ISSN 0709-9231

CAHIER 8724

Some Additional Specification Tests for Generalized Method of Moments Estimators with Macro-Economic Applications Part I : Theory

by

Eric Ghysels[1]
and
Alastair Hall[2]

[1] Département de sciences économiques and Centre de recherche et développement en économique, Université de Montréal.

[2] Department of Economics and Business, North Carolina State University.

September 1987

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche. Ce cahier a également été publié par le Centre de recherche et développement en économique (Publication #2687).

RÉSUMÉ

Plusieurs nouveaux tests de spécification pour la méthode généralisée des moments sont présentés dans ce papier. Le premier test est du type Chow pour la stabilité des coefficients des équations d'Euler. Ce test est inspiré par la critique de Lucas. Un deuxième test développé dans le papier est un test d'englobement par une représentation vectorielle autorégressive. Ce test est basé sur une comparaison de la structure sous-jacente estimée par un processus vectoriel autorégressif, sans contrainte, avec ce même processus estimé conjointement avec les équations d'Euler. Plusieurs applications macroéconomiques sont aussi présentées.

Mots-clés: méthode généralisée des moments, test du type Chow, modèles vectoriels autorégressifs, principe d'englobement.

* * * ABSTRACT * * *

In this paper several additional GMM specification tests are studied. A first test is a Chow-type test for structural parameter stability of GMM estimates. The test is inspired by the fact that "taste and technology" parameters are uncovered. The second set of specification tests are VAR Encompassing tests. It is assumed that the DGP has a finite VAR representation. The moment restrictions which are suggested by economic theory and exploited in the GMM procedure represent one possible characterization of the DGP. The VAR is a different but compatible characterization of the same DGP. The idea of the VAR Encompassing tests is to compare parameter estimates of the Euler conditions and VAR representations of the DGP obtained separately with parameter estimates of the Euler conditions and VAR representations obtained jointly. There are several ways to construct joint systems which are discussed in the paper. Several applications are also discussed.

Key words: generalized method of moments, Chow-type tests, Vecot autoregressive models, encompassing principle.

1. INTRODUCTION

Both Lucas (1976) and Sims (1980) have had a profound impact on the way macroeconomists conduct empirical research. Lucas criticized the existing strategies for econometric analysis of macroeconomic time series and argued that parameters of traditional econometric models are not invariant with respect to shifts in policy regimes. In response to that criticism several inference strategies for "taste and technology " structural parameters were suggested. Hansen's (1982) generalized method of moments (henceforth GMM) instrumental variables procedure is among the most notable ones. Sims (1980), on the other hand, contended that standard strategies for econometric analysis make incredible identification assumptions and suggested finite order vector autoregressive (henceforth VAR) models as an alternative style of empirical macroeconomics. In this paper several specification tests are introduced for GMM Euler conditions estimators which are built on the insights and arguments presented by Lucas and Sims in their seminal papers.

A first test is very much in the spirit of Chow's (1960) classical test for the equality between sets of coefficients in linear regressions. A Chow-type test for structural stability of parameter estimates obtained via the GMM procedure is presented. The desire to uncover structural parameters which are assumed to be invariant across time, regimes, etc. inspired the development of the test statistic. The testing procedure may also be useful for verifying commonly made assumptions about temporal aggregation across agents imbedded in the representative agent model.

The second test is a VAR-Encompassing test and is designed according to the Encompassing Principle discussed by Mizon and Richard (1986). It is assumed that $\{x_t\}$ is a stationary and ergodic multivariate data generating process (henceforth DGP). This assumption corresponds to Hansen's (1982) original development of GMM. It is also assumed that economic theory suggests moment restrictions of the following nature $E(f(x_t, \beta)) = 0$. Since $\{x_t\}$ is ergodic and stationary it is known to have a Wold decomposition moving average representation. The DGP has a VAR representation when some additional assumptions are satisfied. The moment restrictions, if they are valid, represent one possible characterization of the DGP, while the VAR is a characterization of the linearly indeterministic part of the same DGP. The idea of the VAR Encompassing tests is to compare parameter estimates of Euler conditions and VAR representations of the DGP obtained separately with parameter estimates of the Euler conditions and VAR representations obtained jointly. There are several ways to construct joint systems which are discussed in this paper. The major motivation for using a VAR representation is, as Sims (1980) emphasized, to benefit from an inference strategy which does not strictly involve any strong identification assumptions which are typical for "structural" approaches. Specification tests solely depending on Euler Conditions moment functions are conditional on the model specification. Newey (1985), for instance, showed that tests for violation of Euler conditions may fail to detect misspecification asymptotically. In contrast, a VAR representation always exists, given a set of well-known assumptions are satisfied, and will be consistently estimated without misspecification provided the error process is general enough.

A comparison of the VAR-Encompassing testing procedure with some of the recently developped calibration methods and "backwards" solution methods for nonlinear rational expectations models proposed by Sims (1984) is also presented. It is important to emphasize that a VAR Encompassing test does not "attempt to solve" Euler conditions. A direct link between moment restrictions and VAR representations is usually difficult to make, unless Euler conditions are linear and can be solved explicitly as Hansen and Sargent (1980) demonstrated. In general, when Euler conditions are nonlinear one expects that the DGP has nonlinear time series properties. In such cases one can view the VAR as the linear part of a Volterra series expansion of the DGP.

The organization of the paper is as follows: in section 2 the Chow-type test is presented. The VAR Encompassing test is discussed in section 3. Section 4 concludes the paper with a discussion of the possible extensions of the proposed tests and a discussion of the macroeconomic applications which will be presented in Part II of this paper. A mathematical appendix contains the proofs.

2. A CHOW-TYPE GMM TEST STATISTIC

The Lucas critique has led some econometricians to redirect their attention toward estimation of structural "taste and technology" parameters. The idea to make inference about such parameters led to the formulation of Euler conditions estimators such as GMM. A natural question to consider, although ignored so far, is to test whether the parameters uncovered from such models are indeed invariant across policy regimes, time, etc. We propose a Chow-type GMM test for structural stability of parameter esttimates.

The sample is split into two parts, namely:

Sample 1: $t = k - n_1 + 1, k - n_1 + 2, \ldots , k$

Sample 2: $t = k + 1, k + 2, \ldots , k + n_2$

The subsamples are assumed to be sufficiently large for asymptotic theory to be valid. As in Hansen (1982), it is assumed that the following orthogonality conditions suggested by economic theory hold:

$$E(f(x_t, \beta)) = 0 \qquad (2.1)$$

The regularity conditions on f and the stochastic process x_t listed in Hansen (1982) are assumed to hold, except for stationarity and ergodicity. Instead one of the following three regularity conditions will be used in this paper:

Assumption 2.1: (M-dependence (Anderson, 1958)) The process x_t is an M-dependent process.

Assumption 2.2: (Mixing conditions (White and Domowitz, 1984)).
The process x_t satisfies:

a) $\exists$ finite constants $\bar{D}$, $\delta > 0$ and $r > 1$ such that $\forall t$:

$E\left(\left|f(x_t, \beta)\right|^{4(r+\delta)}\right) < \bar{D}$

b) x_t is mixing with either $\emptyset(m)$ of size

$2r/(2r-1)$ or $c(\alpha)$ of size $2r(r-1)$, $r>1$

Assumption 2.3: (Finite order AR) The process x_t is stationary and ergodic and has a finite order vector autoregressive representation.

Assumptions 2.1 and 2.2 will be used in this section while 2.3 will be in the next section. With the regularity conditions holding in each sub-sample one obtains estimates $\hat{\beta}_1$ from sample 1 and $\hat{\beta}_2$ from sample 2. If $\beta_i = \text{plim } \hat{\beta}_i$, $n_i \to \infty$ then the structural constancy can be expressed as follows

$$H_o: \beta_1 = \beta_2 \tag{2.2a}$$

$$H_a: \beta_1 \neq \beta_2 \tag{2.2b}$$

Following Hansen's notation one can define, under the null hypothesis, the matrices A*, D, S (see appendix) as well as the covariance matrix across samples:

$$\bar{S} = E\left(\left[n_1^{-\frac{1}{2}} \sum_{t=k-n_1+1}^{k} f(x_t,\beta)\right] \left[n_2^{-\frac{1}{2}} \sum_{t=k+1}^{k+n_2} f(x_t, \beta)\right]\right) \tag{2.3}$$

The Chow-type GMM test is based on the following theorem:

Theorem 2.4: Under the null hypothesis (2.2a) $\hat{\beta}_1 - \hat{\beta}_2 \xrightarrow{d} N(o,V)$ where the covariance matrix V is given by:

$$V = 2\left\{[A^*D]^{-1}A^*\left[S + \bar{S}\right]A^{*\prime}[A^*D]^{-1\prime}\right\} \tag{2.4}$$

When x_t satisfies assumption 2.1 the test statistic

$$C_o = (\hat{\beta}_1 - \hat{\beta}_2)' V^{-1} (\hat{\beta}_1 - \hat{\beta}_2) \qquad (2.5)$$

is an asymptotically exact test with χ^2 distribution and ℓ degrees of freedom where ℓ = dim β. The test statistic is based on consistent estimates of A*, D and S over the entire sample and $\bar{S} = 0$. Under assumption 2.2 the test with $\bar{S} = 0$ is asymptotically conservative, i.e. the test statistic is always at least as large as the true statistic using the actual value of $\bar{S}$.

Proof: see appendix.

The results stated in theorem 2.4 indicate that the covariance across samples vanishes when the x_t process satisfies the conditions formulated under assumption 2.1 The matrix $\bar{S}$ will be bounded but nonzero under the less restrictive mixing conditions in assumption 2.2. The relative magnitudes of S and $\bar{S}$ in finite samples is worthy of further investigation. The latter is a subject of future research which should determine how conservative the test is under assumption 2.2 (see Ghysels and Hall (1987b)).

The Chow-type GMM test could be used to test stability when a major policy shift occured during the sample or simply as a routine test by splitting the actual samples in two halfs. It is important to have sufficient data left in the subsamples to rely on the asymptotic properties of the test. For that matter it is also important to better understand the features of the test statistic in small samples. A study in the spirit of Tauchen's (1986) small sample evidence on GMM estimators could provide insights on this matter. Some of this research is currently in progress (see Ghysels and Hall (1987b)). It is also important to emphasize that there is more to the Chow-type test than just a test for policy-invariance. Most of the applications of GMM estimators were in the context of representative agent models. Unless some specific preference structures apply, such as suggested by Gorman (1953), we know that the parmeters of representative models depend on the heterogeneity of preferences and endowments across agents.[1] To the extent that there is that heterogeneity, a Chow-type test may be able to reveal it. Hence, a Chow-type test is a GMM specification test for structural invariance with regard to changing policy regimes, income distributions, distribution of preferences, etc. In that regard the Chow-type test distinguishes a GMM procedure where the estimated parameters are purely "nuissance parameters", as Hansen (1986) recently suggested, from a GMM set-up where beyond the validity of the moment restrictions there is a structural interpretation of the parameters.

3. VAR ENCOMPASSING TESTS

To derive the VAR Encompassing tests it will be assumed that the DGP satisfies assumpton 2.3. Consequently, x_t has a finite order VAR representation:

$$x_t = A(L)\, x_{t-1} + \varepsilon_t \tag{3.1}$$

where A(L) is a matrix polynomial of order p in the lag operator L and ε_t is an uncorrelated sequence of innovations. As will be made clear later in this section, it should be noted here that although the ε_t process is linearly indeterministic, since it is the Wold decomposition innovation process, it may still have a nonlinearly predictable part. In fact the inverse of the matrix polynomial A(L), defined in (3.1), may be considered as the linear part of a Volterra series expansion (see Volterra (1959)) for the nonlinear time series representation of the vector process x_t. The dimension of the x_t process will be denoted d_x and it will be assumed that the vector α represents the unknown parameters of A(L) in (3.1) which has dimension $d_\alpha = p \times d_x^2$. Equation (2.1) which represents the moment restrictions emerging from the Euler conditions, and equation (3.1) are two different but compatible characterizations of the DGP. The following notation will be introduced to facilitate the presentation:

$$f(x_t, \beta) = e_{1t}(x_{t+k}, \beta) \otimes Z_t' \tag{3.2}$$

where Z_t is the set of instruments used in the GMM procedure, e_{1t} is the error process generated by the Euler conditions and $k > 1$. Consider now the following three GMM procedures:

(1) Disjoint orthogonality conditions:

$$g_{1t}^{D}(x_{t+k}, \beta) \equiv e_{1t}(x_{t+k}, \beta) \otimes X_{1t} \tag{3.3}$$

$$g_{2t}^{D}(x_t, \alpha) \equiv (x_t - A(L)\, x_{t-1}) \otimes X_{2t-1} \tag{3.4}$$

with $E\, g_{1t}^{D}(x_{t+k}, \beta) = E\, g_{2t}^{D}(x_t, \alpha) = 0$

$X_{1t}' = (1 \; x_t' \; x_{t-1}' \; \ldots . \; x_{t-r}')$ for some r such that $1 + r \times d_x > d_\beta$

$X_{2t}' = (1 \; x_t' \; x_{t-1}' \; \ldots . \; x_{t-s}')$ for some $s > p$

and in addition with

$$g_t^{D'} \equiv (g_{1t}^{D'} \; g_{2t}^{D'})$$

$$g_n^{D} \equiv n^{-1} \sum_{t=1}^{n} g_t^{D} \; ; \; g_{in}^{D} \equiv n^{-1} \sum_{t=1}^{n} g_{it}^{D}$$

(2) Joint orthogonality conditions with VAR instruments:

$$g_{1t}^{V}(x_{t+k}, \alpha, \beta) \equiv e_{1t}(x_{t+k}, \beta) \otimes A^*(L)\, X_{1t} \tag{3.5}$$

$$g_{2t}^{V}(x_t, \alpha) \equiv (x_t - A(L)\, x_{t-1}) \otimes X_{2t-1} \tag{3.6}$$

with $E g_{1t}^{V}(x_{t+k}, \alpha, \beta) = E\, g_{2t}^{V}(x_t, \alpha) = 0$

$$A^*(L) = \text{diag}\,(1 \; A(L), \ldots, A(L))$$

and in addition with definitions for g_t^V, g_n^V and g^V_{in} which are similar to the disjoint system.

(3) Joint orthogonality conditions with nonlinear innovation restrictions:

$$g_{1t}^N(x_{t+k}, \alpha, \beta) \equiv e_{1t}(x_{t+k}, \beta) \otimes A^*(L)\, X_{1t} \tag{3.7}$$

$$g_{2t}^N\,(x_{t+1},\alpha) \equiv (x_{t+1} - A(L)\,x_t) \otimes X_{2t} \tag{3.8}$$

with $Eg_{1t}^N\,(x_{t+k}, \alpha, \beta) = Eg_{2t}^N\,(x_{t+1}, \alpha) = 0$

and in addition with definitions for g_t^N, g_n^N and g_{in}^N which are similar to the disjoint system.

The first system is called disjoint because it consists of the vector functions of two separate GMM procedures, the first yielding Euler conditions parameter estimates $\hat{\beta}_D$ and the second yielding VAR parameter estimates $\hat{\alpha}_D$. The vector $\hat{\Theta}_D$ stacks both together, i.e. $\hat{\Theta}_D' = (\hat{\beta}_D', \hat{\alpha}_D')$. The second system, in contrast, is not disjoint since the VAR polynomial matrices appear in the instrument set of the Euler conditions. The second system yields parameter estimates $\hat{\Theta}_V' = (\hat{\beta}_V', \hat{\alpha}_V')$. Finally, the third system has a structure similar to the second except that the VAR orthogonality conditions are moved one period ahead. It will become clear later why it is justified to call the third system joint with nonlinear innovation restrictions.

Using again Hansen's (1982) notation and theoretical results, it follows that $\hat{\Theta}_D$ is obtained by solving:

$$\underset{\beta}{\text{Min}}\ (g_{1n}^D)'\ W_{1n}^D\ g_{1n}^D \tag{3.9}$$

$$\underset{\alpha}{\text{Min}}\ (g_{2n}^D)'\ W_{2n}^D\ g_{2n}^D \tag{3.10}$$

Moreover, Hansen showed that the optimal weighting matrices minimizing, in the matrix sense, the asymptotic variance-covariance matrix of the parameter estimates are:

$$W_{10}^{D} = (E[\,g_{1t}^{D}(\beta_o)\; g_{1t}^{D}(\beta_o)'])^{-1} \qquad (3.11)$$

$$W_{20}^{D} = (E[\,g_{2t}^{D}(\alpha_o)\; g_{2t}^{D}(\alpha_o)'])^{-1} \qquad (3.12)$$

where $\Theta_o' = (\beta_o'\,\alpha_o')$ represent the true values. In addition:

$$\hat{\Theta}_D - \Theta_o = -[A_D^{*} D_D]^{-1} A_D^{*} g_n^{D} \qquad (3.13)$$

The following theorem establishes the relation between the first two GMM systems:

Theorem 3.1: The GMM estimator of the joint orthogonality conditions with VAR instruments $\hat{\Theta}_V$ is asymptotically equivalent to the estimator solving the following optimization problems:

$$\underset{\alpha}{\text{Min}}\; (g_{2n}^{D})'\, W_{2n}^{D}\, g_{2n}^{D}$$

so that $\hat{\alpha}_V$ and $\hat{\alpha}_D$ are asymptotically equivalent, and

$$\underset{\beta}{\text{Min}}\; (g_{1n}^{V})'\, W_{1n}^{V}\, g_{1n}^{V} \qquad (3.14)$$

where $W_{1n}^{V} \rightarrow W_{10}^{V}$ almost surely with

$$W_{10}^{V} = (E[g_{1t}^{V}(\alpha_o, \beta_o)\, g_{1t}^{V}(\alpha_o, \beta_o)'])^{-1}$$

$$(W_{10}^{V})^{-1} = (W_{10}^{D})^{-1} + W_{D}^{V}$$

$$W_{D}^{V} \neq 0$$

Proof: see appendix.

The result in theorem 3.1 shows that the difference between the disjoint system and the joint system with VAR instruments is essentially the set of instruments used to estimate the β_o vector. The VAR parameters remain asymptotically unaffected although they also appear in the g_{1t}^{V} vector function. It should be pointed out, however, that the result in theorem 3.1 is not "merely" a result emerging from a different set of instruments, namely the VAR instruments. In order to establish the result in theorem 3.1 it was assumed that the orthogonality conditions (3.2) hold for instruments Z_t which may be complicated nonlinear functions of the information set I_t. In particular it is assumed that:

$$E\, g_{1t}^{V}\, (g_{2t}^{V})' = 0$$

because of the timing of the innovation process of the VAR representation. Hence, the comparison of $\hat{\Theta}_V$ with $\hat{\Theta}_D$ implicitly invalidates the orthogonality of the error process e_{1t} to a string of complicated functions of elements of the information set I_t outside the range of linear instruments, i.e. X_{1t} or $A^*(L)X_{1t}$. Using theorem 3.1 and the theoretical results in Hansen (1982) which establish the asymptotic normality of GMM estimators; one can define the first VAR Encompassing test for the hypothesis:

$$H_0: \hat{\beta}_V = \hat{\beta}_D$$

$$H_A: \hat{\beta}_V \neq \hat{\beta}_D$$

The test statistic is:

$$VE_1 = (\hat{\beta}_D - \hat{\beta}_V)' [(T_1 + R_1)(W^V_{10})^{-1}(T_1+R_1)']^{-1}(\hat{\beta}_D - \hat{\beta}_V)$$

which is asymptotically χ^2 distributed with d_β degrees of freedom and where T_1 and R_1 are the upper left $d_\beta \times d_\beta$ submatrices respectively

$$T_V = -[A^*_V D_V]^{-1} A^*_V$$

and

$$R = - [A^*_D D_D]^{-1} A^*_D$$

and finally

$$\hat{\beta}_D - \hat{\beta}_V \overset{d}{\to} N(0, (T_1 + R_1)(W^V_{10})^{-1}(T_1 + R_1)')$$

The test statistic is based on the Encompassing Principle, discussed by Mizon and Richard (1986). Indeed, under the null hypothesis that the Euler conditions hold the joint orthogonality system with VAR instruments should encompass the disjoint system.

It was emphasized that the results stated in theorem 3.1 crucially depend on the timing of the VAR innovations process $(x_t - A(L) x_{t-1})$. The third system has the innovation process moved one period ahead of the information set I_t. This implies that:

$$Eg^N_{1t} (g^N_{2t})' = E [e_{1t}(x_{t+k},\beta) \otimes A^*(L) X_{1t}][(x_{t+1} - A(L) x_t) \otimes X_{2t}]' \neq 0$$

Each element of this matrix is assumed, under the GMM regularity conditions, to exist and to be finite. When Euler conditions are linear this amounts to variance-covariance matrix restrictions. When Euler conditions are nonlinear this amounts to saying that there are nonlinear forecast functions for the VAR innovations. While, in the latter case, this does not solve explicitly for nonlinear forecast functions for the linearly indeterministic innovation process of the VAR, it imposes comovement restrictions between the VAR innovations and nonlinear functions of x_t. In general, under the null hypothesis that the economic model yields valid moment restrictions one expects that the Euler conditions will "contaminate" the VAR estimates by incorporating extra information into the estimation of the VAR innovation process.

The case of linear Euler conditions has received considerable attention already in the work of Hansen and Sargent (1980) who show that there is a well-defined mapping between the parameter vector β_0 governing the Euler conditions and the VARMA representation of the DGP. In contrast, when the Euler conditions are nonlinear one expects a nonlinear time series model for the x_t process. Hence, as pointed out before, the VAR representation of x_t only charaterizes the linearly indeterministic part of the DGP, i.e. the linear part of the Volterra series expansion of the x_t process. It is important to note that the analysis of the nonlinear innovations restrictions is closely related to the work by Gallant and Tauchen (1986). In their work the (nonlinear) deviations from the linearly indeterministic (and normally distributed) VAR innovations are explicitly estimated via semi-nonparametric (SNP) estimation techniques. Hence Gallant and Tauchen's approach yields an explicitly formulated probability model for the VAR innovations approximated via an SNP polynomial expansion. While the approach discussed here is similar but less ambitious in goal it has the main advantage, in contrast to SNP expansions, to be simpler to implement. The argument is simply: since the DGP is a nonlinear time series, when Euler conditions are nonlinear, there are gains to be made from nonlinear moment restrictions when estimating the linearly indeterministic part. The following is assumed to hold in order to derive the next VAR Encompassing tests:

Assumption 3.2: The joint system with nonlinear innovation restrictions is overidentified.

This assumption yields:

Theorem 3.3: Under assumption 3.2 the asymptotic distribution of the GMM estimator of the joint orthogonality conditions with nonlinear innovation restrictions is equivalent to the asymptotic distribution of the solution to:

$$\begin{bmatrix} \partial Eg^N_{1t}(\Theta_0)/\partial\beta' & 0 \\ 0 & \partial Eg^N_{2t}(\Theta_o)/\partial\alpha' \end{bmatrix} \begin{bmatrix} W^N_{011} & W^N_{012} \\ W^N_{021} & W^N_{022} \end{bmatrix} \begin{bmatrix} g_{1n}(\Theta) \\ g_{2n}(\Theta) \end{bmatrix} = 0$$

where

$$W^N_{oij} \neq 0 \qquad i,j = 1,\ 2 \text{ and } i \neq j$$

$$W^N_{oii} \neq W^D_{oi};\ W^N_{oii} \neq W^V_{oi} \quad i = 1,\ 2$$

$$\text{and } W^N_o = (E[\,g^N_t\,(g^N_t)'\,])^{-1}$$

Proof: see appendix.

The result in theorem 3.3 demonstrates that the VAR parameter estimates depend on β and vice versa. The optimal weighting matrix is no longer block-diagonal. Hence, the system of first-order conditions is no longer block-recursive. Using Hansen's notation one has:

$$\hat{\Theta}_N - \Theta_o = -[\,A^*_N\,D_N]^{-1}\,A^*_N\,g^N_n \equiv T_N g^N_n$$

This result yields the following two VAR Encompassing tests for the null hypothesis:

$$H_o:\ \hat{\Theta}_N = \hat{\Theta}_D \qquad\qquad H_o:\ \hat{\Theta}_N = \hat{\Theta}_V$$

$$H_A:\ \hat{\Theta}_N \neq \hat{\Theta}_D \qquad\qquad H_A:\ \hat{\Theta}_N \neq \hat{\Theta}_V$$

$$VE_2 = (\hat{\Theta}_D - \hat{\Theta}_N)'\,[\,(T_N + R)(W^N_o)^{-1}(T_N + R)'\,]^{-1}(\hat{\Theta}_D - \hat{\Theta}_N)$$

$$VE_3 = (\hat{\Theta}_V - \hat{\Theta}_N)'[(T_N + T_V)(W_o^N)^{-1}(T_N + T_V)']^{-1}(\hat{\Theta}_V - \hat{\Theta}_N)$$

since

$$\hat{\Theta}_D - \hat{\Theta}_N \xrightarrow{d} N(0, (T_N + R)(W_o^N)^{-1}(T_N + R)')$$

$$\hat{\Theta}_V - \hat{\Theta}_N \xrightarrow{d} N(0, (T_N + T_V)(W_o^N)^{-1}(T_N + T_V)')$$

The test statistic VE_2 and VE_3 are both asymptotically χ^2 distributed with $d_\beta + d_\alpha$ degrees of freedom. The interpretation of the test statistics in the encompassing sense is as follows: the joint system with nonlinear innovative restrictions should encompass the system with VAR instruments as well as the disjoint system.

The two joint systems discussed so far have one essential link across equation systems. The link is the set of instruments used to estimate the Euler conditions parameter vector β. The VAR instrument choice is not necessarily an optimal instrument choice, however, in the sense that it does not necessarily attain the lower bound, characterized by Hansen (1985), over an admissible set of instruments measurable with respect to the information set I_t. In fact the VAR representation is present in the joint systems only as a particular characterization of the DGP which always exists and, as Sims (1980) emphasized, represents an inference strategy which does not involve any strong identification assumptions which are typical for "structural" approaches. The introduction of the VAR representation into the joint systems comes at a "cost" of a loss of degrees of freedom, since the number of parameters to be estimated increases, and it does not yield the most efficient instrumental variables inference strategy for Euler conditions parameters. In contrast, however, the "strength" or "benefit" of using a VAR representation is that one knows for sure that the parameter estimates of the VAR will be consistently estimated since there is never any source of misspecification. Specification tests solely depending on Euler conditions moment functions, i.e. $f(x_t, \beta)$, do not have such properties because they are conditional on the model specification. Newey (1985), for instance, showed that tests for violation of Euler conditions moment functions, such as Hansen's (1982) test of overidentifying restrictions, may fail to detect misspecification asymptotically against general misspecification.

With some extra complications it is possible to have VAR encompassing testing strategies which allow for optimal instrument choice. The idea is to substitute the VAR representation of x_t into the Euler conditions instead of using the VAR representation as a set of instruments. Consider the following joint system:

(4) Joint orthogonality conditions with VAR substitution:

$$g^S_{1t}(x_{t+k}, z_t, \alpha, \beta) \equiv e_{1t}(A(L)\, x_{t+k-1} + \varepsilon_{t+k}, \beta) \otimes Z'_t \quad (3.15)$$

$$g^S_{2t}(x_t, \alpha) = (x_t - A(L)\, x_{t-1}) \otimes X_{2t} \quad (3.16)$$

with $E\, g^S_{1t}(x_{t+k}, z_t, \alpha, \beta) = Eg^S_{2t}(x_t, \alpha) = 0$

and in addition with definitions for g^S_t, g^S_n and g^S_{in} which are similar to the previous systems.

The instrument vector z_t in (3.15) is left unspecified. Hence one can use an optimal set of instruments attaining the efficiency lower bound, or one may also use the VAR representation or any other instruments. It should also be noted that the timing of the VAR innovations in (3.16) coincides with the information set. Extensions of the derivaton with (3.16) moved one period ahead are straightforward but will not be explicitly discussed here. The presence of the VAR innovation process in e_{1t} implies that (3.15) does not represent a system of equations which can be estimated via a GMM procedure. Under the null hypothesis that the Euler moment restrictions are valid the use of estimates of ε_t on the basis of $\hat{\alpha}_D$, i.e. the disjoint VAR estimates, can be justified. Replacing ε_t by its estimate, denoted $\hat{\varepsilon}_t(\hat{\alpha}_D)$, yields again a well-defined

GMM procedure. Hence the estimates $\hat{\Theta}_S' = (\hat{\beta}_S' \; \hat{\alpha}_S')$ are obtained the usual way. The asymptotic properties of $\hat{\Theta}_S$, however, are somewhat different from the previously discussed estimators. The mean value theorem applied to $\hat{\Theta}_s$ has an additional term because of the presence of $\hat{\varepsilon}_t(\hat{\alpha}_D)$; namely:

$$\hat{\Theta}_s - \Theta_o = -[A_S^* D_S]^{-1} A_S^* [g_n^S + \frac{\partial g_n^S}{\partial \alpha_D'} (\hat{\alpha}_D - \alpha_o)] \qquad (3.17)$$

The notation in (3.17) is again borrowed from Hansen (1982). Furthermore, let the vector $(0 \; \iota)$ be defined such that $(\hat{\alpha}_D - \alpha_o) = (0 \; \iota)\,(\hat{\Theta}_D - \Theta_o)$. From (3.13) it follows that

$$\begin{aligned}\hat{\alpha}_D - \alpha_o &= -(0 \; \iota)[A_D^* D_D]^{-1} A_D^* g_n^D \\ &= -(0 \; \iota)[A_D^* D_D]^{-1} A_D^* g_n^S \equiv -Pg_n^S \qquad (3.18)\end{aligned}$$

Substituting (3.18) into (3.17) finally yields:

$$\hat{\Theta}_s - \Theta_o = -\{[A_S^* D_S]^{-1} A_S^* + \frac{\partial g_n^S}{\partial \alpha_D'} P\} g_n^S \equiv T_S g_n^S \qquad (3.19)$$

The null hypothesis formally tested by the VAR Encompassing test is:

$$H_o: \; \hat{\Theta}_S = \hat{\Theta}_D \qquad (3.20)$$

$$H_A: \; \hat{\Theta}_S \neq \hat{\Theta}_D \qquad (3.21)$$

The following theorem justifies the test statistic:

The following theorem justifies the test statistic:

Theorem 3.4: Under the null hypothesis (3.20):

$$\hat{\Theta}_S - \hat{\Theta}_D \xrightarrow{d} N(0, (T_S + R)(W_o^S)^{-1} (T_S + R)')$$

The proof of the above theorem follows from Hansen's theorem 3.1. Although the matrix W_o^S is block-diagonal, like W_o^V because of the innovation timing, it should be noted that $\hat{\alpha}_D$ and $\hat{\alpha}_S$ are no longer asymptotically equivalent estimators. The latter follows from the fact that, contrary to previous systems with VAR instruments, $\partial Eg_{1t}^S/\partial\alpha_S$ is no longer identically equal to zero. The test statistic is then defined as:

$$VE_4 = (\hat{\Theta}_S - \hat{\Theta}_D)'[(T_S + R)(W_o^S)^{-1}(T_S + R)']^{-1}(\hat{\Theta}_S - \hat{\Theta}_D)$$

which has the same asymptotic distribution as VE_2 and VE_3.

Several VAR Encompasing tests have been discussed now which have more or less the same asymptotic distribution. As with the Chow test more research needs to be done to learn more about their properties (see Ghysels and Hall (1987b)). A few additional observations should be made before leaving the subject, however. First, it should be noted that the "spirit" of the VAR Encompassing test is very similar to an approach put forward by Sims (1984) for solving nonlinear stochastic equilibrium models "backwards". Sims in fact suggests, among other things to "estimate by simulation". The latter is a strategy consisting of generating solution paths of a nonlinear stochastic equilibrium model in order to have a long series of simulated data to estimate VAR models. The VAR obtained this way is then compared to a VAR model estimated with actual data. Our procedure is, in comparison computationally more efficient since it does not involve solving Euler conditions. Sims also points out that this approach is linked to the calibration techniques used in Prescott and Kydland (1982), among others. The VAR Encompassing tests presented here may also be considered in the same "spirit" as calibration techniques but are statistically more sophisticated. Second, and finally, it should also be noted that VAR representations are used as a vehicule to accomplish model validation tests. It is not the purpose, however, to find an "economic" interpretation for VAR estimates as Sims (1980) and others have suggested. One could argue that VAR estimates "contaminated" by the presence of Euler conditions could serve a similar purpose, however.

4. SOME MACROECONOMIC APPLICATIONS

Several applications of the model validation tests presented here will be presented in part II of the paper. There will be two models in particular which will be to explored. The first is a real business cycle model, the second a monetary equilibrium business cycle model. The purpose is not so much to develop new models, but instead to focus on existing models where certain structural parameters play a prominent role in testing certain theoretical equilibrium business cycle models.

The first model presented in part II is a representative agent model with non-time-separable preferences, introduced by Eichenbaum, Hansen and Singleton (1986), henceforth EHS, which can accommodate equilibrium laws of motion for labor supply, consumption, real wages and stochastic interest rates. The non-time-separability introduces nontrivial endogenous sources of dynamics which according to some real business cycle proponents, e.g. Kydland and Prescott (1982) and Kydland (1983), are an important ingredient in explaining the co-movements in aggregate compen- sation and hours worked. The Chow-type of test is particularly useful for addressing two questions, namely (1) to what extent will the assumptions to rationalize a representative agent in the presence of heterogeneous labor supply affect the interpretation of parameters as structural and (2) to what extent are the key parameters which were found to be significantly different from the values assumed in several business cycle models be invariant across subsamples. The second test, the VAR Encompassing tests are also appropriate in the context of the real business cycle model presented by EHS. In many cases results from VAR models have been given a real business cycle interpretation. Examples are Sims (1980a, 1980b) who interpreted the contribution of nominal interest rates in predicting industrial production as capturing expectations about the future productivity of capital, which is in the spirit of RBC analysis. Litterman and Weiss (1985) also give an RBC interpretation to their VAR model results. Moreover, Eichenbaum and Singleton (1986) also link VAR evidence to the validation of RBC models. The idea that stylized facts reported via VAR parameter estimates should drive the search for an appropriate equilibrium business cycle model can be formulated in a formal and testable way via the VAR Encompassing tests.

The second model studied in part II is a monetary consumption asset price model. The money-output linkage is introduced via a cash-in-advance constraint very much in the spirit of theoretical models studied by Lucas (1980, 1984), Lucas and Stokey (1984), Svensson (1985), Townsend (1987) among others. The multi-good models of asset pricing are analysed which were pursued empirically by Dunn and Singleton (1985), Eichenbaum and Hansen (1985) who considered extensions of the model in Hansen and Singleton (1982). The model specifically concentrated on is presented in Singleton (1986) where a non-durable versus durable goods asset pricing model with a Clower constraint is derived and estimated. Key parameters in such models govern substitutability of consumption (durables, nondurables, services) across goods and over time. The Chow test is again relevant to uncover the potential misspecification due to aggregation of heterogeneous consumers. It is also useful to test invariance with regard to changes in policy regime by taking the October 1979 monetary policy shift as a benchmark. Since money plays a determining role in real allocations in the model it is natural to use the Chow-type test in this context. The VAR encompassing tests will also be useful tests to uncover the potential failure of equilibrium monetary asset pricing models.

APPENDIX

Proof of theorem 2.4: Based on Hansen's proof of his theorem 3.1 one has the following result:

$$\hat{\beta}_1 - \beta_1 = -[A^*_{n_1} Dg_{n_1}]^{-1} A^*_{n_1} g_{n_1}(\beta_1) \tag{A.1}$$

$$\hat{\beta}_2 - \beta_2 = -[A^*_{n_2} Dg_{n_2}]^{-1} A^*_{n_2} g_{n_2}(\beta_2) \tag{A.2}$$

where $\hat{\beta}_i$ minimizes $g_{n_i}(\beta)' \; W_{n_i} \; g_{n_i}(\beta)$ and

$$Dg_{n_i} = \partial g_{n_i}(\beta)/\partial\beta'$$

$$W_{n_i} = A'_{n_i} \; A_{n_i}$$

$$A^*_{n_i} = (Dg_{n_i})' \; A'_{n_i} \; A_{n_i}$$

for $i = 1, 2$ and:

$$g_{n_1}(\beta) = n_1^{-1} \sum_{t=k-n_1+1}^{k} f_t(\alpha_t, \beta)$$

$$g_{n_2}(\beta) = n_2^{-1} \sum_{t=k+1}^{k+n_2} f_t(x, B)$$

Under Hansen's regularity conditions, it follows that:

$$\hat{\beta}_1 - \hat{\beta}_2 \xrightarrow{d} N(\beta_1 - \beta_2, V) \tag{A.3}$$

The covariance matrix of the asymptotic distribution is:

$$V = [A_1^* D_1]^{-1} A_1^* S_1 A_1^{*'} \{[A_1^* D_1]^{-1}\}' + [A_2^* D_2]^{-1} A_2^* S_2 (A_2^*)'$$
$$\times \{[A_2^* D_2]^{-1}\}' + [A_1^* D_1]^{-1} A_1^* S_{12}(A_2^*)' ([A_2^* D_2]^{-1})'$$
$$+ [A_2^* D_2]^{-1} A_2^* S_{21}(A_1^*)' ([A_1^* D_1]^{-1})' \qquad \text{(A.4)}$$

Where $A_i^* = \text{plim } A_{n_i}^* \qquad n_i \to \infty \qquad i = 1,2$

$D_i = \text{plim } Dg_{n_i} \qquad n_i \to \infty \qquad i = 1,2$

$S_i = E[g_{n_i} g_{n_i}'] \qquad i = 1,2$

$S_{ij} = E[g_{n_i} g_{n_j}'] \qquad i,j = 1,2 \quad i \neq j$

Under H_o appearing in (2.2a) the following holds:

$A_1^* = A_2^* \qquad S_1 = S_2$

$D_1 = D_2 \qquad S_{12} = S_{21}$

Hence, under H_o equation (2.4) in theorem 2.4 is valid, namely:

$$V = 2\{ [A^*D]^{-1} A^* [S+\bar{S}] (A^*)' [A^*D]^{-1'}\} \qquad \text{(A.5)}$$

where $A_1^* = A_2^* = A^*$, $S = S_1 = S_2$, $D = D_1 = D_2$ and $\bar{S} = S_{12} = S_{21}$.

Equation (2.3) gives the expression for $\bar{S}$ assuming $\beta_i = \beta$; $i=1,2$. In the remainder of the proof, we derive the properties of $\bar{S}$. First, for arbitrary finite n_1, n_2 we have:

$$\bar{S} = (n_1 n_1)^{-\frac{1}{2}} \sum_{s=1}^{n-1} w_s \, E \left[f(x_t, \beta) \, f(x_{t+s}, \beta)' \right]$$

Where $n = n_1 + n_2$

$m = \min \{n_1, n_2\}$

$w_s = s \qquad 1 < s < m$

$w_s = n \qquad m < s < n - m$

$w_s = n_2 - s \qquad n - m < s < n - 1$

Let s* be a finite integer less than n. Then we can write:

$$S = (n_1 n_2)^{-\frac{1}{2}} \left\{ \sum_{s=1}^{s^*} w_s E \left[f(x_t, \beta) \, f(x_{t+s}, \beta)' \right] + \sum_{s=s^*+1}^{n-1} w_s E \left[f(x_t, \beta) \, f(x_{t+s}, \beta)' \right] \right\}$$

As $n_1 n_2$ go to infinity:

$$\lim_{n_1 n_2 \to \infty} \bar{S} = \lim_{n_1 n_2 \to \infty} \left\{ (n_1 n_2)^{\frac{1}{2}} \sum_{s=s^*+1}^{n-1} w_s E \left[f(x_t, \beta) \, f(xl_{t+s}, \beta)' \right] \right\}$$

Since $w_s / (n_1 n_2)^{\frac{1}{2}} < 1$ one has

$$\lim_{n_1 n_2 \to \infty} \left| \bar{S}_{ij} \right| < \lim_{n_1, n_2 \to \infty} \sum_{s=s^*+1}^{n-1} \left| E \left[f_{it} f_{jt+s} \right] \right| = \bar{S}^o_{ij} \qquad \forall \, i$$

and $\bar{S}_{ij}$ is the i-j th element of $\bar{S}$

Under assumption 2.1 $\bar{S}^o_{ij} = o \; \forall i_1 j$. This follows from the definition of an M-dependent process. The mixing conditions presented under assumption 2.2 imply that $\bar{S}^o_{ij}$ is finite $\forall i,j$. To see this the following steps are needed (1) from White and Domowitz (1984) we have that if x_t is mixing of size b, then $f(x_t, \beta)$ is also mixing of that size, furthermore (2) White and Domowitz's results also imply that the autocovariances of x_t decay at this rate, hence $E[f_{it} f_{jt+s}]$ is $O(s^{-\lambda})$ where size in his context stands for $O(n^{-\lambda})$ for $\lambda > 2r/(2r-1)$ for ϕ (n) mixing and $O(n^{-\lambda})$ for $\lambda > 2r/(r-1)$ for α (n) mixing (see assumption 2.2). Finally, the convergence of $\bar{S}^o_{ij}$ to a finite limit as $n_1, n_2 \to \infty$ follows directly from Gallant and White (1986) corollary 3.11.

Given that $E(f_{it} f_{jt+s^*+1})$ is $O((s^*+1)^{-\lambda})$ and $E(f_{it} f_{jt+s^*+k})$ is a decreasing function of k, it may reasonably be expected that $\bar{S}$ is small relative to the variance S. Hence, we suggest to perform "conservative inference" if x_t is mixing as in White (1983) by constructing

$$CV = \{ [A^* D]^{-1} A^* S (A^*)' ([A^* D]^{-1})'$$

Note that as $\bar{S}$ is symmetric, (V - CV) is positive semi-definite. Furthermore, since CV is positive definite and we assume that V is positive definite, it follows that $CV^{-1} - V^{-1}$ is positive semi-definite. Then

$$(\hat{\beta}_1 - \hat{\beta}_2)' CV^{-1} (\hat{\beta}_1 - \hat{\beta}_2) - (\hat{\beta}_1 - \hat{\beta}_2)' V^{-1} (\hat{\beta}_1 - \hat{\beta}_2) < 0$$

Hence the test statistic; using CV, is always at least as large as the true statistic with V. Consequently, rejecting with the CV matrix certainly means rejecting with the matrix V as covariance matrix. The converse is of course not true, i.e., accepting with CV does not mean that the null hypothesis is accepted with V. Q.E.D.

Proof of theorem 3.1: The estimator $\hat{\Theta}_V$ is obtained by minimizing the quadratic form:

$$\underset{\Theta}{\text{Min}} \ (g_n^V(\Theta))' \ W_n^V \ g_n^V(\Theta)$$

where $W_n^V \rightarrow W^V$ almost surely. This is asymptotically equivalent to solving:

$$\underset{\Theta}{\text{Min}} \ (g_n^V(\Theta))' \ W^V \ g_n^V(\Theta)$$

The first-order conditions for this quadratic form are:

$$(\partial g_n^V(\Theta)'/\partial\Theta) \ W^V \ g_n^V(\Theta) = 0$$

The asymptotic distribution of the solution to the first-order conditions will be the same as the asymptotic distribution of the solution to:

$$(\partial E g_t^V(\Theta_o)'/\partial\Theta) \ W_o^V \ g_n^V(\Theta) = 0$$

where

$$W_o^V = (E[g_t^V(\Theta_o) \ g_t^V(\Theta_o)'])^{-1}$$

Since the Euler conditions imply that:

$E[\ell_{1t}(x_{t+k}, \beta_o) \otimes Z_t]$ for any $Z_t \varepsilon I_t$ it follows that:

$$E[g_{1t}^V(\Theta_o) \ g_{2t}^V(\alpha_o)'] = 0$$

because $g_{2t}^V(\alpha_o) \ \varepsilon \ I_t$. Hence W_o^V is the inverse of a block-diagonal matrix. Consequently:

$$W_o^V = \begin{bmatrix} W_{10}^V & 0 \\ 0 & W_{20}^D \end{bmatrix}$$

Since

$$(E[g_{2t}^V(\alpha_o) \ g_{2t}^V(\alpha_o)'])^{-1} = W_{20}^D$$

Furthermore, the gradient of the g_t^V function can be written as:

$$\partial Eg_t^V(\Theta_o)/\partial\Theta' = \begin{bmatrix} \partial Eg_{1t}^V(\Theta_o)/\partial\beta' & \partial Eg_{1t}^V(\Theta_o)/\partial\alpha' \\ 0 & \partial Eg_{2t}^D(\alpha_o)/\partial\alpha' \end{bmatrix}$$

The Euler conditions yield the following result:

$$\partial Eg_{1t}^V(\Theta_o)/\partial\alpha' = Ee_{1t}(x_{t+k}, \beta_o) \otimes (\partial A^*(L)\ X_{1t}/\partial\alpha') = 0$$

Hence the first-order conditions can be written as:

$$\partial Eg_{1t}^V(\Theta_o)/\partial\beta'\ W_{10}^V\ g_{1n}^V\ (\Theta) = 0$$

$$\partial Eg_{2t}^D\ (\alpha_o)/\partial\alpha' W_{20}^D\ \ g_{2n}^D(\alpha) = 0$$

The latter yields the asymptotic equivalence of $\hat{\alpha}_D$ and $\hat{\alpha}_V$. Finally, it should be noted that:

$$(W_{10}^V)^{-1} = E\ g_{1t}^V\ (\Theta_o)\ g_{1t}^V\ (\Theta_o)'$$

Because of the definition of the g_{1t}^V function, it follows that

$$(W_{10}^V)^{-1} = (W_{10}^D)^{-1} + W_D^V$$

where $W_D^V \neq 0$ Q.E.D.

Proof of theorem 3.3: The proof is along the same lines as theorem 3.1's proof. The weighting matrix is not block-diagonal because the covariance between g_{1t}^N and g_{2t}^N is non-zero. It should be noted, however, that

$$\partial Eg_{1t}^N(\Theta_o)/\partial\alpha' = 0$$

because $\partial A^*(L)\ X_{1t}/\partial\alpha'$ remains a function defined on the information set. The overidentifying restrictions assumption is necessary for the first-order conditions to hold without $g_n^N \equiv 0$ at $\hat{\Theta}_N$.

FOOTNOTES

[1] Besides conditions for exact aggregation some recent work by Polemarchakis et al (1986) has focused on approximate aggregation. Their goal is to investigate the degree to which the demand of a collection of von Neumann-Morgenstern agents can be approximated by the demand of a single von Neumann-Morgenstern agent. One of the results states that the risk tolerance of the approximate aggregator is equal to the sum of the individual agent risk tolerances at prices which yield constant, "riskfree", contingent consumption. This and other results show how parameters in a GMM procedure, where exact aggregation conditions do not hold, can depend on the distribution of preferences and income.

[2] It should parenthetically be noted that this procedure does not correspond to the equation-by-equation OLS estimation of a VAR as is usually the case.

REFERENCES

Anderson, T.W., 1958, Introduction to Multivariate Statistical Analysis, (Wiley, New York).

Dunn, K. and K.J. Singleton, 1986, Modeling the Term Structure of Interest Rates Under Nonseparable and Durable Goods, Journal of Financial Economics.

Eichenbaum, M.S., Hansen L.P., 1985, Estimating Models with Intertemporal Substitution using Aggregate Time Series Data, manuscript, Carnegie-Mellon University.

Eichenbaum, M.S., Hansen L.P. and K.J. Singleton, 1986, A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty, N.B.E.R. Working Paper 1981.

Eichenbaum, M.S. and K.J. Singleton, 1986, Do Equilibrium Real Business Cycle Theories Explain Post-War U.S. Business Cycles? N.B.E.R. Working Paper 1932.

Gallant, A.R. and G. Tauchen, 1986, Seminonparametric Estimation of Conditionally Constrained Heterogenous Processes: Asset Pricing Applications, Paper presented at the 1987 Summer Meeting of the Econometric Society, 1987.

Gallant, A.R. and H. White, 1986, A Unified Theory of Estimation and Inference for Non-linear Dynamic Models, (Basil Blackwell, Oxford).

Ghysels, E. and A. Hall, 1986, Testing Non-Nested Euler Conditions with Quadrature-Based Methods of Approximation, Paper to be presented at the 1987 Summer Meeting of the Econometric Society, Berkeley.

Ghysels, E. and A. Hall, 1987a, Some Additional Specification Tests for GMM Estimators with Macroeconomic Applications - Part II: Applications, Work in Progress.

Ghysels, E. and A. Hall, 1987b, Properties of GMM estimators and tests in artificically generated economies not necessarily corresponding to the specified moment restrictions (work in progress).

Gorman, W.M., 1953, Community Preference Fields, Econometrica 21, 63-80.

Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, Econometrica 50, 1029-1054.

Hansen, L.P., 1985, A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators, Journal of Econometrics 30, 203-238.

Hansen, L.P., 1986, Comments on "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data" by George Tauchen, Journal of Business and Economic Statisitcs 4, 418-421.

Hansen, L.P. and T.J. Sargent, 1980, Formulating and Estimating Dynamic Linear Rational Expectations Models, Journal of Economic Dynamics and Control 2, 7-46.

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

Kydland, F.E., 1983, Nonseparable Utility and Labor Supply, Working Paper E-83-16, Hoover Institution, Standford University.

Kydland, F.E., and E.C. Prescott, 1982, Time to Build and Aggregate Fluctuations, Econometrica 50, 1345-1370.

Litterman, R. and L. Weiss, 1985, Money, Real Interest Rates, and Output: A Re-interpretation of Postwar U.S. Data, Econometrica 53, 129-156.

Lucas, R.E. Jr., 1980, Equilibrium in a Pure Currency Economy, Economic Inquiry 18, 203-220.

Lucas, R.E. Jr., 1984, Money in a Theory of Finance, in K. Brunner and A. Meltzer (eds.), Essays on Macroeconomic Implications of Financial and Labor Markets and Political Processes, Carnegie-Rochester Conference Series on Public Policy 21, 9-46.

Lucas R.E. Jr., and N.L. Stokey, 1984, Money and Interest in a Cash-in Advance Economy, Econometrica 55, 491-515.

Mizon, G. and J.F. Richard, 1986, The Encompassing Principle and its Application to Testing Non-Nested Hypotheses, Econometrica 54, 657-678.

Newey, W.K. 1985, Generalized Method of Moments Specification Testing, Journal of Econometrics 29, 229-256.

Piessens, R. et al., 1983, QUADPACK, A Subroutine Package for Automatic Intregration, (Springer-Verlag, Berlin).

Polemarchakis, H.M., L. Selden, P. Zipkin and L. Pohlman, 1986, Approximate Aggregation under Uncertainty, Journal of Economic Theory 38, 189-210.

Sims, C.A., 1980a, Comparison of Interwar and Postwar Business Cycles: Moneterism Reconsidered, American Economic Review 70, 250-257.

Sims, C.A., 1980b, Macroeconomics and Reality, Econometrica 48, 1-48.

Sims, C.A., 1984, Solving Nonlinear Stochastic Equilibrium Models "Backwards", Manuscript, University of Minnesota.

Singleton, K.J., 1986, Testing Specifications of Economic Agent's Intertemporal Optimum Problem Against Non-Nested Alternatives, Journal of Econometrics 30, 391-413.

Svensson, L.E.O., 1985, Money and Asset Prices in a Cash-in Advance Economy, Journal of Political Economy 93, 919-944.

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

Tauchen, G., 1986b, Quadrature Based Methods of Approximating the Solutions to the Integral Equations of Nonlinear Rational Expectations Models, unpublished manuscript, Duke University.

Townsend, R. 1987, Asset Return Anomalies in a Monetary Economy, Journal of Economic Theory (forthcoming).

Volterra, V., 1959, Theory of Functionals and of Integro-Differential Equations (Dover, New York).

White, H. and I. Domowitz, 1984, Nonlinear Regression with Dependent Observations, Econometrica 52, 143-162.

White, H. Corrigendum, Econometrica 51, 513.