On a test for structural stability of euler conditions parameters estimated via the generalized method of moments estimator: small sample properties

Hamori, Shigeyuki
Kitasaka, Shin-ichi
Tanizaki, Hisashi

(Citation)
Econometric Reviews,15(1):97-114

(Issue Date)
1996

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90000569

Econometric Reviews

ON A TEST FOR STRUCTURAL STABILITY OF EULER CONDITIONS PARAMETERS ESTIMATED VIA THE GENERALIZED METHOD OF MOMENTS ESTIMATOR: SMALL SAMPLE PROPERTIES

Shigeyuki HAMORI
Faculty of Economics
Kobe University
Rokkodai, Nada-Ku, Kobe, 657
JAPAN

Shin-ichi KITASAKA
School of International Cooperation Studies
Kobe University
Rokkodai, Nada-Ku, Kobe, 657
JAPAN

and

Hisashi TANIZAKI
Faculty of Economics
Kobe University
Rokkodai, Nada-Ku, Kobe, 657
JAPAN

Keywords
GMM, predictive test, J test, quadrature method

JEL classification number: C12

ABSTRACT

This paper presents Monte Carlo experiments on the small sample performance of the predictive test for structural change proposed by Ghysels and Hall. The predictive test was found to be more powerful than the overidentifying restrictions test in terms of size-corrected power when a shift in parameter has occurred. Also, it was found that the power of the predictive test decreases drastically as the number of the out-sample data decreases.

1. Introduction

We analyze the small sample properties of the predictive test developed by Ghysels and Hall (1990a,1990b). Because of the Lucas critique [Lucas (1976)], recent econometric work has focused on the underlying parameters that are assumed to remain invariant across time and policy regimes. These parameters can be estimated by the generalized method of moments (GMM) developed by Hansen (1982). Whether or not "taste and technology" parameters are actually invariant with respect to time and policy regimes, however, is still an open question.

To analyze the structural stability of these parameters, Ghysels and Hall (1990a,1990b) developed a predictive test in the GMM framework. Structural stability tests are natural diagnostics for Euler equation models and a valuable complement to local alternatives to the overidentifying restrictions test (J test). The J test is known to have power equal to size against a class of local alternatives to parameter constancy [Ghysels and Hall (1990b)]. In consequence, it is becoming more common to test the structural constancy of Euler equation models with applied studies [Ghysels and Hall (1990b), Epstein and Zin (1991), and Hamori (1992)].

The predictive test uses parameter estimates from one sample to evaluate the moment conditions of another sample, and it has computational advantages. It only requires an estimate of one of the subsamples, and it also explores all orthogonality conditions of the other subsample. The small sample properties of the test, however, are as yet unknown to researchers.

This paper presents Monte Carlo results on the small sample performance of the predictive test and the J test [Hansen (1982)]. The simulation design is identical to the one employed by Tauchen (1986) in his study of the small sample properties of the GMM estimator and the J test.[1] The adoption of the simulation design of Tauchen means that this study complements existing evidence on the small sample behavior of GMM.

The predictive test was found to be more powerful than the J test in terms of size-corrected power when a shift in parameter occurred, and it was also found that the power of the predictive test drastically decreased as the number of the out-sample

[1] For the small sample properties of GMM, see Tauchen (1986), Kocherlakota (1990), and Hamori and Kitasaka (1992).

data decreased.

2. The Basic Model

The representative consumer in the economy chooses consumption and asset holding to maximize the expected utility obtained from consumption over a lifetime subject to budget constraints [Lucas (1978) and Breeden (1979)]. Supposing that the preference of the consumer can be expressed by the constant relative risk aversive type, the maximization problem would be formulated as follows:

$$(2.1) \qquad \max \quad E_0[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma} - 1}{1-\gamma}], \qquad \gamma \neq 1,$$

$$(2.2) \qquad \text{s.t.} \quad p_t Q_{t+1} + c_t = (p_t + d_t) Q_t,$$

where β is the subjective discount rate, γ is the measure of relative risk aversion, c_t is the real consumption at time t, q_t is the price of the asset at time t, d_t is the dividend of the asset at time t, Q_t is the quantity of the asset at time t, and $E_t[\bullet]$ is the conditional expectations operator.

The left-hand side of the budget constraint (2.2) shows the outlay, and the right-hand side of the budget constraint shows the revenue. Solving this problem yields the following Euler equation:

$$(2.3) \qquad E_t[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\left(\frac{q_{t+1} + d_{t+1}}{q_t}\right) - 1] = 0.$$

Many researchers applied this model to analyze the movements of asset returns [Hansen and Singleton (1982,1983) and Mehra and Prescott (1985)].

Let us define the dividend growth, the consumption growth, and the price dividend ratio as follows:

$$(2.4) \qquad x_t = \frac{d_t}{d_{t-1}},$$

$$(2.5) \qquad w_t = \frac{c_t}{c_{t-1}},$$

and

$$(2.6) \qquad v_t = \frac{q_t}{d_t}.$$

Suppose the random variables x_t and w_t follow the jointly stationary first order Markov processes:

(2.7) $Y_{t+1} = AY_t + \varepsilon_{t+1}$,

where

$$Y_t = \begin{bmatrix} \log x_t \\ \log w_t \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \qquad E_{t-1}[\varepsilon_t] = 0, \qquad Var[\varepsilon_t] = \Omega.$$

The random variables ε_{1t}, ε_{2t}, are assumed to be white noise. Given this Markovian structure, when the event $\{w_t = w, x_t = x\}$ occurs, the values w and x completely characterize the state of the system.

Then, the Euler equation (2.3) can be rewritten in terms of these variables as follows:

(2.8) $v_t = \beta E_t[w_{t+1}{}^{-\gamma}(1 + v_{t+1})x_{t+1}]$,

or

(2.8') $\beta E_t[w_{t+1}{}^{-\gamma} R_{t+1}] = 1$,

where R_{t+1} is the gross return on the asset from time t to time t+1,defined as follows:

$$R_{t+1} = \frac{1 + v_{t+1}}{v_t} x_{t+1}.$$

Tauchen and Hussey (1991) developed the method to approximate this system by the following discrete equation:

(2.9) $v(s) = \beta \sum_{s'=1}^{N} \pi(s,s') w(s')^{-\gamma} [1 + v(s')] x(s')$,

where s is the state of space, s' is the state for "one period hence," N is the total number of states, and $v(s)$ is the value of the price dividend ratio in the state of nature s.

In the approximation, each x and w could assume one of eight values, giving sixty-four states of nature in total. Realizations of the discrete process $\{x, w\}$ were generated using the quadrature method of Tauchen and Hussey (1991), and the implied price dividend ratio $\{v\}$ was calculated concurrently.

To apply the GMM, let us define the disturbance term for the econometric estimation as follows:

(2.10) $e_t(\theta) = \beta w_{t+1}{}^{-\gamma} R_{t+1} - 1$,

where $\theta = (\beta, \gamma)'$ is the parameter vector. Then, the Euler equation (2.3) implies

(2.11) $E_t[e_t(\theta)] = 0$.

Let Z_t denote a vector of instrumental variables,

$$Z_t = (1, w_t, w_{t-1}, \ldots, w_{t-L+1}, R_t, R_{t-1}, \ldots, R_{t-L+1})',$$

where $L = NLag \geq 1$. *Nlag* is the number of lags used for instruments. Let $g_t(\theta)$ be defined as

(2.12) $$g_t(\theta) = e_t(\theta) Z_t,$$

where $g_t(\theta)$ is a $(2L+1)$ dimensional vector. Then, it follows from (2.11) and (2.12) that

(2.13) $$E[g_t(\theta)] = 0,$$

where $E[\bullet]$ is an unconditional expectations operator. If the model underlying (2.13) is correctly specified, then

(2.14) $$\overline{g_T(\theta)} = \frac{1}{T}\sum_{t=1}^{T} g_t(\theta)$$

should be close to zero for a large number of T. Therefore, the GMM estimator of θ is obtained as follows:

(2.15) $$\hat{\theta} = \arg\min_{\theta} \quad Q_T(\theta) = \overline{g_T(\theta)}' W_T \overline{g_T(\theta)}$$

where W_T is a $(2L+1)$-by-$(2L+1)$ symmetric, positive definite matrix.[2] Hansen (1982) showed that $J_T = TQ_T(\hat{\theta})$ is asymptotically distributed as a chi-square with $(2L-1)$ degrees of freedom. This J statistic is used to test the overidentifying restrictions of the model, expressed in (2.13). The degrees of freedom are equal to the number of orthogonality conditions used in the estimation minus the number of estimated parameters.

To analyze the structural stability of parameter vector $\theta = (\beta, \gamma)'$, let us divide the total sample into the following two subsamples:

Sample 1: $t = -n_1 + 1, \ldots, 0,$

Sample 2: $t = 1, 2, \ldots, n_2.$

Then, the null and the alternative hypotheses for the test are as follows:

H_0: $E[g_1(\theta_0)] = 0$ and $E[g_2(\theta_0)] = 0,$

H_A: $E[g_1(\theta_0)] = 0$ and $E[g_2(\theta_0)] \neq 0,$

where θ_0 shows the true value of parameters, and the suffices 1 and 2 indicate the subsamples 1 and 2. The null hypothesis can be analyzed by the predictive test. The predictive test for structural change uses estimates from one sample to evaluate the moment conditions of another sample. This approach only requires an estimation of

---

[2] W_T is given by $\left[\frac{1}{T}\sum_{t=1}^{T} g_t(\hat{\theta}) g_t(\hat{\theta})'\right]^{-1}$.

one of the subsamples, and it also explores all orthogonality conditions of the other subsample. The test statistic (*PRN*) is defined as follows:

(2.16) $$PRN = n_2[\frac{1}{\sqrt{n_2}}\sum_{t=1}^{n_2} g_2(\hat{\theta}_1)]'(\hat{V}_2^p)^{-1}[\frac{1}{\sqrt{n_2}}\sum_{t=1}^{n_2} g_2(\hat{\theta}_1)],$$

where

$$\hat{V}_2^p = S_2 + c\tilde{D}_2(D_1'S_1D_1)^{-1}\tilde{D}_2',$$

$$S_i = n_i g_i g_i', \qquad i = 1,2,$$

$$g_1 = \frac{1}{n_1}\sum_{t=-n_1+1}^{0} g_t,$$

$$g_2 = \frac{1}{n_2}\sum_{t=1}^{n_2} g_t,$$

$$D_i = \frac{\partial g_i}{\partial \theta'}, \quad i = 1,2,$$

$$\tilde{D}_2 = \frac{\partial g_2(\hat{\theta}_1')}{\partial \theta'},$$

$$c = \frac{n_2}{n_1}.$$

This test statistic has a chi-square distribution with degrees of freedom equal to the dimensions of g_2, which is equal to $2L+1$.

3. Findings from Sampling Experiments

Let us examine the experimental results and the effects of a parameter shift on the J test and the predictive test. These experiments were executed based on the following assumptions:

1) The replication of the Monte Carlo simulation is one thousand times.

2) The iteration of GMM in the second stage is ten times.[3]

---

[3] We estimated the parameter and constructed the J statistics for each data set using the K-step estimator described in the algorithm below:

Let n be the number of iteration and $W_T^{(0)}$ be the identity matrix.

(1) $n = 1$:

$$\theta^{(n)} = \arg\min_{\theta} \quad Q_T^{(n)} = \overline{g_T(\theta)}' W_T^{(n-1)} \overline{g_T(\theta)}.$$

(2) $n = n + 1$: Set

$$Q_T^{(n)} = Q_T(\theta^{(n)}) = \min \quad \overline{g_T(\theta^{(n)})}' W_T^{(n-1)}(\theta^{(n-1)}) \overline{g_T(\theta^{(n)})}.$$

3) The grid points of the quadrature method are eight for each variable, giving sixty-four states of nature in total.

4) The exogenous variables follow the AR(1) process. The chosen coefficient matrices of the VAR are as follows:

$$\text{case I:}\quad A = \begin{bmatrix} -0.5 & 0.0 \\ 0.0 & -0.5 \end{bmatrix},$$

$$\text{case II:}\quad A = \begin{bmatrix} -0.5 & 0.1 \\ 0.1 & -0.5 \end{bmatrix},$$

$$\text{caseIII:}\quad A = \begin{bmatrix} -0.1 & 0.0 \\ 0.0 & -0.1 \end{bmatrix}.$$

The covariance matrix of the error is as follows:

$$\Omega = \begin{bmatrix} 0.01 & 0.00 \\ 0.00 & 0.01 \end{bmatrix}.$$

We evaluated the performance of each test statistic based on the size and the power of the tests.

Let us examine Tables 1-1, 1-2, and 1-3. Table 1-1 shows the experimental design. Table 1-2 indicates the empirical results of the predictive test and the J test.[4]

If $|Q_T^{(n)} - Q_T^{(n-1)}| < 10^{-4}$ or $n = K$, then go to (3). Otherwise, return to (2).

(3) Take $\theta^{(n)}$ as $\hat{\theta}$, and

$$J_T = TQ_T(\hat{\theta}) = T\overline{g_T(\hat{\theta})}' W_T^{(n-1)} \overline{g_T(\hat{\theta})}.$$

The minimization routine always begins its search at the true parameters. In practice, K is often set equal to 2. As Kocherlakota (1990) demonstrated, however, the small sample properties of the two-step estimator are worse than the properties of the multistep estimator. Thus, K was set to 10 in the present study.

[4] The empirical size in Table 1-2 is obtained based on the critical values of the chi-square distribution. These values are shown as follows:

	5%	10%
$\chi^2(1)$	3.841	2.706
$\chi^2(3)$	7.815	6.251
$\chi^2(5)$	11.070	9.236
$\chi^2(7)$	14.067	12.017

Table 1-2 reports the rejection rates of each test at nominal significance levels of 5% and 10%. Except for experiments A10 and A11, all of the in-sample and out-sample data are kept at fifty to evaluate the performances of the predictive test and the J test. Parameter values and the law of motion are also kept the same for both in-sample data and out-sample data. The subjective discount rate is set at 0.970, and the measure of relative risk aversion is set at 0.5. Because no structural change is made, the rejection rate of the predictive test is the empirical size, which should be close to the corresponding nominal size. However, as shown in Table 1-2, the empirical size of the predictive test is larger than the corresponding nominal size. For example, let us look at the results of experiment A1. For each nominal size of 5% and 10%, the corresponding empirical size is 9.6% and 16.4%, respectively. This tendency continues in the other experiments, except for experiments A10 and A11. Thus, the predictive test tends to over-reject the null hypothesis.[5] Also, as the number of lags used for instruments increases, this tendency becomes stronger.

Alternatively, as indicated in Table 1-2, the empirical size of the J test is relatively close to its nominal size.[6] The experimental results clearly indicate that the J test performs well under these conditions. This result is consistent with the analysis of Tauchen (1986).

Because the size of the tests is distorted, especially for the predictive test, we report the size-adjusted critical values in Table 1-3. These values are based on the distribution created by the simulation technique. Following Tanizaki (1995), we created the distribution of each test statistic via simulation and found the critical values based on the distribution.

Tables 2-1 and 2-2 show the experimental design and the corresponding experimental results for changes in γ. In Table 2-1, we changed the parameter value of γ from 0.5 (in-sample data) to 0.6, 0.7, 0.8, 0.9, 1.0, and 1.5 (out-sample data),

---

[5] As the results of experiments A10 and A11 show, however, the empirical size of the predictive test is smaller than the corresponding nominal size when the number of out-sample is small.

[6] In this experiment, one hundred samples were used to evaluate the J test. This sample size was chosen because applied researchers usually use data of the whole sample period to analyze the specification of the model and to ascertain if the structure of the model changes during the sample period.

keeping β at 0.970. We evaluated the performance of the test statistics based on the rejection rates in Table 2-2. Because a parameter shift occurred in the experimental data, the rejection rate is the experimental value of the power of the predictive test, which should be close to one. Table 2-2 shows that both the size-corrected power and the power without size correction of the predictive test increase as the change in γ increases. To compute the size-corrected power, we used the size-adjusted critical values shown in Table 1-3. Table 2-2 also shows that the power without size correction of the predictive test increases as the number of lags used for instruments increases, whereas the size-corrected power of the predictive test does not necessarily increase as the number of instruments increases. For example, when γ changes from 0.5 to 0.9 (B10, B11, and B12), the power without size correction increases as *Nlag* increases, but the size-corrected power decreases as *Nlag* increases. Thus, we can say that given the same parameter shift, the higher power of the predictive test comes from the size distortion but not from the better power properties with more instruments. Also, the power of the predictive test increases as the change in parameter increases.

Table 2-2 also shows the results of the J test. In these experiments, all one hundred samples are used to evaluate the small sample properties of the J test. This evaluates the effect of the parameter shift during the sample period on the performance of the J test. As Table 2-2 shows, the power of the J test is not high, even if a large parameter shift occurs. For example, when γ changes from 0.5 to 1.5 (B16, B17, and B18), both the size-corrected power and the power without size correction are less than 5% at 5% nominal significance level. The J test clearly tends to accept the model in spite of the parameter shift and thus cannot detect the structural change of the model. This finding supports the results of Ghysels and Hall (1990b) that the J test has no power against local alternatives.

Tables 3-1 and 3-2 indicate the experimental design and the corresponding experimental results for changes in β. In Table 3-1, we changed the parameter value of β from 0.970 (in-sample data) to 0.975, 0.980, 0.985, 0.990, 0.995, and 1.000 (out-sample data), keeping γ at 0.50. Tables 3-1 and 3-2 indicate that similar results can be found for changes in β.

Finally, let us examine Tables 4-1 and 4-2. Table 4-1 shows the experimental design. In these experiments, we changed the value of γ from 0.5 (in-sample data) to

0.6, 0.7, 0.8, 0.9, 1.0, and 1.5 (out-sample data), keeping the same value (0.970) for β. These experiments analyze the effect of the difference in sample size between in-sample and out-sample data on the predictive test and the J test. Table 4-2 indicates the experimental results and shows that as the number of out-sample data decreases, the power of the predictive test drastically decreases. For example, let us observe the empirical size at the 5% nominal size when γ changes from 0.5 to 1.5 (see D16, D17 and D18). The power without size correction is 85.6%, and the size-corrected power is 75.8% when the number of in-sample data and out-sample data is equal at fifty and fifty. However, the power without size correction decreases to 47.9%, and the size-corrected power decreases to 35.1% when the number of in-sample data is equal to seventy-five and the number of out-sample data is equal to twenty-five. Furthermore, the power without size correction decreases to 6.9% and the size-corrected power decreases to 3.9% when the number of in-sample data is equal to ninety and the number of out-sample data is equal to ten. Therefore, when structural change occurs near the end of the sample period, it is difficult to detect this change. Table 4-2 also shows that the J test is not affected by the difference in the timing of the structural change.

Table 1-1
Experimental Design
(β=0.970, γ=0.50)

Experiment	Data (insample, outsample)	Autoregressive Matrix	*Nlag*
A1	(50, 50)	case I	1
A2	(50, 50)	case I	2
A3	(50, 50)	case I	3
A4	(50, 50)	case II	1
A5	(50, 50)	case II	2
A6	(50, 50)	case II	3
A7	(50, 50)	case III	1
A8	(50, 50)	case III	2
A9	(50, 50)	case III	3
A10	(75, 25)	case I	1
A11	(90, 10)	case I	1

Table 1-2
Empirical Size

Experiment	J test 5%	J test 10%	Predictive test 5%	Predictive test 10%
A1	0.046	0.108	0.096	0.164
A2	0.036	0.086	0.174	0.243
A3	0.038	0.089	0.299	0.382
A4	0.056	0.094	0.092	0.147
A5	0.046	0.106	0.205	0.262
A6	0.028	0.067	0.271	0.343
A7	0.045	0.083	0.057	0.095
A8	0.028	0.066	0.150	0.211
A9	0.020	0.051	0.254	0.325
A10	0.044	0.098	0.014	0.029
A11	0.052	0.109	0.001	0.002

Table 1-3
Size-Adjusted Critical Value

Experiment	J test 5%	J test 10%	Predictive test 5%	Predictive test 10%
A1	3.747	2.809	11.385	7.681
A2	7.401	5.931	18.947	14.530
A3	10.516	8.934	29.786	23.829
A4	3.908	2.566	9.684	7.528
A5	7.685	6.325	20.991	16.461
A6	9.638	8.459	28.552	22.952
A7	3.460	2.478	7.996	6.068
A8	6.691	5.454	17.973	13.517
A9	9.314	8.263	26.473	20.929
A10	3.699	2.656	4.533	3.180
A11	3.849	2.937	1.411	0.974

Table 2-1
Experimental Design
(β=0.970)

Experiment	γ	*Nlag*
B1	0.6	1
B2	0.6	2
B3	0.6	3
B4	0.7	1
B5	0.7	2
B6	0.7	3
B7	0.8	1
B8	0.8	2
B9	0.8	3
B10	0.9	1
B11	0.9	2
B12	0.9	3
B13	1.0	1
B14	1.0	2
B15	1.0	3
B16	1.5	1
B17	1.5	2
B18	1.5	3

The chosen coefficient matrix of the VAR is case I.
The number of in-sample data and out-sample data is fifty and fifty.

Table 2-2
Power of Tests
(Changes in γ)

	Without Size Correction				With Size Correction			
	J test		Predictive test		*J* test		Predictive test	
Experiment	5%	10%	5%	10%	5%	10%	5%	10%
B1	0.060	0.118	0.136	0.194	0.064	0.109	0.062	0.138
B2	0.047	0.115	0.222	0.290	0.060	0.131	0.059	0.124
B3	0.042	0.093	0.320	0.391	0.054	0.105	0.061	0.099
B4	0.045	0.108	0.201	0.264	0.047	0.097	0.096	0.203
B5	0.038	0.094	0.284	0.360	0.048	0.116	0.103	0.180
B6	0.031	0.081	0.395	0.483	0.045	0.098	0.097	0.166
B7	0.034	0.083	0.277	0.357	0.039	0.079	0.138	0.279
B8	0.042	0.100	0.349	0.435	0.054	0.113	0.145	0.234
B9	0.031	0.088	0.452	0.534	0.042	0.104	0.123	0.202
B10	0.042	0.094	0.375	0.462	0.043	0.085	0.219	0.381
B11	0.037	0.102	0.453	0.551	0.050	0.123	0.208	0.327
B12	0.032	0.100	0.532	0.613	0.049	0.108	0.169	0.262
B13	0.054	0.110	0.457	0.545	0.057	0.105	0.307	0.465
B14	0.047	0.099	0.571	0.648	0.058	0.114	0.319	0.427
B15	0.028	0.070	0.672	0.741	0.038	0.083	0.297	0.393
B16	0.045	0.102	0.856	0.893	0.049	0.094	0.758	0.858
B17	0.032	0.088	0.915	0.940	0.040	0.112	0.781	0.866
B18	0.029	0.085	0.936	0.957	0.040	0.099	0.754	0.834

Without Size Correction shows the power without size correction.
With Size Correction shows the size-corrected power.

Table 3-1
Experimental Design
(γ=0.50)

Experiment	β	*Nlag*
C1	0.975	1
C2	0.975	2
C3	0.975	3
C4	0.980	1
C5	0.980	2
C6	0.980	3
C7	0.985	1
C8	0.985	2
C9	0.985	3
C10	0.990	1
C11	0.990	2
C12	0.990	3
C13	0.995	1
C14	0.995	2
C15	0.995	3
C16	1.000	1
C17	1.000	2
C18	1.000	3

The chosen coefficient matrix of the VAR is case I.
The number of in-sample data and out-sample data is fifty and fifty.

Table 3-2
Power of Tests
(Changes in β)

	Without Size Correction				With Size Correction			
	J test		Predictive test		*J* test		Predictive test	
Experiment	5%	10%	5%	10%	5%	10%	5%	10%
C1	0.043	0.100	0.122	0.173	0.044	0.096	0.051	0.125
C2	0.054	0.103	0.188	0.264	0.063	0.126	0.062	0.104
C3	0.040	0.087	0.269	0.347	0.050	0.094	0.058	0.102
C4	0.052	0.108	0.121	0.180	0.055	0.099	0.041	0.127
C5	0.045	0.103	0.221	0.297	0.054	0.120	0.068	0.126
C6	0.024	0.070	0.282	0.371	0.034	0.082	0.051	0.101
C7	0.048	0.104	0.164	0.243	0.052	0.094	0.080	0.174
C8	0.034	0.090	0.268	0.353	0.043	0.112	0.092	0.172
C9	0.028	0.073	0.329	0.412	0.038	0.091	0.070	0.116
C10	0.046	0.106	0.240	0.334	0.051	0.098	0.120	0.246
C11	0.044	0.107	0.298	0.381	0.055	0.123	0.107	0.192
C12	0.032	0.082	0.396	0.477	0.046	0.089	0.100	0.173
C13	0.050	0.105	0.336	0.431	0.054	0.099	0.179	0.346
C14	0.040	0.087	0.386	0.469	0.049	0.108	0.144	0.239
C15	0.037	0.094	0.459	0.546	0.048	0.110	0.120	0.202
C16	0.034	0.085	0.409	0.509	0.037	0.079	0.239	0.417
C17	0.041	0.091	0.449	0.534	0.045	0.106	0.200	0.310
C18	0.033	0.076	0.531	0.598	0.044	0.092	0.173	0.250

Without Size Correction shows the power without size correction.
With Size Correction shows the size-corrected power.

Table 4-1
Experimental Design
(β=0.970)

Experiment	γ	Data (insample, outsample)
D1	0.6	(50, 50)
D2	0.6	(75, 25)
D3	0.6	(90, 10)
D4	0.7	(50, 50)
D5	0.7	(75, 25)
D6	0.7	(90, 10)
D7	0.8	(50, 50)
D8	0.8	(75, 25)
D9	0.8	(90, 10)
D10	0.9	(50, 50)
D11	0.9	(75, 25)
D12	0.9	(90, 10)
D13	1.0	(50, 50)
D14	1.0	(75, 25)
D15	1.0	(90, 10)
D16	1.5	(50, 50)
D17	1.5	(75, 25)
D18	1.5	(90, 10)

The chosen coefficient matrix of the VAR is case I.

Table 4-2
Power of Tests
(Different Timing in Structural Change)

	Without Size Correction				With Size Correction			
	J test		Predictive test		*J* test		Predictive test	
Experiment	5%	10%	5%	10%	5%	10%	5%	10%
D1	0.060	0.118	0.136	0.194	0.064	0.109	0.062	0.138
D2	0.049	0.107	0.012	0.022	0.056	0.101	0.004	0.012
D3	0.054	0.105	0.000	0.000	0.057	0.097	0.000	0.000
D4	0.045	0.108	0.201	0.264	0.047	0.097	0.096	0.203
D5	0.053	0.107	0.015	0.028	0.055	0.101	0.006	0.017
D6	0.049	0.105	0.001	0.001	0.049	0.099	0.000	0.001
D7	0.034	0.083	0.277	0.357	0.039	0.079	0.138	0.279
D8	0.040	0.099	0.039	0.050	0.045	0.089	0.015	0.039
D9	0.057	0.105	0.000	0.001	0.059	0.094	0.000	0.000
D10	0.042	0.094	0.375	0.462	0.043	0.085	0.219	0.381
D11	0.051	0.115	0.047	0.079	0.053	0.104	0.022	0.051
D12	0.058	0.125	0.003	0.004	0.063	0.114	0.001	0.003
D13	0.054	0.110	0.457	0.545	0.057	0.105	0.307	0.465
D14	0.048	0.088	0.080	0.129	0.057	0.086	0.038	0.082
D15	0.043	0.106	0.006	0.009	0.046	0.093	0.001	0.006
D16	0.045	0.102	0.856	0.893	0.049	0.094	0.758	0.858
D17	0.064	0.111	0.479	0.568	0.065	0.103	0.351	0.486
D18	0.048	0.094	0.069	0.098	0.052	0.088	0.039	0.072

Without Size Correction shows the power without size correction.
With Size Correction shows the size-corrected power.

4. Concluding Remarks

We studied the small sample properties of the predictive test by Monte Carlo experiments, and we examined the effect of structural change on the J test in small samples. The main findings from the experiments are summarized as follows:

1) The empirical size of the predictive test is biased toward rejecting the null hypothesis. Also, as the number of lags used for instruments increases, this tendency becomes stronger. On the other hand, the empirical size of the J test is relatively close to its nominal size.
2) The power of the predictive test becomes higher as the shift in a parameter increases. However the power of the J test is not high even if a large parameter shift occurs, and thus the J test clearly tends to accept the model in spite of the parameter shift.
3) The power of the predictive test also becomes higher as the number of instruments increases. Given the same parameter shift, this higher power of the predictive test comes from the size distortion but not from the better power properties with more instruments. If the predictive test and the J test are compared in terms of the size-corrected power, however, the power of the predictive test is higher than that of the J test. This finding becomes apparent when a large parameter shift occurs.
4) As the number of out-sample data decreases, the power of the predictive test decreases significantly. Thus, it may not be easy to find the structural change if it occurs near the end of the sample period.

ACKNOWLEDGMENTS

The authors would like to thank the editor Esfandiar Maasoumi; the anonymous referee of this journal; and the participants in the workshop at Nanzan University and the Finance Forum for their helpful comments and suggestions. We also would like to thank George Tauchen, who kindly provided the Markov chain Gauss code.

REFERENCES

Breeden, D., (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, pp. 265-296.

Epstein, L. G. and S. E. Zin, (1991), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, pp. 263-286.

Ghysels, E. and A. Hall, (1990a), "Are Consumption-Based Intertemporal Capital Asset Pricing Models Structural ?," *Journal of Econometrics*, 45, pp. 121-139.

Ghysels, E. and A. Hall, (1990b), "A Test for Structural Stability of Euler Conditions Parameters Estimated via the Generalized Method of Moments Estimator," *International Economic Review*, 31, pp. 355-364.

Hamori, S., (1992), "On the Structural Stability of Preference Parameters Obtained from Japanese Financial Market Data," *Economics Letters*, 40, pp. 459-464.

Hamori, S. and S. Kitasaka, (1992), "Small Sample Properties of Generalized Method of Moments Estimator with MA(1) Error," *Journal of the Japan Statistical Society*, 22, pp. 161-172.

Hansen, L. P., (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, pp. 1029-1053.

Hansen, L. P. and K. J. Singleton, (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, pp. 1269-1286.

Hansen, L. P. and K. J. Singleton, (1983), "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, pp. 249-265.

Kocherlakota, N. R., (1990), "On Tests of Representative Consumer Asset Pricing Models," *Journal of Monetary Economics*, 26, pp. 285-304.

Lucas, L. E. Jr., (1976), "Econometric Policy Evaluation: A Critique," in K.Brunner ed., *The Phillips Curve and Labor Market*, Carnegie-Rochester Conference Series on Public Economics, 1, (North Holland) pp. 19-46.

Lucas, L. E. Jr., (1978), "Asset Prices in a Exchange Economy," *Econometrica,* 46, pp. 1429-1445.

Mehra, R. and E. C. Prescott, (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* 15, pp. 145-161.

Tanizaki, H., (1995), "Asymptotically Exact Confidence Intervals of CUSUM and CUSUMSQ Tests: A Numerical Derivation Using Simulation Technique," *Communications in Statistics: Simulation and Computation,* 24, (forthcoming).

Tauchen, G., (1986), "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data," *Journal of Business and Economic Statistics,* 4, pp. 397-416.

Tauchen, G. and R. Hussey, (1991), "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica,* 59, pp. 371-396.