

# Small-Sample Properties of GMM-Based Wald Tests

**Craig BURNSIDE**

The World Bank, Washington, DC 20433

**Martin EICHENBAUM**

Department of Economics, Northwestern University, Evanston, IL 60208

This article assesses the small-sample properties of generalized-method-of-moments-based Wald statistics by using (a) a vector white-noise process and (b) an equilibrium business-cycle model as the data-generating mechanisms. In many cases, the small-sample size of the Wald tests exceeds its asymptotic size and increases sharply with the number of hypotheses being jointly tested. We argue that this is mostly due to difficulty in estimating the spectral-density matrix of the residuals. Estimators of this matrix that impose restrictions implied by the model or the null hypothesis substantially improve the properties of the Wald statistics.

**KEY WORDS:** Automated bandwidth selection; Business-cycle model; Monte Carlo simulation; Spectral-density matrix.

This article assesses the small-sample properties of generalized method of moments (GMM) based Wald statistics. The analysis is conducted assuming that the data-generating process corresponds to (a) a simple vector white-noise process and (b) the equilibrium business-cycle model considered by Burnside and Eichenbaum (1994). Our key findings are that, in many of the cases that we examine, the small-sample size of the Wald tests exceeds its asymptotic size. In all cases, size increases uniformly with the dimensionality of joint hypotheses. For tests involving even moderate numbers of moment restrictions, the small-sample size of the tests greatly exceeds its asymptotic size. Relying on asymptotic distribution theory leads one to reject joint hypotheses far too often. We argue that the bulk (but not all) of the problem is the difficulty in estimating the spectral-density matrix of the GMM residuals, which is needed to conduct inference in a GMM environment. In both of our examples, imposing restrictions implied by the underlying model being investigated or the null hypothesis being tested on this spectral-density matrix leads to substantial improvements in the small-sample properties of the Wald tests.

A common approach to evaluating quantitative equilibrium business-cycle models is to compare model- and nonmodel-based estimates of the second moments of aggregate time series. No uniform method for making these comparisons has emerged. Many authors in the real business cycle (RBC) literature make these comparisons in a way that abstracts from sampling uncertainty in estimates of models' structural parameters (e.g., Hansen 1985 or Kydland and Prescott 1982). Other authors have estimated and tested RBC models using full information maximum likelihood methods (e.g., Altug 1989; Christiano 1988; Leeper and Sims 1994; McGratten, Rogerson, and Wright 1993).

An intermediate strategy is to simultaneously estimate model parameters and second moments of the data using a variant of Hansen's (1982) GMM procedure. Christiano and Eichenbaum (1992) showed how, in this framework, simple Wald-type tests can be used to test models' implications

for second moments of the data. Three advantages of this approach are that (1) at the estimation stage of the analysis one need not completely specify agents' environments, (2) it is easy to specify which aspects of the data one wishes to concentrate on for diagnostic purposes, and (3) it is substantially less demanding from a computational point of view than maximum likelihood approaches. The properties of this procedure in small samples are not well understood. This is disturbing in light of recent results in the literature casting doubt on the extent to which asymptotic distribution theory provides a good approximation to various aspects of the small-sample behavior of GMM-based estimators. See, for example, Burnside (1992), Christiano and den Haan (1996), Ferson and Foerster (1991), Fuhrer, Moore, and Schuh (1995), Kocherlakota (1990), Neely (1993), Tauchen (1986), and West and Wilcox (1994).

In this article we address four basic questions concerning the performance of GMM-based Wald statistics. First, does the small-sample size of these tests closely approximate their asymptotic size? Second, do joint tests of several restrictions perform as well or worse than tests of simple hypotheses? Third, how can modeling assumptions, or restrictions imposed by hypotheses themselves, be used to improve the performance of these tests? Fourth, what practical advice, if any, can be given to the practitioner?

We answer these questions under two assumptions about the data-generating process. First, we assume that the true process generating the macro time series is the equilibrium business-cycle model developed by Burnside and Eichenbaum (1994). This case is of interest for two reasons, (1) the model generates time series that in several respects resemble U.S. data and (2) we can study issues of size and inference in an applied context. Second, we assume that the data-generating process corresponds to Gaussian vector

white noise. Working with such a simple process allows us to assess whether the findings that emerge with the more complicated data-generating process also arise in simpler environments. In addition we find it easier to build intuition about our results in the simpler environment.

Our main findings can be summarized as follows. First, there is some tendency for GMM-based Wald tests to over-reject individual moment restrictions. Second, the small-sample size of these tests increases uniformly as the dimension of joint tests increases. For even a moderate number of restrictions, the small-sample size is dramatically larger than the asymptotic size of the test. Indeed correcting for the small-sample properties of the Wald test turns out to have a substantive impact on inference about the empirical performance of the equilibrium business-cycle model that is being analyzed. Third, in our specific applications, the bulk of the problem has to do with the difficulty in accurately estimating the spectral-density matrix of the GMM error terms. We investigate various nonparametric estimators of this matrix that have been suggested in the literature. Although there is some sensitivity to which nonparametric estimator is used, these differences do not affect our basic conclusions. Fourth, we argue that the size characteristics of the Wald tests can be improved if the analyst imposes restrictions that emerge from the model or the hypothesis being tested when estimating the covariance matrix component of the Wald statistic. Not only does such information improve the size of simple tests, it significantly ameliorates the problems associated with tests of joint hypotheses.

The remainder of this article is organized as follows. Section 1 considers the case of the Gaussian white-noise-generating process. Section 2 discusses the case in which the data are generated from an equilibrium business-cycle model. Section 3 contains some concluding remarks.

## 1. GAUSSIAN WHITE-NOISE DATA-GENERATING PROCESS

In this section we consider the small-sample properties of GMM-based Wald statistics within the confines of a very simple statistical environment. In particular we suppose that the data-generating process is a mean-zero, unit-variance Gaussian white-noise process. There are several advantages to working with such a simple process. First, we are able to document that the basic problems that arise in the more complex environment considered in Section 2 also arise here. Second, developing intuition for the results is easier in a simpler environment. Third, we can examine the effects of imposing various assumptions about the data-generating processes on our procedures. Fourth, we can compute all relevant population moments exactly. Fifth, simulation is straightforward, and the number of replications can be increased to gain accuracy in our Monte Carlo experiments.

The remainder of this section is organized as follows. Subsection 1.1 describes the data-generating process. Subsections 1.2 and 1.3 discuss the hypothesis tests and different experiments that we conducted. Finally, we report the results of our Monte Carlo experiments in Subsection 1.4.

### 1.1 The Data-Generating Process

We suppose that an econometrician has time series data on  $J = 20$  random variables  $X_{it}, i = 1, \dots, J$ , each of which are iid  $N(0, 1)$  and mutually independent. The econometrician has  $T = 100$  observations on  $X_{it}, i = 1, \dots, J$ . To assess the robustness of our results, we also conducted experiments in which the data were independent first-order moving average [MA(1)] processes with Gaussian innovations and which were either positively or negatively serially correlated. In addition we redid our experiments assuming that the  $X_{it}$  are mean-zero transformations of chi-squared processes. These perturbations yielded results that were qualitatively similar to those obtained in the Gaussian white-noise case.

To simplify the analysis, we assume that the econometrician knows that  $EX_{it} = 0$  for all  $i$  and  $t$ . The econometrician is interested in estimating and testing hypotheses about the standard deviations,  $\sigma_i$ , of  $X_{it}, i = 1, 2, \dots, J$ . To estimate  $\sigma_i$ , he or she uses a simple, exactly identified GMM estimator based on the moment restriction

$$E(X_{it}^2 - \sigma_i^2) = 0, \quad i = 1, 2, \dots, J. \quad (1)$$

This leads to the GMM estimators

$$\hat{\sigma}_i = \left( \frac{1}{T} \sum_{t=1}^T X_{it}^2 \right)^{1/2}. \quad (2)$$

### 1.2 Hypothesis Testing

The econometrician estimates  $\sigma_i$  to conduct inference. The hypotheses of interest pertain to the variability of the series  $X_{it}$ . The specific hypotheses to be tested are of the form  $H_M: \sigma_1 = \sigma_2 = \dots = \sigma_M = 1, M \leq J$ . The specific Wald statistic that we use to test  $H_M$  is given by

$$\mathcal{W}_T^M = T(\hat{\sigma} - 1)' A' (AV_T A')^{-1} A(\hat{\sigma} - 1). \quad (3)$$

Here  $A = (I_M \ 0_{M \times J-M})$  and  $V_T$  denotes a generic estimator of the asymptotic variance-covariance matrix of  $\sqrt{T}(\hat{\sigma} - \sigma_0)$ , where  $\sigma_0$  is the true value of the parameter vector  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_J)'$ . Given well-behaved estimators  $\hat{\sigma}$  and  $V_T$ ,  $\mathcal{W}_T^M \xrightarrow{d} \chi^2(M)$ .

We consider several questions that arise in testing  $H_M$ . First, how does the choice of estimator  $V_T$  affect inference? We are particularly interested in assessing the small-sample implications of using nonparametric estimators of  $V_T$  and understanding the gains to imposing different types of restrictions on  $V_T$ . Particularly important sources of restrictions are the economic theory being investigated and the null hypothesis being tested. For example, intertemporal consumption-based asset-pricing models typically imply restrictions on the degree of serial correlation in the error terms that define  $V_T$  (e.g., see Eichenbaum and Hansen 1990 or Hansen and Singleton 1982). A different example is provided in Section 2, where we can use the structural model itself to generate an estimate of  $V_T$ . Because imposing restrictions on  $V_T$  can often be computationally burdensome and asymptotic inference is not affected, it is impor-

tant to understand the nature of the small-sample gains in doing so.

Second, how does the dimension of the test—that is, the degrees of freedom  $M$ —affect the size of the test? This question is important because, in many applications, the model gives rise to many overidentifying restrictions. The issue is what trade-offs are involved in simultaneously testing more or fewer of these moment restrictions.

### 1.3 Alternative Covariance Matrix Estimators

In this section we discuss our estimators of the asymptotic covariance matrix of  $\hat{\sigma}$ . The moment conditions used to estimate  $\sigma$ , (1), can be written in the form  $E[u(X_t, \sigma)] = 0$ . Here  $u(\cdot, \cdot)$  is the  $J \times 1$  vector valued function whose  $i$ th element is given by  $(X_{it}^2 - \sigma_i^2)$ . Denoting the true value of  $\sigma$  by  $\sigma_0$ , the asymptotic covariance matrix of  $\sqrt{T}(\hat{\sigma} - \sigma_0)$  is given by  $V_0 = (D_0' S_0^{-1} D_0)^{-1}$ , where

$$D_0 = E \frac{\partial u(X_t, \sigma_0)}{\partial \sigma'}$$

and

$$S_0 = \sum_{j=-\infty}^{\infty} E u(X_t, \sigma_0) u(X_{t-j}, \sigma_0)'$$

The corresponding estimator of  $V_0$  is given by  $V_T = (D_T' S_T^{-1} D_T)^{-1}$ , where  $D_T$  and  $S_T$  are consistent estimators for  $D_0$  and  $S_0$ .

We consider several estimators of  $V_0$ . Each is defined in terms of some estimators  $D_T$  and  $S_T$ . The different estimators impose varying amounts of information at the econometrician's disposal. Some of this information is in the nature of the maintained assumptions concerning the serial and mutual independence properties of  $X_{it}$  and Gaussianity. Other information derives from the null hypothesis being tested.

Initially we consider estimators of  $S_0$  that do not exploit any of this information. Instead we estimate  $S_0$  using versions of the nonparametric estimator proposed by Newey and West (1987). A general version of this estimator can be written as

$$S_T = \sum_{j=-(T-1)}^{T-1} k\left(\frac{j}{B_T}\right) \hat{\Omega}_j,$$

where

$$\hat{\Omega}_j = \begin{cases} (1/T) \sum_{t=j+1}^T u(X_t, \hat{\sigma}) u(X_{t-j}, \hat{\sigma})' & \text{for } j \geq 0 \\ (1/T) \sum_{t=-j+1}^T u(X_{t+j}, \hat{\sigma}) u(X_t, \hat{\sigma})' & \text{for } j < 0 \end{cases}$$

and

$$k(x) = \begin{cases} 1 - |x| & \text{for } |x| \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Here  $B_T$  is a scalar that determines the bandwidth of the lag window  $k(\cdot)$ . We consider three variants of this estimator:

1.  $S_T^1$  uses bandwidth  $B_T = 4$ .

2.  $S_T^2$  uses  $B_T = 2$ .

3.  $S_T^3$  has  $B_T$  chosen automatically using a procedure, suggested by Andrews (1991), which is described in more detail in an appendix available on request from the authors.

The next group of estimators that we consider uses additional amounts of information about the underlying data-generating process. The estimator  $S_T^4$  exploits the assumption that the  $X_{it}$  are serially uncorrelated. This implies that  $S_T^4$  has  $ij$ th element given by  $1/T[\sum_{t=1}^T (X_{it}^2 - \hat{\sigma}_i^2)(X_{jt}^2 - \hat{\sigma}_j^2)]$ .

The estimator  $S_T^5$  imposes the mutual independence of the  $X_{it}$ 's as well as their serial independence. This implies that  $S_T^5$  is a diagonal matrix with  $ii$ th element given by  $1/T[\sum_{t=1}^T (X_{it}^2 - \hat{\sigma}_i^2)^2]$ .

Our next estimator,  $S_T^6$ , also exploits the fact that the  $X_{it}$  are Gaussian. Because Gaussianity implies that  $E(X_{it}^4) = 3\sigma_i^4$ ,  $S_T^6$  is a diagonal matrix with  $ii$ th element given by  $2\hat{\sigma}_i^4$ .

Our next two estimators impose additional restrictions derived from the null hypothesis being tested. Under hypothesis  $H_M$ ,  $\sigma_i = 1$  for  $i = 1, \dots, M$ , but  $\sigma_i$  is unrestricted for  $i = M + 1, \dots, J$ . This suggests the estimator  $S_T^7$ , which is a diagonal matrix with  $ii$ th element 2 for  $i \leq M$  and  $2\hat{\sigma}_i^4$  for  $i = M + 1, \dots, J$ , although these higher elements are irrelevant to the test statistic.

Corresponding to each estimator of  $S_0$  just discussed there is an estimator for  $V_0$  given by

$$V_T^k = [D_T'(S_T^k)^{-1} D_T']^{-1},$$

$k = 1, 2, \dots, 7$ , where  $D_T^1$  is a diagonal matrix with  $ii$ th element  $-2\hat{\sigma}_i$ . Because the null hypothesis can also be imposed on  $D_T^1$ , we also consider the estimator

$$V_T^8 = [D_T^2(S_T^7)^{-1} D_T^2]^{-1},$$

where  $D_T^2$  is a diagonal matrix with  $ii$ th element  $-2$  for  $i \leq M$  and  $-2\hat{\sigma}_i$  for  $i = M + 1, \dots, J$ . Here the  $\mathcal{W}$  statistic reduces to  $\sum_{i=1}^M 2(\hat{\sigma}_i - 1)^2$ . Note that, had we imposed  $\sigma_i = 1$  for all  $i$  in the computation of  $\hat{V}_T^7$  and  $\hat{V}_T^8$ , we would have obtained numerically identical results for our test statistics because all the matrices involved in the calculation are diagonal.

### 1.4 Monte Carlo Experiments

Our experiments were conducted as follows. We generated 10,000 sets of synthetic time series on  $\{X_{1t}, X_{2t}, \dots, X_{Jt}\}_{t=1}^T$ , each of length  $T = 100$ . On each artificial dataset, we estimated the parameter vector  $\sigma$  and the different estimators of the variance-covariance matrix and then calculated the Wald test statistic,  $\mathcal{W}_T^M$ , that is relevant for testing hypothesis  $H_M$ ,  $M \in \{1, 2, 5, 10, 20\}$ . This allowed us to generate an empirical distribution function for  $\mathcal{W}_T^M$  under the null hypothesis that  $H_M$  is true, corresponding to the different estimators of  $V_0$ .

Our results are summarized in Table 1, the columns of which correspond to different specifications of  $M$  (which also equals the degrees of freedom of the test). The rows correspond to fixed asymptotic sizes of the tests, and the

Table 1. Small-Sample Performance of Tests Using Gaussian White-Noise Data

Asymptotic size	Small sample size (%)				
	M = 1	M = 2	M = 5	M = 10	M = 20
(a) Estimated $S_T, B_T = 4$					
1%	2.59	3.41	6.99	16.98	58.68
5%	7.49	9.25	15.61	30.92	73.37
10%	12.65	14.93	23.32	40.10	80.29
(b) Estimated $S_T, B_T = 2$					
1%	2.31	2.87	4.83	9.17	28.88
5%	6.90	8.26	12.22	19.91	45.62
10%	12.03	13.62	19.32	28.55	55.88
(c) Estimated $S_T, B_T$ by Andrews procedure					
1%	2.27	2.91	4.71	9.06	26.64
5%	6.94	8.27	11.94	19.27	43.43
10%	11.98	13.50	19.04	27.87	53.83
(d) Estimated $S_T$ , no lags					
1%	2.15	2.73	4.17	6.67	17.31
5%	6.74	7.94	10.82	16.23	32.87
10%	11.79	13.22	17.43	24.10	42.51
(e) Estimated diagonal $S_T$ , no lags					
1%	2.15	2.67	3.33	3.88	4.71
5%	6.74	7.58	9.32	11.04	13.39
10%	11.79	13.04	15.50	17.56	21.20
(f) Gaussianity applied to (e)					
1%	1.67	1.82	2.22	2.40	2.58
5%	5.94	6.08	7.20	7.72	8.53
10%	10.60	11.30	12.50	13.25	14.45
(g) $H_0$ imposed on $S_T$ in (f)					
1%	1.46	1.67	2.03	2.10	2.10
5%	4.61	5.33	5.97	6.58	7.26
10%	9.34	9.55	10.47	11.70	12.05
(h) $H_0$ imposed on $S_T$ in (f) and on $D_T$					
1%	.96	.97	.99	.96	.92
5%	5.16	4.90	5.08	5.01	4.99
10%	10.14	10.13	10.20	10.11	9.99

entries in the table are the percentages of the 10,000 draws in which the  $\mathcal{W}$  statistic exceeded the relevant critical value of the chi-squared distribution.

Several interesting results emerge here. Consider first the distributions of the test statistics generated using  $V_T^1, V_T^2, V_T^3$ , and  $V_T^4$  [see panels (a)–(d) of Table 1]. First, even for  $M = 1$ , the small-sample sizes of the tests exceed their asymptotic sizes. This result is similar to that obtained by Christiano and den Haan (1996) and Newey and West (1993). Second, the small-sample size of each test rises uniformly with  $M$ . Indeed when we use the estimator  $V_T^1$ , the  $\mathcal{W}$  statistic for hypothesis  $H_{20}$  exceeds its asymptotic (1%, 5%, 10%) critical values (59%, 73%, 80%) of the time. For even moderate sizes of  $M$ , relying on asymptotic distribution theory leads one to reject  $H_M$  far more often than is warranted in small samples. It is true that, as the bandwidth decreases, the small-sample performance of the Wald test improves uniformly. But, as panel (d) indicates, even when we impose the white-noise assumption (i.e., we use  $V_T^4$ ), the

small-sample performance of the large joint tests is dismal. For example, with  $M = 20$ , tests with asymptotic size (1%, 5%, 10%) lead to rejection (17%, 33%, 43%) of the time in samples of 100 observations.

The results generated using  $V_T^5$  (which exploits the assumption that the  $X_{it}$  are mutually independent) are presented in panel (e) of Table 1. Comparing panel (e) to panels (a)–(d), we see that the impact of imposing the independence assumption is to move the small-sample sizes of the tests substantially closer to their asymptotic values. Not surprisingly, the impact of this restriction becomes larger as  $M$  increases because there are more off-diagonal elements being set to their population values. (In the case of  $M = 1$ , the two panels are identical.) With  $M = 20$ , the  $\mathcal{W}$  statistic for  $H_M$  exceeds its asymptotic (1%, 5%, 10%) critical values (4.7%, 13.4%, 21.2%) of the time. This represents a substantial improvement relative to the situation when we do not impose the zero off-diagonal element restriction. Even so, the Wald test still rejects too often in small samples. Panel (f), which reports results based on  $V_T^6$ , indicates that imposing the Gaussianity assumption improves the small performance of  $\mathcal{W}_T^M$  even further. To the extent that fourth moments are less accurately estimated than second moments for Gaussian processes, this result is not surprising.

Recall that the estimator  $V_T^7$  exploits information from the null hypothesis regarding  $\sigma_i$  in constructing  $S_T^7$ . The results generated using  $V_T^7$  are reported in panel (g) of Table 1. Comparing panels (f) and (g), we see that the net effect of imposing these additional restrictions is to move the small-sample size of the test even closer to its asymptotic size (except for the 10% critical value for  $M = 1$ ). For example, with  $M = 20$ , the  $\mathcal{W}$  statistic for  $H_M$  exceeds its asymptotic (1%, 5%, 10%) critical values (2.1%, 7.3%, 12.1%) of the time.

Panel (h) of Table 1 reports results based on  $V_T^8$ , in which we impose the null hypothesis on  $D_T$  as well as on  $S_T$ . Now all of the anomalies associated with the small-sample distribution of the  $\mathcal{W}$  statistic disappear. First, the degree to which the small-sample sizes match their asymptotic sizes is not affected by  $M$ . Second, the small-sample size of the test statistic is extremely close to the corresponding asymptotic size. Indeed, this is true even when we fix the asymptotic size of the test at 1%. So, at least for the present example, the parameter estimates appear to have a small-sample distribution that is very well approximated by their large-sample distribution. The anomalies associated with the small-sample distribution of the  $\mathcal{W}$  statistic seem to be closely related to the small-sample distribution of  $S_T$  and, to a much smaller extent,  $D_T$ . In Section 2 we investigate the extent to which these conclusions continue to hold in a more complex statistical environment.

## 2. A REAL-BUSINESS-CYCLE MODEL AS A DATA-GENERATING PROCESS

In this section we consider the small-sample properties of GMM-based Wald statistics, assuming that the data-generating process is given by the business-cycle model de-

veloped by Burnside and Eichenbaum (1994). The model is briefly summarized in Subsection 2.1. Subsection 2.2 describes the way the model's structural parameters were estimated. Subsection 2.3 discusses the hypothesis tests we investigated. Subsection 2.4 presents estimates of the model using aggregate U.S. data. In Subsection 2.5 we present the results of our Monte Carlo experiments, which use the estimated model as the data-generating mechanism.

## 2.1 The Model

The model economy is populated by many infinitely-lived individuals. To go to work, an individual must incur a fixed cost of  $\zeta$  hours. Once at work, an individual stays for a fixed shift length of  $f$  hours. The time  $t$  instantaneous utility of such a person is given by

$$\ln(C_t) + \theta \ln(T - \zeta - W_t f). \quad (4)$$

Here  $T$  denotes the individual's time endowment,  $C_t$  denotes time- $t$  privately purchased consumption,  $\theta > 0$ , and  $W_t$  denotes the time- $t$  level of effort. The time- $t$  instantaneous utility of a person who does not go to work is given by  $\ln(C_t) + \theta \ln(T)$ .

Time- $t$  output,  $Y_t$ , is produced via the Cobb-Douglas production function

$$Y_t = (K_t U_t)^{1-\alpha} (N_t f W_t X_t)^\alpha, \quad (5)$$

where  $0 < \alpha < 1$ ,  $K_t$  denotes the beginning of time- $t$  capital stock,  $U_t$  represents the capital utilization rate,  $N_t$  denotes the number of individuals at work during time  $t$ , and  $X_t$  represents the time- $t$  level of technology. We assume that the time- $t$  depreciation rate of capital,  $\delta_t$ , is given by

$$\delta_t = \delta U_t^\phi, \quad (6)$$

where  $0 < \delta < 1$  and  $\phi > 1$ . The stock of capital evolves according to

$$K_{t+1} = (1 - \delta_t)K_t + I_t, \quad (7)$$

where  $I_t$  denotes time- $t$  gross investment.

The level of technology,  $X_t$ , evolves according to  $X_t = X_{t-1} \exp(\gamma + v_t)$ , where  $v_t$  is a serially uncorrelated process with mean 0 and standard deviation  $\sigma_v$ . The aggregate resource constraint is given by

$$C_t + I_t + G_t \leq Y_t, \quad (8)$$

where  $G_t$  denotes the time- $t$  level of government consumption. We assume that  $G_t$  evolves according to

$$G_t = X_t g_t^*. \quad (9)$$

Here  $g_t^*$  is the stationary component of government consumption and  $g_t = \ln(g_t^*)$  evolves according to

$$g_t = \mu(1 - \rho) + \rho g_{t-1} + \varepsilon_t, \quad (10)$$

where  $\mu$  is a scalar,  $|\rho| < 1$ , and  $\varepsilon_t$  is a serially uncorrelated process with mean 0 and standard deviation  $\sigma_\varepsilon$ .

In the presence of complete markets, the competitive equilibrium of this economy corresponds to the solution

of the social-planning problem

$$E_0 \sum_{t=0}^{\infty} \beta^t [\ln(C_t) + \theta N_t \ln(T - \zeta - W_t f) + \theta(1 - N_t) \ln(T)], \quad (11)$$

subject to (5)–(10) by choice of contingency plans for  $\{C_t, K_{t+1}, N_{t+1}, U_t, W_t: t \geq 0\}$ . We obtain an approximate solution to this problem using King, Plosser, and Rebelo's (1988) log-linear solution procedure (see Burnside 1993 for details). Let  $k_t = \ln(K_t/X_{t-1})$ ,  $h_t = \ln(H_t)$ ,  $c_t = \ln(C_t/X_t)$ ,  $w_t = \ln(W_t)$ ,  $u_t = \ln(U_t)$ ,  $y_t = \ln(Y_t/X_t)$ ,  $a_t = \ln(Y_t/N_t X_t)$ ,  $i_t = \ln(I_t/X_t)$ ,  $h_t^0 = \ln(H_t^0)$ , and  $a_t^0 = \ln(Y_t/H_t^0 X_t)$ . Here  $H_t$  and  $H_t^0$  denote actual and observed time- $t$  hours of work.

We assume, as did Prescott (1986), that

$$\ln(H_t^0) = \ln(H_t) + \xi_t, \quad (12)$$

where  $\xi_t$  is an iid random variable with mean 0 and variance  $\sigma_\xi^2$ . The time- $t$  state of the system is given by

$$s_t = (1 \quad k_t \quad h_t \quad v_t \quad g_t \quad \xi_t)'$$

Define the vector of time- $t$  endogenous variables  $f_t$  as

$$f_t = (c_t \quad w_t \quad u_t \quad y_t \quad a_t \quad i_t \quad h_t^0 \quad a_t^0)'$$

and the vector of time- $t$  shocks as

$$\tilde{\varepsilon}_t = (0 \quad 0 \quad 0 \quad v_t \quad \varepsilon_t \quad \xi_t)'$$

Our assumptions about the exogenous variables and the log-linear approximation to the model imply that the evolution of the system can be summarized as

$$\begin{aligned} s_t &= M s_{t-1} + \tilde{\varepsilon}_t \\ f_t &= \Pi s_t, \end{aligned} \quad (13)$$

where  $M$  and  $\Pi$  are functions of the model's underlying structural parameters. See Burnside (1993) for details. We take (13) to be the data-generating mechanism in our Monte Carlo experiments.

## 2.2 Estimation

With certain exceptions, the parameters of the model were estimated using a variant of the GMM procedure described by Christiano and Eichenbaum (1992). We did not estimate  $\beta$ ,  $T$ ,  $\zeta$ , and  $f$ . Instead we set  $\beta = 1.03^{-1/4}$ ,  $T = 1,369$  hours per quarter, and  $\zeta = 60$  and chose  $f$  so that the nonstochastic steady-state value of  $W_t$  was 1. Rather than estimate  $\delta$ , we estimated  $\bar{\delta} = \delta U^\phi$ , where  $U$  is the nonstochastic steady-state value of  $U_t$ . This seemed natural because the log-linearized model depends only on  $\bar{\delta}$  and not separately on  $\delta$  or  $U$ . Evaluating the planner's Euler equations for  $U_t$  and  $K_{t+1}$  in nonstochastic steady state, we can solve for the parameter  $\phi$  as a function of the parameters  $\beta$ ,  $\gamma$ , and  $\bar{\delta}$ :

$$\phi(\beta, \gamma, \bar{\delta}) = \frac{\beta^{-1} \exp(\gamma) - 1}{\bar{\delta}} + 1.$$

In the data, the series  $g_t$  displays a time trend, so this series was detrended using a linear time trend. To simplify matters we did not include the time trend in the Monte Carlo experiments. To compute the log-linear solution to the model we need an estimate of  $g/y$ , the nonstochastic steady-state value of  $G_t/Y_t$ . The parameter  $\mu$  [the mean of the process  $g_t = \ln(G_t/X_t)$ ] is relevant only to the extent that it helps us compute  $g/y$ . So rather than estimate  $\mu$ , we chose to estimate  $g/y$ .

In light of these decisions, the vector of model parameters to be estimated, denoted by  $\Psi_1$ , is given by

$$\Psi_1 = (\theta \quad \alpha \quad \bar{\delta} \quad \gamma \quad \sigma_v \quad g/y \quad \rho \quad \sigma_\varepsilon \quad \sigma_\xi)'$$

The hypotheses that we investigate involve various second moments of the data. Because many of the relevant series exhibit marked trends, some stationary-inducing transformation of the data must be applied. To facilitate comparisons with the RBC literature, we chose to process the data using the Hodrick and Prescott (1980) filter. Consequently, the second moments to be discussed pertain to those of Hodrick and Prescott (HP) filtered data. We have redone all of the experiments in this article with first-differenced data. A subset of these is discussed later. For a detailed comparison of some of the small-sample properties of GMM with HP-filtered and first-differenced data, see Christiano and den Haan (1996).

We focus on a set of second moments that have received much attention in the RBC literature—the standard deviation of output,  $\sigma_y$ ; the standard deviations of consumption, investment, and hours relative to the standard deviation of output,  $\sigma_c/\sigma_y, \sigma_i/\sigma_y$ , and  $\sigma_h/\sigma_y$ , and the standard deviation of hours worked relative to the standard deviation of average productivity,  $\sigma_h/\sigma_a$ . We also consider the dynamic correlations between average productivity and hours,  $\rho_{ah}^i = \text{corr}(APL_t, H_{t+i}), i = \pm 1, \pm 2, \pm 3, \pm 4$ , and the dynamic correlations between average productivity and output  $\rho_{ay}^i = \text{corr}(APL_t, Y_{t+i}), i = -4, -3, -1, -1$ . The contemporaneous correlations between these variables and  $\rho_{ay}^i, i = 1, 2, 3, 4$ , can be deduced from the other moments that we consider. We denote the vector of diagnostic moments that must be estimated in ways not involving the model by

$$\Psi_2 = (\sigma_y \quad \sigma_c/\sigma_y \quad \sigma_i/\sigma_y \quad \sigma_h/\sigma_y \quad \sigma_h/\sigma_a \quad \rho_{ah}^{-4} \quad \rho_{ah}^{-3} \quad \rho_{ah}^{-2} \quad \rho_{ah}^{-1} \quad \rho_{ah}^1 \quad \rho_{ah}^2 \quad \rho_{ah}^3 \quad \rho_{ah}^4 \quad \rho_{ay}^{-4} \quad \rho_{ay}^{-3} \quad \rho_{ay}^{-2} \quad \rho_{ay}^{-1})'$$

**2.2.1 Moment Conditions Underlying the Estimator of  $\Psi_1$ .** As discussed by Burnside and Eichenbaum (1994), our estimator of  $\Psi_1$  is based on the following moment conditions:

$$E \left\{ \sigma_\xi^2 - \frac{1}{2} [\Delta \ln(H_t^0)]^2 + \frac{1}{2} \Delta \ln(H_t^0) \Delta \ln(\tilde{H}_t^0) \right\} = 0, \quad (14)$$

$$E[\ln(H_t^0) - \ln(Nf)] = 0, \quad (15)$$

$$E[\ln(\delta_t) - \ln(\bar{\delta})] = 0, \quad (16)$$

$$E \left\{ 1 - \beta \left( \frac{C_t}{C_{t+1}} \right) \times \left[ (1 - \alpha)(1 - \phi(\beta, \gamma, \bar{\delta})^{-1}) \frac{Y_{t+1}}{K_{t+1}} + 1 \right] \right\} = 0, \quad (17)$$

$$E[\ln(X_t^0) - \ln(X_{t-1}^0) - \gamma] = 0, \quad (18)$$

$$E[\ln(X_t^0) - \ln(X_{t-1}^0)]^2 - \sigma_v^2 - 2\varphi_3^2 \sigma_\xi^2 = 0, \quad (19)$$

$$E[\ln(G_t) - \ln(Y_t) - \ln(g/y)] = 0, \quad (20)$$

$$E[(g_t^0 - \rho g_{t-1}^0)g_{t-1}^0] + \rho \varphi_3^2 \sigma_\xi^2 = 0, \quad (21)$$

and

$$E[(g_t^0 - \rho g_{t-1}^0)^2] - (1 + \rho^2)\varphi_3^2 \sigma_\xi^2 - \sigma_\varepsilon^2 = 0. \quad (22)$$

In Equation (14),  $H_t^0$  and  $\tilde{H}_t^0$  refer to our two measures of hours worked, Hansen's (1985) corrected measure of household hours and hours worked from the establishment survey conducted by the Bureau of Labor Statistics. More details concerning the data are provided in an appendix available from the authors. Unlike Burnside and Eichenbaum (1994), we abstract from issues concerning the observability of  $\delta_t$  and  $K_t$ . In particular, we assume, for the purposes of our Monte Carlo experiments, that the econometrician observes these series directly. The variables  $N$ , representing the nonstochastic steady-state value of  $N_t$ , and  $\varphi_3$ , a reduced-form parameter, are functions of the underlying parameter vector,  $\Psi_1$ . Furthermore,  $X_t^0$  represents a measurement-error-corrupted signal of the level of technology that can be constructed given the data and a vector of parameters  $\varepsilon_1$ . Similarly,  $g_t^0$  is a signal of  $g_t$  based on the error-ridden measure of technology  $X_t^0$ . See Burnside and Eichenbaum (1994) for details.

**2.2.2 Moment Conditions Underlying the Estimator of  $\Psi_2$ .** Our estimator of  $\Psi_2$  is based on the following moment conditions:

$$E(y_t^2 - \sigma_y^2) = 0$$

$$E[c_t^2 - (\sigma_c/\sigma_y)^2 y_t^2] = 0$$

$$E[i_t^2 - (\sigma_i/\sigma_y)^2 y_t^2] = 0$$

$$E[h_t^2 - (\sigma_h/\sigma_y)^2 y_t^2] = 0$$

$$E[h_t^2 - (\sigma_h/\sigma_a)^2 a_t^2] = 0$$

$$E \left[ a_t h_{t+i} - \rho_{a,h}^i \left( \frac{\sigma_h}{\sigma_y} \right)^2 \sigma_y^2 / \left( \frac{\sigma_h}{\sigma_a} \right) \right] = 0, \quad i = \pm 1, \dots, \pm 4$$

$$E \left[ a_t y_{t+i} - \rho_{a,y}^i \left( \frac{\sigma_h}{\sigma_y} \right) \sigma_y^2 / \left( \frac{\sigma_h}{\sigma_a} \right) \right] = 0, \quad i = 1, \dots, 4,$$

where a lowercase variable—for example,  $z_t$ —is the cyclical component of  $\ln(Z_t)$  as defined by the HP filter.

To define our joint estimator of  $\Psi_1$  and  $\Psi_2$ , consider the following generic representation of our moment conditions:

Table 2. Hypothesis Tests for the RBC Example

Test #	Moment	Included in test of			
		H1	H2	H3	H4
1	$\sigma_y$	✓			✓
2	$\sigma_c/\sigma_y$	✓			✓
3	$\sigma_i/\sigma_y$	✓			✓
4	$\sigma_h/\sigma_y$	✓			✓
5	$\sigma_h/\sigma_a$	✓			✓
6	$\rho_{ah}^4$		✓		✓
7	$\rho_{ah}^3$		✓		✓
8	$\rho_{ah}^2$		✓		✓
9	$\rho_{ah}^1$		✓		✓
10	$\rho_{ah}^0$		✓		✓
11	$\rho_{ah}^1$		✓		✓
12	$\rho_{ah}^2$		✓		✓
13	$\rho_{ah}^3$		✓		✓
14	$\rho_{ah}^4$		✓		✓
15	$\rho_{ay}^4$			✓	✓
16	$\rho_{ay}^3$			✓	✓
17	$\rho_{ay}^2$			✓	✓
18	$\rho_{ay}^1$			✓	✓
19	$\rho_{ay}^0$			✓	✓
20	$\rho_{ay}^1$			✓	✓
21	$\rho_{ay}^2$			✓	✓
22	$\rho_{ay}^3$			✓	✓
23	$\rho_{ay}^4$			✓	✓

$E[u_t(\Psi^0)] = 0, t = 1, \dots, T$ , where  $\varepsilon^0$  is the true value of  $(\Psi_1' \Psi_2')$  and  $u_t$  is a vector-valued function of dimension equal to the dimension of  $\Psi^0$ . Let

$$g_T(\Psi) = \frac{1}{T} \sum_{t=1}^T u_t(\Psi).$$

The GMM estimator,  $\Psi_T$ , minimizes

$$J_T = Tg_T(\Psi)' \Upsilon_T g_T(\Psi), \tag{23}$$

where  $\Upsilon_T$  is a symmetric, positive definite weighting matrix of dimension equal to the dimension of  $g_T(\Psi)$ . Because our GMM estimator is exactly identified,  $\Psi_T$  is independent of  $\Upsilon_T$ . We simply set  $\Upsilon_T$  equal to the identity matrix in (23).

A consistent estimator of the variance-covariance matrix of  $\sqrt{T}(\Psi_T - \Psi_0)$  is given by  $V_\varepsilon = (D_T' S_T^{-1} D_T)^{-1}$ , where  $D_T = \partial g_T(\Psi_T) / \partial \Psi'$  and  $S_T$  is a consistent estimate of  $S_0$ ,  $2\pi$  times the spectral-density matrix of  $u_t(\Psi^0)$  at frequency 0.

### 2.3 Hypothesis Testing

Suppose that we wish to assess the empirical plausibility of the model's implications for a  $q \times 1$  subset of  $\Psi_2$  given by  $\omega$ . Let  $\Phi(\Psi)$  denote the value of  $\omega$  implied by the model, given the structural parameters  $\Psi_1$ . Here  $\Phi$  denotes the (nonlinear) mapping between the model's structural parameters and the relevant population moments. Denote the nonparametric estimate of  $\omega$  obtained without imposing restrictions from the model by  $\Gamma(\Psi)$ . The hypothesis that we

investigate is of the form

$$H_0: F(\Psi^0) = \Phi(\Psi^0) - \Gamma(\Psi^0) = 0. \tag{24}$$

Christiano and Eichenbaum (1992) showed that a consistent estimate of the asymptotic variance-covariance matrix of  $\sqrt{T}[F(\Psi_T) - F(\Psi^0)]$  is

$$V_F = \left[ \frac{\partial F(\Psi_T)}{\partial \Psi'} \right] V_\Psi \left[ \frac{\partial F(\Psi_T)}{\partial \Psi'} \right]'$$

and that the test statistic

$$W_T = TF(\Psi_T)' V_F^{-1} F(\Psi_T) \tag{25}$$

is asymptotically distributed as a  $\chi^2$  random variable with  $q$  df. We consider two types of hypothesis tests. The first type involves tests of individual moments of the data. The test numbers and corresponding moments being tested are summarized in Table 2. The second type of tests involves joint moment restrictions. Hypothesis H1 tests whether the hypotheses incorporated into tests 1-5 hold simultaneously. Hypotheses H2, H3, and H4 are similar to hypothesis H1 but pertain to different sets of moments, as indicated in Table 2. Hypothesis H4 incorporates all nonredundant moments from tests 1-23.

To implement our hypothesis tests, we require an estimator,  $S_T$ , of  $S_0$ . As in Section 1, our estimators are of the form

$$S_T = \sum_{j=-(T-1)}^{T-1} k\left(\frac{j}{B_T}\right) \hat{\Omega}_j,$$

where

$$\hat{\Omega}_j = \begin{cases} (1/T) \sum_{t=j+1}^T \hat{u}_t \hat{u}'_{t-j} & \text{for } j \geq 0 \\ (1/T) \sum_{t=-j+1}^T \hat{u}_{t+j} \hat{u}'_t & \text{for } j < 0. \end{cases}$$

The kernel function  $k$  varies depending on the estimator,  $B_T$  is the bandwidth, and  $\hat{u}_t = u_t(\Psi_T)$ . Our baseline results are generated using the Bartlett kernel function

$$k(x) = \begin{cases} 1 - |x| & \text{for } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and Andrews's (1991) automatic selection procedure for  $B_T$ . In an available appendix, we discuss the other estimators of  $S_0$  that we considered. As it turns out, our basic results are robust across these different estimators of  $S_0$ .

The bandwidth selection procedure that we used can be described as follows. Andrews (1991) provided an expression for the optimal bandwidth corresponding to a given kernel, a process  $u_t$ , and a set of weights on the different elements of  $S_0$ . The bandwidth is optimal in the sense that it leads to minimum mean squared error estimates of a weighted inner product of the elements of  $S_0$ . Andrews's (1991) procedure simplified the dependence of the optimal bandwidth on the entire spectral density of  $u_t$  by assuming

Table 3. Model Parameters: Estimates and Standard Errors

Parameter	Estimate	Std. error
$\theta$	3.5955	(.0377)
$\alpha$	.6422	(.0193)
$\delta$	.0208	(.0002)
$\gamma$	.0038	(.0012)
$\sigma_v$	.0088	(.0002)
$g/y$	.1763	(.0022)
$g_0$	1.7885	(.0809)
$g_1$	-.0019	(.0003)
$\rho$	.9456	(.0299)
$\sigma_\varepsilon$	.0152	(.0012)
$\sigma_\xi$	.0088	(.0011)

NOTE: All standard errors shown in this table are based on estimates of  $S_T$  computed using the Bartlett window suggested by Newey and West (1987) and the automatic bandwidth selection procedure suggested by Andrews (1991).

a simple parametric model for the error term. The choice of model does not affect the consistency of  $S_T$ . The model that we use corresponds to the simplest example of Andrews (1991). Specifically, we treat the elements of  $u_t$  as independent first-order autoregressive [AR(1)] scalar processes. No weight is given to the off-diagonal elements of  $S_0$ . Under these circumstances, the bandwidth selected will depend on the sample size,  $T$ , the weights, and coefficient estimates obtained by fitting AR(1) processes to the elements of  $u_t(\Psi_T)$ . Roughly speaking, the more persistent the errors, the greater the bandwidth.

In the standard case, equal weight is placed on all of the error terms. With this procedure the median bandwidth across the Monte Carlo draws was 40.1. Moreover, the test statistics had extremely poor small-sample properties. For example, for hypothesis H4, the  $\mathcal{W}$  statistics exceeded their (1%, 5%, 10%) critical values in (99.9%, 100%, 100%) of the Monte Carlo draws. Further details are given in an appendix available on request. Andrews (1991) suggested putting zero weight on the error terms corresponding to constant regressors in regression models. Presumably, this is motivated by the fact that the covariance properties of those error terms are qualitatively dissimilar to the covariance properties of the error terms corresponding to nonconstant regressors. Pursuing this logic, we implemented the Andrews procedure in our context by placing zero weight on (15), (16), and (20) along with unit weight on the other

error terms. This constituted our baseline method and led to test statistics with much better small-sample properties. For example, the resulting median bandwidth across the different Monte Carlo draws was 2.78 and the  $\mathcal{W}$  statistics for H4 exceeded their (1%, 5%, 10%) critical values in (36.7%, 50.7%, 57.6%) of the Monte Carlo draws. Further details of our baseline results follow.

#### 2.4 Parameter Estimates and Some Results Based on Asymptotic Theory

Table 3 reports our point estimates of  $\Psi_1$  along with corresponding standard errors. The dataset used to generate these estimates is described in an available appendix. With the exception of the standard deviation of technology,  $\sigma_v$ , and  $\delta$ , these point estimates are very similar to those used in the RBC literature. Table 4 presents the nonmodel- and model-based estimates of  $\{\sigma_y, \sigma_c/\sigma_y, \sigma_i/\sigma_y, \sigma_h/\sigma_y, \sigma_h/\sigma_a\}$ . Numbers in parentheses are the standard errors of the corresponding point estimates. Numbers in parentheses in the column labeled  $\mathcal{W}$  are the asymptotic probability values of the  $\mathcal{W}$  statistics for testing whether the individual model and data population moments are the same. Notice that we cannot reject any of the individual hypotheses in question. Burnside and Eichenbaum (1994) stressed that, although the propagation mechanisms in their model are quite different, many of the second-moment implications are similar to those of the prototypical RBC model.

Figures 1 and 2 summarize the model's implications for the dynamic correlations between hours worked and average productivity, as well as the dynamic correlations between average productivity and output. The dotted lines in row 1 correspond to the nonmodel-based estimates of  $\{\rho_{ah}^i, i = 0, \pm 1, \pm 2, \pm 3, \pm 4\}$ , and  $\{\rho_{ay}^i, i = 0, \pm 1, \pm 2, \pm 3, \pm 4\}$ , and the solid lines denote the moments implied by the model. The solid lines in row 2 graph the differences between the model- and nonmodel-based estimates, and the dotted lines depict an asymptotic two-standard-error band for the differences. According to these figures, the model does quite well at accounting for the individual dynamic correlations between average productivity and output, as well as average productivity and hours

Table 4. Tests of the Models

Moment	HP-filtered data			First differences		
	U.S. data	Model	$\mathcal{W}$	U.S. data	Model	$\mathcal{W}$
$\sigma_y$	.0192 (.0018)	.0167 (.0013)	1.614 (.204)	.0105 (.0008)	.0114 (.0009)	1.018 (.313)
$\sigma_c/\sigma_y$	.0437 (0.029)	.0480 (.009)	2.005 (.157)	.588 (.010)	.0455 (.038)	11.29 (.001)
$\sigma_i/\sigma_y$	2.224 (.068)	2.244 (.072)	.044 (.835)	2.229 (.086)	1.930 (.115)	4.416 (.036)
$\sigma_h/\sigma_y$	.859 (.069)	.795 (.051)	.990 (.320)	1.403 (.134)	1.297 (.146)	.593 (.441)
$\sigma_h/\sigma_a$	1.221 (.115)	1.033 (.037)	2.258 (.133)	.916 (.020)	.979 (.057)	1.167 (.280)

NOTE: Numbers under the heading "U.S. data" are second moments of U.S. data. Numbers under the heading "Model" are the model's implications for the corresponding moments as functions of  $\Psi_1$ . Standard errors for each are in parentheses. The  $p$  values for the corresponding  $\mathcal{W}$  statistics are in parentheses.

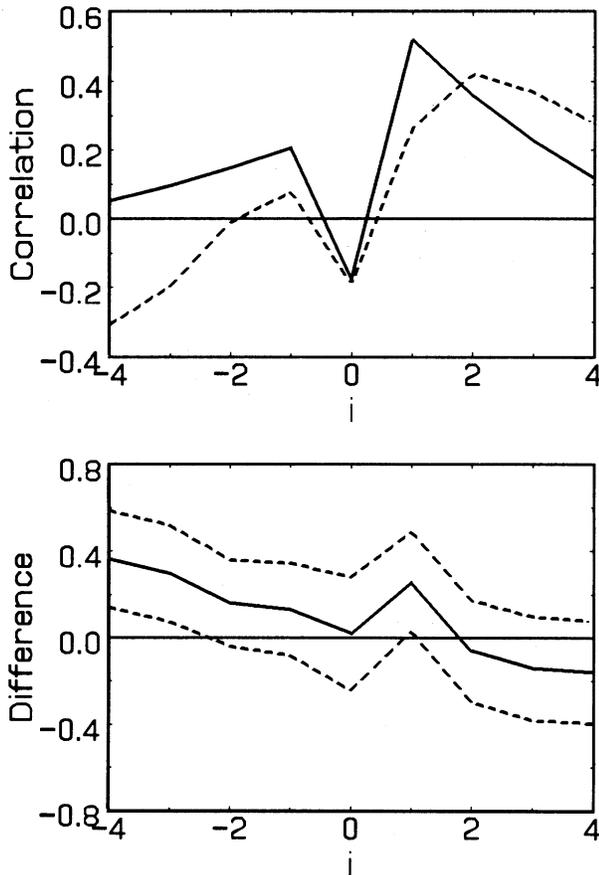


Figure 1. Dynamic Correlations Between Average Productivity and Hours Worked. The data are HP filtered. The figure indicates the correlation between  $APL_t$  (average productivity) and  $H_{t+i}$  (hours worked). In the Correlation panel, the solid line indicates model-predicted correlations, and the dashed line indicates sample correlations. In the Difference panel, the solid line represents the difference between the model and sample correlations, and the dashed lines represent a two-standard-error band around the difference.

worked. Burnside and Eichenbaum (1994) argued that this reflects two features of their model, factor hoarding and their model of measurement error in hours-worked data.

We now turn to our joint hypotheses. Columns 1 and 2 of Table 5 report the  $\mathcal{W}$  statistics for hypotheses  $\{H1, H2, H3, H4\}$  and the corresponding asymptotic probability values. Notice that hypotheses H2, H3, and H4 are all rejected at very low significance levels. To us the strength of these rejections seems at variance with the results of testing the individual components of these hypotheses. One way to reconcile these results is to invoke the pattern of covariances in question. In light of the results in Section 1, however, these strong rejections may simply reflect the small-sample properties of GMM-based Wald statistics as applied to hypotheses involving joint moment restrictions. With this as motivation, we turn to the Monte Carlo experiments.

## 2.5 Monte Carlo Experiments

One natural way to generate artificial datasets for Monte Carlo studies is to bootstrap. Given the type of hypothesis tests we are investigating, what we would need to bootstrap are estimated innovations to technology, government purchases, and the measurement error in hours worked. Be-

cause we estimate the model via GMM and our model is stochastically singular in the observables, we do not have unique estimates of these three variables. Furthermore, we have no measure of the measurement error in hours worked whose variance necessarily corresponds to our estimate of  $\sigma_\xi^2$ . Consequently, we chose not to bootstrap.

Instead, to generate data for our Monte Carlo experiments, we proceeded as follows. Given the estimated value of  $\Psi_1$ , we generated artificial time series according to the following rules:  $C_t = \exp(c_t)X_t$ ,  $Y_t = \exp(y_t)X_t$ ,  $K_t = \exp(k_t)X_{t-1}$ ,  $G_t = \exp(g_t)X_t$ ,  $I_t = \exp(i_t)X_t$ ,  $H_t = f \exp(n_t)$ , and  $\delta_t = \delta \exp(\phi u_t)$ . Here  $c_t, y_t, k_t, u_t, n_t$ , and  $i_t$  are given by (13). The variables  $X_t$  and  $g_t$  were generated according to the laws of motion specified in Section 2.1. One thousand artificial time series datasets, each of length 113, were generated, assuming that the stochastic elements of  $\tilde{\varepsilon}_t$  were Gaussian white noise. We repeated our experiments, doubling the series length to 226. Although the small-sample properties of the test statistics discussed later improved, none of our qualitative conclusions were affected.

To assess the robustness of our results to distributional assumptions, we redid a subset of our Monte Carlo experiments generating the stochastic elements of  $\tilde{\varepsilon}_t$  as follows.

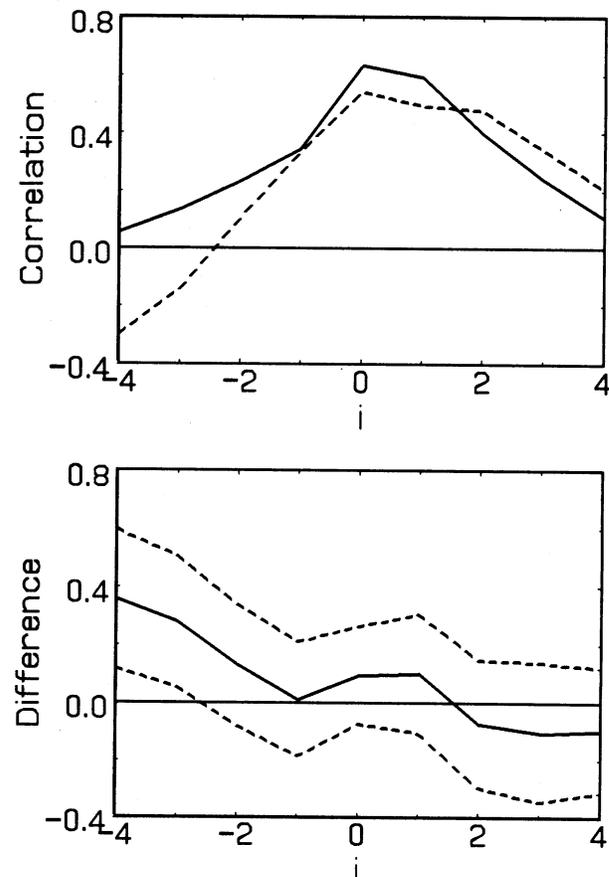


Figure 2. Dynamic Correlations Between Average Productivity and Output. The data are HP filtered. The figure indicates the correlation between  $APL_t$  (average productivity) and  $Y_{t+i}$  (output). In the Correlation panel, the solid line indicates model-predicted correlations, and the dashed line indicates sample correlations. In the Difference panel, the solid line represents the difference between the model and sample correlations, and the dashed lines represent a two-standard-error band around the difference.

Table 5. Small-Sample Performance of the Joint Tests

Hypothesis	Test performed using U.S. data			Size (%) of tests		
	$\mathcal{W}$	$p$ value	MC $p$ value	10%	5%	1%
<i>HP-filtered data</i>						
H1	6.64	.25	.48	31.7	23.0	11.9
H2	43.7	.00	.01	23.6	16.5	7.6
H3	35.5	.00	.01	20.2	13.3	6.5
H4	66.3	.00	.06	57.6	50.7	36.7
<i>First-differenced data</i>						
H1	16.0	.01	.02	12.0	7.6	2.2
H2	30.1	.00	.01	15.6	11.3	5.2
H3	45.9	.00	.00	13.5	9.6	3.9
H4	63.5	.00	.02	37.9	29.8	18.1

NOTE: The numbers under the heading "p value" are the  $p$  values obtained when the  $\mathcal{W}$  statistics for H1, H2, H3, and H4 are compared to  $\chi^2$  distributions with 5, 9, 9, and 17 df, respectively. The numbers under the heading "MC p value" are obtained by comparing these statistics to the distribution of the  $\mathcal{W}$  statistics generated by our Monte Carlo experiments. The numbers under the heading "Size of tests" indicate the frequency (in %) with which the  $\mathcal{W}$  statistics from our Monte Carlo experiments exceeded the 10%, 5%, and 1% critical values of the relevant  $\chi^2$  distributions.

First, we constructed a vector of disturbances whose elements were realizations from a chi-squared distribution with 1 df. We then subtracted one from each of these realizations. Finally, we scaled the random variables so that they had standard deviations equal to the values assumed in our baseline experiments. The reason that these disturbances are of interest is that they are skewed (to the right) and leptokurtic relative to Gaussian white noise. In no case were our results sensitive to this perturbation.

With one exception, all of the moment conditions underlying our estimator of  $\Psi_1$  hold exactly for the artificial data-generating process. The exception is the planner's Euler equation for  $K_{t+1}$ , Equation (17), discussed in the Appendix. To deal with this problem, we computed the expectation in Equation (17) for the true log-linearized model. As it turns out, at these parameter values the error is approximately equal to  $2 \times 10^{-5}$ . To correct for possible bias, we implemented our Monte Carlos, centering Equation (17) around  $2 \times 10^{-5}$  rather than 0. As it turns out, this bias correction had virtually no impact on our results.

We begin by reporting the small-sample behavior of the  $\mathcal{W}$  statistics for hypotheses H1, H2, H3, and H4. The third numerical column of Table 5 reports the fraction of the Monte Carlo draws in which the  $\mathcal{W}$  statistics for these hypotheses were greater than or equal to the corresponding  $\mathcal{W}$  statistic obtained using U.S. data (see col  $\mathcal{W}$ ). We refer to this fraction as the Monte Carlo probability. For hypothesis H1, H2, and H3, the asymptotic and Monte Carlo probabilities are reasonably similar. For hypothesis H4, however, the Monte Carlo probability is much larger than the asymptotic probability (.06 versus .00). According to standard asymptotic distribution theory, the  $\mathcal{W}$  statistic that we obtained for hypothesis H4 would be very unlikely if the model were specified correctly. But according to the small-sample results, one would obtain a  $\mathcal{W}$  statistic this large or larger roughly 6% of the time.

A complementary way to assess the small-sample properties of the Wald tests is to consider the fraction of the Monte Carlo draws in which the  $\mathcal{W}$  statistics exceeded the 1%, 5%,

and 10% critical values of the relevant chi-squared distributions. These are displayed in the last three columns of Table 5. Notice that the small-sample sizes of the test statistics for hypotheses H1 and H4 greatly exceed their asymptotic size. This tendency is particularly dramatic in the case of H4, in which the  $\mathcal{W}$  statistics exceed their asymptotic 1%, 5%, and 10% critical values 37%, 51%, and 58% of the time.

Before analyzing this finding, we briefly discuss the size of the test statistics applied to the individual moments that make up joint hypotheses H1, H2, H3, and H4. Our results are displayed in Figure 3. The height of each bar graph in panels (a), (b), and (c) denotes the fraction of the Monte Carlo draws in which the  $\mathcal{W}$  statistic for a given hypothesis exceeded the 10%, 5%, and 1% critical values of the asymptotic chi-squared distribution.

According to Figure 3, the small-sample sizes of the test statistics for hypotheses 1 and 4–23 are moderately higher than their asymptotic sizes. The small-sample sizes of the test statistics associated with  $\sigma_c/\sigma_y$  and  $\sigma_i/\sigma_y$  are substantially larger than their asymptotic sizes. This is consistent with our finding that Wald tests of hypotheses H1 and H4 overreject in small samples. These effects do not, however, seem large enough to explain the *extent* to which the Wald test overrejects H1 and H4.

Viewed overall, the outstanding feature of our experiments is the excessive small-sample size of the Wald test of hypothesis H4. Inference based on the asymptotic distribution of the  $\mathcal{W}$  statistic leads to a grossly overly critical assessment of the model's performance. In an available appendix we show that this conclusion is robust to various perturbations. First, we consider the effects of different bandwidths when constructing  $S_T$ . These were chosen both on an a priori basis and using the Newey and West (1993) automatic bandwidth procedure. Second, we consider different estimators of  $S_0$  that correspond to different lag windows. Third, we discuss the impact of using a small-sample correction suggested by Andrews (1991).

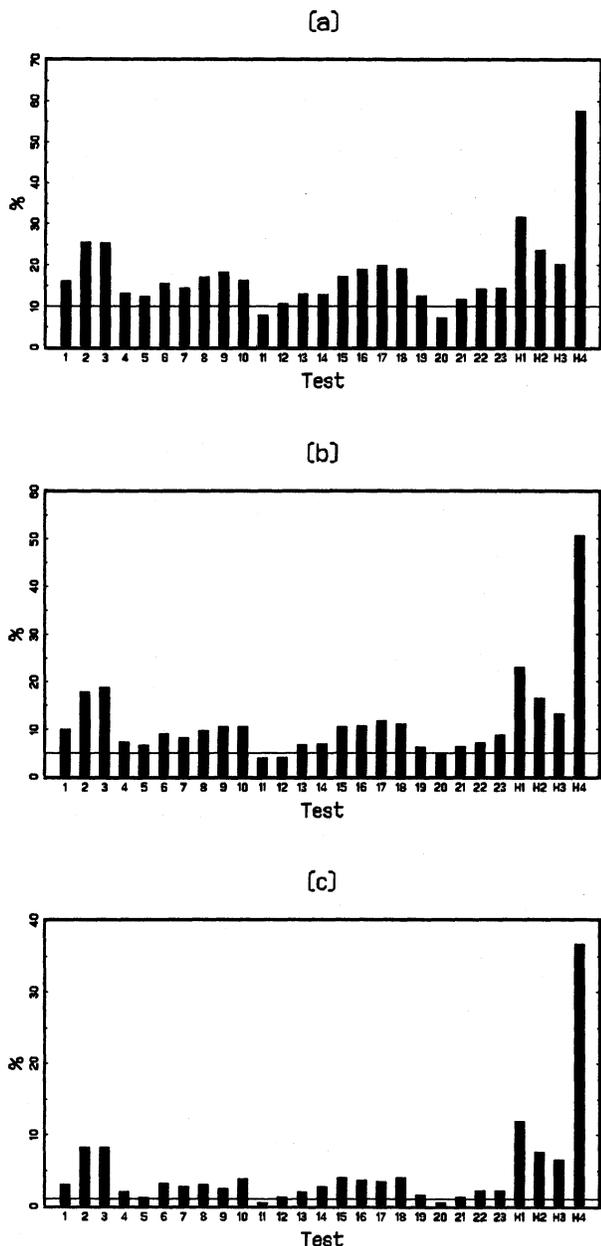


Figure 3. Small-Sample Size of the Wald Tests for the RBC Example With HP-Filtered Data. Each bar in the figure represents the small-sample size measured in percent of the  $\mathcal{W}$  test whose number is indicated on the x axis: (a) Tests with 10% asymptotic size, (b) tests with 5% asymptotic size, (c) tests with 1% asymptotic size.

The previous results pertain to tests of the model's implications regarding the moments of HP-filtered time series. It is of interest to see whether our qualitative results are robust to testing the model's implications for moments of first-differenced data. Table 4 presents the nonmodel- and model-based estimates of  $\{\sigma_y, \sigma_c/\sigma_y, \sigma_i/\sigma_y, \sigma_h/\sigma_y, \sigma_h/\sigma_a\}$ , where  $\sigma_i$  denotes the standard deviation of the log first-difference of variable  $i$ . Notice that the model's performance deteriorates with respect to  $\sigma_c/\sigma_y$  and  $\sigma_i/\sigma_y$  relative to the HP-filter results.

Figure 4 is the analog to Figure 3 for the case in which tests refer to moments of first-differenced data. Table 5 reports the Monte Carlo probability values and small-sample sizes of the test statistics pertaining to hypotheses H1-H4.

The key features of these results are as follows. First, Figure 4 indicates that the small-sample size of a substantial fraction of the Wald statistics for individual moment tests is actually smaller than the corresponding asymptotic size. Second, according to Figure 4 and Table 5, for the joint tests H2, H3, and H4, it is still the case that the small-sample size substantially exceeds the asymptotic size. This is particularly marked for hypothesis H4 in which the  $\mathcal{W}$  statistics exceed their 1%, 5%, and 10% critical values in 18.1%, 29.8%, and 37.9% of the Monte Carlo runs. Overall, we view our basic qualitative results, at least as they pertain to the small-sample size of the joint tests, as being robust to whether the tests pertain to moments of HP-filtered or first-differenced data.

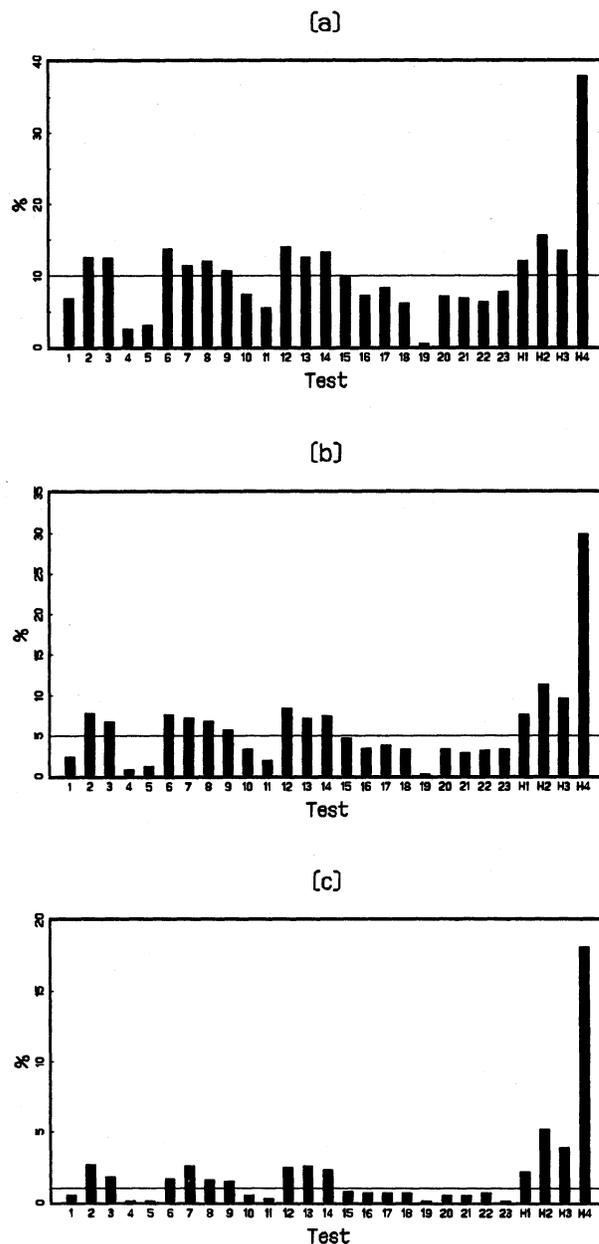


Figure 4. Small-Sample Size of the Wald Tests for the RBC Example With First-Differenced Data. Each bar in the figure represents the small-sample size measured in percent of the  $\mathcal{W}$  test whose number is indicated on the x axis: (a) Tests with 10% asymptotic size, (b) tests with 5% asymptotic size, (c) tests with 1% asymptotic size.

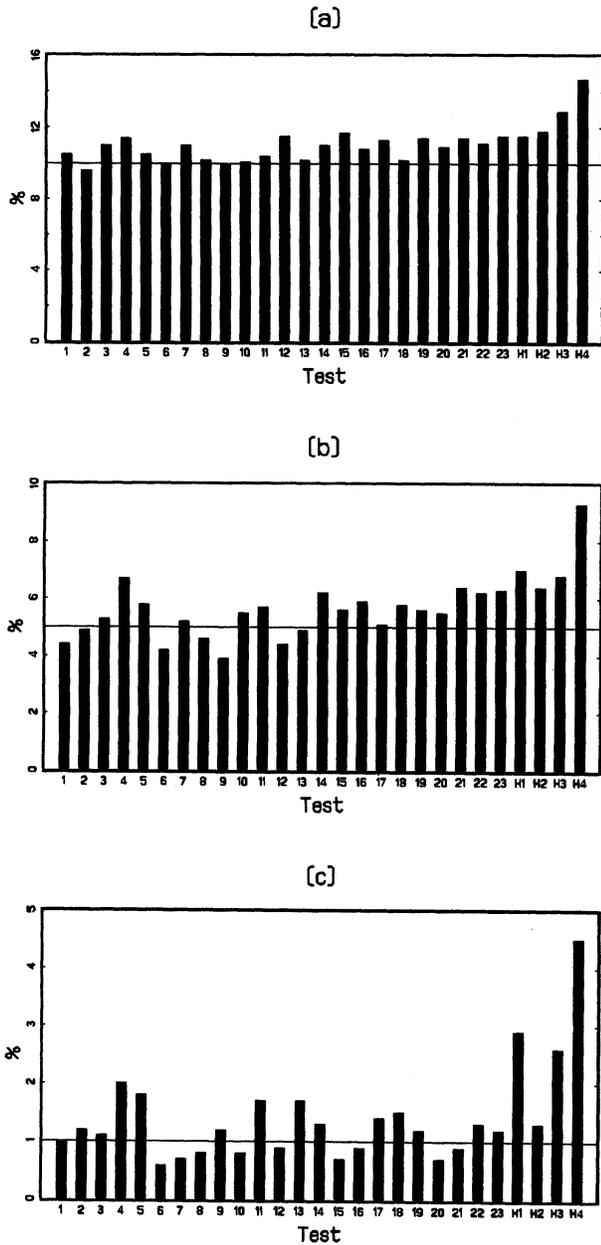


Figure 5. Small-Sample Size of the Wald Tests for the RBC Example With HP-Filtered Data. Each bar in the figure represents the small-sample size measured in percent of the  $\tilde{W}$  test whose number is indicated on the x axis: (a) Tests with 10% asymptotic size, (b) tests with 5% asymptotic size, (c) tests with 1% asymptotic size.

In the remainder of this section, we discuss the factors underlying the excessive small-sample size of our Wald tests, with particular reference to hypothesis H4. The key question is why is the small-sample distribution of  $\mathcal{W}$  so different from estimates of its asymptotic distribution? We consider two basic explanations:

1. The small-sample distribution of  $\sqrt{T}F(\Psi_T)$  is different from its large-sample distribution. This could happen either because of bias in  $F(\Psi_T)$ , because the small-sample covariance matrix of  $\sqrt{T}F(\Psi_T)$  is not equal to its large-sample analog, or because  $\sqrt{T}F(\Psi_T)$  is correlated with estimates of its large-sample covariance matrix.

2. A second possibility is that the small- and large-sample distributions of  $\sqrt{T}F(\Psi_T)$  are the same, but it is difficult to estimate the latter in small samples.

In what follows we present evidence that the second explanation is more relevant in our context. Specifically, we argue that the basic problem is difficulty in estimating the spectral-density matrix of the GMM residuals,  $S_0$ .

We begin by assessing the possible impact of bias in  $F(\Psi_T)$  on the distribution of  $\mathcal{W}$ . To this end we considered the small-sample distribution of the modified Wald statistic  $\tilde{W}_T = T(F_k - \bar{F})'V_{Fk}^{-1}(F_k - \bar{F})$ . Note that  $\tilde{W}_T$  is implicitly indexed by  $k$ . Here  $F_k$  denotes the value of  $F$  in the  $k$ th Monte Carlo run,  $k = 1, \dots, 1,000$ ,  $\bar{F}$  is the average value of  $F_k$  across the 1,000 Monte Carlo runs, and  $V_{Fk}$  is the estimated value of the asymptotic covariance matrix of  $\sqrt{T}F(\Psi_T)$  using data from the  $k$ th Monte Carlo run. The estimator underlying  $V_{Fk}$  is the one discussed in Section 2.3. For any given Monte Carlo run, the only difference between  $\tilde{W}_T$  and the original Wald statistic is the presence of the bias correction term  $\bar{F}$ . Our major finding here was that the distributions of the modified and original Wald statistics across the Monte Carlo runs were very similar. We concluded that bias in  $F(\Psi_T)$  is not the problem.

We now turn to a series of exercises designed to shed light on the way our original estimator of  $V_F$  has an impact on the small-sample properties of the  $\mathcal{W}$  statistics. Consider the test statistic  $\tilde{W}_T = T F_k' \tilde{V}_F^{-1} F_k$ , where

$$\tilde{V}_F = \frac{1}{1,000} \sum_{k=1}^{1,000} (F_k - \bar{F})(F_k - \bar{F})'$$

$\tilde{W}_T$  is implicitly indexed by  $k$ . The weighting matrix  $\tilde{V}_F$  is an estimate of the small-sample covariance matrix of  $\sqrt{T}F(\Psi_T)$  corresponding to the sample covariance matrix of the  $F_k$ 's across the Monte Carlo runs. Figure 5 reports the fraction of the time in the Monte Carlo runs that the  $\tilde{W}$  statistics exceeded the 1%, 5%, and 10% critical values of the relevant chi-squared distributions. The key thing to note is that, although the small-sample properties of the  $\tilde{W}$  statistics are not in perfect accord with their asymptotic behavior, they are much better than those of the  $\mathcal{W}$  statistics. For example, the  $\tilde{W}$  statistics for H4 exceed their asymptotic (1%, 5%, 10%) critical values in (4.5%, 9.3%, 14.7%) of Monte Carlo draws. (Recall that analog percentages for the  $\mathcal{W}$  statistics are 37%, 51%, and 58%). Moreover, for hypotheses involving only one moment restriction, the small-sample sizes of the  $\tilde{W}$  statistics are quite close to their asymptotic sizes.

Why do the  $\tilde{W}$  statistics have much better small-sample properties than the  $\mathcal{W}$  statistics? One possible reason is that the covariance matrix  $\tilde{V}_F$  that is used to construct  $\tilde{W}$  is held fixed across the different Monte Carlo runs. So, by construction, the numerator and denominator terms in  $\tilde{W}$  are uncorrelated with each other. Although this is true asymptotically for the  $\mathcal{W}$  statistics, it need not be true in small samples. To assess the plausibility of this explanation, we constructed an alternative version of  $\mathcal{W}$ , which randomly combines numerator and denominator terms from different Monte Carlo

runs:  $\hat{W}_T = TF'_i V_{F_j}^{-1} F_i$ . Consistent with our previous notation,  $F_i$  denotes the value of  $F$  in the  $i$ th Monte Carlo run, and  $V_{F_j}$  is the estimated value of the asymptotic covariance matrix of  $\sqrt{T}F(\Psi_T)$  in the  $j$ th Monte Carlo run. By choosing  $i$  and  $j$  to be independently and randomly selected numbers between 1 and 1,000, we randomly match  $F_i$ 's and  $V_{F_j}$ 's from different Monte Carlo datasets. This eliminates any correlation between numerator and denominator terms in  $\hat{W}$ . We anticipated that this would lead to a substantial improvement in the small-sample properties of the test statistic. But, to our surprise, we found that the distributions of  $\hat{W}$  and  $W$  across the Monte Carlo runs were very similar. So, at least in our application, the poor small-sample properties of the  $W$  statistic cannot be attributed to

correlation between the constituent numerator and denominator terms.

Evidently, the source of the problem has to do with the small-sample distribution of  $V_F$ , our estimator of the covariance matrix of  $\sqrt{T}(\Psi_T - \Psi^0)$ . Recall that

$$V_F = \left( \frac{\partial F(\Psi_T)}{\partial \Psi'} \right)' V_\Psi \left( \frac{\partial F(\Psi_T)}{\partial \Psi'} \right).$$

In constructing  $W$ , we set  $V_\Psi = (D'_T S_T^{-1} D_T)^{-1}$ . In our next experiment we considered the behavior of the Wald statistic,  $\tilde{W}_T$ , obtained using

$$\tilde{V}_\Psi = \frac{T}{1,000} \sum_{k=1}^{1,000} (\Psi_{Tk} - \bar{\Psi})(\Psi_{Tk} - \bar{\Psi})'$$

in place of  $V_\Psi$ . The estimator  $\tilde{V}_\Psi$  corresponds to the sample covariance matrix of  $\sqrt{T}(\Psi_{Tk} - \bar{\Psi})$  across the Monte Carlo runs. Our key finding here is that the small-sample distributions of the  $\tilde{W}$  and  $W$  statistics are very similar. For example the  $\tilde{W}$  statistics for H4 exceeded their asymptotic (1%, 5%, 10%) critical values in (3.7%, 8.3%, 13.6%) of the Monte Carlo draws. We conclude that the problem with the  $W$  statistic has to do with  $V_\Psi$ , rather than with  $\partial F(\Psi_T)/\partial \Psi'$ .

We can narrow the problem down even further by investigating the behavior of the constituent elements of  $V_\Psi$ . Specifically we considered the small-sample behavior of the Wald statistic,  $\tilde{W}_T$ , constructed by replacing  $S_T$  with the estimate of the small-sample covariance of  $\sqrt{T}g_T(\Psi^0)$  given by

$$\tilde{S} = \frac{T}{1,000} \sum_{k=1}^{1,000} \{[g_{Tk}(\Psi^0) - \bar{g}(\Psi^0)][g_{Tk}(\Psi^0) - \bar{g}(\Psi^0)]'\}.$$

Here  $g_{Tk}(\Psi^0)$  is the value of  $g_T(\Psi^0)$  emerging from the  $k$ th Monte Carlo run and  $\bar{g}(\Psi^0)$  is the average value of  $g_{Tk}(\Psi^0)$  across the Monte Carlo runs. Figure 6 reports the fraction of the Monte Carlo runs in which the  $\tilde{W}_T$  statistics exceeded the 1%, 5%, and 10% critical values of the relevant chi-squared distributions. Notice how similar these are to the analog results for the  $W$  statistics. For example, the  $\tilde{W}$  statistics for H4 exceed their asymptotic (1%, 5%, 10%) critical values in (3.9%, 7.5%, 12%) of the Monte Carlo runs.

This result strongly suggests that the key factor underlying the poor small-sample properties of the original  $W$  statistics is the small-sample properties of  $S_T$ , our original estimator of  $S_0$ . The question is whether the problem is that we cannot accurately estimate  $S_0$  in small samples or that  $S_0$  does not correspond to the small-sample covariance matrix of  $\sqrt{T}g_T(\Psi^0)$ .

To get at this issue we redid our Monte Carlo experiments using the population value of  $S_T, S_0$ , that is implied by the parameters governing the data-generating process. Specifically, on each of the 1,000 datasets, we estimated the parameters of the model but formed the  $W$  statistic using the fixed matrix  $S_0$ . We found that the empirical distribution of the resulting Wald statistics was extremely similar to that of  $W$ . For example the new Wald statistics for H4

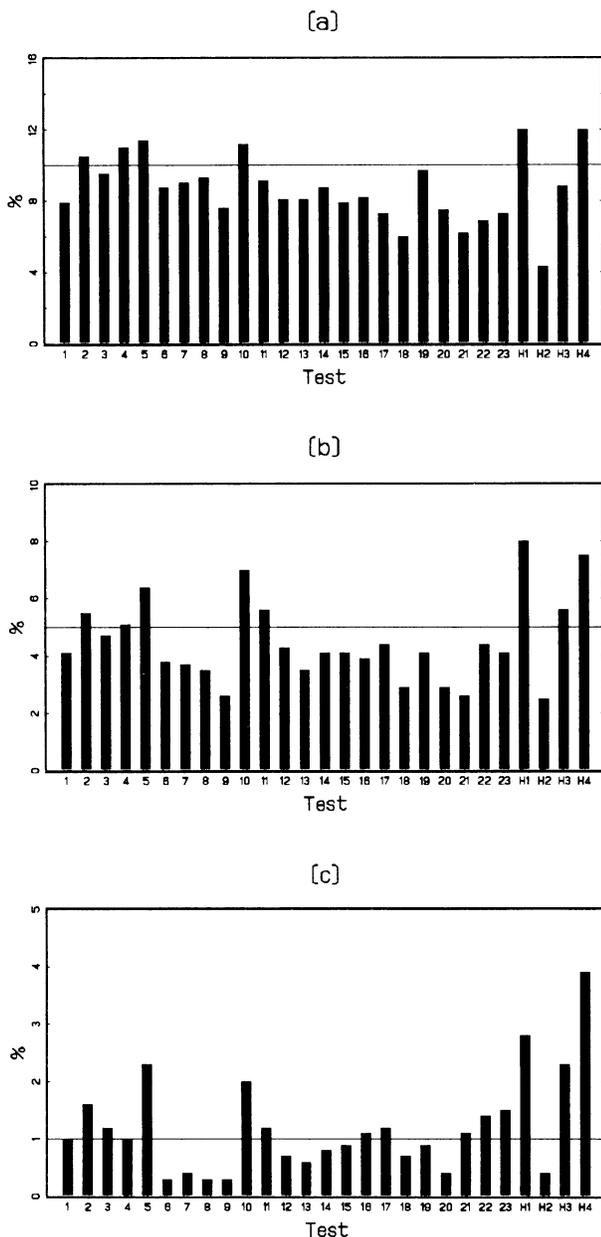


Figure 6. Small-Sample Size of the Wald Tests for the RBC Example With HP-Filtered Data. Each bar in the figure represents the small-sample size measured in percent of the  $\tilde{W}$  test whose number is indicated on the x axis: (a) Tests with 10% asymptotic size, (b) tests with 5% asymptotic size, (c) tests with 1% asymptotic size.

exceed their asymptotic (1%, 5%, 10%) critical values (4%, 8%, 11%) of the time. Recall that the corresponding values of  $\mathcal{W}$  exceeded their asymptotic (1%, 5%, 10%) critical values (3.9%, 7.5%, 12%) of the time. Recall also that in our baseline case the  $\mathcal{W}$  statistics exceeded their asymptotic (1%, 5%, 10%) critical values (37%, 51%, 58%) of the time. Evidently, the fact that we must estimate  $S_0$  accounts for a substantial part of the difficulties with the small-sample properties of our original Wald statistics. Note, though, that even when  $S_0$  is known, relying on asymptotic distribution theory would still lead us to reject hypothesis H4 too often. But with  $S_0$  known the small-sample sizes of the individual hypotheses are quite close to their asymptotic size.

We conclude that sampling error in  $S_T$  plays a substantial role in the excessive small-sample size of Wald tests involving multiple moment conditions. This suggests an alternative way to estimate  $S_T$ . Specifically, the econometrician could calculate the implied population value of  $S_T$  for any given set of parameter estimates when estimating the model. The obvious drawback to this procedure is that, for nontrivial models, it is computationally quite burdensome.

### 3. CONCLUSION

This article examined the small-sample properties of GMM-based Wald statistics. For the data-generating processes considered, we found that the small-sample size of these tests tended to exceed its asymptotic size. The problem became dramatically worse as the dimensionality of the joint tests being considered increased. We offered evidence that the bulk of the problem has to do with difficulty in estimating the spectral-density matrix of the GMM residuals that is needed to conduct inference. Our results lead us to be very skeptical that the problem can be resolved by using any of the alternative nonparametric estimators of this matrix that have been discussed in the literature. Instead we advocate using estimators that impose as much a priori information as possible. Two important sources of such information are the economic theory being investigated and the null hypothesis being tested.

### ACKNOWLEDGMENTS

We thank Lawrence Christiano, Lars Hansen, Narayana Kocherlakota, an associate editor, and two anonymous referees for their helpful comments and suggestions. Eichenbaum is also affiliated with the Federal Reserve Bank of Chicago and the National Bureau of Economic Research.

### APPENDIX: THE EULER EQUATION FOR CAPITAL IN THE RBC EXAMPLE

In this appendix we discuss our procedure for ensuring that the Euler equation for capital holds exactly for the data-generating process underlying our Monte Carlo experiments. The Euler equation for  $K_{t+1}$  does not hold exactly for our linearized representation of the model. This equa-

tion is given by

$$E \left\{ 1 - \beta \left( \frac{C_t}{C_{t+1}} \right) \times \left[ (1 - \alpha)(1 - \phi^{-1}) \frac{Y_{t+1}}{K_{t+1}} + 1 \right] \right\} = 0. \quad (\text{A.1})$$

As a result, when we estimate the model using artificially generated time series from the linearized model, it is important to adjust this moment restriction appropriately. We compute the expectation in Equation (A.1) for our linearized model (it is approximately  $2 \times 10^{-5}$ ) evaluated at the parameter values we use to generate the artificial data. We then center the moment restriction around that value rather than 0. This expectation, denote by  $\varepsilon$ , is computed as follows:

$$\begin{aligned} \varepsilon &= E \left\{ 1 - \beta \left( \frac{C_t}{C_{t+1}} \right) \left[ (1 - \alpha)(1 - \phi^{-1}) \frac{Y_{t+1}}{K_{t+1}} + 1 \right] \right\} \\ &= E \{ 1 - \beta \exp(c_t - c_{t+1}) [(1 - \alpha)(1 - \phi^{-1}) \\ &\quad \times \exp(y_{t+1} - k_{t+1}) + \exp(-\gamma - v_{t+1})] \}. \end{aligned}$$

Let  $s_t = (1 \quad k_t \quad N_t \quad v_t \quad g_t \quad \xi_t)'$  and  $\bar{s}_{t+1} = (s'_{t+1} \quad s'_t)'$ . Any variable in the linearized model, say  $z_t$ , determined at time  $t$  is given as a function  $\pi'_z s_t$ , for some vector  $\pi_z$  determined by the solution to the model. Therefore, we can write the Euler equation error simply as

$$\begin{aligned} \varepsilon &= E \{ 1 - \beta [(1 - \alpha)(1 - \phi^{-1}) \exp(\mu'_1 \bar{s}_{t+1}) \\ &\quad + \exp(-\gamma) \exp(\mu'_2 \bar{s}_{t+1})] \}, \end{aligned}$$

where

$$\mu_1 = \begin{pmatrix} -\pi_c + \pi_y \\ \pi_c - \pi_k \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} -\pi_c - \pi_v \\ \pi_c \end{pmatrix}.$$

In our simulations we assume that the innovations to the exogenous variables are normally distributed. In this case the properties of lognormal random variables can be exploited to show that

$$\begin{aligned} \varepsilon &= 1 - \beta \left[ (1 - \alpha)(1 - \phi^{-1}) \exp \left( \mu'_1 E \bar{s} + \frac{1}{2} \mu'_1 F_s \mu_1 \right) \right. \\ &\quad \left. + \exp \left( -\gamma + \mu'_2 E \bar{s} + \frac{1}{2} \mu'_2 F_s \mu_2 \right) \right], \end{aligned}$$

where  $E \bar{s}$  and  $\bar{\Gamma}_s$  are the mean and unconditional covariance matrix of  $\bar{s}_{t+1}$ . These are both computable as a by-product of the solution method.

[Received May 1994. Revised December 1995.]

### REFERENCES

- Altug, S. (1989), "Time to Build and Aggregate Fluctuations: Some New Evidence," *International Economic Review*, 30, 889–920.  
 Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

- Burnside, C. (1992), "Small Sample Properties of 2-Step Method of Moments Estimators," unpublished manuscript, University of Pittsburgh, Dept. of Economics.
- (1993), "Notes on the Linearization and GMM Estimation of Real Business Cycle Models," unpublished manuscript, University of Pittsburgh, Dept. of Economics.
- Burnside, C., and Eichenbaum, M. (1994), "Factor Hoarding and the Propagation of Business Cycle Shocks," Working Paper 4675, National Bureau of Economic Research, Cambridge, MA.
- Christiano, L. J. (1988), "Why Does Inventory Investment Fluctuate So Much?" *Journal of Monetary Economics*, 21, 247–280.
- Christiano, L. J., and den Haan, W. (1996), "Small-Sample Properties of GMM for Business-Cycle Analysis," *Journal of Business & Economic Statistics*, 14, 309–327.
- Christiano, L. J., and Eichenbaum, M. (1992), "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations," *American Economic Review*, 82, 430–450.
- Eichenbaum, M. S., and Hansen, L. P. (1990), "Estimating Models With Intertemporal Substitution Using Aggregate Time Series Data," *Journal of Business & Economic Statistics*, 8, 53–69.
- Ferson, W., and Foerster, S. E. (1991), "Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," unpublished manuscript, University of Chicago, Graduate School of Business.
- Fuhrer, J., Moore, G., and Schuh, S. (1995), "Estimating the Linear-Quadratic Inventory Model: Maximum Likelihood Versus Generalized Method of Moments," *Journal of Monetary Economics*, 35, 115–158.
- Hansen, G. D. (1985), "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16, 309–328.
- Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.
- Hansen, L. P., and Singleton, K. J. (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.
- Hodrick, R. J., and Prescott, E. C. (1980), "Post-War Business Cycle: An Empirical Investigation," unpublished manuscript, Carnegie Mellon University, Graduate School of Industrial Administration.
- King, R. G., Plosser, C. I., and Rebelo, S. (1988), "Production, Growth and Business Cycles," *Journal of Monetary Economics*, 21, 195–232.
- Kocherlakota, N. (1990), "On Tests of Representative Consumer Asset Pricing Models," *Journal of Monetary Economics*, 26, 285–304.
- Kydland, F. E., and Prescott, E. C. (1982), "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1435–1470.
- Leeper, E., and Sims, C. S. (1994), "Toward a Modern Macroeconomic Model Usable for Policy Analysis," *NBER Macroeconomics Annual*, 9, 81–117.
- McGrattan, E., Rogerson, R., and Wright, R. (1993), "Household Production and Taxation in the Stochastic Growth Model," Staff Report 166, Federal Reserve Bank of Minneapolis, Research Dept.
- Neely, C. J. (1993), "A Reconsideration of Representative Consumer Asset Pricing Models," unpublished manuscript, University of Iowa, Dept. of Economics.
- Newey, W. K., and West, K. D. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.
- (1993), "Automatic Lag Selection in Covariance Matrix Estimation," unpublished manuscript, MIT, Dept. of Economics.
- Prescott, E. C. (1986), "Theory Ahead of Business-Cycle Measurement," *Carnegie-Rochester Conference on Public Policy*, 24, 11–44.
- Tauchen, G. (1986), "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained From Financial Market Data," *Journal of Business & Economic Statistics*, 4, 397–416.
- West, K. D., and Wilcox, D. W. (1994), "Some Evidence on Finite Sample Distributions of Instrumental Variables Estimators of a Linear Quadratic Inventory Model," unpublished manuscript, Board of Governors of the Federal Reserve System.