. These results are solely the product of univariate testing procedures to determine the time series properties of each series and to aid in choosing the best transformations to achieve stationarity.

To begin with, three of the fourteen series would appear to be linear trend stationary. Petroleum inventories, Food sales, and Apparel sales all provided strong evidence against a unit root in the initial ADF, as well as in the more powerful DF-GLS and P_T testing procedures. Tests allowing for a break in trend produced evidence against unit roots at nearly every possible breakpoint for these three series. This can be seen in the plot of Model C for these series.⁸ Additional ADF tests for stationarity on either side of the most likely breakpoints also seem to lend support to this finding. This points to a simple linear detrending procedure as the best transformation for these series.

For Tobacco inventories, while there was evidence of linear trend stationarity in the initial tests, the evidence was somewhat weaker using the DF-GLS and the P_T tests. A visual inspection of the series also gives a clear indication of a break in trend at one or more points. Including the possibility of a single breakpoint resulted in Model A giving the most plausible result, with a break in trend in December of 1993. Allowing for a second break in trend at some point also seemed to give plausible results, with December 1993 again showing up as a possible breakpoint, along with various earlier possibilities. In deciding between the use of a single break model or a two break model in detrending the data it would seem most appropriate to choose the model which requires the least amount of changes to the original series and still attain stationarity. Even though the two break results are somewhat visually appealing, this criterion would lead us to choose the single break model for this series. Tests on the resulting subsamples indicate linear-trend stationarity on the first period, but cannot reject a unit root in the later subsample. This may be due to the short length of this subsample.

Apparel inventories, Chemical inventories, and Petroleum sales all presented little or no evidence against a unit root in the initial tests against linear trend stationarity. The single break tests, however,

⁸Models A and B produced similar results, with t being significant at every breakpoint.

did result in rejections of the null for all three series, with Model B proving the most plausible approach for the two inventory series and Model C for Petroleum sales. Tests allowing for two breaks in the trend function found no evidence against unit roots for Apparel inventories or Chemical inventories, but did for the Petroleum sales series. This leads to the choice of single break models for the two inventory series, with a break in trend in September 1966 for Apparel and March 1969 for Chemicals. While for the Apparel series this choice is well supported by tests of subsamples on either side of the break, the same is not true for the Chemical series. It remains the only choice, however, given the other test results. For Petroleum sales we do have significant two-break tests for each model, but again we have the case where a less intrusive procedure will give the desired result, and so Model C with a single break at February of 1980 is chosen for detrending this series. This choice also is well supported by tests on either side of the break.

There was no evidence against the presence of a unit root for any of the initial tests or for the single break tests on the Rubber inventory series. Only Model AA of the two break tests provided any significant evidence for detrending this series. This leads us to choose this model with breakpoints at November 1967 and October 1979 for transforming the data. Tests of the subsamples on either side of and in between these breakpoints found significant evidence against a unit root only for the subsample after the last breakpoint, and then only at a 10% significance level.

Total sales also presented evidence against a unit root only for one of the double trend break tests. This was model CC, which gave break dates of August 1974 and August 1979. Tests on the subsamples surrounding these dates provided evidence against unit roots for all three at the 5% level.

Tobacco sales and Chemical sales each found little support for a linear trend specification using the first three tests. The single break models also found only weak evidence against the null, with tests on either side of the breakpoints adding little support. Model CC of the two-break tests found strong support at the 1% level for rejection of a unit root for both of these series, however. The indicated breakpoints are October 1982 and January 1993 for the Tobacco series and August 1974 and November 1980 for Chemicals. Note that these breakpoints correspond with those found in the single break tests. Unit root tests about these breakpoints lend some support to these specifications, but not very conclusively. For

Tobacco, both the subsample up to the first break and that after the second found support for linear trend stationarity at the 10% level. The middle subsample could not reject the presence of a unit root. For Chemicals, only the subsample after the second breakpoint was significant in its rejection of a unit root, and this at the 5% level. Despite this the double break specification would seem to be the most appropriate to use in transforming the series.

While there seems to be some support for specifying a linear trend with no breaks for the Rubber sales series, the evidence isn't as strong as it was for the three series above for which we decided upon this specification. In tests for a single break in trend for this series there was strong evidence against a unit root for both Models C and A, with both giving the same break date of April 1979. Tests on either side of this date found only weak evidence of trend stationarity for the subsample prior to this date. Testing for two breaks in trend produced the same break date as above and also the earlier date of November 1974. All tests on the subsamples about the two breaks were significant in their rejections of unit roots at the 5% level. This lends strong support for using model CC for detrending.

This leaves two series for which we have yet to make a determination. Both the Total Inventories and Food inventories series found no evidence for rejecting the unit root in any of the above testing situations, aside from the P_T test on Total Inventories. It is likely that the reason for this is that we haven't found a trend specification which adequately fits the data, since this may lead to the appearance of a unit root process. It is possible that allowing for more breaks in the trend may solve this problem. Unfortunately the computational burden of extending the two-break test above to include three or more possible breaks is prohibitive. In addition, new critical values would need to be calculated. Applying the single-break tests to subsamples on either side of the breakpoints chosen initially also did not result in any significant rejections of the null. Arguably, the best choice to make in this case would be a model that is most likely to be closest to the true trend specification, so we should include as many breaks as possible. Also, since tests on either side of the breakpoints given by the single break models were insignificant, it would seem that a two-break specification is more appropriate. For the Food series, a visual inspection indicates that including a break in the late 1980s is plausible. Model CC gives a break at April 1989, along with May of 1968. Tests on the subsamples resulting from these dates give evidence

at the 5% level against a unit root for the fairly long sample in between. The other two subsamples indicated no rejection of a unit root process. For Total inventories we had two models give the same break dates of September 1966 and April 1989. Tests of the resulting subsamples again found rejection of the null at the 5% level for the middle sample while no evidence for rejection in the others. This would seem to be as good as we can get for these two series.

The models chosen for transformation of the individual series to attain stationarity are summarized in Table 1.11. Graphs of the detrending models and the resulting stationary residuals are presented in Figures 1.5 through 1.7.⁹ ADF tests on the detrended series found all of them to reject the presence of a unit root at a 5% confidence level against a zero mean stationarity alternative.

While prior researchers have generally chosen to apply a linear detrending procedure to this data, they seemed to base this decision on little evidence. The results shown here indicate that, while detrending may be the correct thing to do, simply removing a linear trend from all of the series is inappropriate. Even though inventory and sales behavior may be similar across industries, there is no reason to expect that shocks to these series will happen in all of the industries at the same time or that their trends will be affected in the same way.¹⁰ It is likely that, while things that affect the economy as a whole may affect these industries in similar ways, idiosyncratic shocks may take place within each industry that cause the paths taken in their inventory and sales behavior to be divergent. This indicates that each industry should have its own trend specification.

One of the main results of the unit root literature has been in pointing out the importance of correct specification of trend functions. Lack of this results in the appearance of unit roots and the problems in inference that go along with them. Here we have attempted to discover trend specifications that at least are likely to result in stationary series after their use in detrending. There is fairly strong evidence that the trend specifications chosen here do this. One potential problem with these results is that for many of the industries we found breaks in trend for the inventory series that do not seem to be related in any way to breaks in the sales series. The papers in which the trend break methods were introduced applied

⁹These graphs were generated using the *EasyReg* program of Herman Bierens.

¹⁰Similar issues involving simultaneous structural changes between related series have come up in the common factors literature. While we don't address those issues here, our results may have important implications for this and other areas of research.

them only to single data series, with no mention of the implications when those series are then used in combination to estimate some structural model of the economy. The implications of this for models of inventory behavior are not entirely clear.

1.5 Impact on Parameter Estimates from a Linear-Quadratic Model of Inventory Behavior

1.5.1 The L-Q Model

One model of inventory behavior that has been prevalent in the literature is the linear-quadratic (L-Q) model, so called because it assumes that demand is linear and both production costs and inventory costs are quadratic, indicating rising marginal costs. Generally, the model is expressed in such a manner that it includes several possible motivations for and characteristics of inventories. These may include production smoothing, stock-out avoidance, costs of adjusting production, autocorrelated cost shocks, exogenous or endogenous sales, and demand shocks.¹¹ While the theoretical constructs of the model seem plausible, empirical work has by and large rejected the L-Q model in its different forms. In particular, the many studies using this model have come up with a wide range of parameter estimates that are often at odds with each other and difficult to defend given the empirical facts. The most contentious result has been that the estimated speed of adjustment between actual inventories and desired levels tends to be much too slow given industry production capacity.

A simple form of the linear-quadratic inventory model is based on a representative firm choosing the level of inventories that will minimize the present value of costs, including both costs of production and of holding inventories. This can be represented by the following equations:

$$\min_{I_{t+j}} \sum_{j=0}^{\infty} \beta^j E_t [C_Q(Q_{t+j}) + C_I(I_{t+j}, S_{t+j})] \quad (1.8)$$

s.t.

$$I_{t+j} = I_{t+j-1} + Q_{t+j} - S_{t+j} \quad (1.9)$$

¹¹See Ramey and West (1997) for a good survey.

$$C_Q(Q_{t+j}) = \left(\frac{\delta}{2}\right)Q_{t+j}^2 \quad (1.10)$$

$$C_I(I_{t+j}, S_{t+j}) = \left(\frac{\phi}{2}\right)(I_{t+j} - \omega S_{t+j})^2 \quad (1.11)$$

where I_t is end of period inventories, S_t is exogenous real sales, Q_t is real production, β is a constant discount rate, and $E_t(\bullet) = E_t(\bullet | \Omega_t)$ is the mathematical expectations operator, with Ω_t the information available at time t . The three structural parameters of the model include δ , the marginal cost of production, ϕ , the cost associated with inventories being other than the desired level, and ω , the cost associated with back-orders or stock outs.

All of the parameters are assumed to be non-negative constants. This assumption indicates that production costs are convex. With $\delta > 0$, a production-smoothing motive is implied, meaning firms will meet demand shocks partially out of inventories to avoid additional production costs. For $\omega = 0$, ϕ reflects holding costs of inventories, with costs rising with stock. This also will result in production smoothing, given that all the other parameters are greater than zero. For $\omega > 0$, there is a buffer stock motive — firms will want to hold inventories to fill unexpected orders rather than risk back ordering. The firm must then weigh the cost of holding extra inventory against that of stock outs.

The solution to the above results in the following Euler equation:

$$E_t[\delta Q_t + \phi(I_t - \omega S_t) - \beta \delta Q_{t+1}] = 0 \quad (1.12)$$

This implies that the firm will choose an inventory level such that the cost of making and storing an additional unit this period provides no cost benefit over waiting to produce the unit in the next period.

Substituting in for production yields

$$E_t[\delta(S_t + I_t - I_{t-1}) + \phi(I_t - \omega S_t) - \beta \delta(S_{t+1} + I_{t+1} - I_t)] = 0 \quad (1.13)$$

which can then be solved for I_{t+1} :

$$E_t I_{t+1} = E_t \left[\left(\frac{1+\beta}{\beta} + \frac{\phi}{\beta\delta} \right) I_t - \frac{1}{\beta} I_{t-1} + \frac{1 - \frac{\phi\omega}{\delta}}{\beta} S_t - S_{t+1} \right] \quad (1.14)$$

This can be rearranged and expressed as the second-order difference equation

$$E_t[(1 - \lambda L)(1 - (\lambda\beta)^{-1}L)I_{t+1}] = E_t\left[\frac{\alpha}{\beta}S_t - S_{t+1}\right] \quad (1.15)$$

with $\alpha = 1 - \frac{\phi\omega}{\delta}$, L representing the lag operator, and λ and $(\lambda\beta)^{-1}$ being the roots of the equation: $1 - \Phi L + \beta^{-1}L^2$, where $\Phi = \frac{(1+\beta)}{\beta} + \frac{\phi}{\beta\delta}$. With the structural parameters all non-negative, $0 < \lambda < 1$ and $\alpha \leq 1$.

Solving the equation for optimal I_t yields

$$I_t = \lambda I_{t-1} - S_t + (1 - \alpha\lambda)E_t\left[\sum_{j=0}^{\infty} (\lambda\beta)^j S_{t+j}\right] \quad (1.16)$$

This relation implies several characteristics for firm's inventory decisions. Inventories are positively related to expected future sales, while being negatively related to current sales. This indicates that firms meet current sales out of inventories rather than increase production. It can be shown that the above model accommodates the stock adjustment model sometimes seen in the literature (see Eichenbaum (1989) for details), a basic version of which is: $I_t = (1 - \lambda)[I_t^* - I_{t-1}]$. Here I_t^* is some desired level of inventories, usually dependent on $E_t S_t$, and $1 - \lambda$ the per period speed of adjustment towards that level.

In order to estimate the model, Equation 3.9 is usually expressed as a moment condition and then the Generalized Method of Moments (GMM) is applied to attain parameter estimates. While the structural parameters are not all identified, it is possible to come up with estimates for λ and α . Most of the literature restricts β to be .995 in estimation. The moment condition becomes

$$E_t F(H_{t+1}; \theta) = E_t[(1 - \lambda L)(1 - (\lambda\beta)^{-1}L)I_{t+1} + S_{t+1} - \frac{\alpha}{\beta}S_t] = 0 \quad (1.17)$$

where $H_{t+1} = [I_{t+1} \ I_t \ I_{t-1} \ S_{t+1} \ S_t]$ and $\theta = [\alpha \ \lambda \ \beta]$. Expanding the right hand side yields the following estimable version of the equation:

$$E_t F(H_{t+1}; \theta) = E_t[I_{t+1} - (\lambda + (\lambda\beta)^{-1})I_t + \beta^{-1}I_{t-1} + S_{t+1} - \frac{\alpha}{\beta}S_t] = 0 \quad (1.18)$$

Unfortunately, the various studies using this model have come up with a wide range of parameter estimates that are often at odds with each other and difficult to defend given the empirical facts. It may already be noted that the theory developed in this model is somewhat at odds with the facts mentioned above. To begin with, the data implies that, for the most part, firms do not smooth production, with

output being more variable than sales in most industries. Of course the inclusion of a buffer stock motive may help to explain this. Also the model implies a negative relationship between current sales and inventory investment, which again is at odds with the empirical facts. This also may be partially explained through the positive relationship between inventories and expected future sales. If current sales are unexpectedly high, it is likely that firms will adjust forecasts of future sales to reflect this, thus explaining the pro-cyclicality of inventories. Despite these problems the linear-quadratic model has been used in a great number of studies to date, mostly in attempts to find reasonable speeds of adjustment to desired inventory levels, which can be represented as approximately $1 - \lambda$.

1.5.2 Impact on Parameter Estimates

The main point of unit root testing is to attain stationarity in the data series used in estimation of economic models. This is necessary for the statistical theory underlying most estimation procedures to hold. While the effects of differencing when there is a deterministic trend or, in the opposite case, applying a deterministic detrending procedure to a stochastically trending process are well documented, it may be interesting to look at the effects of different deterministic detrending specifications on the parameter estimates from our inventory model. The following is a comparison of the GMM estimation results from the L-Q inventory model when three different detrending procedures are used on the data.¹² All results are from iterated GMM estimation with instrument set including a constant and two lags each of inventories and sales.¹³ The detrending procedures include a linear trend specification for both sales and inventories, a quadratic trend specification for both series, and finally the piecewise linear detrending implied by the above results. Parameter estimates and the J-test of model specification are shown in Table 1.12 .

Given the imprecision of the parameter estimates, particularly for λ , it is difficult to gauge whether

¹²The model was also estimated with differenced data. However the resulting estimates for α were either greater than one or negative, which is counter to the assumptions of the model. Therefore we do not report these estimates.

¹³In estimating the model, equation 1.18 was multiplied by $-\beta\lambda$ to obtain a different normalization. The result gives a specification that is identical to the stockout avoidance model estimated in Durlauf and Maccini (1995), Table 2. The reason for this change is that estimation using equation 1.18 resulted in implausible parameter estimates and extremely large standard errors. In addition, the model was estimated with a scaled version of the moment condition in equation 1.18 suggested by some researchers, which results from multiplication by $(1 - \lambda)^{-\rho}$, with $\rho = 1$. This version of the model also resulted in estimation problems, in particular an objective function with at least two local minima. Schuh (1996) found the parameter estimates using this normalization to be sensitive to choice of ρ .

the different detrending procedures have a great impact. It is notable, however, that use of piecewise linear detrending results in decidedly larger standard errors in most instances. In addition, there seems to be little impact on the J-statistics testing model specification. Only in the case of the Rubber industry does the change in detrending procedure affect the results of the specification test.

1.5.3 Other Implications

Despite the fact that there seems to be little effect on the outcome for this model, the above trend specifications do imply permanent changes in the growth paths of the sales and inventory series at the break dates. While previous research has indicated persistence in the inventory-sales relationship,¹⁴ if a shift in trend for the sales series is not accompanied by a corresponding shift in inventories, then the relationship between these two variables is permanently changed. Likewise for shifts in inventories that do not induce shifts in sales. In the context of the L-Q model, this implies that the structural parameters may change at these breakpoints. This can be seen in reference to the moment condition in Equation 1.18.

In order for this moment condition to remain true after a permanent shift in the growth path of just one of the series, there must be a change in one or more of the parameters at that point. This means that marginal costs will have changed at these points. One possible way to test for this would be to estimate the model using subsets of the data defined by the breakpoints. If the resulting estimates differ significantly, then this conclusion is supported. Other tests of structural change could also be applied to the model to determine if the parameters do in fact change over the series' histories. However, the current study is intended to determine if different detrending methods will affect the parameter estimates and such structural change tests are left for future consideration.

¹⁴Ramey and West (1997), using quarterly and annual data for the U.S., report impulse response curves indicating that when there is a shock to one or the other of inventories or sales, they generally return back to their initial stable relationship after a period of time. Their work seems to rely upon assuming a cointegrating relationship between sales and inventories, however. Tests on the monthly BEA series and much other work in the literature seems to reject this as a possibility, this study included.

1.6 Forecasting

Another area where choice of model specification in a univariate sense can have an impact is in forecasting models. In this case stationarity is not the goal. Instead we are looking to develop a model that performs well in producing accurate forecasts. It has recently been shown in Diebold and Kilian (2000) that using the results from unit-root tests to decide between use of a random walk with drift or a linearly trending AR(1) process in forecasting a data series can lead to lower prediction mean squared errors (PMSE). Given this, it is also likely that different choices of trend specification will result in markedly different forecasting results.

Using the inventory series from above, we can look at the impact of three different model specifications on out-of-sample forecasts. These models are a random walk with drift, a linearly trending AR(p) process, and an AR(p) with piecewise linear trend using the breakdates specified in Table 1.11.¹⁵ The choice of p is that chosen by the SIC pre-test model selection procedure outlined above. These choices are listed in Table 1.13. The trend specifications given in the table are the result of unit root tests using this selection procedure, so it seems prudent to include those autoregressive terms in our forecasting models.¹⁶ We are interested in looking at the differences in time paths implied by the out-of-sample forecasts from these different models. Determining the correct trend specification ought to greatly improve forecasts from any model. It would be expected that basing out-of-sample forecasts on a model with breaks should give quite a different result than the other models, since the breaks represent a discreet change in the deterministic path of the series which isn't present in the linear model. The question then becomes how far out in the forecast horizon do we get before this difference becomes significant. Figures 1.8 through 1.10 show the out-of-sample forecasts for each model applied to each series.¹⁷ The results are in levels (millions of dollars) so as to get a better idea of the magnitude of the differences between them. The

¹⁵Using a quadratic trend in a forecasting model does not seem to be a reasonable specification. Since the coefficient on the quadratic trend term is generally negative, the longer the forecast horizon, the more dominant this term is likely to be as the forecasts converge to the series' long run trend. We did look at this for the inventory series and for three of the series the quadratic trend model resulted in declining inventories into the future. For the three series in which this specification does not result in a declining forecast, the quadratic trend coefficient is either insignificant or very nearly so. For Total inventories a quadratic trend model did not converge to give parameter estimates.

¹⁶Campbell and Perron (1991) and Cochrane (1991) do simulations which include such pre-determined lagged variables in their look at forecasting and unit root testing. They did not consider different trend specifications or actual data, however.

¹⁷A piecewise linear specification was rejected for the petroleum series and so is not applied here. In addition, for the Chemical inventories series, the linear trend forecasting model did not converge.

forecast horizon goes out to 120 periods, or ten years.

Notably, for Rubber inventories the forecasts of all three models are nearly indistinguishable for the whole forecast horizon. The maximum difference, after 120 periods, is \$52million between the linear forecast and the piecewise linear. Recall, however, that a unit root was not rejected for this series for any of the tests with the exception of Model AA, which rejected at the 10% level. It is interesting to see that the wrong choice of forecasting model in this case would not have a great impact on the results, even at long horizons. Unfortunately, this is unlikely to be true for most other forecasting applications, as we shall see.

Also interesting are the forecasts for the Tobacco industry. The random walk model forecasts declining inventories right from the start. Both the linear and piecewise linear models give a different picture, with forecasts being very close to each other for several periods before diverging. It is difficult to see in the figure, but these two forecasts reach a maximum difference at 68 periods. From there the piecewise forecasts level out somewhat and the difference between the two shrinks as the forecasts from the linear model approach those from the piecewise. Apparently they will cross at some point in the distant horizon. Despite this, it is still obviously important to choose the right model from the start, as the maximum difference in levels between the linear and piecewise models is over \$100 million and the random walk model gives an even greater difference in levels. In this case, the piecewise linear model was chosen in the unit root testing and therefore should give the most reliable forecasts.

For Food inventories we found that a unit root was not rejected in any of the above testing procedures. This suggests that the random walk model should be the one chosen for forecasting. Surprisingly, the piecewise linear model gives forecasts that are fairly close to those from the random walk at short horizons. The difference doesn't reach \$100 million for thirty-two periods, nearly three years. With inventories in the twenty billion dollar range, this may be acceptable. The forecasts from the linear model are lower and quickly diverge from those of the other two.

For Apparel inventories our testing indicates that the piecewise linear model is the best choice. Again we see the random walk model and the piecewise giving fairly close forecasts at short horizons. The difference doesn't reach \$10 million for 16 periods and doesn't hit \$50 million for ten periods after that.

However, at the end of the 120 periods we have a difference of more than \$800 million between these models. Again there is a quick divergence in the linear model's forecasts, resulting in a difference of more than \$1400 million from the piecewise model at the end of the forecast horizon. Again we see that the choice of model clearly will have a significant impact at long horizons.

For Total inventories we had no rejections of a unit root from any of our testing other than the P_T test. This indicates a random walk model for forecasting. In this case, the linear model remains fairly close to the random walk for several periods before reaching a difference of \$100 million at the 12th period. The maximum difference between these two is \$910 million at the end of the forecast horizon, which may not be too bad considering the inventory levels are in the \$100 billion range. However, the piecewise linear model results in a difference of nearly \$5 billion from the random walk forecast at the end of the horizon.

Finally, for both Chemicals and Petroleum we have a quick divergence between the two models applied in each case. For Petroleum, a linear trend model was strongly indicated by the tests. Using a random walk model quickly leads to a difference of more than \$10 million within two periods and a difference of more than \$760 million by the end of the horizon. Clearly the choice of model has an impact in this case regardless of forecast horizon. For Chemicals our tests indicate a single break model as the best fit. Applying a random walk results in extreme differences in forecasts almost immediately, \$48 million in the first period and rising to more than \$.5 billion in nine months. At the end of the forecast horizon we see a difference of more than \$5.6 billion.

While the figures indicate point forecasts that can differ greatly as we move further out in the forecast horizon, by billions of dollars in some cases, only for three of the industries do the point forecasts for the random walk with drift or the linear trend specification follow paths which takes them outside of the forecast interval for the piecewise linear model. These are Food, Apparel, and Chemicals. For Food it takes 74 periods for the linear model to give a forecast result that lies outside of this interval. For Apparel it takes 75 periods for this to happen. For Chemicals, the linearly trending model did not converge, but in this case the random walk model results in forecasts outside the piecewise linear interval beginning with the 44th period, or almost two years out.

These results give an indication that choosing forecasting models on the basis of unit root tests can result in a significant difference from just choosing a linear trend or random walk model. By including the possibility of breaks in trend for these series we get another choice in our modelling. Apparently this also can have a significant impact on forecasting results. Interestingly, at short horizons there was not a great difference in forecasts from some of the models in some of the industries. As Diebold and Kilian (2000) suggest, it would be interesting to expand their study to see whether these types of trend specifications being included in unit root tests actually result in better forecasts when choosing models on that basis.

1.7 Summary

This study has conducted a rigorous analysis of the time series properties of a widely used set of inventory data. Applying recent advances in the unit root testing literature we find that when including the possibility of one or more breaks in trend for each series, we can reject the presence of a unit root in most cases. With these results in hand we choose models for detrending that remove piecewise linear or linear deterministic trends from each of the series to attain stationarity. The resulting transformed data is then used to test if there is any impact on the estimation results from a popular model of inventory behavior when using different trend specifications for detrending. We find that the type of trend specification does not greatly affect parameter estimates from the model or tests of model specification based on GMM. However, standard errors do seem to be significantly affected by the type of detrending procedure used.

In addition, there has been much interest in determining whether unit root tests are useful in the selection of forecasting models. Again using the results from the unit root tests above we look at the impact that different models have on forecasts out to 120 periods. We find that in many cases the models other than the one selected in testing result in quite different forecasts, especially at long horizons. While it gives no indication as to which model is the best, this empirical application indicates the importance of continuing research into the benefits of applying unit root tests of different types when choosing forecasting models.

Table 1.1: Summary Statistics

Inventories							
	Total	Food	Tobacco	Apparel	Chemicals	Petroleum	Rubber
Mean	47084	12967	973	3073	11725	4379	4278
Standard Error	619	119	6	53	199	35	59
Median	47206	13193	968	2935	11897	4321	4296
Mode	26483	10140	957	2636	5563	3373	2016
Standard Deviation	13458	2594	132	1145	4329	767	1289
Minimum	21981	7643	469	1003	3739	2694	1863
Maximum	74700	18994	1379	5265	20431	6187	7089
Sales							
	Total	Food	Tobacco	Apparel	Chemicals	Petroleum	Rubber
Mean	94395	26820	3502	5093	18932	11075	6009
Standard Error	1114	261	20	45	268	99	136
Median	97786	26744	3570	5324	20010	11746	5565
Mode	91592	17758	3706	5933	8001	6910	1975
Standard Deviation	24232	5667	434	987	5834	2158	2957
Minimum	49747	17012	2058	2869	7294	6458	1774
Maximum	137308	37605	5538	6743	28483	15622	12510

Data is in in millions of chained 1992 dollars

Table 1.2: Adjustment Factors

Series	Adjustment Factor
Food	1.3354
Tobacco	1.7044
Apparel	1.3056
Chemicals	1.4998
Petroleum	1.1465
Rubber	1.3185
Total Nondurables	1.3373

Table 1.3: Prior Research

Paper	Data	Years	Transformation	Reported Unit Root Tests
Blinder (1986)	BEA	1959:2 - 1981:12	detrended	none
Blinder and Maccini (1991)	BEA	1959:1 - 1986:10	detrended	none
Durlauf and Maccini (1995)	BEA	1959:1 - 1990:6	detrended	yes, but unspecified
Eichenbaum (1989)	BEA	1959:1 - 1984:12	separately detrended and differenced	none
Fuhrer et al. (1995)	BEA	1967:1 - 1991:8	none	none
McCarthy and Zakrajsek (1998)	BEA Trade Goods	1959:1 - 1997:8	differenced	panel data version of ADF
Miron and Zeldes (1988)	BEA & M3	1967:5 - 1982:12	detrended only	none
Ramey (1991)	BEA & M3	1959:1 - 1979:12	detrended	none
Ramey and West (1997)	BEA & Citibase	1959:1 - 1996:4	none	none
Schuh (1996)	M3LRD	1985 - 1993	detrended	none
West (1990)	Citibase	1947:1 - 1986:4	differenced	Said-Dickey

Table 1.4: ADF Test: Optimal Lags and Results

Series	H_{A1}		H_{A2}		H_{A3}	
	k_{AIC}	k_{SIC}	k_{AIC}	k_{SIC}	k_{AIC}	k_{SIC}
Inventories						
Total	20	1	20	1	20	1
Food	1	1	1	1	1	1
Tobacco	5*	5*	5*	5*	5	5
Apparel	13	1	1**	1**	1**	1**
Chemicals	7**	0*	7	0	7	0
Petroleum	14	1**	3*	3*	3*	3*
Rubber	2	0	2	0	2	0
Sales						
Total	24**	1	24	1	24	1
Food	4	1	1*	1*	1*	1*
Tobacco	12	5	12	5	12	5
Apparel	12	2	2*	1*	2*	1*
Chemicals	12	1**	12	1	12	1
Petroleum	23	2	23	0**	23	0**
Rubber	24	1	24**	0*	24	0*

H_{A1} is against the alternative of single mean stationarity

H_{A2} is against the alternative of linear-trend stationarity

H_{A3} is against the alternative of quadratic-trend stationarity

* indicates significance at the 5% level

** indicates significance at the 10% level

Table 1.5: Elliot et al. Critical Values

T	1%	2.5%	5%	10%
DF-GLS				
50	-3.77	-3.46	-3.19	-2.89
100	-3.58	-3.29	-3.03	-2.74
200	-3.46	-3.18	-2.93	-2.64
∞	-3.48	-3.15	-2.89	-2.57
P_T				
50	4.22	4.94	5.72	6.77
100	4.26	4.90	5.64	6.79
200	4.05	4.83	5.66	6.86
∞	3.96	4.78	5.62	6.89

Table 1.6: Test Results

	DF-GLS	P_T
Inventories		
Total	-0.798	1.148*
Food	-2.058	9.788
Tobacco	-2.462	6.033***
Apparel	-1.409	18.114
Chemicals	-0.542	95.417
Petroleum	-3.509*	1.927*
Rubber	-1.30	26.206
Sales		
Total	-0.939	36.349
Food	-3.501*	3.815*
Tobacco	-1.979	12.542
Apparel	-2.679***	4.435**
Chemicals	-0.799	61.212
Petroleum	-2.716***	6.303***
Rubber	-2.453	5.501*

* indicates significance at the 1% level

** indicates significance at the 5% level

*** indicates significance at the 10% level

Table 1.7: Zivot and Andrews' Critical Values

Model	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
A	-5.34	-5.02	-4.80	-4.58	-3.75	-2.99	-2.77	-2.56	-2.32
B	-4.93	-4.67	-4.42	-4.11	-3.23	-2.48	-2.31	-2.17	-1.97
C	-5.57	-5.30	-5.08	-4.82	-3.98	-3.25	-3.06	-2.91	-2.72

Table 1.8: Lumsdaine and Papell Critical Values

Model	1%	2.5%	5%	10%
AA	-6.94	-6.53	-6.24	-5.96
CA	-7.24	-7.02	-6.65	-6.33
CC	-7.34	-7.02	-6.82	-6.49

Table 1.9: Structural Breakpoints and Test Results for Single-Break Models

Inventories	Model A		Model B		Model C	
	t	Date	t	Date	t	Date
Total	-3.564	Sept. 1978	-3.175	Jan. 1967	-3.515	Nov. 1971
Food	-4.217	July 1979	-3.558	May 1988	-4.391	Dec. 1981
Tobacco	-8.205*	Dec. 1993	-7.281*	July 1992	-8.267*	Dec. 1990
Apparel	-4.770***	Oct. 1964	-5.022*	Sept. 1966	-5.135**	Nov. 1965
Chemical	-3.926	May 1965	-4.449*	March 1969	-4.561	July 1972
Petroleum	-7.087*	April 1962	-7.045*	May 1963	-7.415*	Feb. 1992
Rubber	-4.236	Oct. 1979	-2.881	Dec. 1968	-4.374	Oct. 1979
Sales						
Total	-3.734	April 1979	-3.964	July 1972	-4.740	April 1979
Food	-6.516*	Nov. 1966	-6.122*	Oct. 1972	-6.489*	Nov. 1966
Tobacco	-4.954**	Oct. 1982	-2.925	May 1993	-4.654	Oct. 1982
Apparel	-6.068*	July 1988	-6.162*	Dec. 1967	-6.253*	Dec. 1978
Chemical	-3.975	Feb. 1968	-4.222***	Feb. 1972	-5.000***	Aug. 1974
Petroleum	-5.115**	Feb. 1981	-6.234*	Nov. 1976	-8.346*	Feb. 1980
Rubber	-5.110**	April 1979	-4.298***	April 1972	-5.327*	April 1979

* indicates significance at the 1% level

** indicates significance at the 5% level

*** indicates significance at the 10% level

Table 1.10: Breakpoints and Test Results for Two Breaks in Trend

Inventories	Model AA			Model CA			Model CC		
	t	Break 1	Break 2	t	Break 1	Break 2	t	Break 1	Break 2
Total	-4.482	10/65	9/78	-5.302	9/72	9/78	-5.166	9/66	4/89
Food	-4.750	9/60	7/79	-5.234	9/72	12/81	-5.255	5/68	4/89
Tobacco	-9.476*	12/90	10/93	-10.176*	10/61	12/93	-10.536*	10/61	12/90
Apparel	-5.756	5/62	7/65	-6.300	12/65	12/83	-6.372	10/68	9/83
Chemical	-4.700	5/65	1/79	-6.083	11/72	1/79	-6.180	11/72	1/79
Petroleum	-7.452*	4/62	-1/72	-7.800*	6/72	8/74	-7.956*	1/72	2/92
Rubber	-6.117***	11/67	10/79	-5.977	2/68	10/79	-5.674	4/68	10/79
Sales									
Total	-4.268	4/79	9/90	-5.483	6/74	2/80	-7.270**	8/74	8/79
Food	-6.766**	9/61	11/66	-8.286*	3/72	11/85	-8.577*	8/72	1/76
Tobacco	-5.502	10/82	12/94	-6.874**	10/82	12/90	-8.290*	10/82	1/93
Apparel	-8.320*	12/63	7/88	-8.730*	7/74	6/75	-9.157*	8/72	6/75
Chemical	-4.289	1/63	2/68	-7.777*	8/74	11/80	-7.530*	8/74	11/80
Petroleum	-7.528*	3/81	4/90	-8.880*	1/80	3/81	-9.071*	1/74	1/80
Rubber	-5.750	4/79	8/90	-6.540	3/74	4/79	-6.940**	11/74	4/79

* indicates significance at the 1% level

** indicates significance at the 5% level

*** indicates significance at the 10% level

Table 1.11: Selected Models For Detrending

Inventories	Model	Break Dates
Total	CC	September 1966 and April 1989
Food	CC	May 1968 and April 1989
Tobacco	A	December 1993
Apparel	B	September 1966
Chemical	B	March 1969
Petroleum	Linear	None
Rubber	AA	November 1967 and October 1979
Sales		
Total	CC	August 1974 and August 1979
Food	Linear	None
Tobacco	CC	October 1982 and January 1993
Apparel	Linear	None
Chemical	CC	August 1974 and November 1980
Petroleum	C	February 1980
Rubber	CC	November 1974 and April 1979

Table 1.12: Estimation Results From Deterministic Detrending Procedures

Industry	Linear			Quadratic			Piecewise Linear		
	λ	α	J-test	λ	α	J-test	λ	α	J-test
Total	.867 (.046)	.963 (.017)	1.95	.868 (.046)	.947 (.024)	2.01	.979 (1.345)	.875 (.052)	2.57
Food	.903 (.099)	.893 (.042)	1.26	.921 (.122)	.825 (.057)	.73	.932 (.323)	.897 (.042)	1.92
Tobacco	.853 (.212)	1.077 (.090)	.90	.787 (.170)	1.133 (.128)	1.10	.763 (.278)	1.421 (.556)	2.20
Apparel	.861 (.085)	.897 (.058)	3.86	.890 (.104)	.841 (.076)	3.42	.943 (.301)	.926 (.042)	4.187
Chemicals	.852 (.041)	.962 (.016)	8.57*	.856 (.043)	.960 (.019)	8.56*	.942 (.316)	.932 (.032)	8.93*
Petroleum	.825 (.075)	.951 (.022)	5.25	.845 (.087)	.889 (.043)	4.01	.821 (.070)	.684 (.084)	3.21
Rubber	.845 (.069)	.938 (.023)	6.36	.861 (.076)	.934 (.024)	6.80	.808 (.092)	.867 (.035)	15.82*

* indicates the model specification is rejected at the 95% level

Table 1.13: Choice of p from SIC

Industry	Model	
	Linear	Piecewise Linear
Total	2	1
Food	2	2
Tobacco	6	1
Apparel	2	2
Chemicals	1	2
Petroleum	4	na
Rubber	1	1

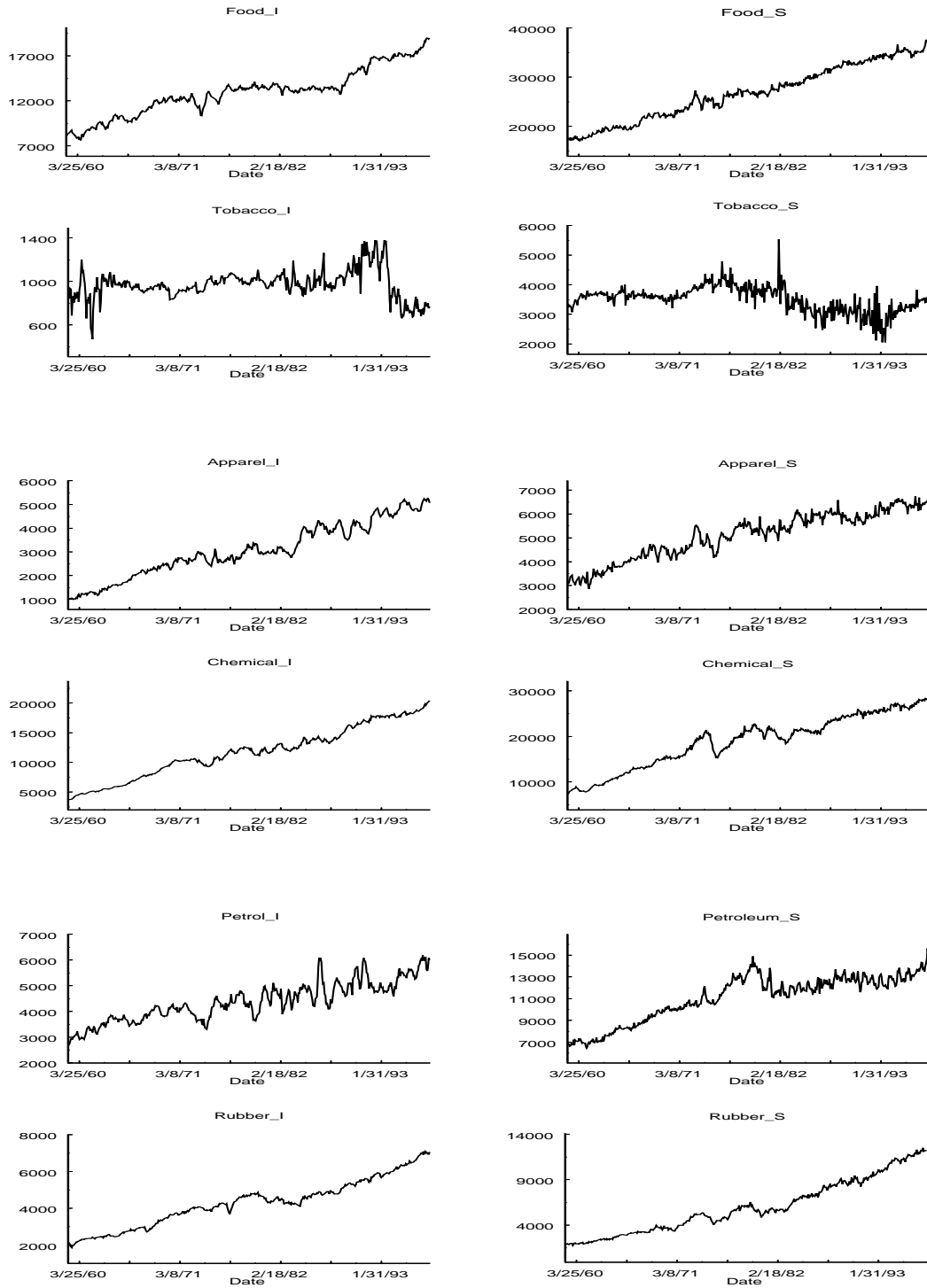


Figure 1.1: Time Series of BEA Inventory and Sales Data

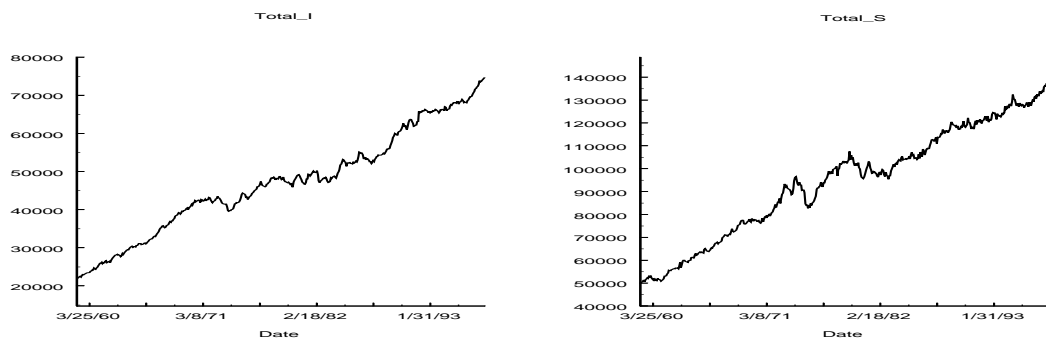
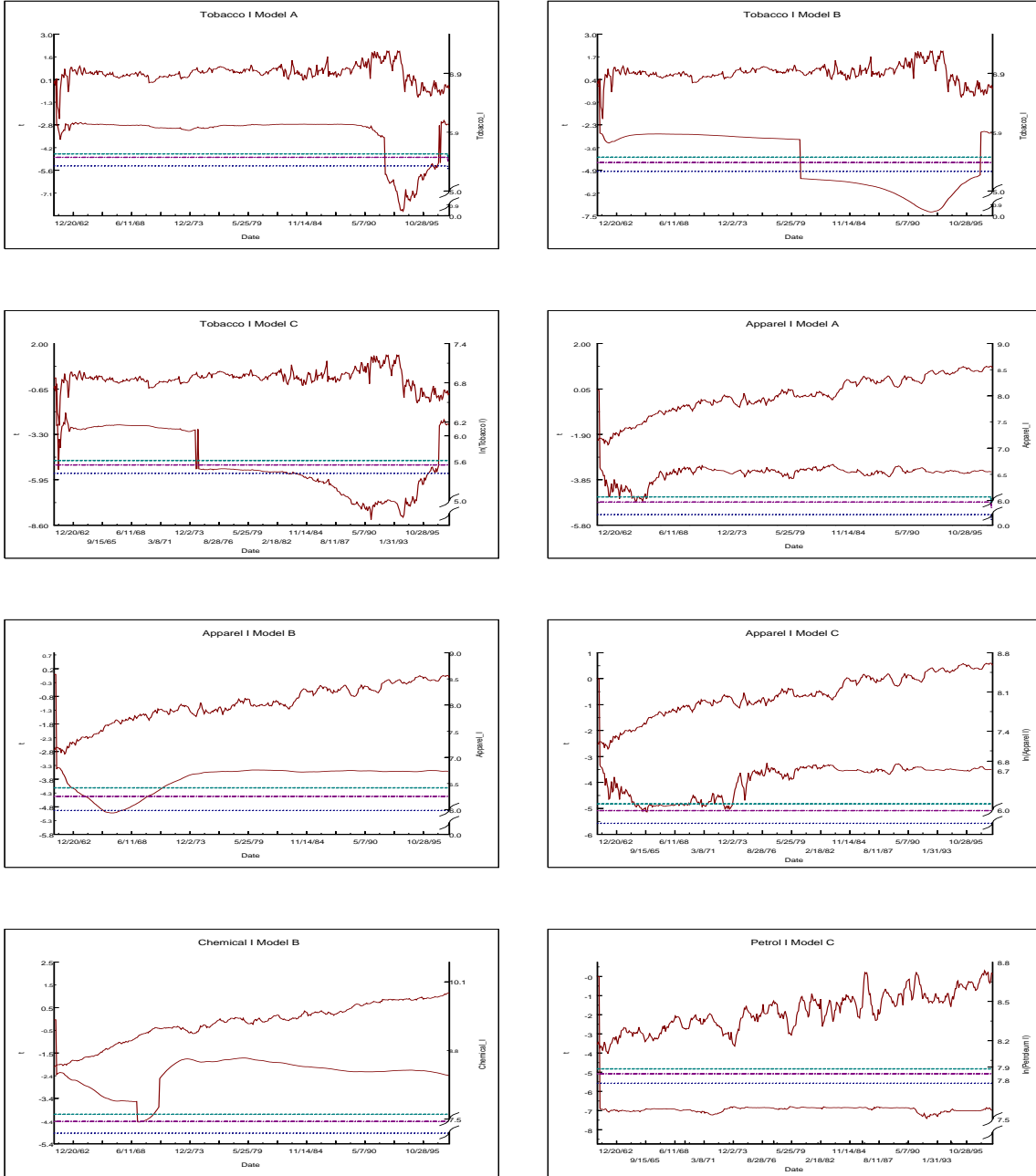
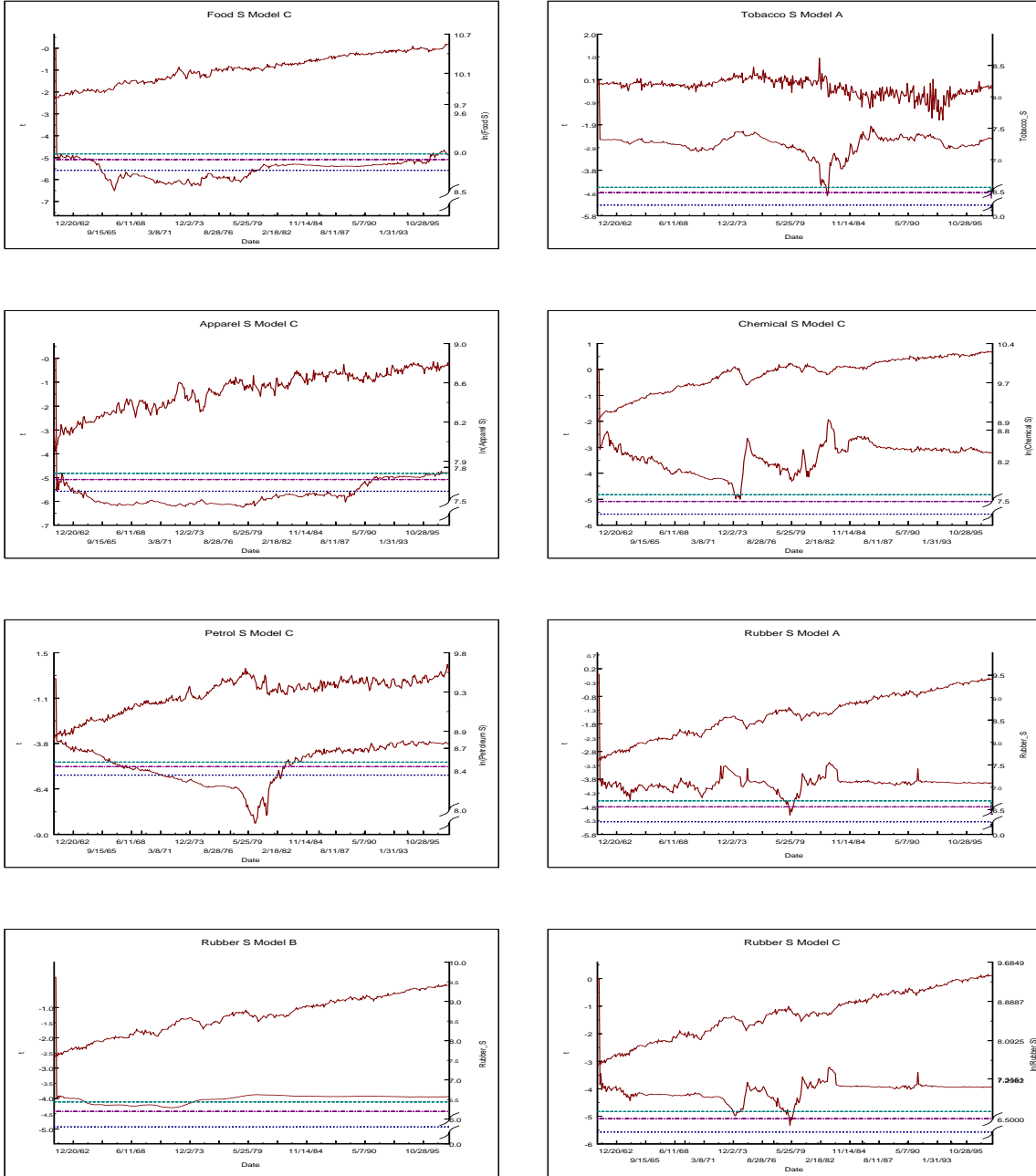


Figure 1.2: Time Series of BEA Inventory and Sales Data



The dashed lines indicate the 10%, 5%, and 1% critical values for the t-tests

Figure 1.3: Sequential t-tests: Selected Models — Inventories



The dashed lines indicate the 10%, 5%, and 1% critical values for the t-tests

Figure 1.4: Sequential t-tests: Selected Models — Sales

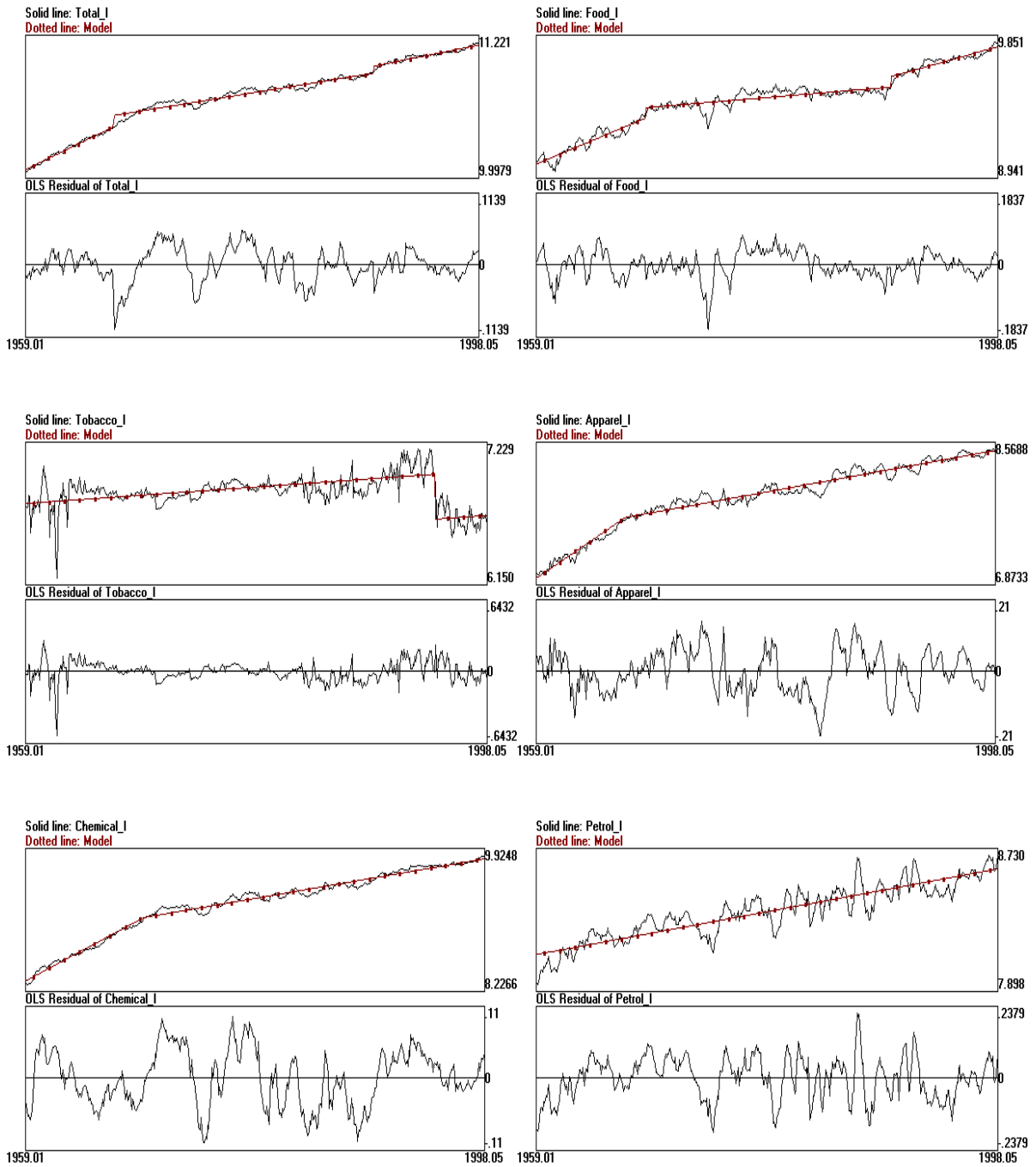


Figure 1.5: Detrending Models and Residuals

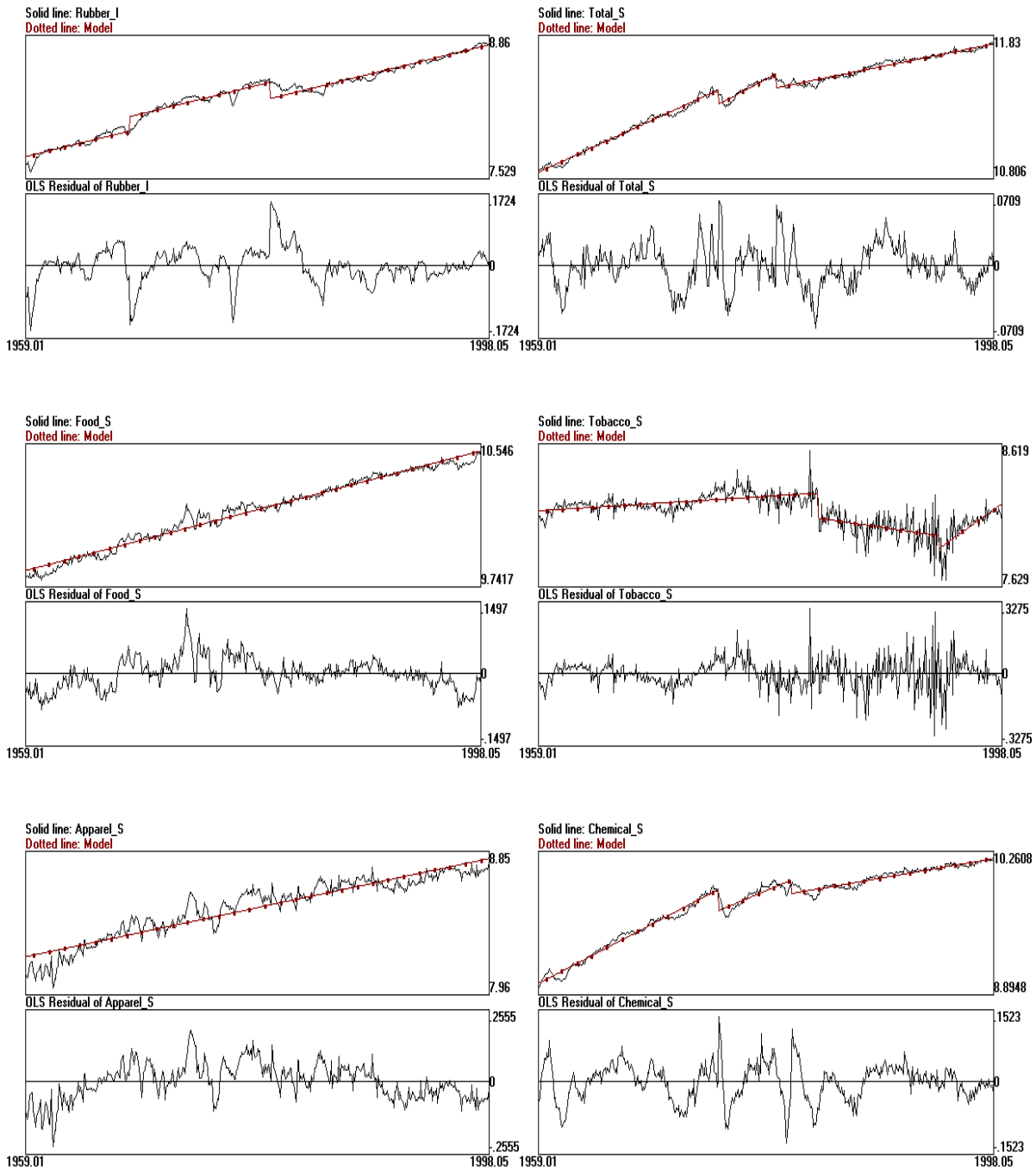


Figure 1.6: Detrending Models and Residuals

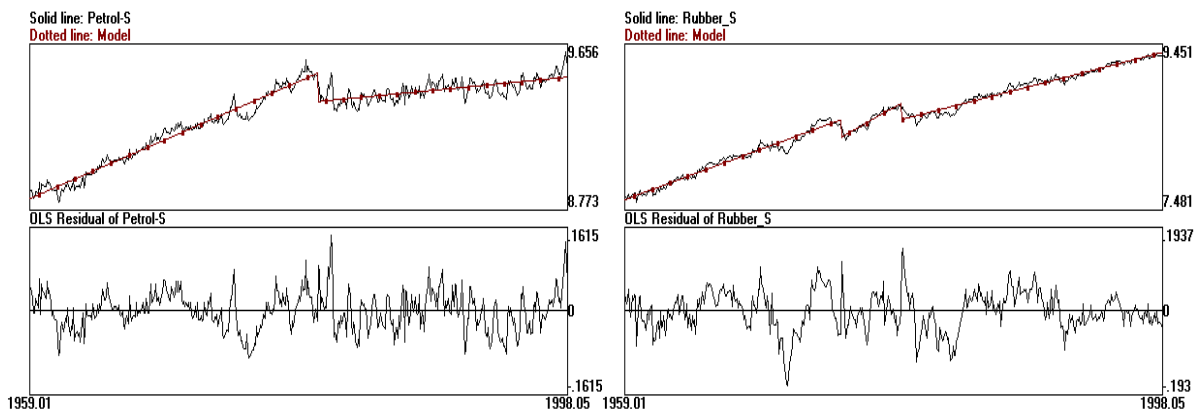


Figure 1.7: Detrending Models and Residuals

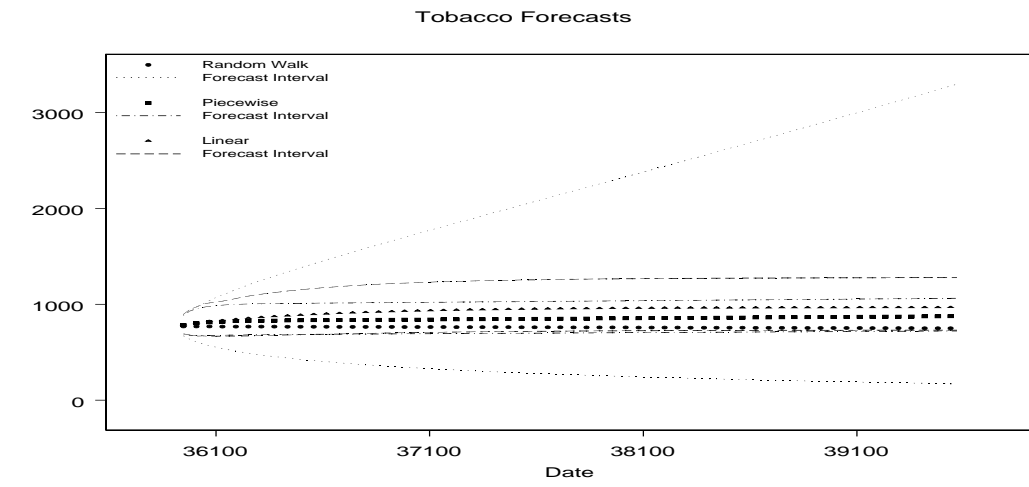
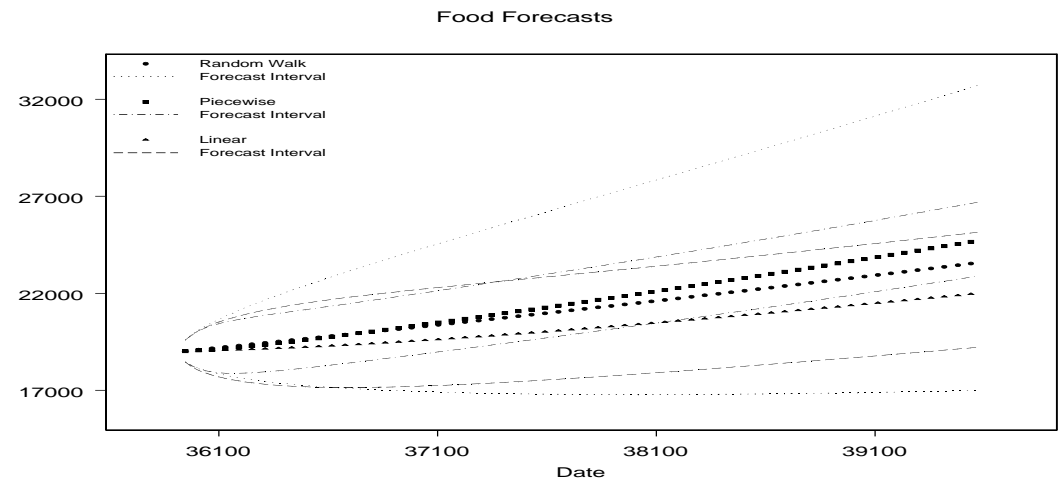
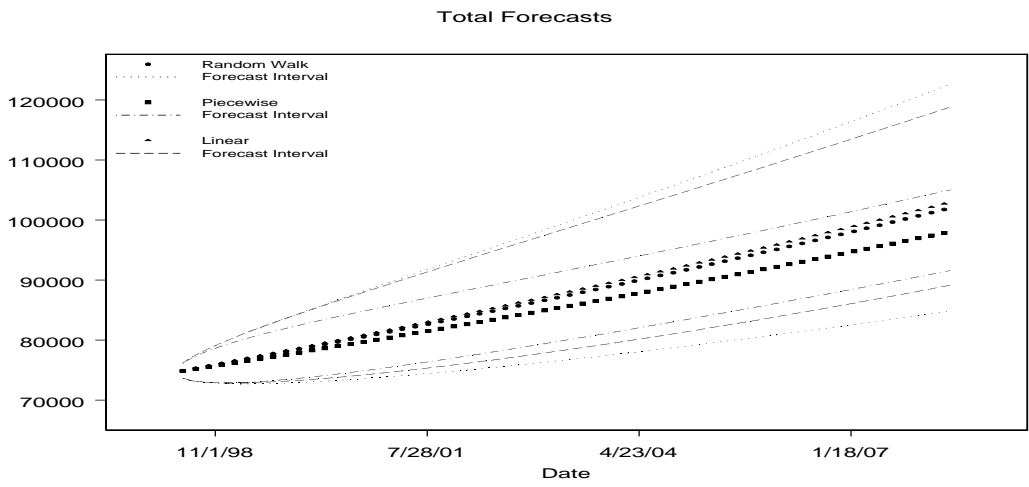
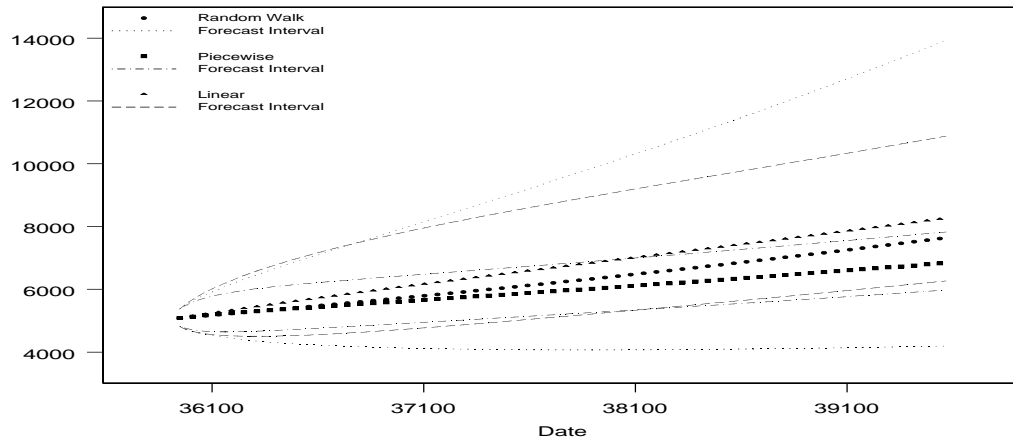
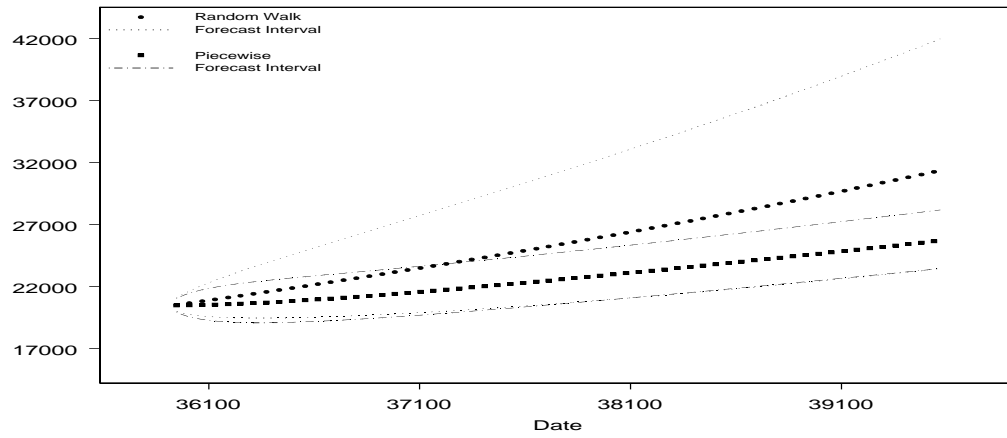


Figure 1.8: Forecasts

Apparel Forecasts



Chemical Forecasts



Petroleum Forecasts

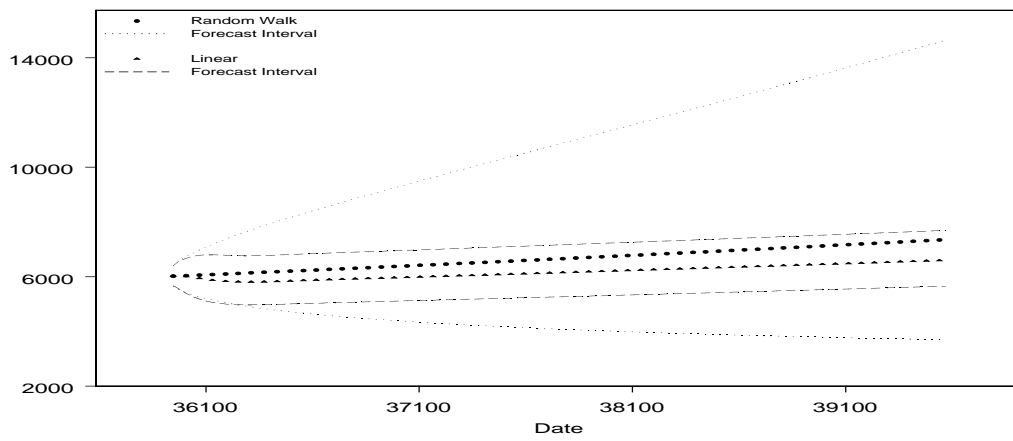


Figure 1.9: Forecasts

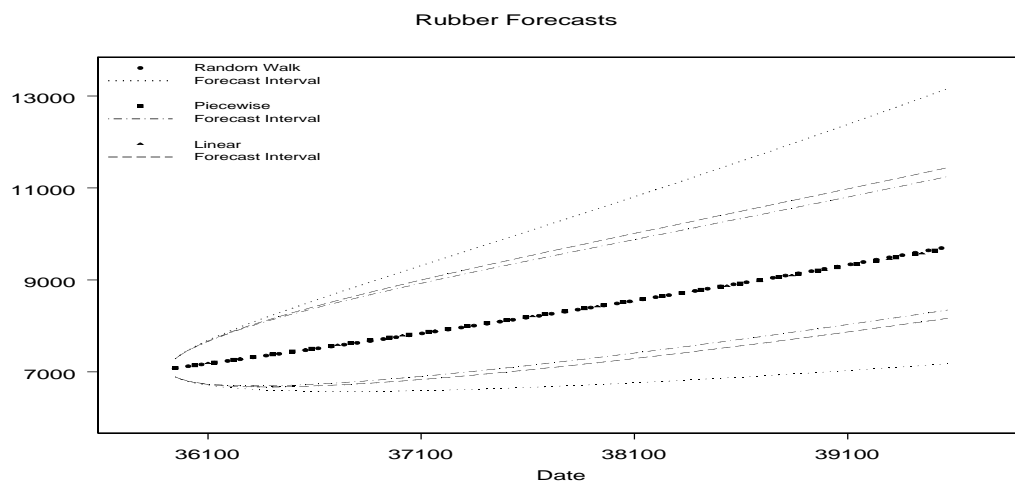


Figure 1.10: Forecasts

Chapter 2

Consequences of Hodrick-Prescott Filtering for Parameter Estimation in a Structural Model of Inventory Behavior

2.1 Introduction

In recent years the Hodrick-Prescott (HP) filter, developed in Hodrick and Prescott (1997), has become a popular method for removing the growth component from data exhibiting long run trends over time. The HP filter has been much applied in the Real Business Cycle (RBC) literature as a method of isolating business cycle frequencies from the long run trend component in both actual and simulated data. The goal in that literature has been to study the business cycle behavior of different macroeconomic variables and to determine the degree of any comovements between them over the course of the cycle. This generally involves a correlation study of the filtered data. Unfortunately for this work, studies of the impact HP filtering can have on the second moments of a data series have shown it to have the

potential to introduce spurious correlations in the filtered data and spurious cross-correlations between series. This may cause the stylized facts gained through the study of HP filtered data to be questionable at best. There have been a number of papers showing this potential for the HP filter to distort results from RBC models and empirical studies, most notably King and Rebelo (1993), Harvey and Jaeger (1993), and Cogley and Nason (1995).

Another application of the HP filter has been in detrending data series which are subsequently used in estimation of parameters from structural econometric models. Most estimation methods require the data used to be covariance stationary in order for the underlying asymptotic theory to hold. While seemingly useful as a tool for detrending and achieving stationarity, the above mentioned problems have the potential to affect estimation using HP filtered data as well. If the cross-correlations between variables are affected by the filtering process, then estimation of a model posited as governing the relationship between those variables will likely be affected as well, along with any inferences based on the estimated parameters. It is the goal of this paper to assess whether there is a significant impact on parameter estimation when using HP filtered data as opposed to other methods of detrending.

To effect our study of the impact HP filtering may have on estimation we conduct a simulation study based on a structural model of inventory behavior that has been widely studied. The simulated data generated by the model contain simple linear trends. The simulated data is then used to estimate the population parameters of the model using the Generalized Method of Moments (GMM) estimator of Hansen (1982). Because the estimated model is linear in the parameters, this method is equivalent to classical instrumental variables estimation. The model is estimated several times using different criteria for detrending the data prior to estimation, including removing a linear trend and HP filtering with different values for the smoothing parameter, discussed below. For comparison, the model is also estimated without detrending the data but instead including time in the instrument set.

The paper proceeds as follows: We first present the Hodrick-Prescott filter and discuss some of the previous results mentioned above. Next we derive a structural model of inventory behavior with which to simulate data. The model is then calibrated to actual data on inventories and sales of non-durable goods. Once this is done, simulated data is used to estimate the parameters of the model

several times using different methods for dealing with the trend. A comparison will be made of the distributional characteristics across the different detrending methods. In addition, we will consider the effect of detrending method on different test statistics resulting from the estimation.

2.2 The Hodrick-Prescott Filter

This derivation of the Hodrick-Prescott filter closely follows that of King and Rebelo (1993). The HP filter is based on the assumption that a time series process can be modeled as the sum of a cyclical component and a growth component:

$$y_t = y_t^g + y_t^c. \quad (2.1)$$

In the real business cycle literature it is desirable to remove the growth component, whether it be a stochastic or deterministic trend, in order to study the behavior of the cyclical component and to compare that behavior between different series. The HP filter achieves this by defining the cyclical component as

$$y_t^c = y_t - y_t^g \quad (2.2)$$

where y_t^c becomes the desired detrended series. The HP filter minimizes the variance of (2.2) while penalizing for excessive changes in the growth component. This can be expressed in terms of the following minimization problem

$$\min_{\{y_t^g\}_{t=1}^T} \sum_{t=1}^T \{(y_t - y_t^g)^2 + \lambda[(y_{t+1}^g - y_t^g) - (y_t^g - y_{t-1}^g)]^2\} \quad (2.3)$$

where λ is termed the "smoothing parameter", since its value determines the penalty applied to excessive changes in the growth rate of the series. The value chosen for λ will differ depending on the sampling frequency of the data and represents the trade-off between smoothness and goodness of fit in the resulting series.¹

¹The smoothing parameter can be viewed as the ratio of the variance of the cyclical component to that of the change in the growth component, although alternative interpretations are available. See, for example, Hodrick and Prescott (1997), Harvey and Jaeger (1993), and Ravn and Uhlig (1997).

The first order condition from the minimization problem is

$$\begin{aligned}
-2(y_t - y_t^g) + 2\lambda [(y_t^g - y_{t-1}^g) - (y_{t-1}^g - y_{t-2}^g)] - 4\lambda [(y_{t+1}^g - y_t^g) - (y_t^g - y_{t-1}^g)] \\
+ 2\lambda [(y_{t+2}^g - y_{t+1}^g) - (y_{t+1}^g - y_t^g)] = 0
\end{aligned} \tag{2.4}$$

Using lag operators and simplifying this becomes

$$\begin{aligned}
y_t - y_t^g &= \lambda [(1-L)y_t^g - (1-L)y_{t-1}^g] - 2\lambda [(1-L)y_{t+1}^g - (1-L)y_t^g] \\
&\quad + \lambda [(1-L)y_{t+2}^g - (1-L)y_{t+1}^g] \\
&= \lambda(1-L)^2 y_t^g - 2\lambda(1-L)^2 y_{t+1}^g + \lambda(1-L)^2 y_{t+2}^g \\
&= \lambda(1-L)^2 [y_t^g - 2y_{t+1}^g + y_{t+2}^g] \\
&= \lambda(1-L)^2 (1 - 2L^{-1} + L^{-2}) y_t^g \\
&= \lambda(1-L)^2 (1 - L^{-1})^2 y_t^g
\end{aligned}$$

or

$$y_t = [1 + \lambda(1-L)^2(1-L^{-1})^2] y_t^g$$

By inversion we can put this in terms of a low frequency filter applied to y_t which yields the growth component

$$y_t^g = H^g(L)y_t \tag{2.5}$$

where

$$H^g(L) = [1 + \lambda(1-L)^2(1-L^{-1})^2]^{-1} \tag{2.6}$$

The cyclical component of the series is then given by

$$y_t^c = y_t - y_t^g = (1 - H^g(L))y_t = H^c(L)y_t \tag{2.7}$$

Where $H^c(L)$ can be expressed in terms of the following filter yielding the cyclical component of the series:

$$H^c(L) = \frac{\lambda[1-L]^2[1-L^{-1}]^2}{1 + \lambda[1-L]^2[1-L^{-1}]^2} = \frac{\lambda L^{-2}(1-L)^4}{1 + \lambda L^{-2}(1-L)^4} = \frac{L^{-2}(1-L)^4}{\frac{1}{\lambda} + L^{-2}(1-L)^4} \tag{2.8}$$

with

$$y_t^c = H^c(L)y_t \tag{2.9}$$

In the frequency domain, the squared gain, or frequency response, of this filter is

$$HP(\omega) = \left| \frac{\lambda[1 - e^{-i\omega}]^2[1 - e^{i\omega}]^2}{1 + \lambda[1 - e^{-i\omega}]^2[1 - e^{i\omega}]^2} \right|^2 = \left| \frac{4\lambda[1 - \cos(\omega)]^2}{1 + 4\lambda[1 - \cos(\omega)]^2} \right|^2 = \left| \frac{4[1 - \cos(\omega)]^2}{\frac{1}{\lambda} + 4[1 - \cos(\omega)]^2} \right|^2 \quad (2.10)$$

where ω is frequency measured in radians. This representation allows us to see the effect the filter has on cycles of different frequencies in the data. Note that $HP(0) = 0$, indicating attenuation of low frequencies, and also that $HP(\pi) = \frac{16}{1/\lambda + 16}$, which is approximately equal to one for large λ .² This implies close to unit gain on high frequencies. In fact, as λ goes to infinity, the filter passes all but the zero frequency, meaning it removes a deterministic linear trend only. In addition, the four differences in the numerator of Equation 2.8 indicate it is capable of rendering stationary integrated processes of up to fourth order.³ This would be the equivalent of differencing the data with $(1 - L)^4$ and then applying the filter $\frac{L^{-2}}{\frac{1}{\lambda} + L^{-2}(1 - L)^4}$ to the result.

Figure 2.1 plots the frequency response of the filter for different values of λ , chosen to reflect values used in the literature for quarterly data ($\lambda = 1600$) and monthly data ($\lambda = 14400, 86400, 129600$). Also depicted are the ideal filters for both monthly and quarterly data, which only pass frequencies above a period of eight years, often chosen to be the cutoff period for business cycle research. On the vertical axis are the weights applied by the filter to different frequencies in the data. Notice that very low frequencies are completely attenuated by the filter while high frequencies are passed without alteration. Also note that regardless of the value of λ , more low frequency content is passed by the HP filter than may be ideal if the goal is to study behavior of a series at business cycle frequencies. In addition, some higher frequency content is attenuated to the right of the cutoff points of the ideal filters. This attenuation tends to decrease as λ is increased, although more undesirable low-frequency content is then left as well.

Notice that, regardless of the value of λ , the HP filter always removes a linear trend, but operates further on the deviations from this trend. Singleton (1988) shows that when applied to stationary data, the HP filter is a very good approximation to an ideal high pass filter, that which only allows cycles of greater frequency than some prespecified value to pass. This implies that if the data is linear trend stationary, the HP filter acts as a high pass filter on the residuals from the deterministic trend.⁴ This

² λ is typically set to 1600 for quarterly data and 14,400 for monthly. See Hodrick and Prescott (1997)

³See King and Rebelo (1993).

⁴See Cogley and Nason (1995).

property makes the filter potentially useful in business cycle research by allowing the isolation of business cycle frequencies from the data. However, if the goal of filtering is primarily to render data stationary prior to use in estimation of econometric models, without regard to business cycle behavior, applying such a high pass filter may be removing more low-frequency content than is perhaps desirable. This may be especially harmful if most of the power of the series lies in the low frequencies, as tends to be the case with many macroeconomic series. An example of this is given in Figure 2.2, which plots the spectral density of monthly aggregate inventories of nondurable goods in the top panel, and then the same series detrended using the HP-filter with different values for λ . Also included is the spectral density of the linearly detrended series. The dashed line in the figure indicates a period of eight years, which is generally considered to be the maximum length of the business cycle. The HP-filter is designed to attenuate any frequencies with period above eight years, which the figure indicates it does fairly well. Again, this seems to be a good procedure if the goal is to study the behavior of a series at business cycle frequencies and to compare that behavior across series. However, as a general detrending method it is clear that HP filtering removes much of the information in the data contained in the lower frequencies. In addition, even at frequencies above eight years, the filter removes more of the variability from the series than does linearly detrending, which is due to the frequency response of the filter as depicted in Figure 2.1. Note, however, that as λ is increased more of the low frequency information is retained, which makes sense since the filter becomes equivalent to linear detrending as λ goes to infinity.

Figures 2.3 and 2.4 illustrate an application of the HP filter to actual data on aggregate nondurable goods inventories. Figure 2.3 plots the HP trend against the actual series, as well as a linear trend fit to the data for comparison. Applying the HP filter to the series results in a trend that very closely follows the time path of the data. While visually appealing, this may not necessarily be a desirable thing.⁵ Figure 2.4 indicates the difference in the residuals from the two detrending procedures. Clearly the HP filter results in a very different detrended series.

The HP filter can be shown to be the optimal linear filter, minimizing the mean square error between the actual cyclical component and its estimate, when Equation 2.1 is given the following structural

⁵As an example, McCallum (2000) contains an application of the HP filter to U.S. real GDP over the 1920s and 1930s. The resulting detrended series gives little evidence of the Great Depression.

representation:⁶

$$y_t = \mu_t + \epsilon_t \quad t = 1, \dots, T \quad (2.11)$$

where μ_t is a stochastic trend and ϵ_t represents the irregular component, with ϵ_t distributed IID normal with variance σ_ϵ^2 .

The stochastic trend can be defined to be a local linear trend:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.12)$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad (2.13)$$

with η_t and ζ_t separately distributed IID normal with zero means and variances σ_η^2 and σ_ζ^2 , respectively. Note that β denotes the slope of the trend, which changes for $\zeta_t \neq 0$. Also, if $\sigma_\zeta^2 = 0$ this becomes a random walk with drift and if $\sigma_\eta^2 = 0$ as well we have a deterministic trend. Assuming $\sigma_\eta^2 = 0$ and $\sigma_\zeta^2 \neq 0$, $HP(\omega)$ is the optimal filter, with $\lambda = \sigma_\epsilon^2 / \sigma_\zeta^2$.⁷

The optimality of the HP filter depends upon a few rather strong assumptions contained in the above model, with the strongest being that the growth component and the cyclical component are uncorrelated. This would indicate that these are generated by completely separate economic events. In addition, the trend component is I(2), causing the process as a whole to be I(2). For many economic time series this is unlikely to be the case. Finally, for the filter to be optimal it is necessary to know *a priori* the variances of the disturbance terms in order to set the smoothness parameter, λ , at its optimal value. Since the purpose of applying the filter to begin with is to isolate the trend and cyclical components, it is unlikely that these will be known ahead of time.⁸ In practice λ is generally set based upon common practice and the sampling frequency of the data.

The above conditions for optimality make it unlikely that the HP filter is an optimal filter in most applications. However, even when optimal the filter tends to cause distortions in the auto-correlations of filtered series. Ehlgen (1998) finds that the autocorrelations and the variances of HP filtered data tend to be too small. Further, he states that this can be a normal characteristic of optimal signal extraction

⁶See King and Rebelo (1993) and Harvey and Jaeger (1993)

⁷See Hodrick and Prescott (1997) and Harvey and Jaeger (1993). Also see Kaiser and Maravall (1999) for alternative interpretations of λ .

⁸Harvey and Jaeger (1993) discuss estimating λ using maximum likelihood methods. Pedersen (1999) proposes other methods to determine the optimal value of λ .

filters and not just a problem with the HP filter. An illustration of this effect is given in Figure 2.5, which shows the impact HP filtering has on the autocorrelation function for aggregate inventories. Note that the HP filter eventually induces negativity in the autocorrelations that doesn't exist in the series in levels or when a linear trend is removed.

Cogley and Nason (1995) look at the effects of HP filtering on difference stationary data and the implications of this for business cycle research. They posit that, if the series is difference stationary then applying the HP filter will likely result in problems akin to linearly detrending a random walk, which has been studied by Nelson and Kang (1981)⁹ This implies that application of the HP filter to difference stationary data will likely cause problems of filter induced cyclicity in the resulting filtered series. Application of the HP filter to an I(1) process is equivalent to applying the filter

$$S(L) = \frac{L^{-2}(1-L)^3}{\frac{1}{\lambda} + L^{-2}(1-L)^4} \quad (2.14)$$

to the stationary differenced series. In this case, $S(L)$ is not a high pass filter. In investigating it's properties Cogley and Nason (1995) find that the filter $S(L)$ will result in spurious cyclicity when applied to stationary data. Furthermore, HP filtering of two independent random walk processes was found to induce cross-correlations between the series that were previously nonexistent. Harvey and Jaeger (1993) also show that the HP filter can induce spurious cycles that are nonexistent in the original series when applied to an ARIMA(p, d, q) process. They found that there is a non-zero probability of inducing spurious cross-correlations between two independent series that have been HP filtered.

The implication of these findings is that the cyclical properties of HP filtered data cannot be relied upon for inference about the business cycle behavior of real economic series, particularly if the series are likely to be integrated of order one or higher. To our knowledge there have been no studies to date on the possible impact these problems may have on estimation of econometric models using HP-filtered data. Although the filter always removes a linear trend from the data, we chose to study the effects of applying the filter to data simulated to contain a linear trend, as this seems the simplest case with which to start. As mentioned above, the HP filter will remove the trend but also act as a high-pass filter on what is left. The result, as seen in Figure 2.2, is that the filter removes more of the variation in the

⁹They found that such detrending caused cycles to appear in the detrended data that were not present in the original series.

data than would just simply detrending, which in turn implies that there is less information left with which to effect estimation. This indicates that the results from estimation performed with filtered data are likely to be impacted. It is the purpose of the rest of this paper to give evidence of this.

2.3 The Model

To get an idea of how the Hodrick-Prescott filter affects estimation results we use a simulation based on the linear-quadratic model of inventory behavior that has been widely studied in the macroeconomics literature. The simulation model used is that of West and Wilcox (1996). In this paper, the model will be calibrated to monthly Bureau of Economic Analysis data on inventories and sales of non-durable goods spanning January 1959 to May 1998.

Consider a representative firm that maximizes the present discounted value of profits subject to a cost function that accounts for production costs, costs of changing production, inventory holding costs, and backlog costs that arise when sales cannot be met out of inventories. This particular model also includes an unobservable cost shock, that may be observable to the firm but not the econometrician. The optimization problem becomes

$$\max_{\{H_{t+j}\}} \lim_{T \rightarrow \infty} E_t \sum_{j=0}^T b^j (p_{t+j} S_{t+j} - C_{t+j})$$

s.t.

$$C_{t+j} = \frac{a_0}{2} \Delta Q_{t+j}^2 + \frac{a_1}{2} Q_{t+j}^2 + \frac{a_2}{2} (H_{t+j-1} - a_3 S_{t+j})^2 + H_{t+j} u_{t+j} + \text{deterministic terms} \quad (2.15)$$

$$Q_{t+j} = S_{t+j} + H_{t+j} - H_{t+j-1} \quad (2.16)$$

where p_{t+j} is price, S_{t+j} is real sales, H_{t+j} real end of period inventories, C_{t+j} real costs, and Q_{t+j} real output. The a terms represent the marginal costs faced by the firm, with a_0 the marginal cost of changing production, a_1 the marginal cost of production, a_2 the marginal cost of holding inventories, and a_3 the cost associated with stock outs. Unobserved cost shocks are represented by u_{t+j} , with u_t assumed to be IID Normal with zero mean. Deterministic terms would account for any linear and trend terms that may exist in the data. The final term, b , is the discount factor and is usually set *a priori*

in estimation of the model. The parameters of the cost function are generally what is estimated, and therefore these are what will need to be calibrated in the simulation.

The first order condition of the model yields:

$$E_t\{a_0(\Delta Q_t - 2b\Delta Q_{t+1} + b^2\Delta Q_{t+2}) + a_1(Q_t - bQ_{t+1}) + ba_2(H_t - a_3S_{t+1}) + \text{deterministic terms} + u_t\} = 0 \quad (2.17)$$

Ignoring deterministic terms, if we substitute in for production from equation 2.16 and then collect terms, we have

$$E_t\{a_0b^2H_{t+2} - [a_0(2b^2 + 2b) + a_1b]H_{t+1} + [a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2]H_t - [a_0(2b + 2) + a_1]H_{t-1} + a_0H_{t-2} + a_0b^2S_{t+2} - [a_0(b^2 + 2b) + a_1b + a_2a_3b]S_{t+1} + [a_1 + a_0(1 + 2b)]S_t - a_0S_{t-1} + u_t\} = 0 \quad (2.18)$$

In order for there to be an optimal solution, the Legendre-Clebsch condition states that the second derivative of the objective function must be non-negative.¹⁰ This yields the requirement that

$$a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2 \geq 0 \quad (2.19)$$

In order to use the model to simulate data we need to solve for a reduced form equation in inventories. This also requires a specification of the sales process. Here we assume sales to be exogenous and described by an AR(2) process:

$$S_t = KS + \alpha t + \phi_1 S_{t-1} + \phi_2 S_{t-2} + \epsilon_{St} \quad (2.20)$$

where KS and αt are constant plus trend terms and ϵ_{St} is assumed IID Normal.

With sales exogenous, we can represent the Euler equation as a fourth order stochastic difference equation in inventories. This is most easily seen in equation 2.18, which can be represented as

$$E_t\{A(L)H_{t+2} = D_t\} \quad (2.21)$$

where

$$A(L) = \frac{1 - a_0^{-1}b^{-2}[a_0(2b^2 + 2b) + a_1b]L + a_0^{-1}b^{-2}[a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2]L^2}{1}$$

¹⁰See Ramey (1991) for a discussion of this optimality condition.

$$-a_0^{-1}b^{-2}[a_0(2b+2)+a_1]L^3+b^{-2}L^4$$

$$D_t = -S_{t+2}+a_0^{-1}b^{-2}[a_0(b^2+2b)+a_1b+a_2a_3b]S_{t+1}-a_0^{-1}b^{-2}[a_1+a_0(1+2b)]S_t+b^{-2}S_{t-1}-a_0^{-1}b^{-2}u_t$$

A stable solution to the model requires the two smallest roots of equation 2.21 to be less than one in modulus. We denote these roots as λ_1 and λ_2 . The stable solution can then be written in terms of these two stable roots as¹¹

$$H_t = (\lambda_1 + \lambda_2)H_{t-1} - \lambda_1\lambda_2H_{t-2} + b^{-1}\lambda_1\lambda_2(\lambda_1 - \lambda_2) \sum_{j=0}^{\infty} \{[(b\lambda_1)^{j+1} - (b\lambda_2)^{j+1}]E_tD_{t+j}\} \quad (2.22)$$

Assuming uncorrelated cost shocks, we can solve for the reduced form equation for inventories as follows.¹² Define the following scalars

$$\begin{aligned} \rho_1 &= \lambda_1 + \lambda_2 \\ \rho_2 &= -\lambda_1\lambda_2 \\ w_1 &= b^2\rho_2 \\ w_2 &= -\rho_2[b^2 + 2b + b(a_1/a_0) + (ba_2a_3/a_0)] \\ w_3 &= \rho_2[2b + 1 + (a_1/a_0)] \\ w_4 &= -\rho_2 \end{aligned}$$

and the vector

$$e' = (1 \quad 0) \quad (2.23)$$

and the matrices

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

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

with I being the 2x2 identity matrix.

¹¹See West (1992), Blanchard (1983), or Fuhrer et al. (1995).

¹²This follows Blanchard (1983) and West (1992).

The reduced form equation for inventories is then

$$H_t = \rho_1 H_{t-1} + \rho_2 H_{t-2} + \pi_1 S_{t-1} + \pi_2 S_{t-2} + \beta t + KH + \epsilon_{Ht} \quad (2.24)$$

with

$$(\pi_1 \quad \pi_2)' = e' D(w_1 \Phi^3 + w_2 \Phi^2 + w_3 \Phi + w_4 I) \quad (2.25)$$

and

$$\epsilon_{Ht} = \frac{\rho_2}{a_0} u_t + \frac{\pi_2}{\phi_2} \epsilon_{St}$$

and $\beta t + KH$ being trend plus constant terms.

In order to simulate the data we need to specify values for the parameters of the exogenous sales process, the reduced form inventory equation, and also the distribution of the sales and cost innovations. Since the parameters in equation 2.24 are functions of the a_i 's they can be determined given some choice for the cost parameters. West and Wilcox (1996) use several sets of values for these based on other work. We chose to instead determine values for the cost parameters that fit the particular data set we are calibrating to. In order to do this it was necessary to jointly estimate the exogenous sales process, equation 2.20, and the reduced form for inventories, equation 2.24. From the parameter estimates for equation 2.24 we can get estimates for combinations of the cost parameters using¹³

$$a_1/a_0 = \rho_1(b - \rho_2^{-1}) - 2b - 2 \quad (2.26)$$

$$a_2/a_0 = -b^{-1}[\rho_2^{-1}(1 + b\rho_1^2) + b^2\rho_2 + (1 + b)a_1/a_0 + (1 + 4b + b^2)] \quad (2.27)$$

and from estimates of the sales process we can also solve for a_3 from equation 2.25.

In calibrating the model, we chose to roughly match estimates from BEA data on total aggregated inventories and sales of non-durable goods. In fixing the a_i 's for the simulation we follow Blanchard (1983) and normalize a_0 to one. Based on our estimates, the resulting values for the a_i 's used in the simulation are given in table 2.1.

With these estimates for the cost parameters in hand we can use them to simulate data for inventories and sales by specifying the parameters of the exogenous sales process along with the distributions of

¹³See Blanchard (1983).

the sales and cost innovations. The autoregressive parameters of the sales process are set to approximately match those of our data. These are given in Table 2.2. Following West and Wilcox (1996), the sales innovation variance is set to normalize the unconditional variance of the sales process to one. In addition we use the same cost shock variance and correlation with sales given in that paper, as these values conveniently result in the variance ratio and the correlation between inventories and sales closely matching that of the data. The sales innovation variance and the variance of the cost shock are given in Table 2.2, along with the correlation between the two.

With the a_i 's fixed, we can then determine implied values for ρ_1, ρ_2, π_1 , and π_2 . Recall that $\rho_1 = \lambda_1 + \lambda_2$ and $\rho_2 = -\lambda_1\lambda_2$. The values for λ_1 and λ_2 can be taken as the two smallest in modulus roots of the characteristic equation for $A(L)$ in equation 2.21 above. This equation is

$$\lambda^4 - a_0^{-1}b^{-2}[a_0(2b^2 + 2b) + a_1b]\lambda^3 + a_0^{-1}b^{-2}[a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2]\lambda^2 - a_0^{-1}b^{-2}[a_0(2b + 2) + a_1]\lambda + b^{-2}$$

The π_i 's are then found from equation 2.25. The resulting values for the parameters of the inventory process are reported in table 2.3.

To determine coefficients on trend terms we again follow West and Wilcox (1996) and set them so that the implied coefficients of variation (cv) for ΔS_t and ΔH_t matches that of our data, which we find to be approximately 14.6 for sales and 9.7 for inventories. The coefficient of variation is

$$cv = \frac{\sqrt{\gamma}}{\mu} \tag{2.28}$$

where γ is the variance of the process and μ the mean, which will involve the trend terms for the process in levels. To solve for these, first express the sales process and the reduced form for inventories in differenced form and solve for their means. The sales process is given by equation 2.20. In differences, this becomes:

$$\Delta S_t = \phi_1 \Delta S_{t-1} + \phi_2 \Delta S_{t-2} + \alpha + \Delta \epsilon_{St} \tag{2.29}$$

The mean is given by:

$$\mu_{\Delta S_t} = E\Delta S_t = \phi_1 E\Delta S_{t-1} + \phi_2 E\Delta S_{t-2} + \alpha \tag{2.30}$$

$$= \phi_1 \mu_{\Delta S_t} + \phi_2 \mu_{\Delta S_t} + \alpha \quad (2.31)$$

$$= \frac{\alpha}{1 - \phi_1 - \phi_2} \quad (2.32)$$

Given the variance and cv, we can then solve for α .

For inventories we have

$$\Delta H_t = \rho_1 \Delta H_{t-1} + \rho_2 \Delta H_{t-2} + \pi_1 \Delta S_{t-1} + \pi_2 \Delta S_{t-2} + \beta + \Delta \epsilon_{Ht} \quad (2.33)$$

The mean of the process is:

$$E\Delta H_t = \rho_1 E\Delta H_{t-1} + \rho_2 E\Delta H_{t-2} + \pi_1 E\Delta S_{t-1} + \pi_2 E\Delta S_{t-2} + \beta \quad (2.34)$$

or

$$\mu_{\Delta H_t} = \rho_1 \mu_{\Delta H_t} + \rho_2 \mu_{\Delta H_t} + \pi_1 \mu_{\Delta S_t} + \pi_2 \mu_{\Delta S_t} + \beta \quad (2.35)$$

$$= \frac{1}{1 - \rho_1 - \rho_2} \{(\pi_1 + \pi_2) \mu_{\Delta S_t} + \beta\} \quad (2.36)$$

$$= \frac{1}{1 - \rho_1 - \rho_2} \left\{ \frac{\pi_1 + \pi_2}{1 - \phi_1 - \phi_2} \alpha + \beta \right\} \quad (2.37)$$

To find the variances, we need to find the covariance matrix of the system of equations given by 2.29 and 2.33. Using methods from Hamilton (1994), first express the system in terms of deviations from mean

$$\begin{aligned} \Delta H_t - \mu_{\Delta H_t} &= \rho_1 (\Delta H_{t-1} - \mu_{\Delta H_t}) + \rho_2 (\Delta H_{t-2} - \mu_{\Delta H_t}) \\ &\quad + \pi_1 (\Delta S_{t-1} - \mu_{\Delta S_t}) + \pi_2 (\Delta S_{t-2} - \mu_{\Delta S_t}) + \Delta \epsilon_{Ht} \end{aligned}$$

$$\Delta S_t - \mu_{\Delta S_t} = \phi_1 (\Delta S_{t-1} - \mu_{\Delta S_t}) + \phi_2 (\Delta S_{t-2} - \mu_{\Delta S_t}) + \Delta \epsilon_{St}$$

where we recall

$$\epsilon_{Ht} = \frac{\rho_2}{a_0} u_t + \frac{\pi_2}{\phi_2} \epsilon_{St} \quad (2.38)$$

so

$$\Delta \epsilon_{Ht} = \frac{\rho_2}{a_0} \Delta u_t + \frac{\pi_2}{\phi_2} \Delta \epsilon_{St} \quad (2.39)$$

Express this system of equations as the first-order VAR

$$Z_t = F Z_{t-1} + C V_t \quad (2.40)$$

where

$$Z_t = \begin{bmatrix} \Delta H_t \\ \Delta H_{t-1} \\ \Delta S_t \\ \Delta S_{t-1} \end{bmatrix} \quad (2.41)$$

$$F = \begin{bmatrix} \rho_1 & \rho_2 & \pi_1 & \pi_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \phi_1 & \phi_2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.42)$$

$$C = \begin{bmatrix} \rho_2/a_0 & \pi_2/\phi_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.43)$$

and

$$V_t = \begin{bmatrix} \Delta u_t \\ \Delta \epsilon_{St} \\ 0 \\ 0 \end{bmatrix} \quad (2.44)$$

To find the covariance matrix, let

$$\Sigma = E(Z_t Z_t') \quad (2.45)$$

Then

$$\begin{aligned} \Sigma &= E(FZ_{t-1} + CV_t)(FZ_{t-1} + CV_t)' \\ &= F\Sigma F' + FE(Z_{t-1}V_t')C' + CE(V_tZ_{t-1}')F' + CE(V_tV_t')C' \\ &= F\Sigma F' + Q \end{aligned}$$

Define

$$\Omega = \begin{bmatrix} \sigma_u^2 & \sigma_{u,\epsilon_S}^2 & 0 & 0 \\ \sigma_{u,\epsilon_S}^2 & \sigma_{\epsilon_S}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.46)$$

Then

$$E(Z_{t-1}V_t') = CV_{t-1}V_t' = -C\Omega \quad (2.47)$$

$$E(V_tZ_{t-1}') = V_tV_{t-1}'C' = -\Omega C' \quad (2.48)$$

$$E(V_tV_t') = 2\Omega \quad (2.49)$$

Q then becomes

$$Q = -FC\Omega C' - C\Omega C'F' + 2C\Omega C' \quad (2.50)$$

To solve for Σ , define

$$A = F \otimes F$$

$$B = I - A$$

Then using the *vec* operator, we have¹⁴

$$vec\Sigma = B^{-1}vecQ \quad (2.51)$$

from which we can reconstruct Σ and retrieve the desired variances for the ΔH_t and ΔS_t processes.

From here we calculate the trend coefficients as

$$\alpha = (1 - \phi_1 - \phi_2)\mu_{\Delta S} = (1 - \phi_1 - \phi_2)\frac{\sqrt{\gamma_{\Delta S}}}{cv_{\Delta S}} \quad (2.52)$$

and

$$\beta = (1 - \rho_1 - \rho_2)\mu_{\Delta H} - \frac{\pi_1 + \pi_2}{1 - \phi_1 - \phi_2}\alpha = (1 - \rho_1 - \rho_2)\frac{\sqrt{\gamma_{\Delta H}}}{cv_{\Delta H}} - \frac{\pi_1 + \pi_2}{1 - \phi_1 - \phi_2} \quad (2.53)$$

The resulting trend coefficients are $\alpha = .001211$ and $\beta = .002685$.

¹⁴See Hamilton (1994), pages 264-266.

For the simulation, the constant terms in equations 2.20 and 2.24 were chosen by fixing the other parameters at the chosen values and using the actual data to solve for KS and KH in

$$\hat{\mu}_{S_t} = KS + \alpha\mu_t + \phi_1\hat{\mu}_{S_{t-1}} + \phi_2\hat{\mu}_{S_{t-2}} \quad (2.54)$$

$$\hat{\mu}_{H_t} = KH + \beta\mu_t + \rho_1\hat{\mu}_{H_{t-1}} + \rho_2\hat{\mu}_{H_{t-2}} + \pi_1\hat{\mu}_{S_{t-1}} + \pi_2\hat{\mu}_{S_{t-2}} \quad (2.55)$$

where $\hat{\mu}_{X_{t-j}} = \frac{1}{T-2} \sum_{t=3-j}^{T-j} X_t$ and $\mu_t = \frac{1}{T-2} \sum_{t=2}^T t$. This yields $KS = 531$ and $KH = 1050$.

The sales and inventory processes specified in equations 2.20 and 2.24 are used to generate the simulated data for sales and inventories. One thousand samples were generated by first obtaining starting values for S_{t-1} , S_{t-2} , H_{t-1} , and H_{t-2} by generating series of length $T = 1000$ from each process while omitting the trend terms. The last two observations from each of these series are then used as the starting values in generating an additional one thousand observations from both the inventory and sales processes, this time including the trend terms. The final five hundred observations from each series is then chosen as the sample.¹⁵

2.4 Estimation Technique

In estimating the model using the simulated data the goal is to retrieve estimates of the population values of the cost parameters. However, it may be noted from (2.17) that the parameters of the cost function are not fully identified by the Euler equation, as any simultaneous scaling of a_0 , a_1 , or a_2 would have no effect on the Euler condition. Because of this we are reduced to estimating combinations of parameters rather than the individuals. This requires some sort of normalization in order to estimate the Euler equation, which amounts to choosing a left hand side variable in (2.17) and solving for that variable. Choice of which variable to solve for has often varied in previous studies, and although it shouldn't affect the Euler equation identities, it does often result in different estimates for the parameter values.¹⁶

However, our goal here is to construct a valid model with which to analyze the effects of filtering techniques, rather than to make any particular observations on the empirical behavior of inventories.

¹⁵A sample size of five hundred is chosen because this is close to the number of observations available in Bureau of Economic Analysis data for monthly inventories and sales at this point.

¹⁶See Fuhrer et al. (1995) for a discussion of different normalizations for the linear-quadratic inventory model.

With this in mind, normalization issues are ignored, and we continue to follow West and Wilcox (1996) in choosing to solve for current period inventories. This is a convenient choice in that it amounts to dividing (2.18) through by the Legendre-Clebsch condition of (2.19). The first order condition can now be expressed as

$$E_t\{H_t - \beta_1 X_{1t+2} - \beta_2 X_{2t+1} - \beta_3 S_{t+1} - u_t\} = 0 \quad (2.56)$$

where

$$\begin{aligned} X_{1t+2} &= -b^2 H_{t+2} + (2b^2 + 2b)H_{t+1} + (2b + 2)H_{t-1} - H_{t-2} \\ &\quad -b^2 S_{t+2} + (b^2 + 2b)S_{t+1} - (2b + 1)S_t + S_{t-1} \end{aligned}$$

$$X_{2t+1} = bH_{t+1} + H_{t-1} + bS_{t+1} - S_t$$

$$\beta_1 = \frac{a_0}{c}$$

$$\beta_2 = \frac{a_1}{c}$$

$$\beta_3 = \frac{ba_2a_3}{c}$$

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

Given the values of the cost parameters in Table 2.1, the implied population values of the regression coefficients are given in Table 2.4.

From estimates of the β_i 's we can derive estimates of the normalized cost parameters, with a_0/c and a_1/c directly estimated by β_1 and β_2 and

$$a_2/c = b^{-1}\{1 - \beta_1(1 + 4b + b^2) - \beta_2(1 + b)\} \quad (2.57)$$

$$a_3 = \frac{\beta_3}{ba_2/c} \quad (2.58)$$

Replacing expectations with realized values in (2.56) and solving for current period inventories gives us a linear equation which can then be used in estimation.

$$H_t = \beta_1 X_{1t+2} + \beta_2 X_{2t+1} + \beta_3 S_{t+1} + v_{t+2} \quad (2.59)$$

where v_{t+2} includes the unobservable cost shock along with expectational errors, or

$$v_{t+2} = u_t - \beta_1(X_{1t+2} - E_t X_{1t+2}) - \beta_2(X_{2t+1} - E_t X_{2t+1}) - \beta_3(S_{t+1} - E_t S_{t+1}) \quad (2.60)$$

Although (2.59) is linear in the parameters, the correlation structure of the regressors rules out ordinary least squares estimation as inconsistent. Instead, an instrumental variables procedure is used to obtain consistent estimates of the regression coefficients. The procedure is implemented using the Generalized Method of Moments (GMM) estimator of Hansen (1982), which is equivalent to classical instrumental variables estimation when the equation is linear in the parameters. This requires (2.59) to be recast as a moment condition

$$E_t\{H_t - \beta_1 X_{1t+2} - \beta_2 X_{2t+1} - \beta_3 S_{t+1}\} = E\{v_{t+2}\} = 0. \quad (2.61)$$

To understand the estimation procedure it is first useful to review the basics of standard GMM theory. Let X_t be an n -dimensional random vector of observed variables, which is assumed to be strictly stationary over the sample space $\mathcal{X} \supseteq \mathfrak{R}^n$. Let θ denote a p -dimensional vector of parameters with true value represented by θ_0 , with $\theta \in \Theta \supseteq \mathfrak{R}^p$. Let the function $G : \mathcal{X} \times \Theta \rightarrow \mathfrak{R}^q$ be a q -dimensional vector-valued function such that $E[G(X_t, \theta_0) \mid \Omega_t] = 0$, with Ω_t the information set at time t . Under certain regularity conditions this represents a population moment condition (PMC) to which we can derive a sample analog.¹⁷ If we let Z be an $(n \times m)$ matrix of variables such that row t of Z is the m -dimensional vector z_t' , which is uncorrelated with $G(X_t, \theta_0)$, then with $z_t \in \Omega_t$, z_t is a vector of instrumental variables that allows us to specify the orthogonality condition $E[z_t G(X_t, \theta_0)'] = 0$. This also can be viewed as a PMC, as it will continue to satisfy the conditions outlined above.

The sample analog to the PMC will then be equivalent to $T^{-1} \sum_{t=1}^T z_t G(X_t, \theta)'$. Letting $\Psi_T(\theta) = z_t G(X_t, \theta)'$, we can define

$$Q_T(\theta) = \left\{ T^{-1} \sum_{t=1}^T \Psi_t(\theta) \right\}' W_T \left\{ T^{-1} \sum_{t=1}^T \Psi_t(\theta) \right\} \quad (2.62)$$

where W_T is a positive definite matrix of constants of the appropriate dimensions. Note that $Q_T(\theta_0) = 0$, so for $\theta \neq \theta_0$, $Q_T(\theta)$ can be thought of as a measure of distance from zero. Using this we can define $\hat{\theta}$ as the minimum distance, GMM estimator of θ_0 , such that

$$\hat{\theta} = \arg \min_{\theta \in \Theta} Q_T(\theta) \quad (2.63)$$

¹⁷See Newey and McFadden (1994).

Standard GMM generally is implemented using either a two-step estimation procedure or an iterated procedure to arrive at an estimate for θ .

The only requirement for the weighting matrix is that it be positive semi-definite and converge to a positive definite matrix of constants, W , as $T \rightarrow \infty$. Hansen (1982) has shown that the optimal weighting matrix will be the inverse of the asymptotic covariance matrix of the sample moment. Of course to calculate this we need a consistent estimate of θ_0 . Because of this, the two-step and iterated estimators use a first-step estimator in which any suitable positive-definite matrix, usually the identity matrix or the inverse of the instrument cross products matrix, is used to get a consistent estimate for θ_0 . This first-step estimate is then used to derive a weighting matrix that is asymptotically efficient which can in turn be used to find a more efficient estimate of θ_0 . The two-step estimator stops here, while the iterated GMM estimator continues, using the new estimates to derive a new weighting matrix at each step until the parameter estimates converge.

For our model we have

$$G(X_t, \theta) = H_t - \beta_1 X_{1t+2} - \beta_2 X_{2t+1} - \beta_3 S_{t+1} \quad (2.64)$$

where $X_t = [H_t \ X_{1t+2} \ X_{2t+1} \ S_{t+1}]'$ and $\theta = [\beta_1 \ \beta_2 \ \beta_3]'$. The vector of instruments used is $z_t = [1 \ H_{t-1} \ H_{t-2} \ S_{t-1} \ S_{t-2}]'$. Because v_{t+2} in (2.60) is MA(2), the weighting matrix, W , used in estimation is the inverse of the Newey and West (1987) heteroskedasticity and autocorrelation consistent covariance matrix, which uses Bartlett weights on the lagged covariances to guarantee positive semi-definiteness. This is

$$\hat{W} = \{\hat{\Gamma}_0 + \sum_{j=1}^m [1 - j/(m+1)](\hat{\Gamma}_j + \hat{\Gamma}'_j)\}^{-1} \quad (2.65)$$

with

$$\hat{\Gamma}_j = T^{-1} \sum_{t=j+1}^T z_t z'_{t-j} v_{t+2} v_{t+2-j} \quad \text{for } j \geq 0 \quad (2.66)$$

and m being the bandwidth, selected using the estimation method of Newey and West (1994).

We estimate the above model five times for each of the 1000 samples, using a different procedure for handling the trend each time. The first method, labeled *LDT*, just removes a linear trend from the data

by regressing the series on a constant and time and retaining the residuals. The next three methods use the HP filter for detrending, but with different values for the smoothing parameter λ , the choices of which have been used or suggested in various studies with monthly data. These will be denoted *HP1*, *HP2*, and *HP3*. *HP1* uses the standard value of $\lambda = 14,400$ generally applied in the literature to monthly data. *HP2* uses a value of $\lambda = 86,400$, used by Bils and Kahn (2000) in their study of inventory behavior. *HP3* uses a value of $\lambda = 129,600$, suggested as optimal for monthly data in Ravn and Uhlig (1997). The final procedure, denoted *NDT*, does not detrend the data prior to estimation, but instead adds additional terms to the moment condition of (2.64) to account for the presence of deterministic terms. This becomes

$$G(X_t, \theta) = H_t - \beta_1 X_{1t+2} - \beta_2 X_{2t+1} - \beta_3 S_{t+1} - \gamma t - k \quad (2.67)$$

with $\theta = [\beta_1 \ \beta_2 \ \beta_3 \ \gamma \ k]'$ now. This specification uses the same instruments given above but with the addition of time in the instrument vector.

2.5 Empirical Example

Before looking at the results from the simulation exercise, we provide empirical evidence that using the HP filter prior to estimation has an impact on the resulting parameter estimates. Table 2.5 reports the results from estimation of the above models using Bureau of Economic Analysis data on inventories and sales of nondurables. This includes the series for total nondurables that is used in calibrating the simulation. Here we estimate the models using data from three additional two-digit industries, food, petroleum, and chemicals. For each industry we apply the same procedure to be used in the simulation, estimating the model five times using the different detrending methods.

Clearly the detrending method has a significant impact on the results. For total nondurables, the biggest impact on point estimates is for β_3 , which range in value from 0.0114 for *NDT* 0.0607 for *HP1*. There does seem to be little impact on estimation of β_1 and only slightly more for β_2 . For the other industries, however, the range of the estimates for all three parameters tends to vary greatly across detrending methods, with HP filtering significantly altering the results from estimation. In addition, in

nearly all cases the parameters are less precisely estimated when the data is prefiltered with HP, with the standard errors of the estimators being larger than those from estimation using data with a linear trend removed or with no detrending prior to estimation.

Also reported in the table is the value of the J_T statistic testing whether the model is correctly specified. While the model is only rejected in one case, petroleum with NDT, there does seem to be some impact on the value of the statistic across detrending methods, although the impact seems to be slight in most cases.

While this exercise does nothing to indict the use of HP filtering in practice, it does provide motivation for our simulation study by indicating that there are real effects on estimation results when using the filter. In the next section we report the results from the simulation and show that there are indeed adverse consequences when using the HP filter.

2.6 Simulation Results

To assess the impact that HP filtering may have on estimation and test statistics we first look at the distributional characteristics of the parameter estimates resulting from the different detrending methods when estimation is done using simulated data calibrated to inventories and sales of total nondurables. In this section we also consider the impact on another parameter not directly estimated from the model, but still of interest to researchers. An estimate of scaled marginal cost of holding inventories can be constructed from estimates of β_1 and β_2 as

$$\beta_4 = \frac{a_2}{c} = b^{-1} [1 - \beta_1(1 + 4b + b^2) - \beta_2(1 + b)]$$

From the calibrated values used in the simulation, $\beta_4 = 0.0018$. In what follows we report results for this parameter as well.

Figures 2.6 through 2.9 plot the densities of the parameter estimates for each detrending method used. Table 2.6 summarizes their distributional characteristics. From these it is clear that not detrending the data prior to estimation, but instead estimating deterministic terms along with the other parameters, results in estimates that contain less mean bias away from the population values than do the estimates

using detrended data. The use of linearly detrended data results in fairly good performance in terms of bias as well, which is to be expected given that the data was generated to contain a linear trend. Although the HP filter also removes a linear trend, its further action on the data induces a bias away from the population values of the parameters. This is most pronounced when $\lambda=14400$, the most commonly used value when applying the filter to monthly data. As λ is increased to 129600, the bias shrinks toward that of the linearly detrended estimator, but is still considerable, especially for β_3 and β_4 .

The root mean squared error (RMSE), measuring variability of the estimates about the population parameters, indicates similar performance across methods for β_1 and β_2 . For β_3 and β_4 both NDT and LDT result in significantly less volatility in the estimates than does HP filtering.

While the point estimates made using data that is HP filtered tend to be biased, it may be that most reasonably sized confidence intervals for these parameters contain the population values. To check this, we first look at the first and third quartiles of each distribution, given in table 2.6. These give the parameter values between which 50% of the estimates lie, and therefore can be seen as a type of 50% confidence interval.¹⁸

For β_1 and β_2 all methods result in the population values being within the interquartile range. However, for β_3 and β_4 the population values lie outside of the range for all of the estimates obtained from HP filtered data, regardless of the value of λ used in filtering, although the range gets closer as λ increases and the mean bias diminishes.

In addition to bias in the mean, all of the distributions tend to be skewed and have positive kurtosis. For β_1 and β_2 these are only slight deviations from the expected normal distribution when detrending the data prior to estimation, regardless of method. For both β_3 and β_4 the deviation from normality is more pronounced for all methods, with the positive kurtosis particularly extreme for NDT and LDT.

Another area in which the detrending method might have an impact is in performing inference about the parameter of the model based on the estimates. Table 2.7 indicates the rejection rates for two test statistics related to the estimation. The nominal rejection rates for both of these tests is .05. The first test is the t -statistic testing the hypothesis that the estimated parameter value is equal to the

¹⁸This is due to West and Wilcox (1996), who devise such an interval in their simulation study of instrumental variables estimators.

population value. This is equivalent to testing whether the true parameter value lies within a 95% confidence interval constructed about the parameter estimate, so complements the above results using the interquartile range. However, the t -statistics are constructed using asymptotic approximations to the standard errors as generated in the estimation procedure, rather than the actual standard deviations of the 1000 estimates in each case, and so these results are not likely to be consistent with those above. From the graphs of the distributions of the estimates, it is clear that t -statistics constructed using the actual standard deviations would reject far more frequently in most cases. This implies that the approximated standard errors are much larger than the actual standard deviations of our estimates.

In all cases but one the t -test over-rejects, with the exception being the test of β_3 from NDT. The degree of over-rejection is similar between LDT and the three HP specifications for β_1 and β_2 , with slight improvement as λ is increased. For β_3 , HP2 and HP3 outperform LDT, approaching the nominal rejection rate as λ is increased. NDT outperforms the other methods in all cases, with near equality to the nominal rate for β_3 . The poor performance of the t -test when using detrended data is likely due in part to the poor small sample performance of GMM estimation in general, which has been well documented in the econometrics literature. To test this we repeated some of the experiments using a sample size of 10,000. This did improve the performance in all cases.

The second test statistic is the commonly used J-test of instrument-residual orthogonality, which is a test of model specification.¹⁹ This test is distributed χ^2 with degrees of freedom $q - p$, where p is the number of parameters being estimated and q the number of instruments used. For LDT, HP1, HP2, and HP3, we have $q - p = 2$, and for NDT, $q - p = 1$. For this test, all of the methods result in actual size very close to the desired nominal size. There is clearly little impact on the size of this statistic when using HP filtered data.

2.7 Conclusions

The simulation results above provide evidence that the use of the HP filter as a general detrending device is potentially dangerous to estimation results when applied to trend stationary data subsequently used in that estimation. In particular, point estimation was significantly affected. For at least one of

¹⁹See Hansen (1982).

the directly estimated parameters and also for β_4 , constructed from the other two, use of the HP filter induced substantial bias in the parameter estimates away from their true values. While the bias tends to shrink as the smoothing parameter of the filter increases, it is still significant even at the highest value we tested. For the lowest value, $\lambda = 14400$, the one most frequently used in practice, the bias is severe in all of the parameter estimates. We also found increased variability in the parameter estimates, as measured by RMSE, when using HP filtered data, particularly for β_3 and β_4 . This also diminished with increases in the smoothing parameter of the filter.

In addition, the 50% confidence intervals given above only contain the population values within them for two of the parameters when using HP filtered data, regardless of λ . For β_3 and β_4 these confidence intervals were quite distant from the true values. This is again due to the significant bias of these distributions to the right of the true parameter values even for large values of λ . However, this also indicates inference based on estimation using HP filtered data may be questionable. The other two methods of dealing with the trend were clearly superior to HP filtering, in that there was much less bias and also less variability about the population values in most cases. This is, of course, due to the fact that the data were generated with simple linear trends, for which those methods are well suited. An interesting next step, then, would be to repeat the above study using data simulated to contain a unit root or even possibly a piecewise linear trend. The performance of estimators making use of HP filtered data in these instances may be very different from those in this study.

There was little impact on the test statistics we looked at. We found that the use of HP filtered data did not result in the t -test performing significantly worse than with linearly detrended data, particularly for the higher values of λ . However, the size properties of this test were found to be significantly better than either of these when the data was not detrended prior to estimation. The size of the J -test of model specification was virtually identical across detrending methods. Despite this result, it may be interesting in future research to consider the possible impact HP filtering might have on the power of the J -test in detecting misspecification.

The implication of these results is that researchers should take great care in choosing a detrending method to apply to non-stationary data prior to estimation of structural models. Clearly the *ad hoc*

use of the HP filter may adversely affect results gained from estimating such models. We do note, however, that the HP filter performs much better when applied to trend stationary data if the value of the smoothing parameter is set at $\lambda = 129600$, which was suggested as the ideal for monthly data in Ravn and Uhlig (1997). This suggests that making use of unit root testing procedures to determine the stationarity properties of a set of data may help in guiding the choice of λ prior to detrending. This also bears further investigation. In regards to the current study, previous research using the BEA data set we used in calibrating the simulation have tended to remove a linear trend prior to using the data in estimation, although stationarity testing has generated somewhat ambiguous results in some cases.²⁰

²⁰See, for example, Eichenbaum (1989) and Doorn (2001).

Table 2.1: Cost Parameters for Simulation Study

	a_0	a_1	a_2	a_3
Parameter Value	1	5.898	.032	7.191

Table 2.2: Specification of Sales Process and Cost Shock Parameters

ϕ_1	ϕ_2	$\sigma_{\epsilon_S}^2$	σ_u^2	$\rho_{\epsilon_S, u}$
.75	.20	.11625	3.5	-.5

Table 2.3: Implied Parameters of Inventory Process

ρ_1	ρ_2	π_1	π_2
1.06	-.12	.006	-.065

Table 2.4: Implied Values of Regression Coefficients

	β_1	β_2	β_3
Population Value	.056	.332	.013

Table 2.5: Empirical Results

Total					Food			
	J	beta1	beta2	beta3	J	beta1	beta2	beta3
LDT se	0.3617	0.0603 0.0479	0.3178 0.1423	0.0126 0.0071	0.1232	0.0179 0.0693	0.4488 0.2074	0.0797 0.0463
HP1 se	0.5537	0.0568 0.0672	0.3463 0.2184	0.0607 0.0518	0.9762	0.0182 0.1006	0.4312 0.3043	0.1983 0.1792
HP2 se	0.4619	0.0569 0.0553	0.3353 0.1696	0.0347 0.0224	0.4288	0.0048 0.0907	0.4869 0.2774	0.1477 0.1063
HP3 se	0.4519	0.0567 0.0540	0.3348 0.1644	0.0321 0.0199	0.3974	0.0051 0.0938	0.4879 0.2869	0.1408 0.1051
NDT se	0.3929	0.0657 0.0509	0.3015 0.1511	0.0114 0.0078	0.1420	0.0173 0.0672	0.4506 0.2007	0.0800 0.0412

Petroleum					Chemicals			
	J	beta1	beta2	beta3	J	beta1	beta2	beta3
LDT se	5.4690	0.0632 0.0774	0.3012 0.2343	0.0179 0.0158	2.0159	0.0218 0.0432	0.4323 0.1286	0.0123 0.0082
HP1 se	0.5205	0.0465 0.0676	0.3751 0.2217	0.3337 0.1930	1.6980	0.0099 0.0544	0.4856 0.1731	0.0531 0.0350
HP2 se	0.0074	-0.0201 0.1050	0.5769 0.3354	0.2983 0.1806	2.1577	0.0223 0.0483	0.4382 0.1484	0.0309 0.0195
HP3 se	0.0799	-0.0285 0.1082	0.5969 0.3425	0.2680 0.1630	2.3751	0.0280 0.0482	0.4190 0.1475	0.0276 0.0173
NDT se	4.0091	0.0259 0.0808	0.4092 0.2426	0.0216 0.0172	2.1941	0.0302 0.0448	0.4073 0.1328	0.0117 0.0089

Table 2.6: RMSE and Distribution Characteristics of Parameter Estimates

		LDT	HP1	HP2	HP3	NDT
RMSE	β_1	0.0469	0.0473	0.0459	0.0458	0.0459
	β_2	0.1419	0.1462	0.1395	0.1386	0.1318
	β_3	0.0152	0.0654	0.0395	0.0359	0.0149
	β_4	0.0090	0.0352	0.0223	0.0203	0.0090
Mean	$\beta_1 = 0.0560$	0.0594 [0.029, 0.092]	0.0712 [0.045, 0.101]	0.0675 [0.040, 0.098]	0.0666 [0.039, 0.098]	0.0570 [0.031, 0.098]
	$\beta_2 = 0.3320$	0.3213 [0.220, 0.413]	0.2718 [0.184, 0.346]	0.2891 [0.196, 0.369]	0.2927 [0.198, 0.374]	0.3285 [0.200, 0.407]
	$\beta_3 = 0.0130$	0.0163 [0.006, 0.023]	0.0595 [0.029, 0.083]	0.0392 [0.019, 0.055]	0.0361 [0.017, 0.051]	0.0167 [0.003, 0.024]
	$\beta_4 = 0.0018$	0.0042 [0.001, 0.009]	0.0330 [0.024, 0.042]	0.0206 [0.014, 0.027]	0.0187 [0.013, 0.025]	0.0039 [0.000, 0.009]

Table 2.7: Simulation Results

		LDT	HP1	HP2	HP3	NDT
t	β_1	0.1300	0.1407	0.1365	0.1339	0.0800
	β_2	0.1310	0.1711	0.1466	0.1440	0.0800
	β_3	0.0962	0.1022	0.0754	0.0669	0.0460
J		0.0491	0.0445	0.0468	0.0528	0.0580

The t-statistic and J-test entries denote the rejection rate of the statistics, as described in the text. Both have nominal rejection rates of 5%.

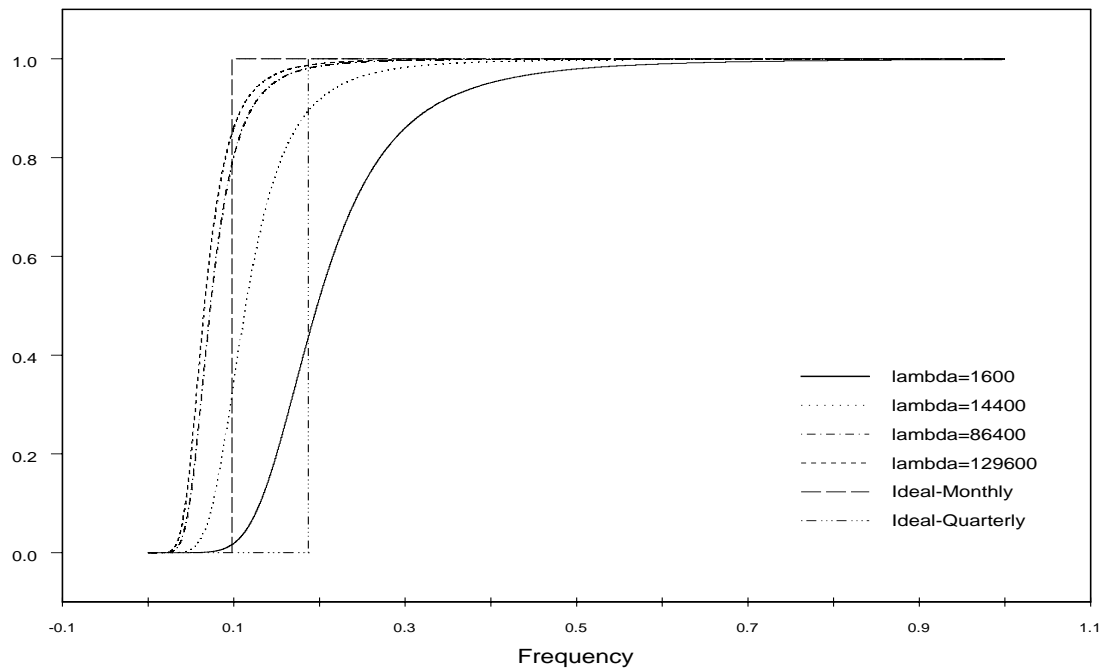


Figure 2.1: **Frequency Response of HP Cyclical Filter** *This plot depicts the squared gain of the filter $HP(\omega)$ for $\lambda = 1600$, $\lambda = 14400$, $\lambda = 86400$, and $\lambda = 129600$. Also shown are the ideal highpass filters for monthly and quarterly data. See text for details.*

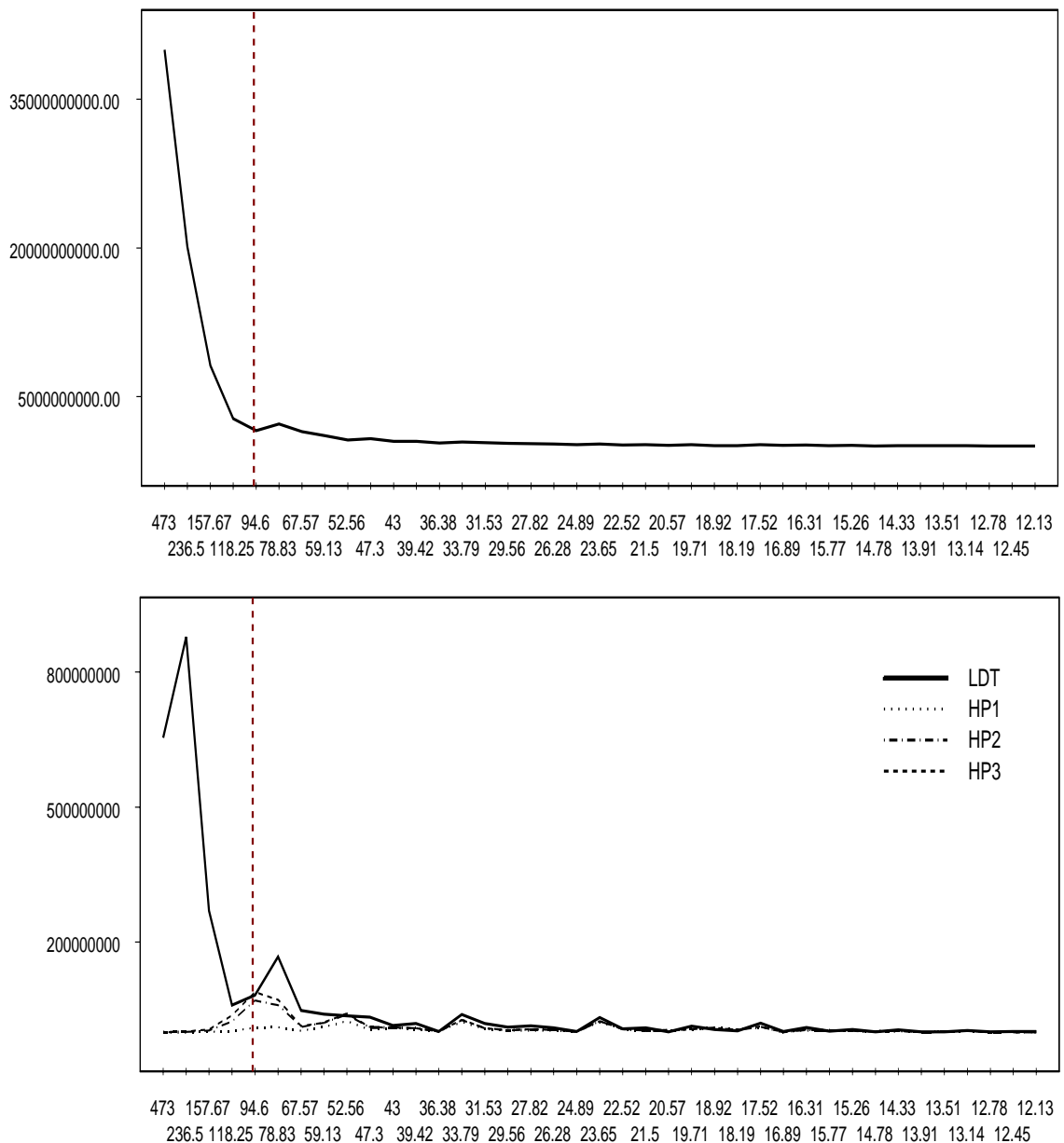


Figure 2.2: **Spectral Density of Total Aggregate Inventories of Non-Durables** *The top panel is the spectral density of the series in levels. The bottom panel shows the spectral densities of the detrended series from the different methods, with LDT indicating linear detrending and HP indicating HP filtering with $\lambda = 14400$ for HP1, $\lambda = 86400$ for HP2, and $\lambda = 129600$ for HP3. The dashed line indicates the cutoff period of eight years, which is generally considered to be the maximum length of the business cycle.*

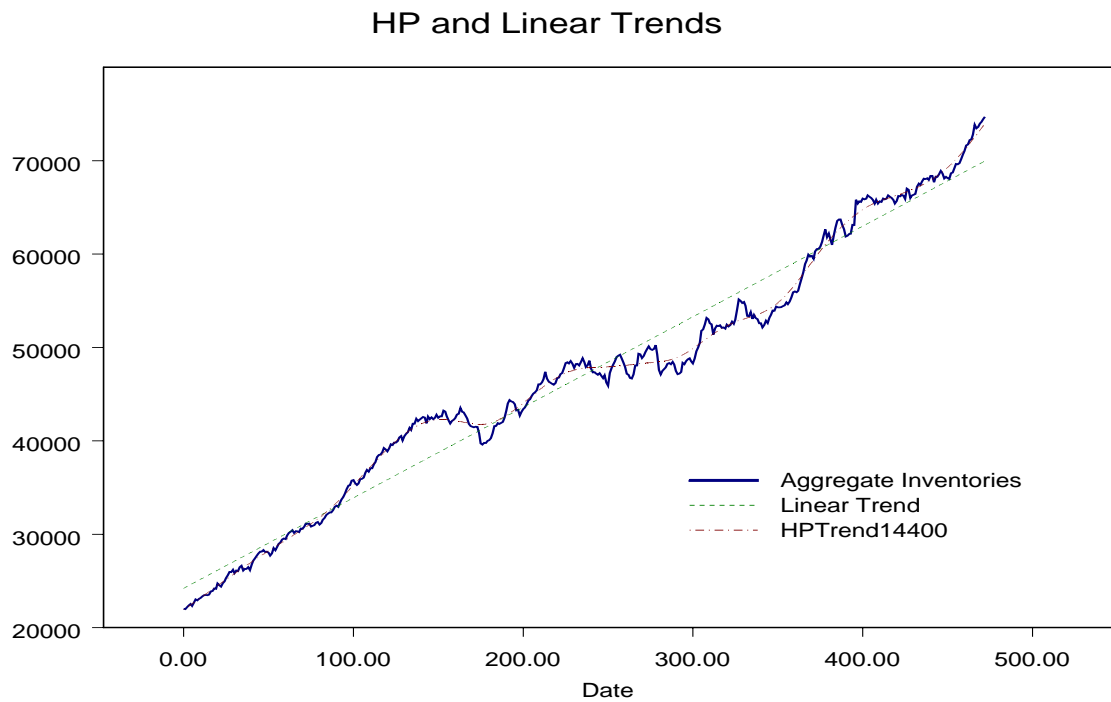


Figure 2.3: **HP and Linear Trends** *This plot superimposes a linear trend along with the HP trend calculated using $\lambda=14400$ over monthly aggregate inventories of nondurable goods. The original series spans January 1959 to May 1998.*

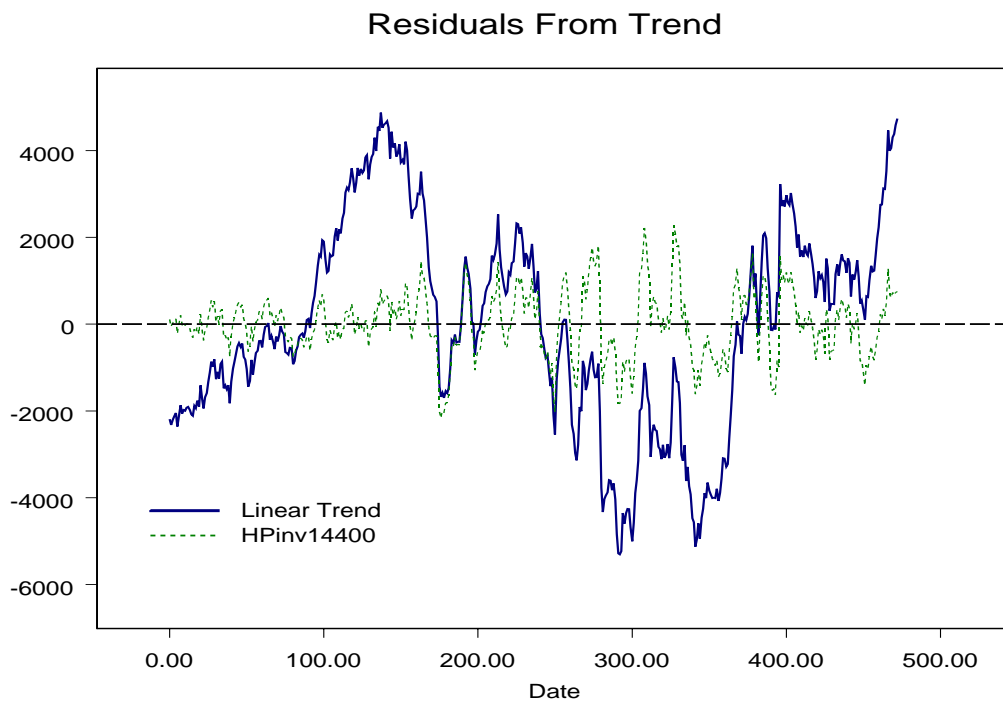


Figure 2.4: **Residuals** A comparison of residuals from detrending monthly aggregate inventories of nondurables with a linear trend as well as with the HP filter using $\lambda=14400$.

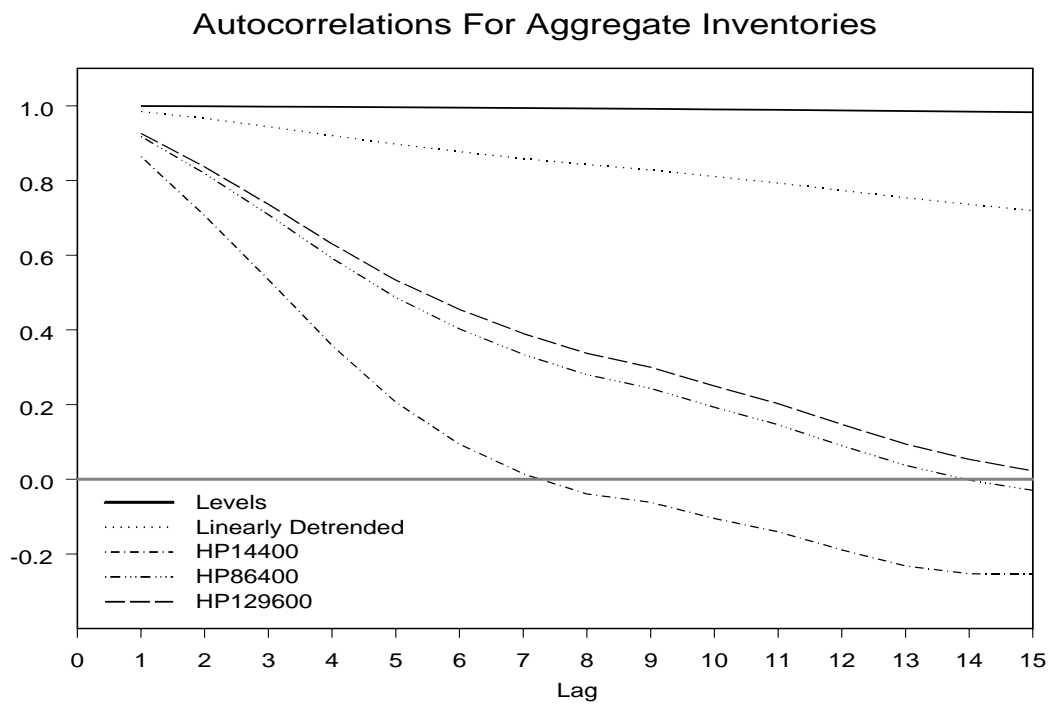


Figure 2.5: **ACF** Autocorrelation function of total aggregate inventories of nondurables for data in levels, linearly detrended, and HP filtered using different values for λ .

Densities for Beta1 Estimates

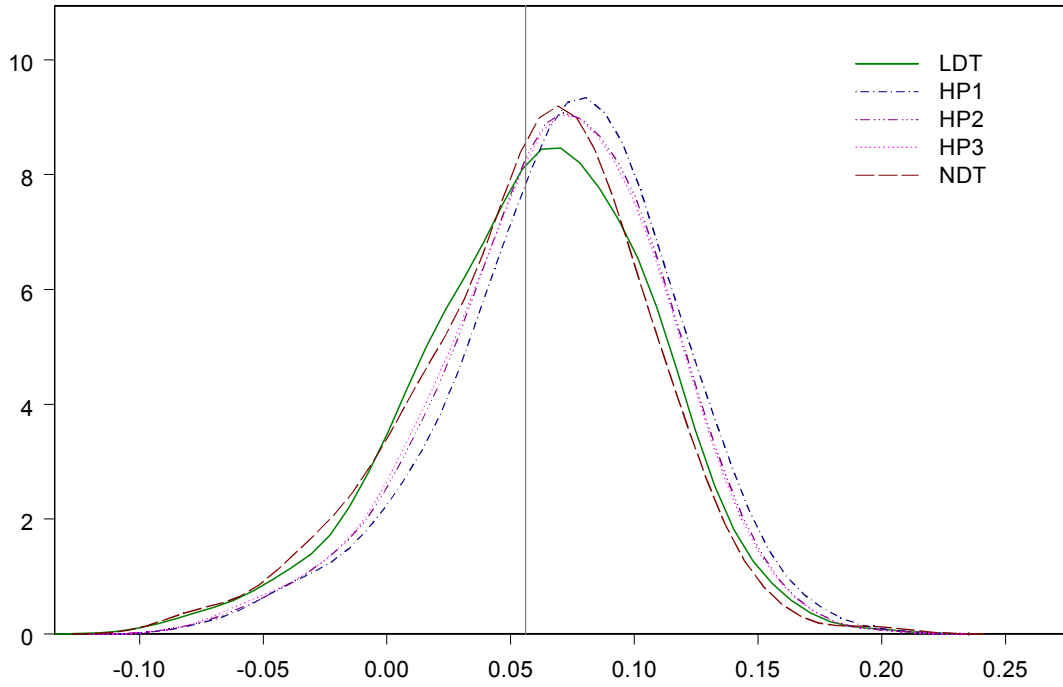


Figure 2.6: **Density of β_1 Estimates.** *The vertical line indicates the population value of 0.056.*

Densities for Beta2 Estimates

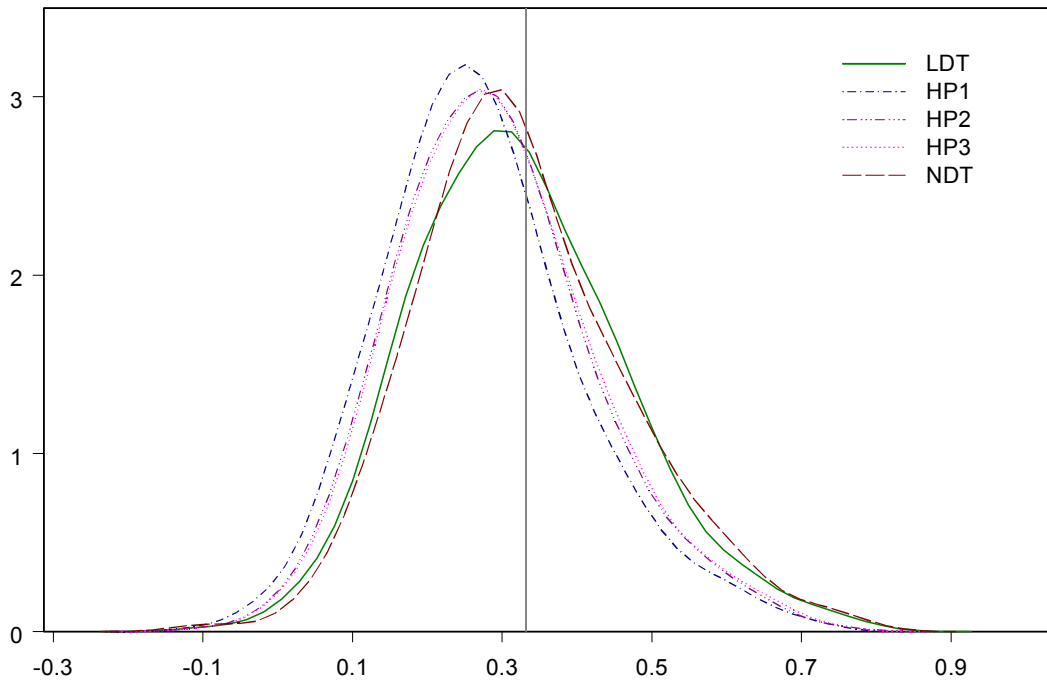


Figure 2.7: Density of β_2 Estimates. The vertical line indicates the population value of 0.332.

Densities for Beta3 Estimates

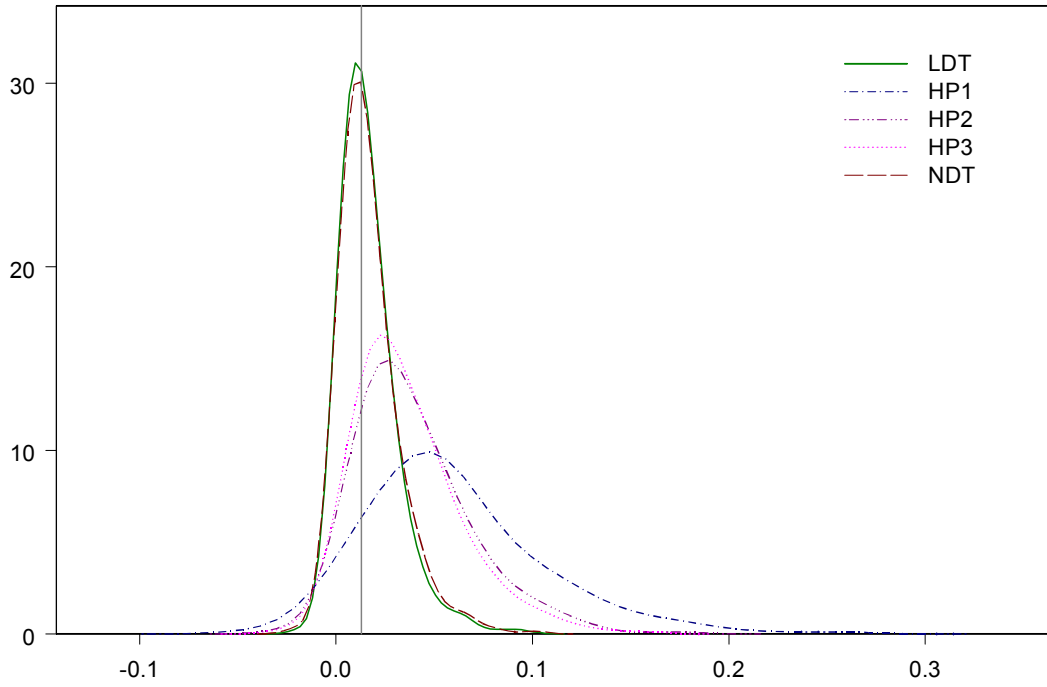


Figure 2.8: Density of β_3 Estimates. The vertical line indicates the population value of 0.013.

Densities for Beta4 Estimates

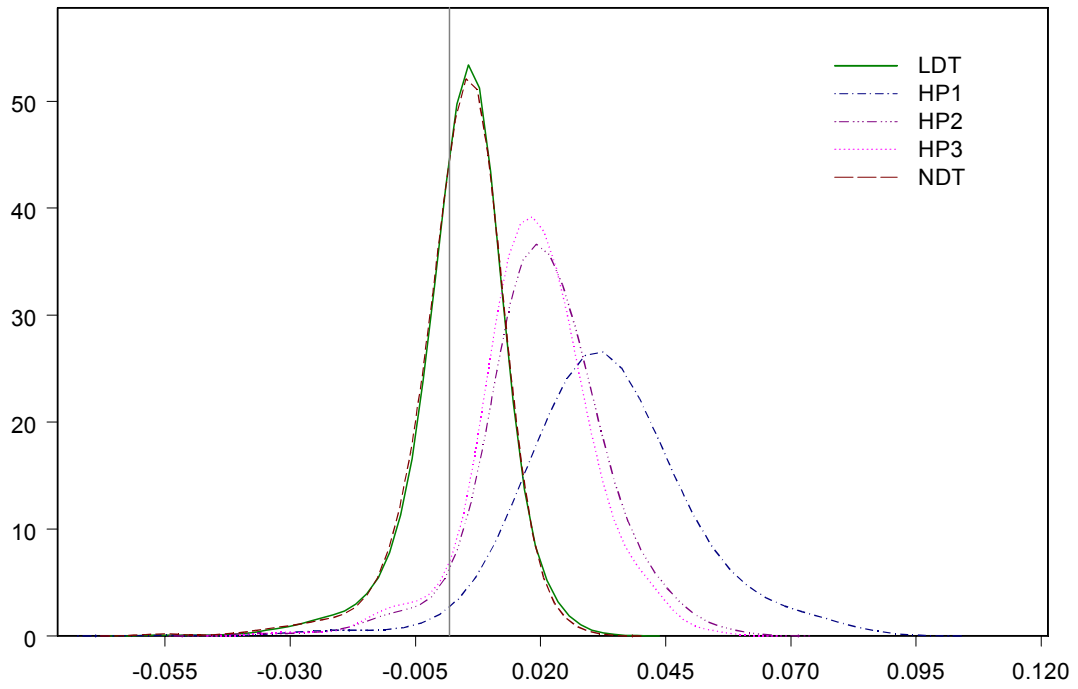


Figure 2.9: **Density of β_4 Estimates.** *The vertical line indicates the population value of 0.0018.*

Chapter 3

Misspecification Issues in Structural Econometric Models

3.1 Introduction

In this paper we study a number of issues pertaining to misspecification in structural econometric models, in particular those estimating Euler equations by instrumental variables techniques. Using a Generalized Method of Moments (GMM) estimation framework, we first present three versions of a macroeconomic model of inventory behavior that has often been studied in the past, with two of the model specifications being restricted versions of the third. The more complete third model is used to generate simulated data which can then be used to investigate several issues of both theoretical and practical interest to researchers. In the sequence of models studied here, the complete model from which the data is generated contains three parameters to be estimated. The two misspecified versions contain only subsets of this complete parameter vector, although these subset parameters represent the same information across models. Because of the non-linear nature of the models studied here, the omission of one or more parameters induces the misspecification we wish to investigate.

The first topic to be considered is the impact such misspecification may have on the probability limits of the parameters to be estimated. Theoretically, misspecification of the type studied here should

induce a bias in the probability limit of the estimator. Using large sample simulated data we estimate the probability limits of the parameters for each of the misspecified models in an effort to investigate the seriousness of such bias. The degree of difference between the true parameter values and their misspecified model probability limits gives us an indication of how misspecification can distort the conclusions that may be arrived at in empirical research.

We next turn to specification testing. We investigate the ability of a commonly used specification test to detect the misspecification present in our models. The J_T test is often the only means of testing model specification that is undertaken by researchers using GMM, with the results being relied upon almost without question in many instances, despite the fact that previous research that has shown the J_T test to have low power against the alternative of misspecification. Using our experimental design to look at the size properties of the J_T statistic, we are able to see how well the statistic performs in rejecting those models which are misspecified. In addition, we look at a recently proposed alternative means of choosing between nested models, the noise ratio method of Durlauf and Maccini (1995). Again our setup allows us to analyze how effective this method is in choosing the correct model.

Finally, the rest of the paper considers recent work by Hall and Inoue (2003). Their paper derives a new asymptotic theory for models which are misspecified. Conventional GMM asymptotic theory is dependent upon the assumption that a model is correctly specified to begin with. If the model is misspecified, then conventional asymptotics do not hold and the ability to conduct inference on model parameters is greatly diminished. We test the impact of using the new asymptotic theory on construction of standard errors and on the performance of t -tests based on these standard errors. This performance is compared to that based on conventionally estimated standard errors in order to determine whether the new methods provide a useful substitute.

To motivate the study, we recognize that, while there are numerous specification tests which can be used to determine whether a particular model being applied is consistent with the empirical behavior of a set of data, results from such tests are often ignored in applications. Ideally, a finding of misspecification should lead to a respecification of the model in an attempt to give it a better fit to the data. In practice, however, many researchers have proceeded to perform inference based on estimation results despite the

rejection of their model by specification tests. Table 3.1 lists a number of papers in which this has been done. For some of these, the authors simply report GMM estimates and standard errors despite evidence that the model is misspecified. However, many of the studies also discuss behavioral implications of the parameter estimates or use such estimates to conduct hypothesis tests about population parameters. In general, such inference is not valid. The distributional characteristics of hypothesis tests based on estimated parameters, i.e. Wald, LM, or Likelihood-ratio type tests, rely upon the estimated model having been correctly specified to begin with. If the model is not correctly specified, then these statistics do not necessarily follow the usual distributions.

In what follows we show that the effects of misspecification make conclusions drawn on estimation results from misspecified models, such as those contained in the papers listed in table 3.1, are suspect at best. The results of our simulation study confirm that misspecification of the type studied here does indeed induce bias in the probability limits of the parameters. Often this bias is severe enough to cause statements about the parameters that are based on point estimates to be very misleading. This is potentially hazardous, especially when the goal is to measure important economic phenomena, such as the speed of adjustment parameter in a flexible accelerator model. Furthermore, we find the J_T statistic to be a poor guide in selecting between alternative models when some of those models are misspecified. Our results provide evidence that the noise-ratio methods of Durlauf and Maccini (1995) perform much better in this regard. Finally, because conventional asymptotic theory does not apply when a model is misspecified, the use of an alternative asymptotic theory may be required to accurately perform inference on parameters from a misspecified model. Our results show that conventional asymptotic approximations for standard errors tend to be grossly inaccurate relative to the true standard deviations of parameter estimates in our simulation. While the alternative asymptotic theory of Hall and Inoue (2003) does seem to provide closer approximations to the actual standard deviations of the estimates in many of our experiments, our results for different calibrations are not consistent enough to provide good evidence in support of its use.

3.2 GMM Background and Definitions

Before outlining the procedures to be undertaken, a brief discussion of the properties of the GMM estimator will be useful. Suppose we have an econometric model which implies a $q \times 1$ vector of population moment conditions based on observable variables and certain parameters defined by the model that can be expressed

$$E[g(x_t, \theta_0)] = 0 \tag{3.1}$$

Here assume x_t is an n -dimensional random vector of observed variables, which are assumed strictly stationary over the sample space $\mathcal{X} \supseteq \mathbb{R}^n$, and θ_0 denotes the population value of a p -dimensional vector of unknown parameters θ , with $\theta \in \Theta \supseteq \mathbb{R}^p$. Then $g : \mathcal{X} \times \Theta \rightarrow \mathbb{R}^q$.

The GMM estimator of θ_0 is defined as

$$\hat{\theta}_T = \arg \min_{\theta \in \Theta} Q_T(\theta) \tag{3.2}$$

where

$$Q_T(\theta) = \left\{ T^{-1} \sum_{t=1}^T g(x_t, \theta) \right\}' W_T \left\{ T^{-1} \sum_{t=1}^T g(x_t, \theta) \right\} \tag{3.3}$$

with $T^{-1} \sum_{t=1}^T g(x_t, \theta)$ being the sample analog to (3.1) and W_T being a weighting matrix of the appropriate dimensions.

The only requirement for the weighting matrix in correctly specified models is that it be positive semi-definite and converge to a positive definite matrix of constants, W , as $T \rightarrow \infty$. Hansen (1982) has shown that the optimal weighting matrix will be the inverse of the asymptotic covariance matrix of the sample moment. Of course to calculate this we need a consistent estimate of θ_0 . Because of this, GMM generally is implemented using either a two-step estimation procedure or an iterative procedure to arrive at an estimate for θ . Both procedures use a first-step estimator in which any suitable positive-definite matrix, usually the identity matrix or the inverse of the instrument cross products matrix, is used to get a consistent estimate for θ_0 . This first-step estimate is then used to derive a weighting matrix that is asymptotically efficient which can in turn be used to find a more efficient estimate of θ_0 . The two-step estimator stops here, while the iterated GMM estimator continues, using the new estimates to derive a new weighting matrix at each step until the parameter estimates converge.

In this context, a model is considered correctly specified when there is a unique value θ_0 for which $E[g(x_t, \theta_0)] = 0$ holds true. Hansen (1982) shows that for correctly specified models, $\hat{\theta}_T$ converges in probability to θ_0 and further that $T^{1/2}(\hat{\theta}_T - \theta_0)$ converges in distribution to a mean zero normal random vector, with covariance matrix denoted Σ_0 . Hall and Inoue (2003) show that these properties do not necessarily hold in the case of misspecified models.

It is useful at this point to clarify the definition of misspecification in terms of the GMM estimator. Misspecification may be of two forms, either local or non-local. Local misspecification occurs when the data do not satisfy the PMC being estimated for any finite sample, but do in the limit. In this case the data is assumed to drift toward the estimated model in such a manner that the misspecification only affects the mean of the asymptotic distribution of the estimator. All other asymptotic properties of the estimator and the weighting matrix are indistinguishable from the correctly specified case. The effect of this is that there is some unique $\tilde{\theta}_T \in \Theta$ such that $\tilde{\theta}_T = \theta_0 + T^{-1/2}\eta$ so that $E[g(x_t, \tilde{\theta}_T)] = 0$ holds for the misspecified model, but the asymptotic distribution of the estimator becomes $T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(\eta, \Sigma_0)$. This results in local misspecification being observationally equivalent to a correctly specified model, since $E[g(x_t, \tilde{\theta})] = 0$ and $E[g(x_t, \theta_0)] = 0$.

Non-local misspecification occurs when there is no value of $\theta \in \Theta$ for which the PMC holds. Hall and Inoue (2003) show that this can only happen if the parameter vector is overidentified in the PMC, or $q > p$. For this situation to be of interest, there must still exist some value of the parameter vector, say θ_* , which minimizes the GMM objective function defined in (3.3), but this minimum is not zero. This can be expressed $E[g(x_t, \theta)] = \mu(\theta)$, where $\mu : \Theta \rightarrow \mathfrak{R}^q$ and $\|\mu(\theta)\| > 0 \forall \theta \in \Theta$, so that $Q_T(\theta_*) < Q_T(\theta) \forall \theta \in \Theta \setminus \theta_*$. This identification condition is unlikely to be satisfied by the same parameter values when using different weighting matrices. This dependence on the weighting matrix then is likely to cause the distributional characteristics of the estimator to depend on choice of W_T in misspecified models. Because of this, Hall and Inoue (2003) derive the limiting distributions of the GMM estimator in misspecified models for a number of choices of W_T that are common in applied work. This study will focus on two of these choices.

Before discussing the details of these, we will first lay out the simulation design and the models

to be estimated. This will provide some context for the methods to be used and also will provide an opportunity to address some additional issues concerning GMM and misspecified models before getting into the results from application of the theory. In addition, it will give us a chance to investigate the performance of a widely used test for misspecification, along with an alternative method of choosing between competing models.

3.3 The Model

It has been posited that firms hold inventories in order to smooth production over time, that is to meet demand shocks out of inventories so as to not incur additional costs, through either changing production or lost sales. This production smoothing motive is the basis for the linear-quadratic model of inventory behavior, which seeks to model such behavior in a way that is consistent with empirical observations on inventory accumulation. In the model's most basic form, usually termed the "pure" production smoothing model, it has been unsuccessful. Model specification tests have found that the pure model does a poor job of explaining the data. Because of this the pure model has been expanded to include other possible motivations for inventory accumulation, along with other factors which may affect it. Some of these new specifications have been found to perform better than others in explaining the data, particularly in specification tests.

We first set up a particular specification of the model, which is essentially the model of West and Wilcox (1996). This initial specification will later be used to simulate the data, thus making this the correctly specified "full" version in our experiments. Various aspects of the model are then modified to give versions that are then misspecified given the data generating process.

3.3.1 The Full Model

The model we use is that of West and Wilcox (1996) and West (1992), which is based on Blanchard (1983). The basic model assumes a representative firm that chooses inventory levels to maximize the present discounted value of profits subject to some cost function. The particular elements of the cost function are what will differ between specifications, reflecting the inclusion or omission of variables

possibly affecting the inventory holding decision. We begin with

$$\max_{\{H_{t+j}\}} \lim_{T \rightarrow \infty} E_t \sum_{j=0}^T b^j (p_{t+j} S_{t+j} - C_{t+j})$$

s.t.

$$C_{t+j} = \frac{a_0}{2} \Delta Q_{t+j}^2 + \frac{a_1}{2} Q_{t+j}^2 + \frac{a_2}{2} (H_{t+j-1} - a_3 S_{t+j})^2 + H_{t+j} u_{t+j} + \text{deterministic terms} \quad (3.4)$$

$$Q_{t+j} = S_{t+j} + H_{t+j} - H_{t+j-1} \quad (3.5)$$

where p_{t+j} is price, S_{t+j} is real sales, H_{t+j} real end of period inventories, C_{t+j} real costs, and Q_{t+j} real output. For simplicity we assume revenues exogenous, so the above is equivalent to a cost minimization problem. The a terms represent the marginal costs faced by the firm, with a_0 the marginal cost of changing production, a_1 the marginal cost of production, a_2 the marginal cost of holding inventories, and a_3 the cost associated with stockouts. Unobserved cost shocks are represented by u_{t+j} , with u_t assumed to be IID Normal with zero mean and uncorrelated over time.¹ Deterministic terms would account for any linear and trend terms that may exist in the data. The final term, b , is the discount factor and is usually set *a priori* in estimation of the model. The parameters of the cost function, or some combination thereof, are generally what is estimated, and therefore these are what will need to be calibrated in the simulation.

Both $a_0 > 0$ and $a_1 > 0$ imply a production smoothing motive in holding inventories, as additional production this period will result in additional costs to the firm, both directly and possibly through having to change production level. Thus demand shocks will be met, at least partially, out of inventories. This motivation is unchanged by assuming $a_0 = 0$, which would imply that shifts in production level do not impose additional costs to the firm. For $a_3 = 0$, $a_2 > 0$ implies positive costs associated with holding additional inventories, which then have to be weighed against additional production costs in the event of a demand shock. When $a_3 > 0$, there is an additional motivation for inventory holding. That is to avoid the risk of having a shortfall in stocks in the face of unexpected demand, which in turn could cause the firm to lose customers or to have to backlog orders. This induces firms to weigh the cost

¹Specifying cost shocks as being uncorrelated is a simplification to facilitate the simulation of the data. This will leave the estimating equations unchanged in what follows from the case where there are no cost shocks, but does affect covariance matrix estimation by inducing particular correlation structures on the residuals from these equations.

of holding inventories against that of stock-outs, and is thus called the stock-out avoidance, or buffer stock, motive. The term a_3S_{t+j} can be seen as some target level of inventories, expressed as a fraction of expected future sales, since H_t is end of period inventories. When $H_{t+j-1} = a_3S_{t+j}$ the target level of inventories is met and no stock-out costs are incurred. When $H_{t+j-1} > a_3S_{t+j}$ additional costs are incurred because inventories are above the desired level. If $H_{t+j-1} < a_3S_{t+j}$ there is a shortfall in stocks and additional costs are again incurred by the firm.

Finally, the existence of cost shocks in the model reflects the possibility that inventory holdings may be affected by unexpected changes in the cost of inputs, with an unexpected increase in costs possibly causing firms to reduce inventories to meet sales rather than incur additional costs of production. We model these cost shocks as unobservable for two reasons. The first is that a number of papers seeking to include costs of specific inputs, such as labor or energy, have found the associated parameter estimates to be insignificant², and the second reason being that the difficulty in modeling additional series in our simulation setting would likely outweigh its usefulness, given the particular aims of this study.

The first order condition of the above model yields:

$$E_t\{a_0(\Delta Q_t - 2b\Delta Q_{t+1} + b^2\Delta Q_{t+2}) + a_1(Q_t - bQ_{t+1}) + ba_2(H_t - a_3S_{t+1}) + \text{deterministic terms} + u_t\} = 0 \quad (3.6)$$

Ignoring deterministic terms, if we substitute in for production from equation 3.5 and then collect terms, we have

$$E_t\{a_0b^2H_{t+2} - [a_0(2b^2 + 2b) + a_1b]H_{t+1} + [a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2]H_t - [a_0(2b + 2) + a_1]H_{t-1} + a_0H_{t-2} + a_0b^2S_{t+2} - [a_0(b^2 + 2b) + a_1b + a_2a_3b]S_{t+1} + [a_1 + a_0(1 + 2b)]S_t - a_0S_{t-1} + u_t\} = 0 \quad (3.7)$$

This first order condition is the basis for data simulation as well as for presenting the alternative specifications given below.

3.3.2 Alternative Specifications

It may be noted from 3.6 that the parameters of the cost function are not fully identified by the Euler equation, as any simultaneous scaling of a_0 , a_1 , and a_2 would have no effect on the Euler condition.

²With the notable exception of Durlauf and Maccini (1995).

Because of this we are reduced to estimating combinations of parameters. We follow Eichenbaum (1989) and Durlauf and Maccini (1995) in expressing the first order condition as an expectational difference equation, which is then used to reparameterize the euler equation using characteristic roots of the appropriate function. This approach allows us to estimate a parameter analogous to the speed of adjustment parameter in a flexible-accelerator model. The resulting equations for each specification are expressed in the form of conditional moment restrictions that will allow the parameters to be estimated by GMM.

Pure Production Smoothing

The pure production smoothing model does not include costs of changing production or a stock-out avoidance motive, thus $a_0 = a_3 = 0$ in the above cost function. This results in the first order condition becoming

$$E_t\{-a_1bH_{t+1} + [a_1(1+b) + ba_2]H_t - a_1H_{t-1} - a_1bS_{t+1} + a_1S_t + u_t\} = 0 \quad (3.8)$$

Rearranging the Euler equation in 3.8 and expressing it as the second order difference equation

$$E_t[(1 - \lambda L)(1 - (b\lambda)^{-1}L)H_{t+1}] = E_t[b^{-1}S_t - S_{t+1} + a_1^{-1}b^{-1}u_t] \quad (3.9)$$

with λ and $(b\lambda)^{-1}$ being the roots of the characteristic equation: $1 - \Phi L + b^{-1}L^2$, where

$$\Phi = \frac{a_1(1+b) + ba_2}{a_1b} \quad (3.10)$$

and

$$\lambda = \frac{1}{2}\{\Phi + b^{-1}(\Phi^2b^2 - 4b)^{1/2}\} \quad (3.11)$$

Expanding the left-hand side of 3.9 and rearranging gives

$$E_t\{[1 - [\lambda + (b\lambda)^{-1}]L + b^{-1}L^2]H_{t+1} + S_{t+1} - b^{-1}S_t\} = E_t\{a_1^{-1}b^{-1}u_t\} = 0 \quad (3.12)$$

which further becomes

$$E_t\{H_{t+1} - [\lambda + (b\lambda)^{-1}]H_t + b^{-1}H_{t-1} + S_{t+1} - b^{-1}S_t\} = 0 \quad (3.13)$$

or

$$E_t\{b\lambda H_{t+1} - [1 + b\lambda^2]H_t + \lambda H_{t-1} + b\lambda S_{t+1} - \lambda S_t\} = 0 \quad (3.14)$$

from which we can estimate λ .

This form of the model is useful in that $1 - \lambda$ can be interpreted as the speed of adjustment of inventories to some desired level in a stock adjustment model for inventory accumulation.³ A large value for λ implies a slow adjustment of stocks toward their desired levels, e.g. a value of $\lambda = .8$ would imply that inventories move toward their target levels at a rate of only 20% per period. Most estimates of this adjustment speed in the literature have found it to imply implausibly low speeds of adjustment given the production capabilities of manufacturers.

Stock-out Avoidance

This specification adds the stock-out avoidance motive to the cost function of the pure model by relaxing the assumption that $a_3 = 0$ in the original cost function in (3.4). This specification is analogous to the production smoothing model of Eichenbaum (1989). With $a_3 \neq 0$ the first order condition becomes

$$E_t\{-a_1bH_{t+1} + [a_1(1+b) + ba_2]H_t - a_1H_{t-1} - [a_1b + a_2a_3b]S_{t+1} + a_1S_t + u_t\} = 0 \quad (3.15)$$

Again we can derive a version of this specification with which to estimate a speed of adjustment parameter. Applying the same methods as for the pure production smoothing model above we get

$$E_t\{b\lambda H_{t+1} - [1 + b\lambda^2]H_t + \lambda H_{t-1} + b\lambda\psi S_{t+1} - \lambda S_t\} = 0 \quad (3.16)$$

where λ and Φ are as defined above and $\psi = 1 + a_2a_3/a_1$. This specification can be used to estimate λ and ψ .

The Full Model Respecified

The full model includes all of the parameters and the unobserved cost shock. The first order condition is thus that given in 3.7. Alternatively, the full model can be expressed in terms of the same λ and ψ as above. To see this, first expand the terms on H_{t+1} , H_t , and H_{t-1} in the Euler condition:

$$E_t\{a_0b^2H_{t+2} - a_0(2b^2 + 2b)H_{t+1} - a_1bH_{t+1} + a_0(1 + 4b + b^2)H_t + [a_1(1+b) + ba_2]H_t - a_0(2b+2)H_{t-1} - a_1H_{t-1}$$

³A basic version of the stock adjustment model is $H_t = (1 - \lambda)[H_t^* - H_{t-1}]$. Here H_t^* is some desired level of inventories, usually dependent on E_tS_t , and $1 - \lambda$ the per period speed of adjustment towards that level. See Eichenbaum (1989) for a discussion of the equivalency of these models.

$$+a_0H_{t-2} + a_0b^2S_{t+2} - [a_0(b^2 + 2b) + a_1b + a_2a_3b]S_{t+1} + [a_1 + a_0(1 + 2b)]S_t - a_0S_{t-1} + u_t\} = 0 \quad (3.17)$$

Divide through by a_1b to get

$$E_t\{\delta bH_{t+2} - \delta(2b+2)H_{t+1} - H_{t+1} + \delta b^{-1}(1+4b+b^2)H_t + \frac{a_1(1+b) + ba_2}{a_1b}H_t - \delta b^{-1}(2b+2)H_{t-1} - b^{-1}H_{t-1} \\ + \delta b^{-1}H_{t-2} + \delta bS_{t+2} - [\delta(b+2) + \psi]S_{t+1} + [b^{-1} + \delta b^{-1}(1+2b)]S_t - \delta b^{-1}S_{t-1} + \frac{u_t}{a_1b}\} = 0 \quad (3.18)$$

where $\delta = a_0/a_1$ and $\psi = 1 + a_2a_3/a_1$.

Noting that

$$-\{H_{t+1} - \frac{a_1(1+b) - ba_2}{a_1b}H_t + b^{-1}H_{t-1}\} = -[1 - [\lambda + (b\lambda)^{-1}]L + b^{-1}L^2]H_{t+1} \quad (3.19)$$

$$= -\{H_{t+1} - [\lambda + (b\lambda)^{-1}]H_t + b^{-1}H_{t-1}\} \quad (3.20)$$

from above, we can substitute this into our equation and multiply through by λb to get

$$E_t\{\lambda b^2\delta H_{t+2} - \lambda b[1 + (2b+2)\delta]H_{t+1} + [1 + b\lambda^2 + \lambda(1+4b+b^2)\delta]H_t - \lambda[1 + (2b+2)\delta]H_{t-1} \\ + \lambda\delta H_{t-2} + \lambda b^2\delta S_{t+2} - \lambda b[(b+2)\delta + \psi]S_{t+1} + \lambda[1 + (1+2b)\delta]S_t - \lambda\delta S_{t-1} + \lambda\frac{u_t}{a_1}\} = 0 \quad (3.21)$$

From this we can estimate λ , δ , and ψ .

This gives three alternative specifications of the model, the correctly specified version, to be used in generating the simulated data, and two misspecified versions. The parameters to be estimated in each specification are the same in that they are equivalent functions of the underlying cost parameters from the true model. The parameter λ appears in all three models, and so ideally we should get the same estimate from each. The parameter ψ shows up in the correctly specified model as well as the stock-adjustment version. Finally, δ is only present in the correct specification of the model.

3.4 A Note on Moment Restrictions

Note that the estimating equations in (3.14), (3.16), and (3.21) imply conditional moment restrictions of the form

$$E_t[f(x_t, \theta_0)] = 0 \quad (3.22)$$

where $E_t[\bullet] = E[\bullet|\Omega_t]$ denotes the conditional expectation based on information available at time t , with Ω_t the information set. The vectors x_t and θ in (3.22) will differ between specifications of the model. The nature of the conditional expectation implies that $f(x_t, \theta_0)$ is orthogonal to any variables in the information set, Ω_t , including any lagged endogenous variables. This implies that a set of unconditional moment restrictions can be written

$$E[z_t e_t] = 0 \tag{3.23}$$

where $z_t \in \Omega_t$ is a vector of instrumental variables and $e_t = f(x_t, \theta_0)$ can be thought of as the expectational error from estimation of the model. This orthogonality condition is then the basis for instrumental variables estimation in the GMM framework, since this now represents a PMC of the form in (3.1) with $E[g(x_t, \theta_0)] = E[z_t e_t]$ now.

From this discussion, it is clear that when a model is misspecified in the non-local sense the conditional moment restrictions of (3.22) are not satisfied. Instead we have

$$E_t[f(x_t, \theta_*)] = \mu(\theta) \tag{3.24}$$

Despite this effect on the conditional expectation, it may be possible that there are situations in which there is non-local misspecification, but the unconditional moment restriction in (3.23) still holds. In this case, $E_t[e_t] = \mu(\theta)$ and $\mu(\theta)$ would contain some function of the omitted parameters and variables in misspecified models like those above. If the chosen instruments in z_t happen to be orthogonal to this function then (3.23) is satisfied. In that event, the asymptotic theory that applies in the correctly specified case will carry through and inference can be carried out in the usual fashion. It would be useful, then, to derive a set of conditions under which this might happen. This will be done in the context of the above model and its alternative specifications.

To begin, we need to attain an explicit expression for $e_t = \mu(\theta)$, the true residual from the conditional expectation, from each of the misspecified models. This is easily done by recognizing that the pure model and the stockout avoidance model are both nested within the true model. It is most straightforward to derive the residual for the stockout avoidance model first. Rearranging (3.21) we have

$$E_t\{\lambda b H_{t+1} - [1 + b\lambda^2]H_t + \lambda H_{t-1} + \lambda b\psi S_{t+1} - \lambda S_t\}$$

$$\begin{aligned}
&= E_t\{\lambda b^2\delta H_{t+2} - \lambda\delta b(2b+2)H_{t+1} + \lambda\delta(1+4b+b^2)H_t - \lambda\delta(2b+2)H_{t-1} \\
&\quad + \lambda\delta H_{t-2} + \lambda b^2\delta S_{t+2} - \lambda\delta b(b+2)S_{t+1} + \lambda\delta(1+2b)S_t - \lambda\delta S_{t-1} + \lambda\frac{u_t}{a_1}\} \\
&= E_t[e_t^{SA}]
\end{aligned}$$

The equation on the left side is just the stockout avoidance model in (3.16), so the residual from this model contains those elements from the full model that are due to the underlying cost parameters unaccounted for in this version. In a similar fashion, for the pure production smoothing model we have

$$E_t\{b\lambda H_{t+1} - [1 + b\lambda^2]H_t + \lambda H_{t-1} + b\lambda S_{t+1} - \lambda S_t\} = E_t[e_t^{SA} + \lambda b(1 - \psi)S_{t+1}] = E_t[e_t^P]. \quad (3.25)$$

For (3.23) to be satisfied in the stockout avoidance model, we must have $E[z_t e_t^{SA}] = 0$. One trivial way this can happen is if $\delta = 0$, causing $E_t[e_t^{SA}] = 0$. Recalling that $\delta = a_0/a_1$, this only happens if $a_0 = 0$, meaning there are no costs associated with changing production. This, of course, is one of the assumptions underlying the stockout avoidance version of the model as specified above, but we know it is not the case for the full model from which the data is generated. Similarly, for the pure model we will have $E[z_t e_t^{SA}] = 0$ if $\delta = 0$ and $\psi = 1$, causing $E_t[e_t^P] = 0$. This will happen if $a_0 = 0$ at the same time that $a_2 = 0$ or $a_3 = 0$ (or both), since $\psi = 1 + \frac{a_2 a_3}{a_1}$. Recall that a_2 is the marginal cost of holding inventories. If this cost were zero there would be no reason to go further in the exercise of estimating the model. Note from (3.10) and (3.11) that setting $a_2 = 0$ results in λ being a function solely of the discount factor b , leaving little reason to do the estimation. The a_3 parameter, on the other hand represents stockout costs and could very well be zero. However, we again know this not to be the case in the current situation where the data generating process is known to us.

Another case in which we could achieve (3.23) even though $E_t[f(x_t, \theta_*)] \neq 0$ in the misspecified versions of the model is for the instruments to be uncorrelated with the residuals, e_t^{SA} or e_t^P . Note that both e_t^{SA} and e_t^P contain within them regressors from the conditional moment restrictions upon which estimation is to be based. This implies that instruments which are orthogonal to, or even weakly correlated with, these residuals would also be poor predictors of the regressors in the moment specifications used in estimation. These weak instruments will then cause numerous other problems in estimation and any inference based upon the results, since the sampling distributions of GMM statistics are no longer

normally distributed in the presence of weak instruments.⁴ The point here is that if we can satisfy (3.23) when the model is misspecified in the way presented here, it is likely due to the use of weak or irrelevant instruments in constructing the unconditional moment restriction. The result is that standard GMM asymptotics do not then apply anyway, and inference done in the normal fashion will still be misleading. This makes standard GMM comparison of estimation results across nested versions of the same model, as done in Durlauf and Maccini (1995), inaccurate at best.

It is often the case that the instruments chosen in GMM estimation include lagged endogenous variables. For the present model, previous work has included two or more lags of both sales and inventories in the instrument set. In what follows we will continue this practice and use $z_t = [1 \ H_{t-1} \ H_{t-2} \ S_{t-1} \ S_{t-2}]$. Because these instruments are clearly in Ω_t they are valid and for correctly specified models we would likely achieve (3.23). However, because both e_t^{SA} and e_t^P contain H_{t-1} , H_{t-2} , and S_{t-1} , it is unlikely that we will achieve orthogonality between z_t and e_t in the models considered. There is no reason to believe that the particular combinations of variables and parameters contained in the residuals would lead to $E_t[e_t^{SA}] = 0$ and $E_t[e_t^P] = 0$. So again (3.23) will not be satisfied and inference based on standard GMM asymptotics will be incorrect.

3.5 Simulating the Data

In order to use the full model to simulate data we need to develop equations describing the data generating process for both inventories and sales that are consistent with the model. Following West and Wilcox (1996), we assume sales to be exogenous and described by an AR(2) process:

$$S_t = KS + \alpha t + \phi_1 S_{t-1} + \phi_2 S_{t-2} + \epsilon_{St} \quad (3.26)$$

where KS and αt are constant plus trend terms and ϵ_{St} is assumed IID Normal.

With sales exogenous, we can represent the Euler equation from the full model as a fourth order stochastic difference equation in inventories. This is most easily seen in equation 3.7, which can be

⁴See Stock et al. (2002) for a good survey of recent work on weak instruments and their impact on estimation and inference in standard GMM and instrumental variables estimation, including methods to circumvent some of these problems.

represented as

$$E_t\{A(L)H_{t+2}\} = E_t\{D_t\} \quad (3.27)$$

where

$$\begin{aligned} A(L) &= 1 - a_0^{-1}b^{-2}[a_0(2b^2 + 2b) + a_1b]L + a_0^{-1}b^{-2}[a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2]L^2 \\ &\quad - a_0^{-1}b^{-2}[a_0(2b + 2) + a_1]L^3 + b^{-2}L^4 \\ D_t &= -S_{t+2} + a_0^{-1}b^{-2}[a_0(b^2 + 2b) + a_1b + a_2a_3b]S_{t+1} - a_0^{-1}b^{-2}[a_1 + a_0(1 + 2b)]S_t + b^{-2}S_{t-1} - a_0^{-1}b^{-2}u_t \end{aligned}$$

A stable solution to the model requires the two smallest roots of AL to be less than one in modulus. We denote these roots as λ_1 and λ_2 . The stable solution can then be written in terms of these two stable roots as⁵

$$H_t = (\lambda_1 + \lambda_2)H_{t-1} - \lambda_1\lambda_2H_{t-2} + b^{-1}\lambda_1\lambda_2(\lambda_1 - \lambda_2) \sum_{j=0}^{\infty} \{[(b\lambda_1)^{j+1} - (b\lambda_2)^{j+1}]E_tD_{t+j}\} \quad (3.28)$$

Assuming uncorrelated cost shocks, as specified in the full model above, we can solve for the reduced form inventory equation as follows.⁶ Define the following scalars

$$\begin{aligned} \rho_1 &= \lambda_1 + \lambda_2 \\ \rho_2 &= -\lambda_1\lambda_2 \\ w_1 &= b^2\rho_2 \\ w_2 &= -\rho_2[b^2 + 2b + b(a_1/a_0) + (ba_2a_3/a_0)] \\ w_3 &= \rho_2[2b + 1 + (a_1/a_0)] \\ w_4 &= -\rho_2 \end{aligned}$$

and the vector

$$e' = (1 \quad 0) \quad (3.29)$$

and the matrices

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

⁵See West (1992), Blanchard (1983), or Fuhrer et al. (1995).

⁶This follows Blanchard (1983) and West (1992).

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

with I being the 2x2 identity matrix. The reduced form equation for inventories is then

$$H_t = \rho_1 H_{t-1} + \rho_2 H_{t-2} + \pi_1 S_{t-1} + \pi_2 S_{t-2} + \beta t + KH + \epsilon_{Ht} \quad (3.30)$$

with

$$(\pi_1 \quad \pi_2)' = e'D(w_1\Phi^3 + w_2\Phi^2 + w_3\Phi + w_4I) \quad (3.31)$$

and

$$\epsilon_{Ht} = \frac{\rho_2}{a_0} u_t + \frac{\pi_2}{\phi_2} \epsilon_{St}$$

and $\beta t + KH$ being trend plus constant terms.

In order to simulate the data we need to specify values for the parameters of the exogenous sales process, the reduced form inventory equation, and also the distribution of the sales and cost innovations. Since the parameters in (3.30) are functions of the a_i 's they can be determined given some choice for the cost parameters. West and Wilcox (1996) use several sets of values for these based on other work. In the present study we chose to generate data separately from several parameter specifications. The parameter values used are based on calibrations to real data on inventories and sales from several two-digit SIC nondurable goods manufacturing industries, including Food, Apparel, Chemicals, Petroleum, and Rubber, as well as for aggregated total nondurables manufacturing.

In order to do this it is necessary to jointly estimate the exogenous sales process in (3.26) and the reduced form for inventories in (3.30) for each of the industries and total nondurables. From the parameter estimates for (3.30) we can get estimates for combinations of the cost parameters using⁷

$$\begin{aligned} a_1/a_0 &= \rho_1(b - \rho_2^{-1}) - 2b - 2 \\ a_2/a_0 &= -b^{-1}[\rho_2^{-1}(1 + b\rho_1^2) + b^2\rho_2 + (1 + b)a_1/a_0 + (1 + 4b + b^2)] \end{aligned}$$

and from estimates of the sales process we can also solve for a_3 from w_2 in (3.31). These values are given in Table 3.2. In fixing the a_i 's for the simulation we follow Blanchard (1983) and normalize a_0 to one.

⁷See Blanchard (1983).

With cost parameter estimates in hand we can simulate data for inventories and sales by further specifying the parameters of the exogenous sales process along with the distributions of the sales and cost innovations. The autoregressive parameters of the sales process in each calibration are set to roughly match those of our data. Following West and Wilcox (1996), the sales innovation variance is set to normalize the unconditional variance of the sales process to one for each of the calibrations. For the cost shock variance and correlation with sales, we use the same values given in West and Wilcox (1996), which are 3.5 and -0.5 respectively. These values conveniently result in the variance ratio and the correlation between inventories and sales closely matching that of the data for total nondurables. For simplicity we use the same values for the industry calibrations as well. These values also are contained in Table 3.2.

With the a_i 's fixed, we can determine implied values for $\rho_1, \rho_2, \pi_1,$ and π_2 . Recall that $\rho_1 = \lambda_1 + \lambda_2$ and $\rho_2 = -\lambda_1\lambda_2$. The values for λ_1 and λ_2 can be taken as the two smallest in modulus roots of the characteristic equation for $A(L)$ in (3.27) above. This equation is

$$\lambda^4 - a_0^{-1}b^{-2}[a_0(2b^2 + 2b) + a_1b]\lambda^3 + a_0^{-1}b^{-2}[a_0(1 + 4b + b^2) + a_1(1 + b) + ba_2]\lambda^2 - a_0^{-1}b^{-2}[a_0(2b + 2) + a_1]\lambda + b^{-2}$$

The π_i 's are then found from equation (3.31). These values also are given in Table 3.2.

In determining coefficients on trend terms we again follow West and Wilcox (1996) and set them so that the implied coefficients of variation (cv) for ΔS_t and ΔH_t matches that of our data for each industry. These values are given in Table 3.2 as well. The coefficient of variation is

$$cv = \frac{\sqrt{\gamma}}{\mu} \tag{3.32}$$

where γ is the variance of the process and μ the mean, which will involve the trend terms for the process in levels. To solve for these, first express the sales process and the reduced form for inventories in differenced form and solve for their means. The sales process is given by equation 3.26. In differences, this becomes:

$$\Delta S_t = \phi_1 \Delta S_{t-1} + \phi_2 \Delta S_{t-2} + \alpha + \Delta \epsilon_{St} \tag{3.33}$$

The mean is given by:

$$\mu_{\Delta S_t} = E\Delta S_t = \phi_1 E\Delta S_{t-1} + \phi_2 E\Delta S_{t-2} + \alpha \quad (3.34)$$

$$= \phi_1 \mu_{\Delta S_t} + \phi_2 \mu_{\Delta S_t} + \alpha \quad (3.35)$$

$$= \frac{\alpha}{1 - \phi_1 - \phi_2} \quad (3.36)$$

Given the variance and cv, we can then solve for α .

For inventories we have

$$\Delta H_t = \rho_1 \Delta H_{t-1} + \rho_2 \Delta H_{t-2} + \pi_1 \Delta S_{t-1} + \pi_2 \Delta S_{t-2} + \beta + \Delta \epsilon_{Ht} \quad (3.37)$$

The mean of the process is:

$$E\Delta H_t = \rho_1 E\Delta H_{t-1} + \rho_2 E\Delta H_{t-2} + \pi_1 E\Delta S_{t-1} + \pi_2 E\Delta S_{t-2} + \beta \quad (3.38)$$

or

$$\mu_{\Delta H_t} = \rho_1 \mu_{\Delta H_t} + \rho_2 \mu_{\Delta H_t} + \pi_1 \mu_{\Delta S_t} + \pi_2 \mu_{\Delta S_t} + \beta \quad (3.39)$$

$$= \frac{1}{1 - \rho_1 - \rho_2} \{(\pi_1 + \pi_2) \mu_{\Delta S_t} + \beta\} \quad (3.40)$$

$$= \frac{1}{1 - \rho_1 - \rho_2} \left\{ \frac{\pi_1 + \pi_2}{1 - \phi_1 - \phi_2} \alpha + \beta \right\} \quad (3.41)$$

To find the variances, we need to find the covariance matrix of the sytem of equations given by 3.33 and 3.37. Using methods from Hamilton (1994), first express the system in terms of deviations from mean

$$\begin{aligned} \Delta H_t - \mu_{\Delta H_t} &= \rho_1 (\Delta H_{t-1} - \mu_{\Delta H_t}) + \rho_2 (\Delta H_{t-2} - \mu_{\Delta H_t}) \\ &\quad + \pi_1 (\Delta S_{t-1} - \mu_{\Delta S_t}) + \pi_2 (\Delta S_{t-2} - \mu_{\Delta S_t}) + \Delta \epsilon_{Ht} \end{aligned}$$

$$\Delta S_t - \mu_{\Delta S_t} = \phi_1 (\Delta S_{t-1} - \mu_{\Delta S_t}) + \phi_2 (\Delta S_{t-2} - \mu_{\Delta S_t}) + \Delta \epsilon_{St}$$

where we recall

$$\epsilon_{Ht} = \frac{\rho_2}{a_0} u_t + \frac{\pi_2}{\phi_2} \epsilon_{St} \quad (3.42)$$

so

$$\Delta \epsilon_{Ht} = \frac{\rho_2}{a_0} \Delta u_t + \frac{\pi_2}{\phi_2} \Delta \epsilon_{St} \quad (3.43)$$

Express this system of equations as the first-order VAR

$$Z_t = FZ_{t-1} + CV_t \quad (3.44)$$

where

$$Z_t = \begin{bmatrix} \Delta H_t \\ \Delta H_{t-1} \\ \Delta S_t \\ \Delta S_{t-1} \end{bmatrix} \quad (3.45)$$

$$F = \begin{bmatrix} \rho_1 & \rho_2 & \pi_1 & \pi_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \phi_1 & \phi_2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3.46)$$

$$C = \begin{bmatrix} \rho_2/a_0 & \pi_2/\phi_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.47)$$

and

$$V_t = \begin{bmatrix} \Delta u_t \\ \Delta \epsilon_{St} \\ 0 \\ 0 \end{bmatrix} \quad (3.48)$$

To find the covariance matrix, let

$$\Sigma = E(Z_t Z_t') \quad (3.49)$$

Then

$$\begin{aligned} \Sigma &= E(FZ_{t-1} + CV_t)(FZ_{t-1} + CV_t)' \\ &= F\Sigma F' + FE(Z_{t-1}V_t')C' + CE(V_tZ_{t-1}')F' + CE(V_tV_t')C' \end{aligned}$$

$$= F\Sigma F' + Q$$

Define

$$\Omega = \begin{bmatrix} \sigma_u^2 & \sigma_{u,\epsilon_S}^2 & 0 & 0 \\ \sigma_{u,\epsilon_S}^2 & \sigma_{\epsilon_S}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.50)$$

Then

$$E(Z_{t-1}V_t') = CV_{t-1}V_t' = -C\Omega \quad (3.51)$$

$$E(V_tZ_{t-1}') = V_tV_{t-1}'C' = -\Omega C' \quad (3.52)$$

$$E(V_tV_t') = 2\Omega \quad (3.53)$$

Q then becomes

$$Q = -FC\Omega C' - C\Omega C'F' + 2C\Omega C' \quad (3.54)$$

To solve for Σ , define

$$A = F \otimes F$$

$$B = I - A$$

Then using the *vec* operator, we have⁸

$$vec\Sigma = B^{-1}vecQ \quad (3.55)$$

from which we can reconstruct Σ and retrieve the desired variances for the ΔH_t and ΔS_t processes.

From here we calculate the trend coefficients as

$$\alpha = (1 - \phi_1 - \phi_2)\mu_{\Delta S} = (1 - \phi_1 - \phi_2)\frac{\sqrt{\gamma_{\Delta S}}}{cv_{\Delta S}} \quad (3.56)$$

and

$$\beta = (1 - \rho_1 - \rho_2)\mu_{\Delta H} - \frac{\pi_1 + \pi_2}{1 - \phi_1 - \phi_2}\alpha = (1 - \rho_1 - \rho_2)\frac{\sqrt{\gamma_{\Delta H}}}{cv_{\Delta H}} - \frac{\pi_1 + \pi_2}{1 - \phi_1 - \phi_2} \quad (3.57)$$

The resulting trend coefficients for each industry calibration are given in Table 3.2.

⁸See Hamilton (1994), pages 264-266.

Finally, the constant terms in equations 3.26 and 3.30 were chosen by fixing the other parameters at the chosen values and using the actual data to solve for KS and KH in

$$\hat{\mu}_{S_t} = KS + \alpha\mu_t + \phi_1\hat{\mu}_{S_{t-1}} + \phi_2\hat{\mu}_{S_{t-2}} \quad (3.58)$$

$$\hat{\mu}_{H_t} = KH + \beta\mu_t + \rho_1\hat{\mu}_{H_{t-1}} + \rho_2\hat{\mu}_{H_{t-2}} + \pi_1\hat{\mu}_{S_{t-1}} + \pi_2\hat{\mu}_{S_{t-2}} \quad (3.59)$$

where $\hat{\mu}_{X_{t-j}} = \frac{1}{T-2} \sum_{t=3-j}^{T-j} X_t$ and $\mu_t = \frac{1}{T-2} \sum_{t=2}^T t$. These values also are contained in Table 3.2.

Given the calibrated values of the parameters in Table 3.2, the implied population values for the three parameters to be estimated in what follows are given in Table 3.3.

3.6 Probability Limits of the Parameter Estimates

The first step in seeing the effects of estimating the pure model or the stockout avoidance model when data is generated from the full model is to determine the degree of impact on the probability limits of the parameter estimates. To estimate the asymptotic limits of the parameter estimates we generated 1,000 samples of size $T = 25,000$ for each of the industry calibrations given in Table 3.2. The models were then estimated using a two-step GMM estimator to determine the large sample limits of the parameter estimates. Because there are several possibilities for choice of weighting matrix used in the first step of GMM estimation, we estimate the models twice for each sample, first setting $W = I_q$ and then setting $W = (Z'Z)^{-1}$, the inverse of the instrument cross-products. It is assumed that e_t follows a martingale difference process in constructing the weighting matrix for the second step estimation. This assumption would be reasonable for a researcher who believed these restricted versions of the model to be the correct specification prior to estimation.⁹ If we were to estimate the full model, this assumption would not hold, since e_t then follows an MA(2) process.¹⁰ The mean estimate for both first step and second step estimation using these large samples are given in Table 3.4.

Looking at the results for λ we see that for some industry calibrations the estimate of the probability limit differs greatly from the true value while for others the misspecification seems to have little impact.

As mentioned above, $1 - \lambda$ can be interpreted as an adjustment speed for inventory accumulation, that

⁹See Eichenbaum (1989), for example.

¹⁰See West and Wilcox (1996).

is the percentage of the gap between beginning of period inventories and desired inventory levels that is closed over a period. In general, this is the parameter of greatest interest to those estimating any of the versions of the model and so it is worth taking a look at what the results for the different models imply for these adjustment speeds.¹¹

For the most part, the pure model leads to the greatest differences in the value of λ , especially for the data calibrated to the food industry. Here the true value of λ implies that roughly 3 percent of the difference in actual and desired inventories is made up over a period, while for the pure model the two-step estimate implies a rate of about 13 percent. The calibration for total nondurables shows similar results, with the pure model implying between 15 and 17 percent of the gap is closed when the true λ implies only 7 percent is. Looking at the rubber industry calibration indicates that the pure model does not have a pattern of overestimating this speed of adjustment. The results here show a much slower rate for the pure model. at about 6 percent, than for the true value of λ , which implies a 15 percent closure of the gap. The other three calibrations show quite a bit less impact on the value of λ that would result from estimating the pure model, with the greatest difference in implied adjustment speeds being 3.5 percent for apparel and the least being about 0.5 percent for the chemical industry calibration. Finally, notice that there is little difference in the results from the use of the alternative first-step weighting matrices. If the model were correctly specified this would not be surprising given the sample size, as choice of first-step weighting matrix should not matter asymptotically in correctly specified models. However, given the misspecification we would expect an impact.

For the stockout avoidance specification there seems to be less of an impact on the values that λ tends to. The petroleum and rubber calibrations show the greatest movement away from the true values of λ , with both calibrations resulting in a difference of around 6.3 in the percent of the gap closed between the true values. The chemical calibration has a difference of about 5 percent, while the rest of the calibrations differ by an amount of 2 percent or less in the implied adjustment speeds. For ψ there is no straightforward economic interpretation of the resulting estimate, other than values close to unity

¹¹Recognizing that all of the population values imply very slow movement towards desired stock levels per period, we note that most work attempting to estimate the adjustment speed have found such implausibly slow rates. This has led many to reject the linear-quadratic inventory model outright, in particular Blinder and Maccini (1991). For our purposes this is unimportant in that our goal is not to come up with a plausible estimate for λ , but instead it is to do a simulation study using a data generating process that is consistent with a given model in order to study the statistical properties of a particular method of conducting inference.

implying support for a stockout avoidance motive.¹² Of course this motive is present in the model used to simulate the data, so the limits resulting from our large sample estimates are very close to unity. From Table 3.4 we see that the estimated probability limits for ψ from the stockout avoidance model differ little from the true values of the parameter in nearly all of the calibrations. The largest difference comes in the apparel calibration, with $\psi = 1.16$ versus $\psi = 1.08$ in truth. This implies that the misspecified model provides slightly less evidence for the stockout avoidance motive than is perhaps present in the simulated data. Nearly all of the other calibrations give estimated probability limits that give slightly greater support to this motive than do the true values.

The results here lend support to the idea that the pure model and the stockout avoidance model are misspecified in the non-local sense as defined above. The estimated probability limits given in Table 3.4 imply parameter vectors, θ_* , that minimize the GMM objective functions being estimated, but deviate from the true parameter vector, θ_0 , in all cases. Thus inference in the standard fashion will not be appropriate for these models and we turn to the asymptotic theory developed in Hall and Inoue (2003) for a remedy.

3.7 Specification Testing

In empirical work, models estimated by GMM are generally tested using Hansen's J -test of over-identifying restrictions¹³ to determine whether or not the model is correctly specified given the data. In this section we look at the J -test's ability to capture the misspecification present in the alternative versions of the model being estimated. In addition, we employ another method developed in Durlauf and Maccini (1995) which allows comparison of different models to determine which provides a better fit to the data.

¹²See footnote 9 in Durlauf and Maccini (1995) for a discussion of this point.

¹³See Hansen (1982).

3.7.1 Hansen's J-Test

When an econometric model is overidentified, as in this study, the GMM estimation method gives us q sample orthogonality conditions which should be close to being satisfied if the model is the true model.

To test these overidentifying restrictions, Hansen's test sets¹⁴

$$J_T = \left\{ T^{-1/2} \sum_{t=1}^T g(x_t, \hat{\theta}) \right\}' W_T \left\{ T^{-1/2} \sum_{t=1}^T g(x_t, \hat{\theta}) \right\} = T Q_T(\hat{\theta}) \quad (3.60)$$

and recognizes that

$$J_T \xrightarrow{d} \chi_{q-p}^2 \quad (3.61)$$

Note that J_T is just T times the GMM objective function calculated at the parameter values which minimize it. Essentially we are testing whether the parameter values obtained in estimation set the orthogonality restrictions close enough to zero that we can reasonably accept the model as being correctly specified.

It has been noted in the literature that the J_T test often does a poor job detecting a misspecified model, so it is worth seeing how it performs in the current context. Before looking at the small sample results, however, we note that the J_T -test rejected the pure model at the 5 percent level 100 percent of the time for all calibrations when doing the large sample ($T = 25000$) estimation used to obtain the probability limit estimates above. In addition, the test rejected the stockout avoidance version of the model 100 percent of the time for all calibrations except food when using the large samples. The rejection rates for the food calibration were about 88.5 percent when using either weighting matrix for the first step estimation. So it appears that in large samples the test works well in capturing the misspecification present in the alternative versions of the model. However, in small samples we get quite a different result from the J_T -test.

Before reporting the small sample results for the J_T test, we make note of a problem that arose in the estimation of these models using the smaller sample size. We have found that the GMM objective function has two local minima in λ , regardless of which specification is used.¹⁵ While in the vast majority of cases, the minimization procedure chose a value that is less than unity for this parameter, it

¹⁴See Hansen (1982) or Hamilton (1994), chapter 14.

¹⁵Hall (2003), chapter 9, discusses a similar problem found in different normalizations of the stockout adjustment model.

occasionally chose a value greater than one. We saw above that the speed of adjustment of inventories toward their desired level is $1 - \lambda$. This means that $\lambda \geq 1$ is not economically meaningful, since $\lambda = 1$ would imply there is no inventory accumulation toward some ideal level and $\lambda > 1$ would indicate inventory disinvestment rather than accumulation. In dealing with this problem we simply discarded those results for which the estimated value of λ had no economic meaning in the context of the model. Of the one thousand replications estimated for each calibration and weight matrix, the maximum number of results discarded for any of them was about seventy-five, and in the majority of cases was far less than this. This small number should have little impact on the conclusions we arrive at using the remaining results. We also note that this problem did not arise in the large sample estimation undertaken to ascertain the probability limits above. As should be expected, the larger samples did a better job of correctly identifying the parameter values.

Table 3.5 gives the rejection rates for a nominal 5 percent test when using the J_T -test for all versions of the model and each calibration. Results are given for both choices of first step weighting matrix for the two-step estimator based on 1000 samples with a sample size of $T = 500$ for each. For the pure production smoothing version of the model, the test only exhibits high rejection rates for the food and petroleum calibrations, with the model being rejected about 99 percent of the time for food and around 95 percent of the time for petroleum. For all other calibrations the test has low rejection rates despite the misspecification, ranging from 7.3 percent for chemicals when $W = I$ to 49.8 percent for apparel when using $W = (Z'Z)^{-1}$. In fact, for the chemicals calibration the pure model is rejected by the test less often than the full model is.

For the stockout avoidance version of the model none of the calibrations generates J_T test results that give a high rejection rate. Only the petroleum calibration has a rejection rate greater than 50 percent. All others have rejection rates lower than 15 percent regardless of first step weighting matrix used. For half of the calibrations the stockout avoidance specification is rejected less often than is the full model for both choices of W . So again we see the J_T statistic to be a poor indicator of misspecification.

While the misspecified models are not rejected by the J_T statistic nearly often enough, in most cases, the full model is rejected too often in at least half of the cases. With the full model the test's

rejection rate should approach the nominal size of the test, which is 5.0 percent here. Only for two of the calibrations, petroleum and rubber, do we come close to the correct size for this statistic, although for a couple of the other calibrations it is not too distant from the correct size. This is likely to improve in larger samples..

From the above results it is clear that the J_T test does a very poor job in rejecting the misspecified models. In many of the cases it is possible, and even likely, that a researcher would not reject either the pure or the stockout avoidance model even though it is misspecified. In some cases the J_T test has been used as a model selection tool, with researchers choosing between alternative specifications based on the test results. Here we have provided evidence that this method of model selection is not likely to work well. In many of the calibrations the results of such testing would lead to selection of the misspecified alternatives over the full model.

3.7.2 Noise Ratios

Rather than reject or accept an econometric model outright based on misspecification testing, Durlauf and Maccini (1995) offer a method of comparing different models based on their ability to explain the behavior of the data used in estimation. They advocate constructing "noise ratios" that indicate how well a particular model fits a given set of data relative to other models attempting to explain the same data. Essentially the noise ratio of a given model is a normalized measure of model noise, defined as the variation in the model's residuals that is due to misspecification rather than expectational errors in forecasting. Durlauf and Maccini (1995) claim this to be a useful measure in that, even if a series of models are rejected in specification testing, we can still make a determination as to which provides a better fit to the data. Using the pure model for illustration, the theory and construction of noise ratios in the context of the models used here follows.

Recall that for the pure model we wish to estimate λ in the conditional euler equation

$$E_t\{b\lambda H_{t+1} - [1 + b\lambda^2]H_t + \lambda H_{t-1} + b\lambda S_{t+1} - \lambda S_t\} = 0 \quad (3.62)$$

In order to make this estimation operational, we use our set of time series data to substitute realized values in for those variables whose values are unknown to the econometrician at time t , which in this

case includes H_{t+1}, H_t, S_{t+1} , and S_t .¹⁶ This leads to the creation of an expectational error resulting from the difference between the value of (3.62) using realized values for these unknown quantities and their actual conditional expectation at t as implied by the model. Durlauf and Maccini (1995) denote this expectational error as g_{t+1} , which gives us

$$g_{t+1} = b\lambda H_{t+1} - [1 + b\lambda^2]H_t + \lambda H_{t-1} + b\lambda S_{t+1} - \lambda S_t \quad (3.63)$$

If the model is correctly specified then g_{t+1} is solely the result of new information which comes about within the period, that is $\Omega_{t+1} - \Omega_t$. Durlauf and Maccini (1995) denote this new information as v_{t+1} and so we have $H_0 : g_{t+1} = v_{t+1}$ for a correctly specified model.

In the case of a misspecified model, not only will there be error due to lack of information available at time t , but also error due to the conditional expectation of the particular combination of unknown inventories and sales at time t posited by the model being different from the true conditional expectation of these variables. To clarify, if (3.62) were true, it implies

$$E_t\{b\lambda H_{t+1} - [1 + b\lambda^2]H_t + b\lambda S_{t+1} - \lambda S_t\} = -\lambda H_{t-1}, \quad (3.64)$$

since H_{t-1} is in Ω_t . However, if the model is misspecified the conditional expectation on the left hand side will differ from $-\lambda H_{t-1}$ at time t . The difference between the actual conditional expectation at t and that posited by the model is realized at time t . This is what Durlauf and Maccini (1995) call model "noise", and we will denote it Υ_t here. This leads to an alternative hypothesis of the form $H_A : g_{t+1} = v_{t+1} + \Upsilon_t$, indicating that in estimation of euler equations, the estimation error will consist of both an expectational error as well as model noise when the model is misspecified.

The goal is then to obtain an estimate of the unobservable model noise for each model estimated. Durlauf and Maccini (1995) pose this as a signal extraction problem from which we can get an estimate of Υ_t from a linear projection of g_{t+1} onto the information set Ω_t using our data. Since v_t is by definition orthogonal to Ω_t , this results in

$$\hat{\Upsilon}_t = proj(\Omega_t)g_{t+1} \quad (3.65)$$

To make this operational we follow Durlauf and Maccini (1995) and generate the sequence g_{t+1} using

¹⁶Recall that H_T indicates end of period inventories, which is unknown at time t because sales for the period, S_t , are unknown.

parameter values obtained from two-step GMM estimation of the appropriate moment condition. Using ordinary least squares, we obtain the projection coefficients by regressing the g_{t+1} sequence on the instrument set used in the GMM estimation. We then use this to generate the forecast \hat{Y}_t .

Calculating the variance of \hat{Y}_t gives us a measure of how well the model fits the data, since a high variance would indicate that the actual conditional expectation of the true model differs significantly from that posited by the model being considered. Rather than compare $var(\hat{Y}_t)$ across models, Durlauf and Maccini (1995) propose normalizing the measure by dividing by $var(g_{t+1})$ to get an estimated noise ratio (NR):

$$NR = \frac{var(\hat{Y}_t)}{var(g_{t+1})} \quad (3.66)$$

This is useful in that NR now expresses the estimated variance of model noise as a proportion of the overall variance in g_{t+1} .

Table 3.5 contains the mean of the noise ratios calculated for each of the model versions above. Note that, regardless of the behavior of the J_T statistic, for all industry calibrations the noise ratio works well in indicating the full model as being that which provides the best fit to the data among the three choices. Moving from the pure model to the stockout avoidance version we see a significant decrease in NR for all calibrations, indicating that the SA model provides a better fit than does the pure model. Also notable is that there is little difference in NR for each model and calibration when alternative weighting matrices are used in the estimation.

The results above indicate that calculation of noise ratios can be a useful tool in choosing between models, especially when results from specification tests are deemed unreliable. Here it is clear that the J_T statistic is not reliable as an indicator of model misspecification, rejecting the true model too often and the misspecified versions not often enough in many cases. Given these results, choosing a model based upon noise ratios may be a good route to take when estimating euler equation models with GMM. However, given the fact that choosing a model based on the minimum noise ratio gives no guarantee that the resulting model is correctly specified in the sense of the definition given above, the additional use of valid asymptotic theory for inference in misspecified models may also be required. This lends support to further investigating the potential value of the theory developed in Hall and Inoue (2003). As pointed

out below, when the asymptotic theory developed therein is applied and the model is not misspecified, in the cases studied here the results should be the same as those for correctly specified models using standard GMM asymptotics.

3.8 Impact of HI Asymptotics

In this section we return to the theory put forth in Hall and Inoue (2003). As mentioned above, in that paper the distribution theory is developed for a number of alternative GMM weighting matrices. Here we look at the impact of their results for two of these. The first case considered in that paper is the use of a fixed weighting matrix, such as the identity matrix, which is often used on the first step in GMM estimation. For this situation, they show that the estimator is still normally distributed and approaches θ_* in the limit, with $T^{1/2}(\hat{\theta}_T - \theta_*) \xrightarrow{d} N(0, \Sigma_I)$. However, Σ_I is not the same covariance matrix that would be obtained in the correctly specified case when the model is misspecified, since it is now a function of $\mu(\theta_*)$, given that the moment condition does not approach zero. Theorem 1 from Hall and Inoue (2003) provides the formula for calculating Σ_I as

$$\Sigma_I = H_*^{-1}(G_*'W\Omega_{11}WG_* + G_*'W\Omega_{12} + \Omega_{21}WG_* + \Omega_{22})H_*'^{-1} \quad (3.67)$$

where $\{\Omega_{ij}\}_{i,j=1,2}$ are the components of the asymptotic covariance matrix in

$$\begin{pmatrix} T^{-1/2} \sum_{t=1}^T (g(x_t, \theta_*) - \mu_*) \\ T^{1/2}(G_T(\theta_*) - G_*)'W\mu_* \end{pmatrix} \xrightarrow{d} N(0, \Omega_I)$$

with

$$\Omega_I = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$$

and

$$\mu_* = E[g(x_t, \theta_*)]$$

$$G_* = E[\partial g(x_t, \theta_*)/\partial \theta']$$

$$H_* = G_*'WG + (\mu_*' \otimes I_p)G_*^{(2)}$$

$$G_*^{(2)} = E[(\partial/\partial \theta') \text{vec}\{\partial g(x_t, \theta_*)/\partial \theta'\}]$$

Standard errors may be obtained from estimates of Σ_I that can then be used in constructing test statistics based on the correct covariance matrix. Note that if $\mu_* = 0$ this reduces to

$$\Sigma_I = \Sigma_0 = (G'_*WG_*)^{-1}(G'_*WVWG_*)(G'_*WG_*)^{-1}, \quad (3.68)$$

the covariance matrix for correctly specified models.

A second case covered in their paper is the use of a weighting matrix that is dependent on T but converges to a positive definite matrix of constants in the limit. That is $\text{plim}_{T \rightarrow \infty} W_T = W$ and W_T converges to a normal distribution at rate \sqrt{T} . This applies to the use of the inverse matrix of instrument cross-products in the first step of an instrumental variables estimation and also to the assumption that $g(x_t, \theta_*) - \mu(\theta_*)$ is a martingale difference sequence (MDS) in constructing a two-step estimator of θ_* . In this case it is shown that $T^{1/2}(\hat{\theta}_T - \theta_*) \xrightarrow{d} N(0, \Sigma_Z)$. Along with $\Sigma_Z \neq \Sigma_I$, Σ_Z will also differ from the correctly specified case when the model is misspecified. Formulas for the correct calculation of Σ_Z are provided in Theorem 2 of Hall and Inoue (2003) as

$$\Sigma_Z = H_*^{-1}\Omega_*H_*'^{-1} \quad (3.69)$$

where

$$\Omega_* = G'_*W\Omega_{11}WG_* + \Omega_{22} + G'_*\Omega_{33}G_* + G'_*W\Omega_{12} + G'_*W\Omega_{13}G_* + \Omega_{21}WG_* + G'_*\Omega_{31}WG_* + \Omega_{23}G_* + G'_*\Omega_{32}$$

and $\{\Omega_{ij}\}_{i,j=1,2,3}$ come from the asymptotic covariance matrix in

$$\begin{pmatrix} T^{-1/2} \sum_{t=1}^T (g(x_t, \theta_*) - \mu_*) \\ T^{1/2}(G_T(\theta_*) - G'_*)'W\mu_* \\ T^{1/2}(W_T - W)\mu_* \end{pmatrix} \xrightarrow{d} N(0_{(p+2q) \times 1}, \Omega_Z).$$

with

$$\Omega_Z = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{13} \\ \Omega_{32} & \Omega_{32} & \Omega_{33} \end{pmatrix}$$

In addition to the above, there are other choices available for the GMM weighting matrix, but these are not considered in the current analysis. A summary of the asymptotic distribution of the GMM

estimator for correctly specified and misspecified models is given in Table 3.6. In what follows we look at the results of the above for first-step estimates only, with the two choices for weighting matrix then being I_q and $(Z'Z)^{-1}$. One reason for this is that the optimality of the two-step estimator only has theoretical support when a model is correctly specified, while the first step estimator still generates a consistent estimate of Θ , although it may not be efficient. In addition, the increased computational burden involved in calculating the covariance matrix for the second step of estimation, due to its dependency on that used in the first step, is likely to outweigh any efficiency gains from doing so in practice.

3.8.1 Estimation Results

For this section, we generated from each calibration additional samples of size $T = 1000$ and $T = 5000$ in order to ascertain the effect of sample size on the results. Estimation was again carried out on each of the 1000 samples from each calibration and sample size using the alternative weighting matrices discussed above. All results are reported for the first-step estimator only, for the reasons mentioned above.

The Full Model

We first report estimation results for the full model, as this is the correctly specified version and will thus serve as a benchmark for the results from the misspecified versions of the model. These are contained in table 3.7. The table has two sections, with the results from the use of $W = I_q$ on the left and those from $W = (Z'Z)^{-1}$ on the right. The first three columns of each section contain the mean, the median, and also the actual standard deviation of the one thousand parameter estimates. Next is the median of the estimated standard errors from each sample calculated using conventional asymptotics, which should be appropriate in this case.¹⁷ The final column reports the actual size of a nominal 5% significance level t -test of $H_0 : \theta = \theta_0$, for $\theta = \lambda, \psi$, and δ .

As it turns out, the one-step estimator performs rather poorly in most of the calibrations, despite the fact that the model is correctly specified. This perhaps should come as no surprise when $T = 500$, as GMM is not well noted for its strong small sample performance, even when using two-step or iterated

¹⁷We use a HAC estimator with estimated bandwidth for the covariance matrix here because, as mentioned previously, the expectational error for this model is MA(2).

procedures. However, performance does not improve by much at the larger sample sizes, and actually deteriorates in some cases. Choice of weighting matrix also seems to have little impact, as the results are qualitatively similar for both $W = I_q$ and $W = (Z'Z)^{-1}$.

Looking at the distributional characteristics of the estimates we see that in a number of cases the distributions are skewed, with a fair amount of distance between the median and the mean. This is particularly a problem for $\hat{\delta}$, for which the mean and median values are reasonably close only for the total and chemicals calibration for either choice of W when $T = 5000$, although in the case of chemicals $\hat{\delta}$ is of the wrong sign. The distributions for $\hat{\lambda}$ are skewed for the apparel, chemicals, and food calibrations at all sample sizes when $W = I_q$, but not for $W = (Z'Z)^{-1}$. The distributions of $\hat{\psi}$ seem to be well centered about the mean in all cases.

Comparing the means of the parameter estimates to their true population values, we see that for $\hat{\lambda}$ there is significant mean bias when $T = 500$ for nearly all of the calibrations. While the bias tends to diminish for most of the calibrations at the larger sample sizes, it actually increases for the apparel, petroleum, and rubber calibrations when $W = I_q$. There is similar bias in the distributions of $\hat{\psi}$ for $T = 500$. In this case the bias increases with sample size for the chemicals and rubber calibrations for both choices of W , but diminishes for the remaining calibrations. For $\hat{\delta}$, the estimates bear little resemblance to their true values in the majority of calibrations. In fact for the chemicals calibration the mean estimates are of the wrong sign for all sample sizes and weight matrices. For apparel and rubber the already severe bias worsens as sample size increases, regardless of W . For the food calibration the bias when $T = 500$ is severe, but diminishes significantly as T increases and nearly disappears when $W = I_q$ for $T = 5000$. Only for the total and the petroleum calibrations do the estimates for δ perform well, with little bias at any of the sample sizes tested here.

The third column of each side of table 3.7 contains the actual standard deviations of the 1000 estimates in each trial. These are reported as a benchmark for the performance of the asymptotic approximations. While the dispersion of the estimates for λ and ψ is reasonable and diminishes with increased sample size, the standard deviations of the estimates of δ are rather large relative to the true values in most cases. This indicates that the estimation procedure had great difficulty in estimating this parameter.

Comparing the median of the asymptotic approximations for the standard errors in the fourth column to the actual standard deviations of the estimates, we find that in the vast majority of cases these are larger than the actual standard deviations for $\hat{\lambda}$ and $\hat{\psi}$, but too small for $\hat{\delta}$. The median approximations are fairly close for ψ , even at the smaller sample sizes in most cases, while those for λ tend to be very large relative to the actual standard deviations. For δ , the approximate median standard errors tend to be of the same magnitude as the large standard deviations reported above, thus reflecting the difficulty in estimating this parameter from the simulated data. Taken as a whole, the asymptotic approximations to the standard errors perform very well for ψ , reasonably well for δ , and somewhat poorly for λ in many cases.

The above results are reflected in the actual size of the t -statistic in our experiments. This is presented in the final column of each side of the table. For $W = I_q$, the t -test for λ rejects too often for the total, chemicals, and food calibrations when $T = 500$, although this over-rejection is not too severe for total and chemicals. There is improvement toward the nominal size as we increase sample size to $T = 1000$ for all of these, but for $T = 5000$ the statistic does not reject often enough for the total calibration, with a rate of just 0.5%. For chemicals and food the statistic is reasonably close to its nominal size for $T = 5000$. When using $W = (Z'Z)^{-1}$, the t -statistics for λ again over-reject for the total, chemicals and food calibrations, but approach the nominal size as sample size increases to $T = 5000$. For the remaining calibrations the test statistic did not reject often enough when $T = 500$ and performance deteriorated further as sample size increased, regardless of choice of W .

The t -statistics for ψ performed very poorly for the apparel and rubber calibrations, with severe over-rejection at $T = 500$ becoming even worse as sample size increases, regardless of choice of W . For chemicals the actual size of the t -test is only slightly below its nominal size for $T = 500$, but this also deteriorates to severe over-rejection when $T = 5000$. In all of these cases the poor size of the statistic at $T = 5000$ is the result of the mean bias mentioned above in combination with fairly small standard errors, so even if the distributions are correct for these parameter estimates, they are centered far enough from their population values that confidence intervals constructed about them are unlikely to contain the true values. For the food and petroleum calibrations these distributions are centered close to their

true values and the median asymptotic standard errors are very close to the actual standard deviations, with the result being that the t -statistics are close to correctly sized in these cases. The same holds the total calibration when $W = (Z'Z)^{-1}$, but for $W = I_q$, the estimated standard errors are too large and the statistic under-rejects when $T = 5000$.

For δ , the t -statistics over-reject when $T = 500$ for the apparel, chemical, and rubber calibrations with either choice of W , again with diminishing performance as sample size is increased. This is not surprising given the wide range of estimates obtained for this parameter for these calibrations, as evidenced by the large standard deviations. For the total, food, and petroleum calibrations, there is over-rejection when $T = 500$, but performance improves with sample size for both choices of W .

Taken as a whole, the results for GMM estimation of the correctly specified model are pretty dismal. Only for the total calibration with $W = (Z'Z)^{-1}$ can the results be considered acceptable performance by an estimator for all of the sample sizes investigated. For the food calibration, either choice of W leads to fairly good performance, but only at the large sample size of $T = 5000$. While some of this may be due to use of a one-step estimator, which is not optimal, there may be other reasons as well. We discuss this further in the conclusion, but for now turn to the estimation results for the misspecified versions of the model to determine the potential impact of using HI asymptotics, rather than conventional asymptotics, in the performance of inference using the estimates.

Results for Misspecified Models

Because we are primarily interested in how well the GMM estimator performs relative to the limiting values of the parameters in misspecified models, the performance in accurately estimating the population values and performing inference concerning them is not a particular concern. This is because the definition of non-local misspecification gives us no reason to expect the resulting parameter estimates to be centered on the population values to begin with, and we have already shown this not to be the case in generating the probability limits of the parameters in these models. Because the probability limits were generated from the misspecified versions of the model using the simulated data, gauging the performance of the estimator based on these limiting values should be a valid exercise, despite the poor

results from the full model. Those results thus serve only as a guide to potential problems that may be encountered in what follows.

The one-step GMM estimation results for the pure production smoothing specification are given in table 3.8, while those for the stockout avoidance version are contained in table 3.9. The first three columns of these tables report the mean, the median, and the actual standard deviation of the one thousand parameter estimates. Next is the median of the estimated standard errors from each sample, calculated using conventional asymptotics, followed by the median standard error from applying HI asymptotics.¹⁸ Finally, we report the percent of the 1000 runs for which the HI standard error is smaller than that calculated in the conventional fashion. We first look at the results for the pure specification, in which λ is the only parameter.

In all cases for the pure model there is significant mean bias in the distribution of parameter estimates with respect to their probability limits. This bias is greatest in the case of $T = 500$ and diminishes as sample size increases until it all but disappears for the larger sample size of $T = 5000$. Neither choice of weight matrix seems to perform better than the other in this respect. The maximum mean bias for $T = 500$ occurs for the rubber calibration with $W = (Z'Z)^{-1}$, where $\bar{\lambda} - \lambda_{PLIM} = .0667$. This diminishes to a maximum of $\bar{\lambda} - \lambda_{PLIM} = .0045$ for $T = 5000$, again for the rubber calibration with $W = (Z'Z)^{-1}$. Such bias is somewhat typical in GMM estimation, which is not noted for its strong small sample performance. That the point estimation improves with sample size is at least an indication that the asymptotic properties do kick in at some point and the estimator does tend toward its probability limit as T increases. The closeness of the means and medians in each distribution indicate that the parameter estimates are fairly evenly distributed about the mean, and what skewness exists at the smaller sample sizes nearly disappears at $T = 5000$.

The actual standard deviations of $\hat{\lambda}$ from the 1000 trials for each calibration and weight matrix are reported as a gauge for how well the asymptotic approximations of the standard error perform when using conventional methods and when using HI asymptotics. Inspection of the medians from these asymptotic

¹⁸In estimating the covariance matrix for the HI asymptotics, a HACCC estimator with Bartlett weights was used in constructing Ω_I and Ω_Z . The bandwidth was in turn estimated using the methods of Newey and West (1994). The reasons for this are two-fold. First, there is no reason to expect that the processes being estimated are not serially dependent, and second, previous results based on martingale difference estimates of these covariances led to results for the final HI-based standard errors that were very erratic, even relative to those we report here.

approximations indicates that the conventional approximation greatly overstates the actual dispersion of the estimates in a number of cases. With $W = I_q$, the standard errors calculated using standard GMM asymptotics are extremely large relative to the actual standard deviations for the apparel, chemicals, and rubber calibrations, with little improvement at increased sample sizes. For the total and petroleum calibrations the median standard errors are about double the standard deviations of the estimates for nearly all sample sizes. Only for the food calibration are the median approximations from conventional GMM smaller than the actual standard deviation of the estimates when $W = I_q$. The results are similar for $W = (Z'Z)^{-1}$, with the exception of the food calibration, for which the standard errors calculated in the conventional fashion are now slightly larger than the actual standard deviations.

Use of HI asymptotics, on the other hand, results in median standard errors that are in most cases very similar in magnitude to the actual standard deviations of the estimates, particularly for $W = I_q$ and at the larger sample sizes. Using (3.67) when $W = I_q$, the HI median standard errors for the total, apparel, chemicals, and food calibrations all are smaller than the actual standard deviations when $T = 500$. This indicates that the asymptotic approximation given in (3.67) may tend to understate the actual variance of the estimator. This tendency seems to diminish as T increases, as the median approximate standard errors in the table are extremely close to the actual standard deviations when $T = 5000$ for the total, apparel, and chemicals calibrations. Only for the food calibration is there still a significant difference at the greater sample size. The median approximations for the petroleum and rubber calibrations were larger than the actual standard deviations of the estimates, but these also become very close as sample size is increased.

Using (3.69) to approximate standard errors when $W = (Z'Z)^{-1}$, we find the median approximation tends to overstate the actual standard deviation of the estimates for all calibrations. Although this overstatement does diminish at increased sample sizes, the median approximations are still significantly larger than the actual standard deviations when $T = 5000$.

In directly comparing the two methods of approximation in the final column of each section of table 3.8, we find that for $W = I_q$, the approximate standard errors based on HI asymptotics are almost always smaller than those based on conventional asymptotics for all calibrations but food, where it is

still smaller in the majority of trials. For $W = (Z'Z)^{-1}$, HI approximations are again smaller in virtually all cases other than for the food calibration. For the food calibration, however, conventional asymptotics yield the smaller standard error in the majority of trials. In addition, only for the food calibration do the conventional asymptotic approximations result in median standard errors that are closer to the actual standard deviations than are the approximations using HI asymptotics, although both are fairly close in value. This is true for all sample sizes and for both choices of weighting matrix.

Other than mean bias in the estimated parameter distributions at small sample sizes, estimation of the pure model does not seem to be plagued by the problems we saw in the full model. The results here indicate that HI asymptotics perform far better than do standard GMM methods in approximating the standard errors of the estimator in the pure model. This implies that inference based on HI asymptotics should also be superior to that based on conventional asymptotics. This will be investigated below, but first we turn to the estimation results for the stockout avoidance model. These are contained in table 3.9.

The estimation results for the stockout avoidance specification are similar to those for the pure model with respect to mean bias in the parameter estimates at the smaller sample sizes for $\hat{\lambda}$. However, the bias in this case is more pronounced at $T = 500$, with the maximum bias being in the estimate from the food calibration, where $\bar{\lambda} - \lambda_{PLIM} = .1017$ for $W = (Z'Z)^{-1}$. The minimum bias at $T = 500$ is in the results for the petroleum calibration with $W = (Z'Z)^{-1}$, where $\bar{\lambda} - \lambda_{PLIM} = .0166$. Again the bias diminishes as sample size increases until the maximum bias at $T = 5000$ is $\bar{\lambda} - \lambda_{PLIM} = .0171$ for the food calibration with $W = I_q$. The minimum bias at this sample size is again with the petroleum calibration, where $\bar{\lambda} - \lambda_{PLIM} = .0015$ for $W = (Z'Z)^{-1}$.

The point estimation results for ψ are much better, with the maximum bias at $T = 500$ being $\bar{\psi} - \psi_{PLIM} = .0529$ for the apparel calibration with $W = I_q$. At $T = 5000$ the maximum bias is only $\bar{\psi} - \psi_{PLIM} = .0058$, also for the apparel calibration with $W = I_q$.

Looking at the standard errors for $\hat{\lambda}$ from the stockout avoidance specification, we again obtain very large median standard errors using the conventional approximations, particularly when $W = I_q$ and regardless of sample size. Many of these are an order of magnitude larger than the actual standard

deviations of the estimates. On the other hand, the median approximations obtained using HI asymptotics to estimate the standard errors are generally too small. These estimates tend to be less than half the size of the actual standard deviations in most cases and at all sample sizes. Only for the rubber calibration with $W = I_q$ do the median HI standard errors come close to the actual standard deviations of the estimates at each sample size. Despite this, the HI median approximations are always closer in size to the actual standard deviations than are those using conventional asymptotics. In addition, in the vast majority of trials the HI standard errors are again smaller than are those calculated in the usual way.

Looking at the standard errors for $\hat{\psi}$ a different picture emerges. For $W = I_q$, the HI median standard errors are extremely close to the actual standard deviations of the estimates, even when using the small sample size $T = 500$ for many of the calibrations. In contrast, the conventional approximations lead to median standard errors that generally are much larger at all samples sizes and for most of the calibrations. When using $W = (Z'Z)^{-1}$, the HI median approximations are again very close to the actual standard deviations, except for perhaps with the apparel and the petroleum calibrations, in which cases they are slightly large.

Interestingly, for the food calibration the HI median approximations and those from the conventional approximation are very close to each other for both weighting matrices for ψ . This implies that μ_* is very close to zero here, so perhaps the model is close to being correctly specified. This may be due to the fact that for this calibration the estimated probability limits for both parameters in this model are very close to their true values from the full model. This could happen if the omitted parameter, δ , plays only a small role in the data generating process. In fact, considering the noise ratio results above, this may very well be the case. Recall that for the food calibration the average noise ratio for the stockout avoidance model was only slightly larger than that for the full model, and in fact the SA model was chosen by the noise ratio as providing a better fit to the data close to a third of the time for each choice of weighting matrix.

From the estimation results we see that conventional asymptotic approximations for the standard errors do a poor job when the model is misspecified. Use of the alternative HI asymptotics for misspecified

models greatly improves the accuracy of the approximations toward the actual standard errors of the estimates in our trials. However, even these improved approximate standard errors are often fairly distant from the true standard deviations. This will impact the accuracy of inference performed on model parameters that is based upon these standard errors. This is what we consider in the next section.

Inference in Misspecified Models

In order to ascertain the impact of misspecification on inference in these models, we report the rejection rates for two separate 5% significance level t -tests. The first, which we will denote t_{PLIM} , is a test of the hypothesis $H_0 : \theta = \theta_{PLIM}$ and the second, denoted t_0 , tests the hypothesis $H_0 : \theta = \theta_0$. These t -statistics are constructed both using standard errors derived from conventional asymptotics and using standard errors based on HI asymptotics. The rejection rates are calculated from the 1000 trials for each of the calibrations, weight matrices, and sample sizes. Since the nominal size of each test is five percent, comparison of rejection rates allows us to see how well the HI asymptotics perform relative to conventional methods. Given the above results concerning estimated standard errors, we should expect inference based on HI standard errors to be more appropriately sized.

Results for the pure model are contained in Table 3.10, while those for the stockout avoidance specification are in table 3.11. The rejection rates for t_{PLIM} will indicate the actual size of the test for each version of the statistic in our experiments, so these are the main results of interest. Although there is no reason to expect the true parameter values to be accurately estimated, given that the PLIM values are generally not close to these, we report the results for t_0 in order to ascertain how often these are contained within 95 percent confidence intervals constructed about the parameter estimates using the standard errors from each approximation method. This is essentially done to see if there is any hope for gaining information about the true parameter values when the model is misspecified.

Looking at the results for the conventional construction of the t -statistic, we see it has very poor size properties. When testing $H_0 : \lambda = \lambda_{PLIM}$ for the pure model, the nominal five percent level test severely under-rejects for five of the six calibrations for all sample sizes, regardless of which weight

matrix is used in estimation. Only for the Food calibration is this not the case. When using $W = I_q$ in the food calibration, the statistic over-rejects, with a rejection rate ranging from 34.3 percent for $T = 500$, to an improved 15.5 percent for $T = 5000$. For $W = (Z'Z)^{-1}$ the size of the statistic is much improved, although it still over-rejects for $T = 500$. The size seems to approach the nominal rate at $T = 1000$, with a rejection rate of 7.5 percent. However, at $T = 5000$ the statistic under-rejects, with a rate 2.5 percent. All of this behavior is explained by the poor ability of the conventional asymptotics to correctly approximate the standard error, and these results follow directly from those reported above. As an example, for the food calibration the estimated standard errors were too small for the two smaller sample sizes, but too large for $T = 5000$ when using $W = (Z'Z)^{-1}$. Hence the t -statistic results. Also, the extremely large standard errors for a number of the calibrations implies very wide confidence intervals, so in these cases the t -statistic under-rejects, and in fact accepts the null in nearly all cases.

Using HI asymptotics with the pure model results in over-rejection of H_0 for sample sizes $T = 500$ and $T = 1000$ for the the total, apparel, chemical, and especially the food calibration when using $W = I_q$. Performance does increase with sample size in these calibrations and for $T = 5000$ the rejection rates for the total, apparel, and chemicals calibrations become fairly close to the nominal size, although the result for the food calibration is still too high of a rejection rate to consider good performance. For the petroleum calibration with $W = I_q$ the t -statistic is close to correctly sized for all of the sample sizes, as would be expected given the closeness of the approximations to the actual standard errors in the estimation results for this calibration. The t -statistics for the rubber calibration are correctly sized for $T = 500$, but then under-reject for the larger sample sizes.

When using theorem two of Hall and Inoue (2003), with $W = (Z'Z)^{-1}$, there is a tendency to under-reject, but not to the extent that we see with the conventional asymptotics. The total, food, and rubber calibrations all over-reject by a small amount for $T = 500$, but this moves to under-rejection as sample size increases. For the apparel, chemicals, and petroleum calibrations there is under-rejection at $T = 500$ which gets worse as we increase the sample size to $T = 1000$. However, for the apparel and chemicals calibrations the size improves again for $T = 5000$. It continues to deteriorate for the petroleum calibration, however.

For the pure model, the performance of the t -statistic using HI asymptotic standard errors does seem to offer an improvement over the use of conventional asymptotics, except for in the food calibration. This is particularly true for $W = I_q$. However, it often takes rather large sample sizes for these gains to be realized.

We next look at the results for the stockout avoidance specification. Recall that for λ the median values for the conventionally approximated standard errors were always much larger than the actual standard deviations. This would imply that 95% confidence intervals constructed from these standard errors are likely to be too wide, and so the rejection rates for the t -statistics will be too low. This is indeed the case for most of the calibrations. For the food calibration, however, there is over-rejection at the smaller sample sizes, which improves as sample size increases. This is surprising given the large median for the approximated standard errors. Only for the total calibration does the size of the t -statistic come close to its nominal value of 5%, but this again becomes too small as T increases to 5000.

The rejection rates for the λ -based t -statistics when using HI asymptotics are nearly always too large for the stockout avoidance specification, well over 50% in some cases. In addition, size does not improve as we increase the sample size, but in fact deteriorates in many cases. Only for the petroleum calibration with $W = I_q$ does the size approach its nominal value, and actually looks like it performs well. On the other hand, the performance of the HI t -statistics testing $H_0 : \psi = \psi_{PLIM}$ looks very good in the majority of cases when using $W = I_q$. Actual size is close to 5% for the total, apparel, food, and petroleum calibrations with this choice of weight matrix, even for $T = 500$. For chemicals, the statistic over-rejects at $T = 500$ but improves to be accurately sized when $T = 5000$. For rubber there is slight improvement at $T = 1000$, but this turns to under-rejection when $T = 5000$. Use of (3.69) when $W = (Z'Z)^{-1}$ leads to actual size being close to correct for only two of the calibrations, total and food, although there is improvement with sample size for the chemicals and rubber calibrations as well.

Use of conventional asymptotics results in severe under-rejection by the t -statistics in nearly all cases when testing $H_0 : \psi = \psi_{PLIM}$. Even for the food calibration, where the median standard errors from both approximations were close to each other in size, the conventional t -statistics tended to under-reject the null hypothesis. There was, however, some improvement at increased sample size in this case.

As a whole, the results for the stockout avoidance model are somewhat contradictory with respect to lending support to the superiority of HI asymptotics. On the one hand the performance is abysmal for inference on λ and conventional asymptotics performed better in many cases. On the other hand, inference on ψ was far better when using HI approximations for the standard errors. It seems that estimation of this model may have been plagued by some of the same problems we had with the correctly specified model. We attempt to investigate this further in the next section of the paper, but first we take a quick look at the results for t_0 .

While the rejection rates for t_0 do not have the same meaning as those for t_{PLIM} , as the population values are not the limiting values for the parameters when the model is misspecified, they are still useful in telling us the percentage of 95% confidence intervals about the parameter estimates which do not contain the true value. For the pure model, we see that there is a wide disparity among these results. When the PLIM value is far from the true value we would expect to see high rejection rates, and this is indeed the case for the total and food calibrations regardless of which method is used for calculating standard errors. However, for all other calibrations we see that nearly all of the confidence intervals contain the true parameter when constructed using conventional standard errors, even though the PLIM value and the true value are fairly distant from each other for some. This isn't likely to be very useful in practice because the confidence intervals constructed using the conventional approximations are likely to be much too wide to be very informative about the true parameters. Alternatively, the rejection rates for t_0 tend to be substantially higher for most calibrations and for either choice of W when using HI standard errors, implying that the true parameter values are unlikely to be covered by confidence intervals constructed using HI asymptotics. We conclude from this that use of neither asymptotic approximation in estimating standard errors is likely to yield very useful information about the population values of the parameters.

3.9 Application

Despite the problems with the simulation results, it would be interesting to see the results from an empirical application of the above methods. We estimate each of the three models using the six BEA

series that we calibrated to in the simulation, with $T = 473$. We report one-step parameter estimates and standard errors based on both conventional and HI approximations in table 3.12. Also reported are J_T statistics and noise ratios for each model and weight matrix calculated from two-step estimation.

Focussing on the statistics for model selection from the two-step procedure, we see that the full model is not always chosen by either method. While there is no reason to expect that it would be in all cases, this may be one of the sources of the problems in the simulation, since we calibrated to this data assuming all three parameters of the full model pertained. The pure model is rejected by J_T for four of the six industries with either choice of W . These are apparel, food, petroleum, and rubber. In all cases the noise ratio is largest for the pure model. Together both statistics indicate the pure model does a poor job in explaining the inventory behavior of these industries.

For two of the six industries, chemicals and petroleum, the stockout avoidance model is chosen by the noise ratios, although for chemicals with $W = (Z'Z)^{-1}$, the noise ratios are virtually the same. For the rubber industry, all of the models are rejected by the J_T statistic, but the full model would be chosen by the NR as that which provides the best fit to the data. In all but one case, choosing the model with the minimum J_T statistic would result in the same model as basing the decision on noise ratios. The exception is again for chemicals with $W = (Z'Z)^{-1}$, in which case the full model has the lower value for J_T . Furthermore, the noise ratios for the food industry indicate that the stockout adjustment model very nearly does as good a job of explaining the data as does the full model. This is also reflected in the closeness of the J_T statistics for these models. This result does not lend much support to the use of noise ratios in conjunction with J_T , as they are essentially telling us the same things here. However, it is good to have additional evidence in support of J_T , given its poor performance in the simulations and also in other research.

Looking at the one-step estimation results, we first note that for the chemicals industry and the food industry we have $\hat{\lambda} \geq 1$ for some of the models, which is economically implausible. This is likely due to the existence of more than one local minimum for the objective function, as we found in the simulation. While there are ways to re-estimate and control for this, we do not do that here. In addition, the one-step estimates for λ tend to be quite different for each specification depending on which choice of W is used.

This is not uncommon in GMM and the estimates tend to converge quickly to similar values when using the two-step or iterated procedures.¹⁹ However, in almost all cases the estimates for ψ are very similar for both choices of W within the SA and full models for each industry. For δ the estimates again differ significantly across choices of W for most industries .

Despite the indication by both the J_T test and the noise ratios that the full model provides the best fit in the majority of industries, we note that the estimates for δ are not significantly different from zero except for in the petroleum industry when using $W = (Z'Z)^{-1}$. This result is consistent with previous research, which generally has not found a role for costs of changing production in explaining inventory behavior.²⁰ A researcher faced with this result would likely opt for the stockout avoidance model for making statements about inventory behavior. In any event, the parameter estimates for ψ are very close to each other for each of these specifications for nearly all industries and weight matrices. The estimates for λ are often close as well.

Of primary interest to us is the behavior of the asymptotic standard errors from each method of calculation. According to the J_T statistic, none of the models is misspecified for total nondurables or the chemicals industry, and only the pure model is for apparel and food. For the rubber industry, all of the models are misspecified according to J_T . The question becomes, then, which asymptotics are appropriate for these models? If the models are truly not misspecified, then either should give the same asymptotic approximation to the standard errors for the weight matrices studied here. Clearly this is not the case for the majority of the estimates, although in the case of $\hat{\psi}$ there are a number of cases in which these are very close to each other for both the full model and the stockout avoidance specification. It may be that in practice it is highly unlikely researchers are ever able to devise a model that fully accounts for the behavior of the data. This was the rationale in Durlauf and Maccini (1995) for devising the noise ratios as a measure of fit. If models are not likely to ever be correctly specified given a particular set of data, this would indicate that HI asymptotics should always be used, regardless of results from specification testing. However, for certain methods of calculating covariance matrices and subsequent weighting matrices in two-step or iterated GMM, which were not studied here, Hall and Inoue (2003)

¹⁹For the two-step estimation results from which the J_T and NR values were obtained, the parameter estimates are very similar for each choice of W .

²⁰See Ramey and West (1997) for a survey.

find that it is possible for the GMM estimator to converge to a degenerate limiting distribution or a constant in some cases when the model is misspecified. In addition, for some methods HI asymptotics do not become equivalent to conventional asymptotics when the model is correctly specified. This implies that it is best to make a determination as to correct specification before relying on any of these methods. Use of more reliable methods than the J_T test would likely be in order in doing this.

For the pure model, which J_T tells us is misspecified for four of the industries, the conventional standard errors for λ tend to be very large when $W = I_q$ for all but petroleum. The HI standard errors, on the other hand, are relatively conservative and their use would lead to much tighter confidence intervals for λ . Of course given the likely bias in the probability limit for the estimator, there is no reason to believe that the true λ will fall within these limits at anywhere near the rate specified by the confidence level, as we saw in the simulation. When using $W = (Z'Z)^{-1}$, the conventional standard errors still tend to be larger than their HI counterparts, but not to as much of an extreme.

In the cases where the stockout avoidance model is chosen over the full model, i.e. the chemicals and petroleum industries, the standard errors for $\hat{\psi}$ are very similar from each asymptotic approximation, as would be expected if this were the correct specification. However, this is not the case for $\hat{\lambda}$. We see similar results for some industries when the full model is determined to be the best specification. For total nondurables with $W = I_q$ the standard errors are equivalent across methods, and nearly so for food. However, when $W = (Z'Z)^{-1}$ these aren't even close.

Given the simulation results above it is difficult to arrive at any concrete conclusions from this empirical application. In the majority of cases HI asymptotics produce smaller estimated standard errors, but even this doesn't necessarily follow from the theory and in a few cases the HI standard errors are much larger than the conventional. It is certain, however, that in most cases the standard errors from each method tend to be very different, even when the model is determined to be correctly specified by J_T , which most are. This implies that HI asymptotics will certainly have an impact in practice, particularly on inference about model parameters. The difficulty lies in determining when it is appropriate to use these methods.

3.10 Summary and Conclusions

In this paper we study the effects of misspecification on GMM estimation of a structural model of inventory behavior. By generating synthetic data from a complete version of the model we attempted to ascertain the impact of estimating two alternative and incomplete versions of the model that contained subsets of the full model parameter vector. In theory, non-local misspecification should result in parameter estimates that minimize the GMM objective function, but do not necessarily approach the true parameter vector that sets the correctly specified GMM objective function to zero. In generating large-sample probability estimates for the parameters in the two misspecified versions of the model we showed that the limiting values for the parameters were biased away from the true values to varying degrees, with often substantial consequences for statements about the parameters of the model.

We next looked at the performance of the J_T statistic in detecting model misspecification and its ability to correctly choose between models based on a two-step estimation procedure. This statistic was shown to do a poor job of detecting misspecification in the majority of cases and also exhibited poor size properties with the correctly specified model in a number of the cases we looked at. As an alternative means of choosing between nested models, we looked at the noise ratio methods of Durlauf and Maccini (1995). This method performed very well in choosing the correctly specified model in the vast majority of cases.

The final portion of our study looked at the impact on the ability to perform inference based on estimation results from misspecified models. Conventional asymptotic theory does not apply when a model is misspecified, and in our experiments performed poorly in approximating standard errors of the parameter estimates in both misspecified versions of the model. This in turn led to poor size properties of t -statistics constructed from these approximations. We then considered the alternative asymptotics of Hall and Inoue (2003) for models which are misspecified. Applying the alternative asymptotic approximations to the pure model resulted in vastly improved performance relative to conventional asymptotics in the majority of cases for $W = I_q$. However, applying Theorem 2 from Hall and Inoue (2003) when $W = (Z'Z)^{-1}$, while still leading to better performance than did the conventional approximations, did not provide satisfactory results, particularly at the larger sample sizes.

For the stockout avoidance specification, results were much worse. While the approximate standard errors resulting from application of the alternative theory were much closer to the actual standard errors from our experiments than were the conventional estimates, the t -statistics built upon them did not approach correct size in the majority of cases. We found no improvement, and in fact deterioration in many cases, when increasing sample size incrementally from $T = 500$ to $T = 5000$ for most of the calibrations we studied. This was primarily the result of bias in the parameter estimates in combination with inaccurate approximations to the standard errors. The poor performance of the GMM estimator in even the correctly specified version of the model, however, leads us to question the validity of the results as an indictment of the theory.

In the empirical application we found virtually no difference in model selection between the J_T test and use of noise ratios to make this determination. Despite this result, the noise ratio calculation does seem to be a good addition in providing support to the evidence given by the J_T test. We also saw one-step estimation results that mirrored those in the simulation, with HI standard errors generally differing dramatically from the conventional approximations. This perhaps is not surprising considering the synthetic data was calibrated to these industries using a version of the full model. However, only for a few of the models chosen as correctly specified did both conventional and HI asymptotics arrive at similar estimates for the standard errors for any of the model parameters. In most cases the results were very different. This may be a problem with the theory, but also may be due to other factors, such as the possibility that none of the models is sufficiently well specified to explain the data sets used here. About the only conclusion we can derive from the results is that HI asymptotics will likely make a significant difference in the standard errors one obtains for their parameter estimates. The simulation results are not strong enough to allow us make any relevant empirical statements about inventory behavior based on HI asymptotics as opposed to conventional.

The failure of our experiments to adequately provide evidence in support of or in rejection of the usefulness of alternative HI asymptotics may have several sources. As mentioned above, the reliance on first-step estimates may be the source of much of the poor performance. However, even the one-step GMM estimator is theoretically consistent, if not efficient, and should perform better than it did

here, particularly at the larger sample sizes. Clearly the poor performance of the estimator in even the correctly specified version of the model indicates there may be deeper problems.

One possibility is that one or more of the parameters of the model is weakly identified by the generated data, which may in turn be due to the particular normalization chosen for the experiments. Recall that the simulated data for inventories was generated based on calibrated values for four underlying cost parameters from the base inventory model. While the data is consistent with the original model, we are attempting to estimate particular combinations of these underlying parameters. It is well known that GMM results tend to be very sensitive to normalization²¹ and it may be that the combinations of parameters we are attempting to estimate are not well identified by the simulated data. If weak identification is the problem then we began with a model for which the asymptotic theory is not likely to perform well.²² As yet there is no theory available to aid in dealing with misspecification in combination with weak identification of the parameters.²³ These and other possibilities require additional investigation that is beyond the scope of this paper.

Unfortunately, these estimation problems cast some doubt on the previous results as well, even though those results were largely as expected given the underlying theory. Despite this, we would recommend the use of noise ratios in addition to specification testing when choosing between competing models. The well documented low power of the J_T statistic in detecting misspecification warrants this conclusion at the very least.

²¹See Fuhrer et al. (1995) for discussion of this.

²²Stock and Wright (2000) discuss this problem and suggest alternative methods of constructing confidence intervals in the event of weak identification.

²³We did conduct some preliminary investigation into whether weak identification may be a cause of our problems. For all of the models and calibrations we checked to see if the true values or the PLIM values of the parameters jointly satisfied the identifying restrictions of the GMM estimator, which are a necessary condition in estimation (see Hall (2003)). The results thus far are inconclusive.

Table 3.1: Examples of Papers Performing Inference Based on Misspecified Models

Report GMM Estimates and Standard Errors	Conduct Hypothesis Testing on Parameters
<p>Durlauf and Maccini 1995</p> <p>Eichenbaum 1989</p> <p>Cochrane 1996</p> <p>Epstein and Zin 1991</p> <p>Yashiv 2000</p> <p>Jacobs 1999</p>	<p>Meghir and Weber 1996</p> <p>Ferson and Constantinides 1991</p> <p>Hyslop 2001</p>

Table 3.2: Calibrated Parameter Values By Industry

Parameter	Total	Apparel	Chemical	Food	Petroleum	Rubber
a_0	1	1	1	1	1	1
a_1	5.8980	2.5390	13.2180	7.7091	2.1203	6.6796
a_2	0.0320	0.0218	0.0077	0.0071	0.0630	0.1738
a_3	7.1910	9.4317	89.4608	186.6700	2.4756	3.3489
ρ_1	1.0600	1.1466	1.0445	1.0766	1.1045	0.9704
ρ_2	-0.1200	-0.2128	-0.0646	-0.1013	-0.2204	-0.1000
π_1	0.0060	-0.0240	-0.0442	0.0477	0.0467	0.0176
π_2	-0.0650	0.0564	0.0467	-0.0549	-0.0491	0.0015
ϕ_1	0.7500	0.6565	0.7629	0.6410	0.8787	0.8152
ϕ_2	0.2000	0.2035	0.2142	0.1845	0.0637	0.1604
$var(\epsilon_S)$	0.11625	0.30736	0.05480	0.36916	0.11877	0.05580
$var(u)$	3.5	3.5	3.5	3.5	3.5	3.5
$\rho_{\epsilon_S, u}$	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
$CV(\Delta S)$	14.60	24.54	8.80	10.80	21.83	7.35
$CV(\Delta H)$	9.70	9.42	4.64	8.02	21.20	5.38
α	0.0012	0.0012	0.0014	0.0030	0.0008	0.0008
β	0.0027	0.0033	0.0001	0.0007	0.0021	0.0020
K_S	531	553	324	3541	571	1068
K_H	1050	83	239	484	535	476

Table 3.3: Implied Population Values of Model Parameters

	Total	Apparel	Chemicals	Food	Petroleum	Rubber
λ	0.9310	0.9139	0.9785	0.9724	0.8441	0.8535
ψ	1.0390	1.0810	1.0523	1.1724	1.0736	1.0871
δ	0.1695	0.3939	0.0757	0.1297	0.4716	0.1497

Table 3.4: Large Sample Probability Limits By Specification & Weight Matrix

Model		W	λ	ψ	δ	W	λ	ψ	δ
		Total Calibration				Apparel Calibration			
Pure	One-step	I	0.8270	-	-	I	0.9721	-	-
		$(Z'Z)^{-1}$	0.8473	-	-	$(Z'Z)^{-1}$	0.9478	-	-
	Two-step	I	0.8469	-	-	I	0.9478	-	-
		$(Z'Z)^{-1}$	0.8473	-	-	$(Z'Z)^{-1}$	0.9478	-	-
Stockout Avoidance	One-step	I	0.9878	1.0444	-	I	0.9677	1.1585	-
		$(Z'Z)^{-1}$	0.9503	1.0444	-	$(Z'Z)^{-1}$	0.9192	1.1606	-
	Two-step	I	0.9504	1.0444	-	I	0.9193	1.1607	-
		$(Z'Z)^{-1}$	0.9503	1.0444	-	$(Z'Z)^{-1}$	0.9192	1.1606	-
Full	Population Values	-	0.9310	1.0390	0.1695	-	0.9139	1.0810	0.3939
		Chemical Calibration				Food Calibration			
Pure	One-step	I	0.9874	-	-	I	0.7411	-	-
		$(Z'Z)^{-1}$	0.9729	-	-	$(Z'Z)^{-1}$	0.7705	-	-
	Two-step	I	0.9729	-	-	I	0.7684	-	-
		$(Z'Z)^{-1}$	0.9729	-	-	$(Z'Z)^{-1}$	0.7705	-	-
Stockout Avoidance	One-step	I	0.9759	1.0241	-	I	0.9793	1.1700	-
		$(Z'Z)^{-1}$	0.9273	1.0256	-	$(Z'Z)^{-1}$	0.9600	1.1707	-
	Two-step	I	0.9274	1.0256	-	I	0.9601	1.1707	-
		$(Z'Z)^{-1}$	0.9273	1.0256	-	$(Z'Z)^{-1}$	0.9601	1.1707	-
Full	Population Values	-	0.9785	1.0523	0.0757	-	0.9724	1.1724	0.1297
		Petroleum Calibration				Rubber Calibration			
Pure	One-step	I	0.8356	-	-	I	0.9569	-	-
		$(Z'Z)^{-1}$	0.8605	-	-	$(Z'Z)^{-1}$	0.9413	-	-
	Two-step	I	0.8597	-	-	I	0.9413	-	-
		$(Z'Z)^{-1}$	0.8605	-	-	$(Z'Z)^{-1}$	0.9412	-	-
Stockout Avoidance	One-step	I	0.9786	1.0632	-	I	0.9324	1.0274	-
		$(Z'Z)^{-1}$	0.9063	1.0637	-	$(Z'Z)^{-1}$	0.9181	1.0283	-
	Two-step	I	0.9067	1.0636	-	I	0.9182	1.0283	-
		$(Z'Z)^{-1}$	0.9063	1.0637	-	$(Z'Z)^{-1}$	0.9182	1.0283	-
Full	Population Values	-	0.8441	1.0736	0.4716	-	0.8535	1.0871	0.1497

Table 3.5: Rejection Rates of \mathbf{J}_T -test (%), Average Noise Ratios ($\bar{\mathbf{N}}\mathbf{R}$), and Noise Ratio Model Selection by Specification and Weight Matrix.

		Calibration								
		Total			Apparel			Chemicals		
Model	W	\mathbf{J}_T	$\bar{\mathbf{N}}\mathbf{R}$	Select	\mathbf{J}_T	$\bar{\mathbf{N}}\mathbf{R}$	Select	\mathbf{J}_T	$\bar{\mathbf{N}}\mathbf{R}$	Select
Pure	I	32.80	0.0171	-	43.00	0.0210	-	7.30	0.0113	-
	$(Z'Z)^{-1}$	33.40	0.0170	-	49.80	0.0210	-	12.20	0.0113	-
SA	I	10.50	0.0087	12.0	12.10	0.0090	26.9	7.40	0.0075	22.0
	$(Z'Z)^{-1}$	14.80	0.0087	10.6	16.10	0.0091	27.8	11.50	0.0075	22.8
Full	I	10.00	0.0048	88.0	16.10	0.0072	73.1	9.50	0.0053	78.0
	$(Z'Z)^{-1}$	8.40	0.0041	89.4	22.20	0.0079	72.2	14.10	0.0055	77.2
		Food			Petroleum			Rubber		
Model	W	\mathbf{J}_T	$\bar{\mathbf{N}}\mathbf{R}$	Select	\mathbf{J}_T	$\bar{\mathbf{N}}\mathbf{R}$	Select	\mathbf{J}_T	$\bar{\mathbf{N}}\mathbf{R}$	Select
Pure	I	99.20	0.0489	-	95.20	0.0396	-	10.80	0.0126	-
	$(Z'Z)^{-1}$	99.70	0.0484	-	94.10	0.0396	-	16.40	0.0126	-
SA	I	4.50	0.0061	36.1	64.10	0.0250	0.3	7.90	0.0071	9.2
	$(Z'Z)^{-1}$	7.00	0.0061	29.4	69.80	0.0251	0.8	10.20	0.0071	11.4
Full	I	18.70	0.0056	63.9	5.20	0.0036	99.7	4.40	0.0036	90.8
	$(Z'Z)^{-1}$	12.40	0.0056	70.6	8.90	0.0046	99.2	7.80	0.0039	88.6

The nominal rejection rate for \mathbf{J}_T is 5 percent. The χ^2 critical values are 9.49, 7.81, and 5.99 for the pure model, the stockout avoidance model, and the full model, respectively. The column labeled **Select** indicates the percent of the 1000 trials for which the given specification had the minimum noise ratio. The pure model was never selected by this method.

Table 3.6: Limiting Distributions for Alternative GMM Estimators and the J_T Test Statistic

Case	\mathbf{W}_T	Distribution	Distribution of \mathbf{J}_T
Correctly Specified	All	$T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, \Sigma_0)$	χ_{q-p}^2
Locally Misspecified	All	$T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(\eta, \Sigma_0)$	Non-central χ_{q-p}^2
Non-Locally Misspecified			
First-step estimator:	\mathbf{I}_q	$T^{1/2}(\hat{\theta}_T - \theta_*) \xrightarrow{d} N(0, \Sigma_I)$	NA
	$(Z'Z)^{-1}$	$T^{1/2}(\hat{\theta}_T - \theta_*) \xrightarrow{d} N(0, \Sigma_Z)$	NA
Two-step estimator:	MDS	$T^{1/2}(\hat{\theta}_T - \theta_*) \xrightarrow{d} N(0, \Sigma_Z)$	Diverges

Table 3.7: Estimation Results for the Full Model

W=I						W=inv(Z'Z)					
$\lambda = 0.9310$	mean	median	Total st.dev.	Conv. SE	t(0)	$\lambda = 0.9310$	mean	median	Total st.dev.	Conv. SE	t(0)
T=500	0.9092	0.9078	0.0659	0.1210	10.34%	T=500	0.8873	0.8832	0.0605	0.0936	13.1%
T=1000	0.9247	0.9246	0.0541	0.0927	8.0%	T=1000	0.9095	0.9065	0.0490	0.0745	10.8%
T=5000	0.9318	0.9238	0.0380	0.0760	0.5%	T=5000	0.9333	0.9285	0.0313	0.0408	3.2%
$\psi = 1.0390$						$\psi = 1.0390$					
T=500	1.0476	1.0460	0.0225	0.0257	2.07%	T=500	1.0454	1.0443	0.0247	0.0257	3.0%
T=1000	1.0430	1.0419	0.0147	0.0177	2.8%	T=1000	1.0418	1.0417	0.0157	0.0176	2.7%
T=5000	1.0392	1.0395	0.0065	0.0143	0.0%	T=5000	1.0395	1.0397	0.0064	0.0074	2.4%
$\delta = 0.1695$						$\delta = 0.1695$					
T=500	0.1539	0.1067	0.2279	0.2057	14.37%	T=500	0.1621	0.1107	0.2358	0.1697	14.3%
T=1000	0.1642	0.1483	0.1482	0.1580	10.2%	T=1000	0.1651	0.1453	0.1415	0.1317	10.3%
T=5000	0.1729	0.1702	0.0696	0.0746	5.2%	T=5000	0.1702	0.1652	0.0620	0.0620	5.9%
Apparel						Apparel					
$\lambda = 0.9139$	mean	median	st.dev.	Conv. SE	t(0)	$\lambda = 0.9139$	mean	median	st.dev.	Conv. SE	t(0)
T=500	0.9382	0.9712	0.0618	0.6287	1.7%	T=500	0.9050	0.9105	0.0557	0.1658	2.3%
T=1000	0.9514	0.9769	0.0475	0.4191	0.8%	T=1000	0.9185	0.9192	0.0434	0.1170	0.6%
T=5000	0.9702	0.9831	0.0271	0.2152	0.0%	T=5000	0.9234	0.9228	0.0231	0.0506	0.1%
$\psi = 1.0810$						$\psi = 1.0810$					
T=500	1.1928	1.1872	0.0754	0.0947	13.9%	T=500	1.1807	1.1731	0.0721	0.0852	16.8%
T=1000	1.1744	1.1727	0.0432	0.0590	30.8%	T=1000	1.1624	1.1612	0.0444	0.0541	30.6%
T=5000	1.1572	1.1572	0.0166	0.0231	94.8%	T=5000	1.1500	1.1494	0.0188	0.0220	91.0%
$\delta = 0.3939$						$\delta = 0.3939$					
T=500	0.0686	0.0122	0.2231	0.1295	63.8%	T=500	0.0580	0.0134	0.2028	0.1141	67.5%
T=1000	0.0347	0.0123	0.1095	0.0877	82.1%	T=1000	0.0415	0.0220	0.1172	0.0838	82.9%
T=5000	0.0263	0.0211	0.0429	0.0398	100.0%	T=5000	0.0417	0.0363	0.0494	0.0400	99.9%
Chemicals						Chemicals					
$\lambda = 0.9785$	mean	median	st.dev.	Conv. SE	t(0)	$\lambda = 0.9785$	mean	median	st.dev.	Conv. SE	t(0)
T=500	0.9442	0.9739	0.0571	1.1112	8.0%	T=500	0.9182	0.9284	0.0554	0.1737	12.69%
T=1000	0.9610	0.9875	0.0455	1.0056	8.3%	T=1000	0.9360	0.9403	0.0423	0.1231	11.9%
T=5000	0.9791	0.9966	0.0281	0.6442	4.1%	T=5000	0.9531	0.9543	0.0236	0.0587	8.0%
$\psi = 1.0523$						$\psi = 1.0523$					
T=500	1.0457	1.0429	0.0183	0.0313	2.0%	T=500	1.0478	1.0444	0.0210	0.0248	3.10%
T=1000	1.0354	1.0337	0.0107	0.0161	13.5%	T=1000	1.0359	1.0345	0.0109	0.0139	19.1%
T=5000	1.0271	1.0268	0.0037	0.0053	93.6%	T=5000	1.0271	1.0268	0.0037	0.0051	99.6%
$\delta = 0.0757$						$\delta = 0.0757$					
T=500	-0.0454	-0.0692	0.1032	0.0826	45.1%	T=500	-0.0633	-0.0781	0.0806	0.0664	56.55%
T=1000	-0.0610	-0.0742	0.0649	0.0523	65.8%	T=1000	-0.0643	-0.0747	0.0577	0.0462	74.8%
T=5000	-0.0691	-0.0706	0.0228	0.0225	99.1%	T=5000	-0.0604	-0.0616	0.0232	0.0223	99.7%

This table contains the estimation results for λ , ψ , and δ from 1000 samples of size T for each choice of weighting matrix. T is given in first column of each side. The average, median, and standard deviation of the 1000 estimates comprise the next three columns, followed by the median value of the 1000 asymptotic standard errors calculated using conventional asymptotics. The final column in each section gives the rejection rate of the t-statistic testing that the parameter is equal to its population value.

Table 3.7 (continued): Estimation Results for the Full Model

W=I						W=inv(Z'Z)					
$\lambda = 0.9724$	Food					$\lambda = 0.9724$	Food				
	mean	median	st.dev.	Conv. SE	t(0)		mean	median	st.dev.	Conv. SE	t(0)
T=500	0.8974	0.9163	0.0868	0.2086	21.0%	T=500	0.8699	0.8742	0.0826	0.1346	26.0%
T=1000	0.9329	0.9697	0.0672	0.4277	14.3%	T=1000	0.9137	0.9246	0.0648	0.1564	14.8%
T=5000	0.9591	0.9817	0.0426	0.2496	9.5%	T=5000	0.9570	0.9622	0.0382	0.1290	5.7%
$\psi = 1.1724$						$\psi = 1.1724$					
T=500	1.1844	1.1798	0.0472	0.0495	2.9%	T=500	1.1801	1.1755	0.0498	0.0485	3.5%
T=1000	1.1736	1.1721	0.0310	0.0346	3.4%	T=1000	1.1722	1.1710	0.0307	0.0339	2.5%
T=5000	1.1672	1.1670	0.0146	0.0151	3.9%	T=5000	1.1679	1.1672	0.0141	0.0146	6.6%
$\delta = 0.1297$						$\delta = 0.1297$					
T=500	-0.0411	-0.1419	0.3978	0.3087	31.1%	T=500	0.0626	-0.0128	0.5253	0.2082	23.6%
T=1000	0.0629	-0.0078	0.3630	0.2703	27.2%	T=1000	0.1096	0.0574	0.2581	0.1798	15.1%
T=5000	0.1307	0.1036	0.2060	0.1519	15.8%	T=5000	0.1217	0.1070	0.1016	0.0913	9.3%
Petroleum						Petroleum					
$\lambda = 0.8441$	mean	median	st.dev.	Conv. SE	t(0)	$\lambda = 0.8441$	mean	median	st.dev.	Conv. SE	t(0)
T=500	0.8592	0.8622	0.0976	0.1508	0.8%	T=500	0.8400	0.8436	0.0838	0.1263	0.54%
T=1000	0.8618	0.8561	0.0805	0.0923	1.0%	T=1000	0.8475	0.8452	0.0694	0.0847	1.0%
T=5000	0.8581	0.8545	0.037	0.0374	0.6%	T=5000	0.8497	0.8472	0.3180	0.0359	0.6%
$\psi = 1.0736$						$\psi = 1.0736$					
T=500	1.0800	1.0760	0.0313	0.0284	2.6%	T=500	1.0817	1.0770	0.0336	0.0294	3.16%
T=1000	1.0717	1.0704	0.0176	0.0180	3.7%	T=1000	1.0745	1.0734	0.0195	0.0192	4.6%
T=5000	1.0666	1.0661	0.0071	0.0076	16.5%	T=5000	1.0702	1.0698	0.0079	0.0082	7.3%
$\delta = 0.4716$						$\delta = 0.4716$					
T=500	0.4948	0.3765	0.5298	0.3055	10.8%	T=500	0.4544	0.3350	0.6887	0.2603	16.97%
T=1000	0.4645	0.3953	0.3117	0.2132	11.4%	T=1000	0.4441	0.3791	0.3164	0.1963	15.6%
T=5000	0.4262	0.4168	0.1006	0.0991	11.6%	T=5000	0.4392	0.4304	0.0996	0.0990	10.0%
Rubber						Rubber					
$\lambda = 0.8535$	mean	median	st.dev.	Conv. SE	t(0)	$\lambda = 0.8535$	mean	median	st.dev.	Conv. SE	t(0)
T=500	0.8554	0.8597	0.0916	0.1570	2.8%	T=500	0.8378	0.8375	0.0804	0.1310	3.0%
T=1000	0.8741	0.8642	0.0790	0.1049	2.6%	T=1000	0.8559	0.8482	0.0688	0.0912	2.9%
T=5000	0.8857	0.8782	0.0466	0.0477	0.2%	T=5000	0.8726	0.8691	0.0372	0.0448	0.2%
$\psi = 1.0871$						$\psi = 1.0871$					
T=500	1.0490	1.0461	0.0188	0.0208	47.6%	T=500	1.0496	1.0472	0.0188	0.0206	47.1%
T=1000	1.0382	1.0367	0.0106	0.0123	90.2%	T=1000	1.0384	1.0368	0.0106	0.0122	91.6%
T=5000	1.0303	1.0300	0.0041	0.0047	9960.0%	T=5000	1.0301	1.0298	0.0039	0.0048	99.9%
$\delta = 0.1497$						$\delta = 0.1497$					
T=500	0.1280	0.0789	0.2004	0.1717	14.9%	T=500	0.1075	0.0584	0.2082	0.1356	20.4%
T=1000	0.1002	0.0698	0.1267	0.1110	19.7%	T=1000	0.0969	0.0714	0.1200	0.0982	23.0%
T=5000	0.0804	0.0753	0.0509	0.0477	38.7%	T=5000	0.0946	0.0878	0.0480	0.0463	30.4%

This table contains the estimation results for λ , ψ , and δ from 1000 samples of size T for each choice of weighting matrix. T is given in first column of each side. The average, median, and standard deviation of the 1000 estimates comprise the next three columns, followed by the median value of the 1000 asymptotic standard errors calculated using conventional asymptotics. The final column in each section gives the rejection rate of the t-statistic testing that the parameter is equal to its population value.

Table 3.8: Estimation Results for the Pure Model

W=I	Total					
$\lambda = 0.8270$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8015	0.8019	0.0285	0.0481	0.0270	99.9%
T=1000	0.8186	0.8207	0.0206	0.0337	0.0184	100.0%
T=5000	0.8252	0.8254	0.0082	0.0150	0.0078	100.0%
$\lambda = 0.9721$	Apparel					
T=500	0.9662	0.9672	0.0097	0.6825	0.0079	100.00%
T=1000	0.9697	0.9702	0.0058	0.5198	0.0053	100.0%
T=5000	0.9718	0.9719	0.0025	0.2381	0.0024	100.0%
$\lambda = 0.9874$	Chemicals					
T=500	0.9759	0.9786	0.0127	0.8591	0.0081	100.0%
T=1000	0.9819	0.9831	0.0067	0.6472	0.0051	100.0%
T=5000	0.9866	0.9868	0.0023	0.3388	0.0020	100.0%
$\lambda = 0.7411$	Food					
T=500	0.6893	0.6915	0.0558	0.0373	0.0343	73.7%
T=1000	0.7175	0.7200	0.0394	0.0271	0.0243	84.8%
T=5000	0.7370	0.7373	0.0167	0.0122	0.0112	91.2%
$\lambda = 0.8356$	Petroleum					
T=500	0.8318	0.8318	0.0418	0.0868	0.0464	100.0%
T=1000	0.8338	0.8330	0.0306	0.0610	0.0331	100.0%
T=5000	0.8360	0.8356	0.0139	0.0273	0.0147	100.0%
$\lambda = 0.9569$	Rubber					
T=500	0.9361	0.9406	0.0245	0.4878	0.0218	100.0%
T=1000	0.9472	0.9490	0.0138	0.3816	0.0143	100.0%
T=5000	0.9557	0.9559	0.0052	0.1941	0.0058	100.0%
W=inv(Z'Z)	Total					
$\lambda = 0.8473$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8213	0.8225	0.0239	0.0507	0.0361	99.8%
T=1000	0.8346	0.8355	0.0166	0.0367	0.0250	100.0%
T=5000	0.8451	0.8456	0.0070	0.0166	0.0103	100.0%
$\lambda = 0.9478$	Apparel					
T=500	0.9278	0.9307	0.0232	0.3001	0.0297	100.0%
T=1000	0.9388	0.9402	0.0150	0.2408	0.0196	100.0%
T=5000	0.9462	0.9468	0.0066	0.1199	0.0082	100.0%
$\lambda = 0.9729$	Chemicals					
T=500	0.9369	0.9431	0.0322	0.2741	0.0423	100.0%
T=1000	0.9526	0.9558	0.0216	0.2288	0.0282	100.0%
T=5000	0.9701	0.9713	0.0089	0.1425	0.0104	100.0%
$\lambda = 0.7705$	Food					
T=500	0.7299	0.7320	0.0363	0.0417	0.0462	18.0%
T=1000	0.7522	0.7535	0.0252	0.0302	0.0333	13.1%
T=5000	0.7671	0.7672	0.0109	0.0137	0.0158	1.5%
$\lambda = 0.8605$	Petroleum					
T=500	0.8439	0.8448	0.0283	0.0915	0.0466	100.0%
T=1000	0.8527	0.8535	0.0215	0.0679	0.0328	100.0%
T=5000	0.8593	0.8592	0.0092	0.0315	0.0146	100.0%
$\lambda = 0.9413$	Rubber					
T=500	0.8880	0.8912	0.0397	0.2343	0.0566	100.0%
T=1000	0.9113	0.9145	0.0292	0.2119	0.0388	100.0%
T=5000	0.9368	0.9380	0.0116	0.1349	0.0145	100.0%

This table contains the estimation results for λ from 1000 samples of size T for each choice of weighting matrix. T is given in first column of each side. The average, median, and standard deviation of the 1000 estimates comprise the next three columns, followed by the median value of the 1000 asymptotic standard errors calculated using conventional asymptotics and then using HI asymptotics. The final column gives the percentage of the 1000 samples for which the HI standard error is less than that calculated in the conventional fashion.

Table 3.9: Estimation Results for the Stockout Avoidance Model

W=I						
	Total					
$\lambda = 0.9878$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.9250	0.9323	0.0597	0.1911	0.0422	97.9%
T=1000	0.9477	0.9642	0.0475	0.2609	0.0288	98.4%
T=5000	0.9708	0.9869	0.0294	0.2705	0.0091	100.0%
$\psi = 1.0444$						
T=500	1.0517	1.0495	0.0201	0.0328	0.0180	98.1%
T=5000	1.0442	1.0439	0.0051	0.0093	0.0048	100.0%
Apparel						
$\lambda = 0.9677$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.9255	0.9497	0.0620	0.4994	0.0385	95.7%
T=1000	0.9407	0.9591	0.0479	0.4124	0.0292	99.5%
T=5000	0.9596	0.9703	0.0304	0.2330	0.0122	100.0%
$\psi = 1.1585$						
T=500	1.2114	1.2019	0.0611	0.1271	0.0543	94.9%
T=5000	1.1643	1.1637	0.0144	0.0351	0.0140	100.0%
Chemicals						
$\lambda = 0.9759$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.9304	0.9507	0.0568	0.4902	0.0367	98.7%
T=1000	0.9488	0.9651	0.0443	0.3690	0.0355	99.4%
T=5000	0.9689	0.9793	0.0268	0.2427	0.0234	99.8%
$\psi = 1.0241$						
T=500	1.0456	1.0431	0.0179	0.0459	0.0152	98.7%
T=5000	1.0259	1.0255	0.0035	0.0099	0.0034	100.0%
Food						
$\lambda = 0.9793$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8850	0.8888	0.0860	0.1608	0.0808	84.7%
T=1000	0.9250	0.9477	0.0683	0.2354	0.0601	90.1%
T=5000	0.9622	0.9891	0.0417	0.4410	0.0305	98.0%
$\psi = 1.1700$						
T=500	1.1811	1.1766	0.0440	0.0508	0.0445	85.2%
T=5000	1.1695	1.1688	0.0133	0.0150	0.0129	98.6%
W=inv(Z'Z)						
	Total					
$\lambda = 0.9503$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8975	0.8968	0.0475	0.1255	0.0093	98.5%
T=1000	0.9213	0.9225	0.0372	0.1163	0.0056	99.7%
T=5000	0.9428	0.9435	0.0191	0.0688	0.0022	100.0%
$\psi = 1.0444$						
T=500	1.0519	1.0498	0.0216	0.0320	0.0194	96.2%
T=5000	1.0448	1.0445	0.0055	0.0093	0.0051	100.0%
Apparel						
$\lambda = 0.9192$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8859	0.8864	0.0522	0.2052	0.0260	92.3%
T=1000	0.9038	0.9043	0.0393	0.1694	0.0128	97.5%
T=5000	0.9164	0.9167	0.0205	0.0838	0.0040	99.9%
$\psi = 1.1606$						
T=500	1.1965	1.1879	0.0588	0.1209	0.0889	63.5%
T=5000	1.1633	1.1633	0.0149	0.0349	0.0203	96.7%
Chemicals						
$\lambda = 0.9273$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8800	0.8796	0.0576	0.1617	0.0374	87.9%
T=1000	0.9008	0.8981	0.0439	0.1230	0.0343	85.1%
T=5000	0.9251	0.9244	0.0246	0.0658	0.0177	85.9%
$\psi = 1.0256$						
T=500	1.0484	1.0448	0.0202	0.0420	0.0158	86.2%
T=5000	1.0272	1.0268	0.0038	0.0098	0.0041	86.7%
Food						
$\lambda = 0.9600$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8583	0.8615	0.0737	0.1199	0.0269	89.0%
T=1000	0.9053	0.9115	0.0588	0.1295	0.0147	95.5%
T=5000	0.9460	0.9487	0.0335	0.0978	0.0052	99.6%
$\psi = 1.1707$						
T=500	1.1781	1.1732	0.0431	0.0493	0.0499	46.4%
T=5000	1.1711	1.1702	0.0131	0.0146	0.0142	62.3%

This table contains the estimation results for λ and ψ from 1000 samples of size T for each choice of weighting matrix. T is given in first column of each side. The average, median, and standard deviation of the 1000 estimates comprise the next three columns, followed by the median value of the 1000 asymptotic standard errors calculated using conventional asymptotics and then using HI asymptotics. The final column gives the percentage of the 1000 samples for which the HI standard error is less than that calculated in the conventional fashion.

Table 3.9 (continued): Estimation Results for the Stockout Avoidance Model

W=I							W=inv(Z'Z)						
Petroleum							Petroleum						
$\lambda = 0.9786$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:	$\lambda = 0.9063$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.9459	0.9664	0.0449	0.5698	0.0264	99.9%	T=500	0.8897	0.8910	0.0375	0.1543	0.0172	99.2%
T=1000	0.9592	0.9728	0.0321	0.4516	0.0148	100.0%	T=1000	0.8990	0.8991	0.0282	0.1159	0.0117	100.0%
T=5000	0.9749	0.9780	0.0101	0.2322	0.0054	100.0%	T=5000	0.9048	0.9044	0.0118	0.0539	0.0050	100.0%
$\psi = 1.0632$							$\psi = 1.0637$						
T=500	1.0766	1.0736	0.0231	0.0346	0.0226	99.3%	T=500	1.0773	1.0736	0.0242	0.0341	0.0315	63.6%
T=1000	1.0696	1.0681	0.0145	0.0225	0.0146	99.8%	T=1000	1.0700	1.0690	0.0152	0.0223	0.0202	71.2%
T=5000	1.0643	1.0640	0.0060	0.0096	0.0061	100.0%	T=5000	1.0648	1.0645	0.0063	0.0095	0.0084	92.2%
Rubber							Rubber						
$\lambda = 0.9324$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:	$\lambda = 0.9181$	mean	median	st.dev.	Conv. SE	HI SE	HI<Conv:
T=500	0.8838	0.8861	0.0698	0.2565	0.0700	91.9%	T=500	0.8529	0.8515	0.0598	0.1785	0.0406	87.1%
T=1000	0.9029	0.9055	0.0615	0.2066	0.0570	88.0%	T=1000	0.8780	0.8741	0.0506	0.1518	0.0259	92.4%
T=5000	0.9277	0.9261	0.0398	0.1179	0.0379	89.3%	T=5000	0.9112	0.9102	0.0267	0.0965	0.0097	96.6%
$\psi = 1.0274$							$\psi = 1.0283$						
T=500	1.0484	1.0458	0.0180	0.0356	0.0167	94.6%	T=500	1.0501	1.0478	0.0181	0.0343	0.0158	85.4%
T=1000	1.0373	1.0357	0.0102	0.0215	0.0103	92.3%	T=1000	1.0385	1.0367	0.0103	0.0212	0.0090	89.9%
T=5000	1.0289	1.0286	0.0037	0.0086	0.0041	91.9%	T=5000	1.0297	1.0295	0.0036	0.0086	0.0032	95.0%

This table contains the estimation results for λ and ψ from 1000 samples of size T for each choice of weighting matrix. T is given in first column of each side. The average, median, and standard deviation of the 1000 estimates comprise the next three columns, followed by the median value of the 1000 asymptotic standard errors calculated using conventional asymptotics and then using HI asymptotics. The final column gives the percentage of the 1000 samples for which the HI standard error is less than that calculated in the conventional fashion.

Table 3.10: t-statistic Rejection Rates: Pure Model

W=I	Total				W=inv(Z'Z)	Total			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	16.8%	2.0%	99.4%	83.5%	T=500	7.8%	1.0%	89.9%	60.9%
T=1000	11.3%	0.5%	99.9%	99.5%	T=1000	2.7%	0.3%	99.5%	87.4%
T=5000	6.1%	0.3%	100.0%	100.0%	T=5000	1.1%	0.0%	100.0%	100.0%
	Apparel					Apparel			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	11.3%	0.0%	98.7%	0.0%	T=500	2.3%	0.0%	11.1%	0.0%
T=1000	7.0%	0.0%	100.0%	0.0%	T=1000	0.7%	0.0%	30.7%	0.0%
T=5000	6.4%	0.0%	100.0%	0.0%	T=5000	1.7%	0.0%	95.4%	0.0%
	Chemicals					Chemicals			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	16.3%	0.0%	10.3%	0.0%	T=500	2.9%	0.5%	3.4%	0.5%
T=1000	12.6%	0.0%	22.6%	0.0%	T=1000	1.4%	0.0%	1.4%	0.0%
T=5000	7.7%	0.0%	92.1%	0.0%	T=5000	3.8%	0.0%	1.5%	0.0%
	Food					Food			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	38.2%	34.3%	100.0%	100.0%	T=500	10.5%	14.5%	100.0%	100.0%
T=1000	29.4%	25.3%	100.0%	100.0%	T=1000	3.5%	7.6%	100.0%	100.0%
T=5000	19.4%	15.5%	100.0%	100.0%	T=5000	0.7%	2.5%	100.0%	100.0%
	Petroleum					Petroleum			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	7.2%	0.3%	8.3%	0.4%	T=500	1.9%	0.0%	1.0%	0.0%
T=1000	5.8%	0.2%	6.8%	0.5%	T=1000	1.0%	0.0%	0.9%	0.0%
T=5000	4.3%	0.1%	7.9%	0.6%	T=5000	0.4%	0.0%	7.6%	0.0%
	Rubber					Rubber			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	4.9%	0.0%	83.3%	0.0%	T=500	6.7%	0.3%	12.9%	0.0%
T=1000	2.9%	0.0%	99.2%	0.0%	T=1000	3.8%	0.1%	35.8%	0.0%
T=5000	2.6%	0.0%	100.0%	0.0%	T=5000	2.2%	0.0%	100.0%	0.0%

Table 3.11: t-statistic Rejection Rates: Stockout Avoidance Model

W=I					W=inv(Z'Z)				
λ	Total				λ	Total			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	30.5%	11.2%	37.6%	3.0%	T=500	76.4%	6.6%	71.5%	3.7%
T=1000	26.8%	7.0%	47.6%	1.3%	T=1000	78.8%	2.3%	78.2%	0.9%
T=5000	60.4%	2.7%	61.1%	0.1%	T=5000	83.3%	0.1%	88.2%	0.0%
ψ					ψ				
T=500	4.3%	0.4%	5.5%	0.4%	T=500	5.7%	0.3%	6.2%	0.6%
T=1000	6.2%	0.2%	8.4%	0.3%	T=1000	5.7%	0.2%	8.1%	0.6%
T=5000	6.5%	0.1%	15.7%	0.7%	T=5000	6.9%	0.1%	17.4%	1.1%
λ	Apparel				λ	Apparel			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	16.8%	1.8%	42.2%	0.4%	T=500	31.6%	0.7%	30.9%	0.5%
T=1000	29.3%	1.0%	47.8%	0.1%	T=1000	46.6%	0.1%	44.7%	0.1%
T=5000	46.6%	0.5%	65.4%	0.0%	T=5000	68.9%	0.0%	69.1%	0.0%
ψ					ψ				
T=500	5.9%	0.0%	60.7%	1.3%	T=500	2.3%	0.0%	12.7%	1.2%
T=1000	4.7%	0.0%	88.9%	1.6%	T=1000	1.2%	0.0%	30.2%	1.8%
T=5000	5.6%	0.0%	100.0%	84.9%	T=5000	0.9%	0.0%	99.7%	83.7%
λ	Chemicals				λ	Chemicals			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	18.5%	1.4%	17.9%	1.6%	T=500	35.5%	2.6%	50.5%	8.4%
T=1000	25.5%	1.0%	24.5%	1.0%	T=1000	29.8%	1.0%	48.4%	6.1%
T=5000	37.0%	0.4%	34.6%	0.4%	T=5000	21.6%	0.1%	59.5%	7.8%
ψ					ψ				
T=500	16.5%	0.0%	18.4%	0.0%	T=500	21.3%	0.0%	19.2%	0.0%
T=1000	12.4%	0.0%	50.5%	0.5%	T=1000	11.9%	0.0%	41.6%	0.6%
T=5000	5.1%	0.0%	99.9%	92.7%	T=5000	3.2%	0.0%	87.4%	88.6%
λ	Food				λ	Food			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	27.7%	23.4%	26.0%	22.0%	T=500	61.9%	23.9%	67.5%	26.6%
T=1000	19.8%	15.5%	19.5%	15.0%	T=1000	59.5%	13.7%	62.4%	15.4%
T=5000	26.2%	7.6%	38.8%	6.4%	T=5000	78.9%	2.6%	75.7%	5.6%
ψ					ψ				
T=500	2.7%	1.4%	3.0%	1.2%	T=500	3.2%	2.1%	3.5%	2.3%
T=1000	4.4%	1.8%	4.8%	2.0%	T=1000	2.4%	1.8%	2.6%	1.9%
T=5000	5.4%	2.6%	6.3%	2.3%	T=5000	3.8%	2.7%	3.8%	3.0%
λ	Petroleum				λ	Petroleum			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	9.1%	1.4%	65.9%	0.0%	T=500	37.3%	0.1%	60.5%	0.0%
T=1000	7.0%	0.9%	81.6%	0.0%	T=1000	42.8%	0.1%	84.2%	0.0%
T=5000	5.7%	0.0%	100.0%	0.0%	T=5000	40.9%	0.0%	100.0%	0.0%
ψ					ψ				
T=500	3.4%	0.4%	4.9%	0.2%	T=500	1.1%	0.5%	1.7%	0.2%
T=1000	4.0%	0.0%	7.9%	0.5%	T=1000	1.3%	0.1%	3.7%	0.7%
T=5000	5.0%	0.0%	36.4%	8.4%	T=5000	0.7%	0.1%	15.2%	8.1%
λ	Rubber				λ	Rubber			
	HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)		HI: t(PLIM)	conv: t(PLIM)	HI: t(0)	conv: t(0)
T=500	21.2%	2.8%	26.4%	0.2%	T=500	37.8%	2.7%	24.3%	0.3%
T=1000	29.2%	2.0%	29.2%	0.2%	T=1000	42.4%	1.6%	34.1%	0.1%
T=5000	24.6%	0.1%	36.7%	0.0%	T=5000	42.4%	0.0%	81.4%	0.0%
ψ					ψ				
T=500	12.7%	0.1%	59.3%	12.2%	T=500	15.7%	0.2%	57.1%	13.0%
T=1000	9.8%	0.0%	90.5%	67.5%	T=1000	12.8%	0.0%	87.4%	69.3%
T=5000	2.9%	0.1%	98.3%	100.0%	T=5000	9.0%	0.0%	98.1%	100.0%

Table 3.12: Estimation Results - Actual Data

Total	W	One-step									Two-step	
		λ	conv. se(λ)	HI se(λ)	ψ	conv. se(ψ)	HI se(ψ)	δ	conv. se(δ)	HI se(δ)	J_T	NR
Pure	I	0.9272	0.1069	0.0402	-	-	-	-	-	-	7.78	0.0191
	(Z'Z) ⁻¹	0.8847	0.0653	0.0552	-	-	-	-	-	-	8.51	0.0191
SA	I	0.8378	0.0474	0.0409	1.0402	0.0200	0.0155	-	-	-	2.80	0.0073
	(Z'Z) ⁻¹	0.8648	0.0574	0.0028	1.0406	0.0201	0.0131	-	-	-	2.64	0.0073
Full	I	0.8633	0.0469	0.0575	1.0421	0.0168	0.0168	0.0500	0.2338	0.8083	0.79	0.0018
	(Z'Z) ⁻¹	0.8609	0.0502	0.0169	1.0387	0.0159	0.0869	0.1871	0.2338	0.4628	0.35	0.0019
Apparel	W	λ	conv. se(λ)	HI se(λ)	ψ	conv. se(ψ)	HI se(ψ)	δ	conv. se(δ)	HI se(δ)	J_T	NR
Pure	I	0.9506	0.3169	0.0301	-	-	-	-	-	-	10.17*	0.0247
	(Z'Z) ⁻¹	0.9082	0.1683	0.0445	-	-	-	-	-	-	11.08*	0.0247
SA	I	0.9897	1.4130	0.0082	1.1067	0.0530	0.0332	-	-	-	4.58	0.0117
	(Z'Z) ⁻¹	0.9283	0.2309	0.0069	1.1037	0.0529	0.0330	-	-	-	5.21	0.0117
Full	I	0.9887	1.4720	0.0518	1.0865	0.0627	0.0508	0.2329	0.2859	0.3735	1.38	0.0031
	(Z'Z) ⁻¹	0.9300	0.2666	0.0228	1.0798	0.0620	0.1139	0.2232	0.2804	1.0291	1.61	0.0031
Chemicals	W	λ	conv. se(λ)	HI se(λ)	ψ	conv. se(ψ)	HI se(ψ)	δ	conv. se(δ)	HI se(δ)	J_T	NR
Pure	I	0.7965	0.0383	0.0378	-	-	-	-	-	-	8.09	0.0191
	(Z'Z) ⁻¹	0.8721	0.0609	0.0726	-	-	-	-	-	-	6.92	0.0190
SA	I	1.1692	0.0771	0.0682	1.0342	0.0180	0.0178	-	-	-	3.06	0.0115
	(Z'Z) ⁻¹	0.8561	0.0553	0.0088	1.0314	0.0179	0.0188	-	-	-	3.32	0.0115
Full	I	1.1652	0.0620	0.0407	1.0346	0.0167	0.0183	-0.0674	0.0872	0.3863	3.68	0.0128
	(Z'Z) ⁻¹	0.8539	0.0494	0.0045	1.0311	0.0179	0.0407	0.0313	0.1069	0.6771	2.13	0.0116
Food	W	λ	conv. se(λ)	HI se(λ)	ψ	conv. se(ψ)	HI se(ψ)	δ	conv. se(δ)	HI se(δ)	J_T	NR
Pure	I	0.9371	0.1709	0.0356	-	-	-	-	-	-	10.97*	0.0305
	(Z'Z) ⁻¹	0.8855	0.0913	0.0511	-	-	-	-	-	-	12.22*	0.0305
SA	I	0.9547	0.2737	0.1190	1.1860	0.0597	0.0538	-	-	-	0.25	0.0005
	(Z'Z) ⁻¹	0.9886	1.0356	0.0052	1.1817	0.0588	0.0457	-	-	-	0.24	0.0005
Full	I	1.0000	5.0714	0.0896	1.1797	0.0564	0.0577	0.0539	0.1893	0.4178	0.15	0.0003
	(Z'Z) ⁻¹	0.9887	0.8933	0.0154	1.1770	0.0557	0.0988	0.0348	0.1708	0.9862	0.17	0.0003
Petroleum	W	λ	conv. se(λ)	HI se(λ)	ψ	conv. se(ψ)	HI se(ψ)	δ	conv. se(δ)	HI se(δ)	J_T	NR
Pure	I	0.6522	0.0774	0.0947	-	-	-	-	-	-	14.59*	0.0473
	(Z'Z) ⁻¹	0.7503	0.0904	0.0653	-	-	-	-	-	-	11.91*	0.0470
SA	I	0.8570	0.1883	0.0415	1.0538	0.0231	0.0223	-	-	-	3.51	0.0189
	(Z'Z) ⁻¹	0.7962	0.1216	0.0121	1.0607	0.0233	0.0248	-	-	-	3.94	0.0190
Full	I	0.8508	0.1322	0.0063	1.0592	0.0211	0.0121	-0.1413	0.1007	0.1831	6.89*	0.0441
	(Z'Z) ⁻¹	0.9554	0.4575	0.0404	1.0541	0.0196	0.0204	-0.2888	0.0800	0.1189	4.15	0.0313
Rubber	W	λ	conv. se(λ)	HI se(λ)	ψ	conv. se(ψ)	HI se(ψ)	δ	conv. se(δ)	HI se(δ)	J_T	NR
Pure	I	0.9704	0.4582	0.0135	-	-	-	-	-	-	14.00*	0.0339
	(Z'Z) ⁻¹	0.8976	0.1052	0.0474	-	-	-	-	-	-	16.12*	0.0339
SA	I	0.9871	0.8464	0.0084	1.0237	0.0157	0.0100	-	-	-	11.19*	0.0265
	(Z'Z) ⁻¹	0.8757	0.0887	0.0275	1.0273	0.0159	0.0070	-	-	-	14.00*	0.0266
Full	I	0.9909	0.8983	0.0404	1.0252	0.0115	0.0071	-0.1849	0.0388	0.4742	6.49*	0.0179
	(Z'Z) ⁻¹	0.9638	0.2615	0.0349	1.0254	0.0114	0.0058	-0.2066	0.0362	0.3710	7.28*	0.0180

This table contains one-step GMM estimates of model parameters from each of the three model specifications using BEA data on inventories and sales. Also reported are the conventional and HI based asymptotic standard errors for each parameter estimate. The final two columns report the JT-statistic and noise ratio from two-step GMM estimation of the models.

* indicates rejection of the model.

Bibliography

- A. Banerjee, R. L. Lumsdaine, and J. L. Stock. Recursive and sequential tests of the unit-root and trend-break hypothesis: Theory and international evidence. *Journal of Business and Economic Statistics*, 10(3):271–287, July 1992.
- Dan Ben-David, Robin L. Lumsdaine, and David H. Papell. Unit roots, postwar slowdowns and long-run growth: Evidence from two structural breaks. *Unpublished Manuscript*, 1996.
- Mark Bils and James A. Kahn. What inventory behavior tells us about business cycles. *The American Economic Review*, 90(3):458–481, June 2000.
- Olivier J. Blanchard. The production and inventory behavior of the american automobile industry. *Journal of Political Economy*, 91:365–400, 1983.
- Alan S. Blinder. Can the production smoothing model of inventory behavior be saved? *The Quarterly Journal of Economics*, 101(3):431–453, August 1986.
- Alan S. Blinder and Louis J. Maccini. Taking stock: A critical assessment of recent research on inventories. *Journal of Economic Perspectives*, 5(1):73–96, Winter 1991.
- John Y. Campbell and Pierre Perron. Pitfalls and opportunities: What macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, pages 141–201, 1991.
- J. H. Cochrane. A cross-sectional test of an investment asset pricing model. *Journal of Political Economy*, 104:572–621, 1996.
- John H. Cochrane. Comment on pitfalls and opportunities: What macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, pages 201–219, 1991.

- Timothy Cogley and James M. Nason. Effects of the hodrick-prescott filter on trend and difference stationary time series: Implications for business cycle research. *Journal of Economic Dynamics and Control*, 19:253–278, 1995.
- David A. Dickey and Wayne Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Society*, 74:427–431, 1979.
- David A. Dickey and Wayne Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49:1057–1072, 1981.
- Francis X. Diebold and Lutz Kilian. Unit-root tests are useful for selecting forecasting models. *Journal of Business and Economic Statistics*, 18(3):265–273, July 2000.
- David J. Doorn. Trend breaks in finished goods inventories and sales of non-durables. *Unpublished Manuscript, North Carolina State University*, 2001.
- Steven N. Durlauf and Louis J. Maccini. Measuring noise in inventory models. *Journal of Monetary Economics*, 36:65–89, 1995.
- Jürgen Ehlgen. Distortionary effects of the optimal hodrick-prescott filter. *Economics Letters*, 61: 345–349, 1998.
- Martin Eichenbaum. Some empirical evidence on the production level and production cost smoothing models of inventory investment. *The American Economic Review*, 79(4):853–864, September 1989.
- Graham Elliot, Thomas J. Rothemberg, and James H. Stock. Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836, July 1996.
- L. G. Epstein and S. E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99:263–286, 1991.
- W. E. Ferson and G. M. Constantinides. Habit persistence and durability in aggregate consumption. *Journal of Financial Economics*, 29:199–240, 1991.

- Jeffrey Fuhrer, George Moore, and Scott Schuh. Estimating the linear - quadratic inventory model: Maximum likelihood versus generalized method of moments. *Journal of Monetary Economics*, 35: 115–157, 1995.
- Alastair Hall. Testing for a unit root in time series with pretest data-based model selection. *Journal of Business and Economic Statistics*, 12(4):461–470, October 1994.
- Alastair R. Hall. Generalized method of moments. *manuscript*, 2003.
- Alastair R. Hall and Atsushi Inoue. The large sample behavior of the generalized method of moments estimator in misspecified models. *Journal of Econometrics*, 114:361–394, 2003.
- James D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, New Jersey, 1994.
- Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.
- A.C. Harvey and A. Jaeger. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–247, Jul.-Sep. 1993.
- Robert J. Hodrick and Edward C. Prescott. Post-war u.s. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, 29:1–16, 1997.
- Dean R. Hyslop. Rising u.s. earnings inequality and family labor supply: The covariance structure of intrafamily earnings. *The American Economic Review*, 91(4):755–777, September 2001.
- Kris Jacobs. Incomplete markets and security prices: Do asset-pricing puzzles result from aggregation problems? *Journal of Finance*, 54(1), February 1999.
- Regina Kaiser and Agustín Maravall. Estimation of the business cycle: A modified hodrick-prescott filter. *Spanish Economic Review*, 1:175–206, 1999.
- Robert G. King and Sergio T. Rebelo. Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17:207–231, 1993.

- Robin L. Lumsdaine and David H. Papell. Multiple trend breaks and the unit root hypothesis. *The Review of Economics and Statistics*, pages 212–218, 1997.
- James G. MacKinnon. Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business and Economic Statistics*, 12(2):167–176, April 1994.
- Bennett T. McCallum. Alternative monetary policy rules: A comparison with historical settings for the united states, the united kingdom, and japan. *Economic Quarterly of the Federal Reserve Bank of Richmond*, 86:49–79, 2000.
- Jonathan McCarthy and Egon Zakrajsek. Trade inventories. *Staff Reports: Federal Reserve Bank of New York*, (53), December 1998.
- C. Meghir and G. Weber. Intertemporal nonseparability or borrowing restrictions? a disaggregate analysis using a u.s. consumption panel. *Econometrica*, 64:1151–1181, 1996.
- Jeffrey A. Miron and Stephen P. Zeldes. Seasonality, cost shocks, and the production smoothin model of inventories. *Econometrica*, 56(4):877–908, July 1988.
- Charles R. Nelson and Heejoon Kang. Spurious periodicity in inappropriately detrended time series. *Econometrica*, 49(3):741–751, May 1981.
- Whitney K. Newey and Dennis McFadden. Large sample estimation and hypothesis testing. In R. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4, pages 2113–2247. Elsevier Science Publishers, Amsterdam, The Netherlands, 1994.
- Whitney K. Newey and Kenneth D. West. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708, 1987.
- Whitney K. Newey and Kenneth D. West. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61:631–654, 1994.
- Torben Mark Pedersen. Spectral analysis, business cycles, and filtering of economic time series: A survey. *Unpublished Ph.D. dissertation, Institute of Economics, University of Copenhagen*, 1999.

- P. Perron. The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, 57:1361–1401, 1989.
- Valerie A. Ramey. Nonconvex costs and the behavior of inventories. *Journal of Political Economy*, 99(2):306–334, 1991.
- Valerie A. Ramey and Kenneth D. West. Inventories. *Working Paper: National Bureau of Economic Research*, (6315), December 1997.
- Morten O. Ravn and Harald Uhlig. On adjusting the hp-filter for the frequency of observations. *Unpublished Manuscript, Center for Economic Research, Tilburg University, The Netherlands*, May 1997.
- Scott Schuh. Evidence on the link between firm-level and aggregate inventory behavior. *Working Paper*, August 1996.
- G. Schwartz. Estimating the dimension of a model. *The Annals of Statistics*, 6:461–464, 1978.
- G.W. Schwert. Tests for unit roots: A monte carlo investigation. *Journal of Business and Economic Statistics*, 7:147–59, 1989.
- R. Shibata. Selection of the order of an autoregressive model by akaike’s information criterion. *Biometrika*, 63:117–126, 1976.
- Kenneth Singleton. Econometric issues in the analysis of equilibrium business cycle models. *Journal of Monetary Economics*, 21:361–386, 1988.
- James H. Stock and Jonathan Wright. Gmm with weak identification. *Econometrica*, 68:1055–1096, 2000.
- James H. Stock, Jonathan H. Wright, and Motohiro Yogo. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, 20(4): 518–529, October 2002.
- Kenneth D. West. A note on the econometric use of the constant dollar inventory series. *Economics Letters*, 13:337–341, 1983.

Kenneth D. West. The sources of fluctuations in aggregate inventories and gnp. *The Quarterly Journal of Economics*, pages 939–971, November 1990.

Kenneth D. West. Inventory models. *Unpublished Manuscript, University of Wisconsin*, 1992.

Kenneth D. West and David W. Wilcox. A comparison of alternative instrumental variables estimators of a dynamic linear model. *Journal of Business and Economic Statistics*, 14(3):281–293, July 1996.

Eran Yashiv. The determinants of equilibrium unemployment. *The American Economic Review*, 90(5): 1297–1322, 2000.

Eric Zivot and Donald W.K. Andrews. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, 10(3):251–270, July 1992.