NBER TECHNICAL PAPER SERIES

ESTIMATION AND HYPOTHESIS TESTING WITH RESTRICTED SPECTRAL DENSITY MATRICES: AN APPLICATION TO UNCOVERED INTEREST PARITY

Danny Quah

Takatoshi Ito

Technical Working Paper No. 50

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1985

The authors are indebted to Thomas Sargent for suggesting the approach, and to Mark Watson for conversations that have influenced the course of the discussion in this paper. An earlier version of this paper was presented at the Econometric Society Meetings in Dallas, December 1984. The first author thanks Robert Litterman and Thomas Sargent for permitting the use of results obtained in an ongoing joint research project. The second author is grateful for financial support from the National Bureau of Economic Research (the U.S. - Japan Interdependence Project) and the Hoover Institution (National Fellowship). The research reported here is part of the NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Technical Paper #50
August 1985

Estimation and Hypothesis Testing with Restricted Spectral Density Matrices: an Application to Uncovered Interest Parity

ABSTRACT

This paper explores an econometric estimation technique for dynamic linear models. The method combines the analytics of moving average solutions to dynamic models together with computational advantages of the Whittle likelihood. A hypothesis of interest to international and financial economists is represented in the form of cross-equation restrictions and tested under the technique. This paper employs data on Japanese yen- and U.S. dollar-denominated interest rates and yen/dollar exchange rates to examine the hypothesis of uncovered interest parity under rational expectations.

Danny Quah
Harvard University
Department of Economics
Littauer Center 200
Cambridge, MA 02138

Takatoshi Ito
National Bureau of
Economic Research
204 Junipero Serra Boulevard
Palo Alto, CA 94305

1. INTRODUCTION

The purpose of this paper is two-fold. First, it explores an econometric estimation technique for dynamic linear models. Secondly, a hypothesis of interest to international and financial economists is represented in the form of cross-equation restrictions and tested under the technique. This paper employs data on Japanese yen- and U.S. dollar-denominated interest rates and yen/dollar exchange rate to examine the hypothesis of uncovered interest parity under rational expectations.

Linear models form a convenient framework to examine the explicit dynamics of various macroeconomic and expectational hypotheses. Time domain vector autoregression models, in particular, have been widely used to obtain explicit empirical representations of the dynamics implied by distinct economic hypotheses. (See, among others, Sims (1980a,b, 1982).) This paper explores similar ideas using a frequency domain technique to articulate and examine cross-equation restrictions. These are often the interesting observable implications of hypotheses regarding market behavior and expectations formation.

Sargent (1979) and Ito (1984) used the vector autoregression representation as a framework for testing hypotheses. The former examined the rational expectations hypothesis with regard to the term structure of interest rates; the latter tested the hypothesis of uncovered interest parity relevant to open economy finance. In both of those earlier papers, the testable restrictions implied by the hypotheses of interest took the form of cross-equation restrictions on a system of stochastic difference equations. Since reasonable theories imply constraints on the *interdependence* of different variables, and rarely on the dynamic properties of any single variable, it is almost always the case that system methods are desirable for examining the validity of these hypotheses.

Further, the optimal predictor of a process is explicitly calculable when the model takes a linear form. As expectations formation remains an issue in many different research areas, this is an attractive property of any empirical implementation. For these reasons and others, we believe a strong case can be made for the usefulness of adopting the approach we have taken here.

The procedure used in this paper has certain advantages over conventional time domain vector autoregression methods. The lag structures that are feasible in estimation are somewhat more general than those permitted in time domain work. This is not emphasized in what follows because in the absence of *a priori* knowledge regarding the appropriateness of any one lag distribution, we do not believe that simplicity is necessarily undesirable. In the empirical work below however, we do estimate a matrix rational lag structure that is not easily implementable in purely time domain work.

A closely related point is that we show that a closed form representation for the equilibrium of a model often is not necessary for the application of " full information " methods of estimation. Thus, the method here shares some of the same *computational* advantages as instrumental variable estimation in Generalized Method of Moments applications to rational expectations models.

Uncovered interest parity is the hypothesis that the expected relative depreciation of a domestic currency is equal to the differential in interest rates between two countries. Ito (1984) expressed that hypothesis in terms of cross-equation restrictions on a vector autoregressive representation of the exchange rate and domestic and foreign interest rates. That earlier work also tested that restricted representation of the dynamics of the yen/dollar exchange rate, the Japanese domestic interest rate and the Euro-dollar interest rate. This paper tests the same hypothesis within a more flexible parametrization afforded by the frequency domain approach. It is an advantage of the procedure that it is easily modified for application to data sampled finely in time. This is exploited below in the estimation using weekly data.

The rest of the paper is organized as follows. The uncovered interest parity relation is developed as a testable hypothesis using optimal prediction formulas for processes with rational spectral densities in Section 2. Section 3 demonstrates the validity of the simplifying computational techniques used here within the framework of a general discussion of identification subject to, and the statistical testing of certain familiar key cross-equation restrictions. Section 4 then makes explicit the exact nature of the deviation from exact likelihood techniques, in small sample, of the method that is used here. Section 5 makes concrete the claim that estimating the model without calculating the closed form does not prevent the analyst from obtaining correct representations for the dynamic properties of the model. This is important for the proper simulation of, and for the calculation of forecasts using such models. As should be clear in what follows, the results here are broadly applicable to the empirical analysis of a general class of models. An important class of applications naturally lies in studies of financial markets that take expectations to be central in the analysis. Section 6 presents empirical results and Section 7 concludes the paper.

2. TESTABLE RESTRICTIONS IMPLIED BY UNCOVERED INTEREST ARBITRAGE

In this section, we develop the restrictions implied by the relation of uncovered interest parity. Consider forward contracts for ν-period ahead foreign exchange transactions, with the observation interval normalized to one period. As usual, s_t is the logarithm of the observed current spot rate (the value of foreign currency in terms of domestic currency), $r_t(\nu)$ and $r_t^*(\nu)$ are respectively the domestic and foreign ν-period interest rates observed at time t. Uncovered interest parity holds if the expected depreciation of the currency is equal to the observed interest rate spread. To put it another way, uncovered interest parity holds when the expected yields for comparable assets denominated in different currencies are equal. The hypothesis is succinctly stated as

$$\widehat{E}\left[s_{t+\nu} \,\middle|\, H_y^-(t) \right] = s_t + r_t(\nu) - r_t^*(\nu)$$

where $\{ y_t \}$ is some vector random sequence with finite second moments, and $y_t, t \leq N$ is assumed to be observed at time N. We denote by $H_y^-(t)$ the Hilbert space spanned by infinite sequences of linear combinations of the random variables $\{ y_u, u \leq t \}$ complete under mean square norm. The symbol $\widehat{E}$ denotes linear least squares projection. We hold ν fixed and assume that $x_t = \left[s_t, r_t(\nu), r_t^*(\nu) \right]^T$ is contained in y_t. By this assumption $H_x^-(t)$ is contained in $H_y^-(t)$, that is, economic agents are assumed to have at least as much information as that generated by observations on exchange and interest rates. To summarize the above, we make operational the uncovered interest arbitrage hypothesis by assuming agents form expectations using optimal prediction formulas, and that the spot exchange, and domestic and foreign interest rates are publicly observed.

In the absence of capital controls, arbitrage between assets with different currency denomination forces *covered* interest parity :

$$s_t + r_t(\nu) - r_t^*(\nu) = f_t(\nu),$$

where $f_t(\nu)$ is the forward exchange rate. Combining these 2 relations, we have

$$\hat{E}\left[s_{t+\nu} \middle| H_y^-(t)\right] = f_t(\nu)$$

that is, the forward rate is the rationally expected future spot rate. All information available at period t that may be optimally used in predicting the future spot rate is embodied in the movements of the forward rate.

This observable implication has been extensively studied in the literature. (See for example, Cumby and Obstfeld (1984) and Hansen and Hodrick (1980) and the references in their papers.) A single equation method has been favored in the literature. The unexpected change in the spot rate is regressed on elements identified with current and lagged information: the appropriate test becomes one of orthogonality. When the observation period is finer than the length of the forward contract (that is, ν exceeds 1), the error term is known to have a finite order moving average structure. Hansen and Hodrick (1980) and Cumby and Obstfeld (1984) have implemented solutions to this nonstandard inference problem in regression analysis.

In this paper, we estimate the dynamic structure of these variables directly, so that the moving average properties are jointly estimated with all the other parameters of the restricted representation. If one were only interested in tests of the restrictions *per se*, one might proceed as in Ito (1984), where the unrestricted vector autoregression is estimated by ordinary least squares and a Wald-type test is employed in examining the validity of the hypothesis. Here, however, we are also concerned with exploring the explicit dynamic properties associated with the restricted representation. There are reasons to examine statistical representations of economic data even if the economic hypotheses used to obtain those parametrizations are rejected at normal (classical) levels of significance. For now, we continue with the formal development.

If $\{ y_t \}$ is covariance stationary and linearly regular, it possesses a Wold moving average representation that is time invariant:

$$y_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}, \text{ where } E\varepsilon_t\varepsilon_{t-u} = I\,\delta(u).$$

For notational simplicity define the vectors

$$a_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix},\ b = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Then the parity relation above can alternatively be expressed as

$$\left[a_1 z^{-\nu} C(z) \right]_+ = bC(z)$$

when the determinant of $C(z)$ has no zeroes inside the unit circle, and the equality is taken in the sense of element-by-element equality as analytic functions. The restrictions above are a special case of the characteristic cross-equation restrictions that arise in a wide class of models. These will be studied more carefully in the next sections.

For now we need to express these restrictions in a form convenient for estimation. First notice that there are no restrictions across all but the first 3 rows of the matrix function $C(z)$. Call these rows $\gamma_j(z)$, $j = 1,2,3$. Then the restrictions are exactly

$$\left[z^{-\nu}\gamma_1(z) \right]_+ = \gamma_1(z) + \gamma_2(z) - \gamma_3(z).$$

Hence imposing the cross-equation restrictions in estimation is achieved by constructing the matrix function $C(z)$ with

$$\gamma_3(z) = \gamma_1(z) + \gamma_2(z) - \left[z^{-\nu} \gamma_1(z) \right]_+$$

and using this constrained $C(z)$ in building the spectral density matrix $C(e^{-i\omega})C(e^{-i\omega})^*$.

Note that estimation of the constrained model under the null hypothesis may be performed for a variety of configurations of forward contract lengths and sampling intervals simply by altering the value of ν in the above expression. This corresponds to using alternative principal roots of unity in evaluating a frequency domain approximation to the (quasi-)likelihood of the given model. The ease with which one may adapt the given estimation routine to different timing intervals reflects an oft-quoted adage about the use of time and frequency domain procedures. The reader should compare the manipulations here to those in Sargent (1979), Hakkio (1981), Baillie *et al* (1983) and Ito (1984), who estimate similar models in the time domain.

The trick here of building up the spectral density matrix using the cross-equation restrictions is a useful one for estimation. The reader who is uneasy about the propriety of such a mechanism for the analysis of these models is correct in the conjecture that some aspects of the dynamics are lost in the construction.[1] However the following sections demonstrate that in this instance all of these aspects are unnecessary in the *estimation*.

To continue with the development, it is useful to present a method for the explicit computation of the Wiener-Kolmogorov expressions for processes with rational spectral density.

Consider the random sequence $\{ \xi_t \}$ possessing a rational spectral density. By factoring the spectral density, we may represent

---

[1]In particular, the notion that " stable roots are solved backwards, unstable roots forwards " in optimizing models seems to be ignored here.

$$\xi_t = \frac{\alpha(L)}{\beta(L)} \eta_t \text{ , where } E\ \eta_t \eta_{t-u} = 0 \text{ for } u \neq 0,$$

and $\frac{\alpha(z)}{\beta(z)}$ is a rational function with the common factors in $\alpha(z)$ and $\beta(z)$ removed. Further the function $\frac{\alpha}{\beta}$ is analytic in a domain containing the unit circle and has no zeroes inside the unit circle. We wish to calculate $\hat{E}\left[\xi_{t+\nu} \,\middle|\, H_\xi^-(t)\right]$. This is simply the optimal linear least squares forecast of ξ ν periods from t conditional on the information in current and past realizations of ξ. We know by our assumptions that

$$\hat{E}\left[\xi_{t+\nu} \,\middle|\, H_\xi^-(t)\right] = \hat{E}\left[\xi_{t+\nu} \,\middle|\, H_\eta^-(t)\right] = \left[L^{-\nu} \frac{\alpha(L)}{\beta(L)}\right]_+ \eta(t).$$

To calculate $\left[L^{-\nu} \frac{\alpha(L)}{\beta(L)}\right]_+$, we consider $\left[z^{-\nu} \frac{\alpha(z)}{\beta(z)}\right]_+$, which is that analytic function obtained from $z^{-\nu} \frac{\alpha(z)}{\beta(z)}$ by removing its principal or singular component. By the assumptions on α and β , the function $\frac{\alpha}{\beta}$ has a convergent one-sided Laurent series expansion in nonnegative powers of z about 0 :

$$\frac{\alpha(z)}{\beta(z)} = \sum_{j=0}^{\infty} \gamma_j z^j.$$

Notice that the function $z^{-\nu} \frac{\alpha(z)}{\beta(z)}$ has a removable pole of order ν at the origin. This is also the only singularity of the function inside the unit circle. The singular part of $z^{-\nu} \frac{\alpha(z)}{\beta(z)}$ at the origin is therefore given simply by $\sum_{j=0}^{\nu-1} \gamma_j z^{j-\nu}$. Then

$$\left[z^{-\nu} \frac{\alpha(z)}{\beta(z)}\right]_+ = z^{-\nu} \frac{\alpha(z)}{\beta(z)} - \sum_{j=0}^{\nu-1} \gamma_j z^{j-\nu}$$

$$= z^{-\nu} \left[\frac{\alpha(z)}{\beta(z)} - \sum_{j=0}^{\nu-1} \gamma_j z^j\right]$$

$$= \frac{z^{-\nu}}{\beta(z)} \left[\alpha(z) - (\sum_{j=0}^{\nu-1} \gamma_j z^j) \beta(z) \right].$$

Since we assume ξ_t has a spectral density that is rational, α and β may be chosen to be polynomials in nonnegative powers of z. Therefore

$$\left[z^{-\nu} \frac{\alpha(z)}{\beta(z)} \right]_+ = \frac{z^{-\nu}}{\beta(z)} \left[\alpha(z) - (\sum_{j=0}^{\nu-1} \gamma_j z^j) \beta(z) \right]$$

remains a rational function possessing a convergent one-sided Laurent series expansion at the origin.

To state these results differently, notice that we can write the decomposition

$$\frac{\alpha(z)}{\beta(z)} = \sum_{j=0}^{\infty} \gamma_j z^j \implies \alpha(z) = (\sum_{j=0}^{\nu-1} \gamma_j z^j) \beta(z) + (\sum_{j=\nu}^{\infty} \gamma_j z^j) \beta(z).$$

Now note that the second term on the right hand side $(\sum_{j=\nu}^{\infty} \gamma_j z^j) \beta(z)$ has no nonzero coefficient on any power of z strictly less than ν. This implies that the first $\nu - 1$ coefficients of $\alpha(z)$ match exactly the first $\nu - 1$ coefficients of $(\sum_{j=0}^{\nu-1} \gamma_j z^j) \beta(z)$.

Hence in forming $z^{-\nu} \left[\alpha(z) - (\sum_{j=0}^{\nu-1} \gamma_j z^j) \beta(z) \right]$, all coefficients on strictly negative powers of z are guaranteed to vanish.

The outcome of the discussion in this section is that when the Wiener-Kolmogorov formula is applied to a process with rational spectral density, the calculation is easily performed using only polynomial long division and multiplication.

In the next sections of the paper, we show precisely the necessary results that make the proposed estimation technique an analytically and computationally attractive one.

3. REMARKS ON IDENTIFICATION AND TESTING

In this section, we address identification in the general model amenable to the estimation procedure. A slight generalization of the framework employed in the rest of the paper is useful for the discussion. The modification is chosen to allow the reader to compare the results here with those in Hansen and Sargent (1981). While some facts established in that paper are repeated here, the statements and proofs are sufficiently different to motivate our presentation.

There are two identification issues relevant for the kinds of models considered in the current work. The first arises only in the multivariate case; the second relates to the notion of a *fundamental representation* for dynamic models.

3.1 Orthogonalization

The identification issue specific to multivariate timeseries models also surfaces in the guise of orthogonalization orders for interpreting vector autoregressions. It is perhaps best understood in that context. When a relatively unrestricted representation for a multivariate time series model is analyzed, the vector of residuals correspond to different shocks that impinge on the system. Without *a priori* identifying restrictions, the residuals for each equation do not necessarily represent shocks to the variable on the left hand side of that equation. Researchers employ different *a priori* assumptions to identify individual elements of the residual vector with recognizable economic shocks.

Such association however is not necessary for the purpose of testing the cross-equation restrictions of interest in the current work. We do need to note exactly the restrictions imposed by the economic hypotheses over and above those used to just-identify a statistical alternative. Proposition 1 below states precisely the characteristics of such a just-identified alternative. This well-known result is employed either implicitly or explicitly in any discussion of orthogonalization orders for relatively unrestricted vector autoregressions.

For ease of exposition we first make exact some terms to be used in what follows. An array of rational functions in the complex indeterminate z is a *matrix rational function.* Unless otherwise noted, we consider only matrix rational functions that are square arrays. When we take series expansions for each element of an array, the resulting array is called the series expansion for that matrix function. A square matrix rational function is *regular* if it has full rank almost everywhere. A regular matrix rational function is *fundamental* if it is analytic and vanishes nowhere in an open domain containing the unit disk.

We can now state Proposition 1.

Proposition 1 Let $S(\omega)$, ω in $(-\pi,\pi]$, be a given rational spectral density matrix that has full rank almost everywhere. Then there exists a factorization

$$S(\omega) = C(e^{-i\omega})C(e^{-i\omega})^*$$

where $C(z)$ is fundamental. Further, if $C(0)$ is chosen to be real and lower triangular with all positive elements on the diagonal, then the factorization is unique.

Proof The first part of the proposition is simply the classic Spectral Factorization Theorem.[2] There exists a decomposition

$$S(\omega) = F(e^{-i\omega})\ G\ F(e^{-i\omega})^*$$

where $F(z)$ is fundamental and G is real symmetric positive definite. Further

[2]Key references for the multivariate Spectral Factorization Theorem are Rozanov (1960, pp.368-374; 1967, Ch.I, pp.43-48) and Hannan (1970, Ch.II, pp.64-67; Ch.III, p.128).

the representation is unique if $F(0)$ is chosen to be the identity matrix. Using the Choleski factorization, G may be written as $G = UU^T$ where now the factorization is unique when U is restricted to be a square matrix that is real and lower triangular with all positive elements on the diagonal. Setting $C(z) = F(z)U$ completes the demonstration.

Q.E.D.

Note that the class of fundamental representations may be generated from that with $F(0) = I$ by post-multiplying by some matrix U where $UU^T = G$. Further an autoregressive representation can be computed by applying a vector generalization of Levinson's algorithm to the matrix covariogram.[3] The covariogram, in turn, may be calculated by inverse Fourier transforming the spectral density matrix, or by taking the Laurent series expansion of the rational matrix covariance generating function. In either case, that representation is *constrained* to be fundamental and corresponds to the choice of $F(0)$ as the identity matrix. Different orthogonalization orders are simply alternative decompositions of the symmetric positive definite matrix G. When $F(0)$ is the identity, the matrix G is the prediction error covariance matrix that the Levinson algorithm computes simultaneously with $F(z)^{-1}$.

3.2 Fundamental Representation

We now discuss identification in the presence of cross-equation restrictions such as those in the earlier section. Recall the restrictions implied by the hypothesis of uncovered interest parity may be stated as

---

[3]This is a fast algorithm to solve the Yule-Walker equations for dynamic regressions. When the "number of regressors" is k, the Levinson algorithm is an $O(k^2)$ operation whereas usual algorithms employed by least squares routines for inverting XX^T are $O(k^3)$. This saving results from exploiting the Toeplitz properties of XX^T matrices in dynamic models.

$$\left[a_1 z^{-\nu} C(z) \right]_+ = bC(z).$$

We generalize this to

$$\left[a(z) C(z) \right]_+ = b(z) C(z) \qquad (\mathbf{R}^\dagger)$$

where $C(z)$ and $b(z)$ have one-sided series expansions convergent in nonnegative powers of z. The series expansion of $a(z)$ is generally two-sided in z. Hansen and Sargent (1981) show that a large class of dynamic linear models may be written in this form. This class includes, but is not restricted to, models whose equilibria are the result of an optimization problem. Examples of such models are stationary versions of the hyperinflation model in Sargent and Wallace (1984), or aggregative models like those of staggered wage contract economies in Taylor (1980) where the equilibria are not the outcome of a maximization program but may be described by ($\mathbf{R}^\dagger$).

We now present a relation between fundamental and nonfundamental representations that satisfy the cross-equation restrictions $\mathbf{R}^\dagger$. This result is a trivial generalization of Hansen and Sargent's Lemma 2 (1981). The method of proof, however, is different. In the following, a diagonal matrix Blaschke factor is a matrix whose diagonal comprises only 1's and Blaschke functions, and is zero elsewhere.

Proposition 2 Let $C(z)$ satisfy $\mathbf{R}^\dagger$. The matrix functions $C(z)$ and $b(z)$ are assumed to have one-sided Laurent series expansions in nonnegative powers of z; $a(z)$ is arbitrary. The spectral density matrix generated by $C(z)$ is

$$S(\omega) = C(e^{-i\omega})C(e^{-i\omega})^*, \text{ for } \omega \text{ in } (-\pi,\pi].$$

Then there exists a $\varphi(z)$ that is fundamental and satisfies the boundary condition $\varphi(e^{-i\omega})\varphi(e^{-i\omega})^* = S(\omega)$ for all ω in $(-\pi,\pi]$ so that $\varphi(z)$ generates the same spectral density as $C(z)$. Further, that $\varphi(z)$ satisfies

$$\left[a(z)\varphi(z) \right]_+ = b(z)\varphi(z).$$

Proof By the proof of the Spectral Factorization Theorem, the matrix function $C(z)$ determines a finite sequence of unitary matrices $\left\{ U_j \right\}_{j=1}^{h}$ and diagonal matrix Blaschke factors $\left\{ B_j(z) \right\}_{j=1}^{h}$ such that

$$\varphi(z) = C(z) \prod_{j=1}^{h} U_j B_j(z) = C(z)V(z),$$

is fundamental. Note that $V(z)$ is constructed to have a convergent one-sided series expansion in strictly negative powers of z. By hypothesis,

$$\left[a(z)C(z) \right]_+ - b(z)C(z) = 0$$

==>

$$a(z)C(z) - Q(z) - b(z)C(z) = 0,$$

where $Q(z)$ is one-sided in strictly negative powers of z, and corresponds to the principal component of $a(z)C(z)$. Postmultiplying by $V(z)$ yields

$$a(z)C(z)V(z) - Q(z)V(z) - b(z)C(z)V(z) = 0.$$

Apply the $\left[\quad \right]_+$ operator on both sides of the equation to obtain

$$\left[a(z)C(z)V(z) \right]_+ - \left[b(z)C(z)V(z) \right]_+ = 0.$$

The term $\left[Q(z)V(z) \right]_+$ vanishes because $Q(z)V(z)$ is one-sided in strictly negative powers of z by construction. Now note that

$$b(z)C(z)V(z) = b(z)\varphi(z)$$

has a convergent power series expansion in nonnegative powers of z; $b(z)$ being convergent in nonnegative powers of z by assumption and $\varphi(z)$ being fundamental by construction. Hence $\left[b(z)C(z)V(z)\right]_+$ is simply $b(z)\varphi(z)$. This completes the demonstration for we have shown that

$$\left[a(z)\varphi(z)\right]_+ = b(z)\varphi(z).$$

Q.E.D.

Hansen and Sargent (1981) use an iterated projection argument in their proof. They exploit the identification between the Hilbert space associated with a random process and the location of its transfer function singularities and determinantal zeroes. The proof above makes no reference to those results but assumes the Spectral Factorization Theorem. The practical implications of this proposition can be stated briefly. Suppose we impose the cross-equation restrictions in the estimation procedure, and estimation yields a function $C(z)$ with determinantal zeroes inside the unit circle. Then there is an alternative $\varphi(z)$ that *is* fundamental, satisfies exactly the restrictions implied by the hypothesis of interest, and generates precisely the same second order statistics as the estimated function $C(z)$. Hence statistical tests of the model that rely only on second order statistics of the data do *not* depend on the econometrician delivering an estimate that is fundamental. Examples of such tests are the (quasi-) likelihood ratio statistics usually computed in this literature. To restate this somewhat, if an estimated $C(z)$ is not fundamental, we can obtain an estimated $\varphi(z)$ that *is* fundamental and yields exactly the same (quasi-) likelihood value. (Estimation criteria have not yet been presented so this anticipates somewhat the discussion below.) The procedure to achieve this involves only multiplication of unitary matrices and diagonal matrix Blaschke factors, all of which may be computed from the initial $C(z)$. This result is also useful for calculation of impulse response functions when the information content of shocks to the system are embodied in movements of the observed variables.[4]

3.3 *Estimation*

We turn to a preliminary discussion of estimating a representation for the matrix transfer function $C(z)$. Consider a class of criterion functions $l(S, X^N)$; S denotes the second order statistics generated by the model and X^N is $(x_1, x_2, \cdots, x_N)$ the observed partial realization of the model vector random process. We study estimators of $C(z)$ that maximize l subject to the restrictions $\mathbf{R}^\dagger$.

Proposition 3 Given X^N, let $\hat{C}(z)$ maximize l subject to the restrictions $\mathbf{R}^\dagger$, where $\hat{C}(z)$ is not necessarily fundamental. Then there exists a matrix transfer function $\hat{\varphi}(z)$, that is fundamental and may be constructed from knowledge of $\hat{C}(z)$ alone. Further $\hat{\varphi}(z)$ maximizes l in the class of fundamental factors satisfying $\mathbf{R}^\dagger$.

Proof This follows almost immediately from Proposition 2. Since X^N is fixed throughout the discussion, it is omitted without ambiguity. Let Φ_R be the collection of fundamental matrix functions satisfying $\mathbf{R}^\dagger$. Similarly let C_R denote the collection of square matrix functions that satisfy $\mathbf{R}^\dagger$. Obviously C_R contains Φ_R. Therefore

$$\sup_{C \text{ in } C_R} l(C(e^{-i\omega})C(e^{-i\omega})^*) \geq \sup_{\varphi \text{ in } \Phi_R} l(\varphi(e^{-i\omega})\varphi(e^{-i\omega})^*).$$

Hence we have

$$l(\hat{\varphi}(e^{-i\omega})\hat{\varphi}(e^{-i\omega})^*) = l(\hat{C}(e^{-i\omega})\hat{C}(e^{-i\omega})^*)$$

$$= \sup_{C \text{ in } C_R} l(C(e^{-i\omega})C(e^{-i\omega})^*)$$

[4]For instance, the procedure is used in calculating fundamental vector moving average representations from spectral density matrices restricted to be rank deficient. See Litterman, Quah and Sargent (1985).

$$\geq \sup_{\varphi \text{ in } \Phi_R} l(\varphi(e^{-i\omega})\varphi(e^{-i\omega})^*).$$

Q.E.D.

This proposition allows us to use estimates of fundamental factors *constructed* from possibly nonfundamental factors that maximize such criterion functions. As discussed more fully below, this class of permissible objective functions includes the Whittle likelihood or Hannan's approximation to the likelihood function for normally distributed vector processes. Estimators that maximize these criterion functions are known to be asymptotically quasi-maximum likelihood (see for example Kohn (1979)).

4. ESTIMATION

When ε_t has the normal distribution, x_t is similarly normally distributed as it is simply a linear combination of ε_{t-j}. The (conditional) log-likelihood function is therefore

$$l_N(\vartheta, X^N) = -\tfrac{1}{2}\left\{ \ln \det \Gamma_\vartheta + X_N^T \Gamma_\vartheta^{-1} X_N \right\},$$

$$\text{with } X_N := (x_1^T, x_2^T, \cdots, x_N^T),$$

where Γ_ϑ is that symmetric block Toeplitz matrix constructed from the theoretical (predicted) covariogram:

$$\Gamma_\vartheta = \mathrm{SBT}\,(\,\Gamma(0), \Gamma(-1), \cdots, \Gamma(-(N-1))\,), \qquad \Gamma(j) = E_\vartheta\, x_t x_{t-j}^T.$$

As usual, ϑ denotes the vector of parameters, and X_N is the N-term data sample. Denote by p the dimension of the observed vector x_t. The function l_N is repeatedly evaluated at different values of the parameter vector ϑ in estimation. Since matrix inversion is $O(n^3)$, and Γ_ϑ is $Np \times Np$, direct function evaluation quickly gets prohibitively expensive.[5] In this section, we formulate an operationally efficient approximation to the log-likelihood function. Estimates are obtained as maximizing elements of an objective function that is sometimes

---

[5]In most cases, l_N is decomposed first into a product of conditional distributions. This is easily done when the model has a simple autoregressive form. One can calculate a sequence of conditional distributions by forming prediction error decompositions for a given value of ϑ. The matrices inverted are then reduced in size. However recall that we wish to estimate models without necessarily first calculating the closed form representation. The conditional decompositions may be fairly expensive to compute in some models, while those models are easily handled by the techniques here. The cost may take the form of deterioration in numerical precision. This problem is well-known in eigenvalue or polynomial root location. The reader may note that since the matrix Γ_ϑ is block Toeplitz, one can exploit Levinson's algorithm in its inversion. It is easy to verify, however, that that is $O(N^2p^3)$ while the frequency domain criterion function evaluation method used here is $O(N \log Np^3)$ and so remains an order of magnitude faster.

referred to as the Whittle likelihood or as Hannan's frequency domain approximation. The development below makes explicit the approximation error *in the function.* Whether (constrained) maximizers of the approximation are "close" to maximizers of the original function is a different question. For that we appeal to Kohn (1979). When the degrees of all polynomial functions used are assumed known *a priori,* our assumptions ensure that maximizers of the frequency domain criterion function are strongly consistent for the same parameter vector as are maximizers of the conditional log-likelihood l_N.[6] Further the "likelihood ratio" statistics asymptotically have the usual χ^2 distribution under the null. All the efficiency properties of quasi-maximum likelihood estimators are inherited by the estimator here.

We approximate l_N by

$$l_N^F(\vartheta, X^N) = -\tfrac{1}{2}\sum_{j=1}^{N}\left\{ \ln\det S_\vartheta(\omega_j) + \operatorname{tr}\left[S_\vartheta(\omega_j)^{-1} I(\omega_j) \right] \right\},$$

$$\omega_j = \frac{2\pi}{N}(j-1)$$

where $S_\vartheta(\omega) = C_\vartheta(e^{-i\omega})C_\vartheta(e^{-i\omega})^*$ is the theoretical spectral density and $I(\omega) = \frac{1}{N}(\sum_{t=1}^{N} x_t e^{-i\omega t})(\sum_{t=1}^{N} x_t e^{-i\omega t})^*$ is the empirical periodogram. This criterion function is in the class discussed in the section on identification, so that all the results there apply. Without restrictions on the theoretical spectral density, an extremal point of l_N^F sets $C(z)$ so that $S(\omega)$ approaches $I(\omega)$ for all ω. (In the limit $S(\omega)$ is singular so that the function is not defined.) Maximizing l_N^F subject

[6] We do not mention the usual regularity assumptions. Simulation results of Phadke and Kedem (1978) for multivariate moving average models suggest deterioration when a determinantal zero is close to the unit circle. Kohn's theorems apply when ε_t is a vector sequence of second order stationary square integrable martingale differences, which weakens the requirement of *iid* normality.

to constraints may be viewed as matching the theoretical spectral density to an (admittedly inconsistent) estimate of the true spectral density.

To see how l_N^F approximates l_N, form the block (inverse) Fourier transform matrix

$$F_{(N,p)} = \frac{1}{\sqrt{N}} \left[I_p e^{i\omega_k(j-1)} \right]_{j,k}, \quad \omega_k = \frac{2\pi}{N}(k-1)$$

with the (j,k) $p \times p$ block the identity matrix multiplied by the complex sinusoid $e^{i\omega_k(j-1)}$. By construction $F_{(N,p)}$ is unitary. Define the block diagonal matrix

$$S_{(N,p)} = block\ diag \left\{ S_\vartheta(\omega_1), S_\vartheta(\omega_2), \cdots, S_\vartheta(\omega_N) \right\}.$$

Then

$$F_{(N,p)} S_{(N,p)} F_{(N,p)}^* = \left[N^{-1} \sum_{l=1}^{N} S_\vartheta(\omega_l) e^{i(j-k)\omega_l} \right]_{j,k},$$

that is, $F_{(N,p)} S_{(N,p)} F_{(N,p)}^*$ is symmetric block Toeplitz composed of N^2 $p \times p$ blocks. Each $p \times p$ block is a term in the discrete inverse Fourier transform of the spectral density. Since

$$\Gamma(j-k) \approx N^{-1} \sum_{l=1}^{N} S_\vartheta(\omega_l) e^{i(j-k)\omega_l}$$

we have

$$\begin{aligned} \Gamma_\vartheta &= \mathrm{SBT}\ (\ \Gamma(0), \Gamma(-1), \cdots, \Gamma(-(N-1))\ ) \\ &\approx F_{(N,p)} S_{(N,p)} F_{(N,p)}^* \end{aligned}$$

It follows that

$$\ln \det \Gamma_\vartheta = \sum_{k=1}^{N} \ln \det S_\vartheta(\omega_k)$$

and

$$X_N{}^T \Gamma_\vartheta^{-1} X_N \approx X_N{}^T F_{(N,p)} S_{(N,p)}{}^{-1} F_{(N,p)}{}^* X_N,$$

$$= \sum_{k=1}^{N} \tilde{x}(\omega_k)^* S_\vartheta(\omega_k)^{-1} \tilde{x}(\omega_k), \text{ with } \tilde{x}(\omega_k) := \frac{1}{\sqrt{N}} \sum_{j=1}^{N} x_j e^{-i\omega_k(j-1)}$$

$$= \sum_{k=1}^{N} \text{tr} \left\{ S_\vartheta(\omega_k)^{-1} I(\omega_k) \right\},$$

where $I(\omega_k) := N^{-1} \left[\tilde{x}(\omega_k)\tilde{x}(\omega_k)^* \right]$ is the periodogram. Therefore

$$l_N(\vartheta, X^N) \approx -\frac{1}{2} \sum_{j=1}^{N} \left\{ \ln \det S_\vartheta(\omega_j) + \text{tr } S_\vartheta(\omega_j)^{-1} I(\omega_j) \right\}.$$

The right hand side is simply $l_N^F(\vartheta, X_N)$. Note the data enters the criterion function only through the periodogram. The periodogram is a sufficient statistic and is formed only once in the estimation.[7] The approximation arises from the substitution of $\frac{1}{2\pi} \int_0^{2\pi} S(\omega) e^{ij\omega} d\omega$ for $N^{-1} \sum_{l=1}^{N} S(\omega_l) e^{ij\omega_l}$. This suggests possible improvements in the estimator considered here.[8]

---

[7]The determinant and inverse of the Hermitian arrays $S_\vartheta(\omega)$ are easily computed using the isomorphism between $p \times p$ complex and $2p \times 2p$ real matrices. See for example Hannan (1970, pp.224). By exploiting the conjugate symmetry about π of the spectral density matrix, only approximately half of these on $(-\pi, \pi]$ need to be calculated.

[8]We do not pursue this but leave it as a topic for future research. Adrian Pagan informs us that this approximation error argument applies to every term in the covariogram expansion: the aggregate error when summed to form the Whittle likelihood may be substantial.

5. FUNDAMENTAL FACTORIZATION OF A SPECTRAL DENSITY MATRIX

Let $C(z)$ be a factorization of $S(\omega)$ on the unit circle, not necessarily fundamental, $C(e^{-i\omega})C(e^{-i\omega})^* = S(\omega)$ for ω in $(-\pi,\pi]$, where the series expansion $C(z) = \sum_{k=-\infty}^{\infty} C_k z^k$ has all real coefficients.

We seek to obtain a fundamental factor $\varphi(z)$ that generates the given spectral density, that is,

$$S(\omega) = \varphi(e^{-i\omega})\varphi(e^{-i\omega})^*, \text{ for } \omega \text{ in } (-\pi,\pi].$$

We assume $S(\omega)$ is $p \times p$, rational and has full rank almost everywhere.[9]

First we derive from $C(z)$ a matrix rational function with all entries analytic in a domain containing the unit circle. This factor will then admit a convergent one-sided power series expansion in nonnegative powers of z. For each column j of $C(z)$, collect the denominator function zeroes $\{\rho_{jk}\}_k$ with $\left|\rho_{jk}\right| < 1$ in elements of the column vector. Form Blaschke factors, one for each column, $j = 1,2,\cdots,p$,

$$d_j(z) = \prod_k \frac{1-\overline{\rho}_{jk} z}{z-\rho_{jk}}, \quad \text{where } \left|\rho_{jk}\right| < 1$$

letting d_j be unity when column j has no element with denominator function vanishing inside the unit circle. Since the zeroes come in complex conjugate pairs, $d_j(z)$ is simply a rational function where both numerator and denominator polynomials can be chosen to have all real coefficients. Further for all j d_j

---

[9]See Litterman, Quah and Sargent (1985) for the modifications required when the spectral density matrix is rank deficient. Typically some $C(z)$ is readily available. When one is not, it may be obtained from the spectral density matrix by a generalized Choleski factorization. Litterman, Quah and Sargent implement this modified Choleski algorithm.

has modulus 1 on the unit circle: $d_j(z)\overline{d_j(z)} = 1$ for $\left| z \right| = 1$.

Construct the diagonal matrix function

$$D(z) = diag \left\{ d_1(z), d_2(z), \cdots, d_p(z) \right\},$$

and define

$$\varphi_1(z) = C(z)\, D(z)^{-1}.$$

The unit modulus property of $d_j(z)$ on $\left| z \right| = 1$ ensures that $D(z)$ is unitary on the unit circle so φ_1 generates the same spectral density as C. The matrix function φ_1 has elements that are ratios of polynomials with all real coefficients. Now no entry of φ_1 has singularities inside the unit circle so that φ_1 admits a one-sided Laurent series expansion convergent in nonnegative powers of z :

$$\varphi_1(z) = \sum_{k=0}^{\infty} \psi_{1k} z^k.$$

Note however that even had det $C(z)$ vanish nowhere on the unit disk, post multiplication by $D(z)^{-1}$ will, in general, introduce such zeroes into det $\varphi_1(z)$. These zeroes will need to be removed to construct a fundamental factor. We again employ Blaschke factors but now their use is a little more subtle.

Write $\varphi_1(z) = B(z)^{-1} A(z)$, where $B(z)$ is the diagonal matrix polynomial whose $j-th$ nontrivial element $B_{jj}(z)$ is the product of the denominator functions in the $j-th$ row of $\varphi_1(z)$; $A(z)$ is the matrix polynomial whose $j-th$ row is the product of the $j-th$ row of $C(z)$ with $B_{jj}(z)$. Then det $\varphi_1(z)$ is the ratio of det $A(z)$ to det $B(z)$.[10] After canceling common factors, write $\left| \varphi_1(z) \right| = \frac{P(z)}{Q(z)}$,

[10] The function det $A(z)$ may be formed using an extension of the Faddeev - Sominskii algorithm. Gantmacher (1959, pp.87-89) presents the algorithm for $A(z) = A$ an array of complex numbers. The modification straightforwardly replaces complex numbers by scalar polynomials.

and locate those zeroes of the numerator polynomial that are strictly inside the unit circle. Call these λ_j, $j = 1,2, \cdots ,h$. We seek to modify $\varphi_1(z)$ so that these zeroes are removed from the determinant, while keeping all entries in the factor analytic on the unit disk *and* maintaining the boundary condition $\varphi(e^{-i\omega})\varphi(e^{-i\omega})^* = S(\omega)$.

A simple example will illustrate the issues involved:

Example 1 Let

$$\varphi_1(z) = \begin{bmatrix} 1-2z & 0 \\ a(z) & 1 \end{bmatrix}$$

with $a(z)$ analytic in a domain containing the unit disk. For definiteness, assume that $a(½) = 2$. Then det $\varphi_1(z) = 1-2z$ has a zero inside the unit circle at $z = ½$ Although $\varphi_1(z)$ admits a one-sided Laurent series expansion convergent in nonnegative powers of z on the unit disk, $\varphi_1(z)$ is not fundamental for the spectral density matrix $S(\omega) = \varphi_1(e^{-i\omega})\varphi_1(e^{-i\omega})^*$. Another way to see this is to construct a time domain representation

$$x_t = \varphi_1(L)\varepsilon_t, \quad E\varepsilon_t\varepsilon_{t-u} = I\ \delta(u).$$

Clearly x_t may be formed from current and lagged ε_t so that $H_\varepsilon^-(t)$ contains $H_x^-(t)$. However formally inverting $\varphi_1(z)$,

$$\varphi_1(z)^{-1} = (1-2z)^{-1} \begin{bmatrix} 1 & 0 \\ -a(z) & 1-2z \end{bmatrix}$$

$$= \begin{bmatrix} (1-2z)^{-1} & 0 \\ -a(z)(1-2z)^{-1} & 1 \end{bmatrix}$$

and we note that the entry $(1-2z)^{-1}$ is not convergent in a one-sided series expansion in nonnegative powers of z. We see from

$$\varepsilon_t = \varphi_1(L)^{-1}x_t$$

that ε_t cannot be formed as a square summable linear combination of current and lagged x_t's (nor as the mean square limit of such a linear combination). However the first element of the vector ε_t can be obtained as a square summable linear combination of *future* realizations of x_t, while the second element can be formed only by taking linear combinations of current, lagged and future x_t's. Hence we say that $H_\varepsilon^-(t)$ is *not* contained in $H_x^-(t)$, and that such a $\varphi_1(z)$ is not fundamental for the spectral density matrix of the vector process $\{ x_t \}$.

To obtain a fundamental representation, we first try to multiply $\varphi_1(z)$ by a diagonal matrix of Blaschke factors. Such a matrix must satisfy two conditions: 1° the Hermitian product remains at the given spectral density matrix, and 2° the determinant of the resulting product matrix function vanishes nowhere inside the unit circle. Both conditions hold for

$$\tilde{\varphi}_1(z) = \varphi_1(z)B(z) = \varphi_1(z)\begin{bmatrix} 1 & 0 \\ 0 & \dfrac{1-\frac{1}{2}z}{z-\frac{1}{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 1-2z & 0 \\ a(z) & \dfrac{z-2}{1-2z} \end{bmatrix}.$$

Clearly $\tilde{\varphi}_1(e^{-i\omega})\tilde{\varphi}_1(e^{-i\omega})^* = \varphi_1(e^{-i\omega})\varphi_1(e^{-i\omega})^*$, and $\left| \tilde{\varphi}_1(z) \right| = z-2$ has neither zeroes nor singularities inside the unit circle. However the bottom right-hand element of $\tilde{\varphi}_1$ does have a singularity at $z = \frac{1}{2}$ *inside* the unit circle, so that $\tilde{\varphi}_1$ cannot be fundamental for the given spectral density matrix.[11] We conclude that

[11]Write $x_t = \tilde{\varphi}_1(L)\tilde{\varepsilon}(t)$, and note that the second element of x_t cannot be formed from (mean square limits of) square summable linear combinations of current and lagged $\tilde{\varepsilon}(t)$'s.

it is easy to find diagonal matrices of Blaschke factors that modify the determinant of the product matrix by removing zeroes of the determinant located inside the unit circle, while satisfying the given boundary condition. However because the Blaschke factors have singularities inside the unit circle, it is just as easy to introduce entries into the product matrix function that are not analytic throughout the unit disk.

■

Discussion of example 1 will continue below. For now we proceed with the formal development of the method.

Recall that λ_j, $j = 1,2,\cdots,h$ are the zeroes of $\left|\varphi_1(z)\right|$ inside the unit circle.

(i) Compute the singular value decomposition (Klema and Laub (1980))

$$\varphi_1(\lambda_1) = U_1 D_1 V_1^*.$$

Both U_1 and V_1 are unitary and D_1 has positive elements in the first r_1 places along the diagonal and is zero elsewhere. Since $\left|\varphi_1(z)\right|$ vanishes at $z = \lambda_1$, we know that $r_1 \leq p-1$. This factorization is exactly to prevent introduction into the resulting array elements that are not analytic. Using the Blaschke factor $b_1(z) = (z-\lambda_1)^{-1}(1-\overline{\lambda}_1 z)$, form the diagonal matrix function

$$B_1(z) = \begin{bmatrix} I_{r_1} & 0 \\ 0 & b_1(z) I_{p-r_1} \end{bmatrix}$$

Now let

$$\varphi_2(z) = \varphi_1(z)\, V_1\, B_1(z)$$

(ii) In the iterative step, for $j = 2, \cdots, h$, we construct in turn

$$\varphi_j(\lambda_j) = U_j\, D_j\, V_j^* \qquad \textit{(singular value decomposition)}$$

$$b_j(z) = \frac{1-\bar{\lambda}_j z}{z-\lambda_j} \qquad \textit{(Blaschke function)}$$

$$B_j(z) = \begin{bmatrix} I_{r_j} & 0 \\ 0 & b_j(z) I_{p-r_j} \end{bmatrix}$$

$$\varphi_{j+1}(z) = \varphi_j(z)\, V_j\, B_j(z).$$

The singular value decompositions are performed on *different* matrix functions evaluated at the *original* set of zeroes $\{ \lambda_j \}_{j=1}^{h}$. To see that the construction yields the appropriate result, we verify some of its properties.

a. *For* $j = 2, \cdots, h$, $\varphi_j(\lambda_j)$ is singular.

By construction,

$$\varphi_j(z) = \varphi_1(z) \prod_{k=1}^{j-1} V_k\, B_k(z).$$

Hence

$$\left|\varphi_j(z)\right| = \pm \left|\varphi_1(z)\right| \prod_{k=1}^{j-1} \left[\frac{1-\bar{\lambda}_k z}{z-\lambda_k}\right]^{p-r_k} ==> \left|\varphi_j(\lambda_j)\right| = 0 \text{ if } \left|\varphi_1(\lambda_j)\right| = 0.$$

b. For $j = 1,2,\cdots,h$, the matrix function $\varphi_{j+1}(z) = \varphi_j(z)\, V_j\, B_j(z)$ satisfies the boundary condition $\varphi_{j+1}(e^{-i\omega})\, \varphi_{j+1}(e^{-i\omega})^* = S(\omega)$. It has all entries analytic in a region containing the unit disk, and $\left|\varphi_{j+1}(z)\right|$ has at least 1 fewer zero inside the unit circle than does $\left|\varphi_j(z)\right|$.

Since

$$\varphi_{j+1}(z)\varphi_{j+1}(z)^* = \varphi_1(z) \left[\prod_{k=1}^{j} \left\{V_k\, B_k(z)\, B_k(z)^*\, V_k^*\right\}\right] \varphi_1(z)^*$$

and $B_k(z)$ is unitary on the unit circle,

$$\varphi_{j+1}(e^{-i\omega})\varphi_{j+1}(e^{-i\omega})^* = \varphi_1(e^{-i\omega})\varphi_1(e^{-i\omega})^* = S(\omega).$$

Further,

$$\left|\varphi_{j+1}(z)\right| = \pm\left|\varphi_j(z)\right| \left|\left[\frac{1-\bar{\lambda}_j z}{z-\lambda_j}\right]^{p-r_j}\right.$$

By construction $\left|\varphi_j(z)\right|$ has a zero of multiplicity $p-r_j$ at λ_j. These zeroes are removed at the $j-th$ iteration above and no new zeroes inside the unit circle are introduced.

Now we need to verify that φ_{j+1} remains analytic on the unit disk if the elements of φ_j are analytic. This is not obvious as the multiplying Blaschke factors have a singularity inside the unit circle at λ_j.

Assuming φ_j has all elements analytic in a domain containing the unit disk, it admits a one-sided series expansion, convergent in nonnegative powers of z, about any point on the unit disk. Further, its value at λ_j is

$$\varphi_j(\lambda_j) = U_j\, D_j\, V_j^*.$$

Hence taking the Laurent series expansion about λ_j, $\left|\lambda_j\right| < 1$,

$$\varphi_j(z) = U_j\, D_j\, V_j^* + \sum_{k=1}^{\infty} \psi_{jk}(z-\lambda_j)^k .$$

So

$$\varphi_{j+1}(z) = \varphi_j(z)\, V_j\, B_j(z)$$

$$= U_j\, D_j\, V_j^*\, V_j\, B_j(z) + \sum_{k=1}^{\infty} \psi_{jk}\, V_j \begin{bmatrix} (z-\lambda_j)^k I_{r_j} & 0 \\ 0 & (1-\overline{\lambda}_j z)(z-\lambda_j)^{k-1} I_{p-r_j} \end{bmatrix}$$

$$= U_j \begin{bmatrix} \Delta_j & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_{r_j} & 0 \\ 0 & \left[\dfrac{1-\overline{\lambda}_j z}{z-\lambda_j}\right] I_{p-r_j} \end{bmatrix}$$

$$+ \sum_{k=1}^{\infty} \psi_{jk}\, V_j \begin{bmatrix} (z-\lambda_j)^k I_{r_j} & 0 \\ 0 & (1-\overline{\lambda}_j z)(z-\lambda_j)^{k-1} I_{p-r_j} \end{bmatrix}$$

$$= U_j \begin{bmatrix} \Delta_j & 0 \\ 0 & 0 \end{bmatrix} + \sum_{k=1}^{\infty} \psi_{jk}\, V_j \begin{bmatrix} (z-\lambda_j)^k I_{r_j} & 0 \\ 0 & (1-\overline{\lambda}_j z)(z-\lambda_j)^{k-1} I_{p-r_j} \end{bmatrix}$$

Hence φ_{j+1} remains analytic on the unit disk, and in particular at λ_j :

$$\lim_{z\to\lambda_j} \varphi_{j+1}(z) = \varphi_{j+1}(\lambda_j) = U_j \begin{bmatrix} \Delta_j & 0 \\ 0 & 0 \end{bmatrix} + \psi_{j1}\, V_j \begin{bmatrix} 0 & 0 \\ 0 & (1-\left|\lambda_j\right|^2) I_{p-r_j} \end{bmatrix}.$$

c. *The* ***endpoint*** $\boldsymbol{\varphi} = \boldsymbol{\varphi}_{k+1}$ is a fundamental factor of $\boldsymbol{S}(\omega)$, and admits the convergent power series expansion

$$\varphi(z) = \sum_{k=0}^{\infty} \psi_k z^k$$

where for all k, ψ_k is real.[12]

It follows from *b.* that $\varphi = \varphi_{h+1}$ generates S. Since we iterate over the zeroes of the original factor, at each step removing that multiplicity of zeroes, the end-point is a factor with a determinant that has no zeroes inside the unit circle. Also from *b.*, $\varphi = \varphi_{h+1}$ has all elements analytic in a domain containing the unit disk. It only remains to check that the sequence of matrix coefficients $\{ \psi_k \}_{k=0}^{\infty}$ is real.

Collect the zeroes $\{ \lambda_1, \lambda_2, \cdots, \lambda_h \}$ so that the complex values, which always come in conjugate pairs, appear last. Recall that φ_1 is a matrix of rational functions with all real coefficients. Then in the first few iterations of the algorithm, associated with those zeroes that are real, V_j is real, and the coefficients in the power series expansion of $B_j(z)$ in a region containing the unit disk are similarly entirely real.

Take the complex zeroes in conjugate pairs. Suppose at the end of the $(j-1)-th$ step, $\varphi_j(z)$ is a matrix function with all real coefficients, and $\lambda_j = \overline{\lambda_{j+1}}$. Continuing the algorithm

$$\varphi_{j+2}(z) = \varphi_j(z)\, V_j\, B_j(z)\, V_{j+1}\, B_{j+1}(z), \qquad V_j,\ V_{j+1}\ \textit{unitary}.$$

Since the ordering of the zeroes is otherwise arbitrary, we also have

$$\varphi_{j+2}(z) = \varphi_j(z)\, W_j\, B_{j+1}(z)\, W_{j+1}\, B_j(z)$$

for W_j, W_{j+1} unitary constructed by the singular value decomposition. Further,

---

[12] The sequence of matrix coefficients in the expansion is the moving average response to a fundamental innovation process and therefore has to be real.

$$\varphi_j(\lambda_j) = U_j\, D_j\, V_j^*, \quad \varphi_j(z)\, V_j\, B_j(z)\, \Big|_{z=\overline{\lambda_j}} = U_{j+1}\, D_{j+1}\, V_{j+1}^*$$

and

$$\varphi_j(\overline{\lambda_j}) = T_j\, F_j\, W_j^*, \quad \varphi_j(z)\, W_j\, B_{j+1}(z)\, \Big|_{z=\lambda_j} = T_{j+1}\, F_{j+1}\, W_{j+1}^*.$$

As φ_j has all real coefficients, $\varphi_j(\overline{\lambda_j}) = \overline{\varphi_j(\lambda_j)}$. Hence in the singular value decompositions of $\varphi_j(\lambda_j)$ and $\varphi_j(\overline{\lambda_j})$

$$W_j = \overline{V_j}$$

Also it is easy to verify that $B_{j+1}(\lambda_j) = \overline{B_j(\lambda_{j+1})}$. This implies

$$\varphi_j(z)\, W_j\, B_{j+1}(z)\, \Big|_{z=\lambda_j} = \varphi_j(z)\, \overline{V_j}\, B_{j+1}(z)\, \Big|_{z=\lambda_j}$$

$$= \overline{\varphi_j(z)\, V_j\, B_j(z)}\, \Big|_{z=\overline{\lambda_j}}$$

that is,

$$T_{j+1}\, F_{j+1}\, W_{j+1}^* = \overline{U_{j+1}}\, \overline{D_{j+1}}\, V_{j+1}^T,$$

so that we can choose

$$W_{j+1} = \overline{V_{j+1}}.$$

Since $\varphi_j(z)$ has full rank almost everywhere,

$$V_j\, B_j(z)\, V_{j+1}\, B_{j+1}(z) = \overline{V_j}\, B_{j+1}(z)\, \overline{V_{j+1}}\, B_j(z).$$

The coefficients of $B_j(z)$ are simply complex conjugates of those of $B_{j+1}(z)$. Hence the coefficients of $V_j\, B_j(z)\, V_{j+1}\, B_{j+1}(z)$ are complex conjugates of themselves. It follows that the coefficients of $\varphi_{j+2}(z)$ are real. Proceeding thus, we show that $\varphi(z)$ has real coefficients in the Laurent series expansion $\varphi(z) = \sum_{k=0}^{\infty} \psi_k\, z^k$.

We now complete the discussion of the Example 1 by implementing this algorithm to obtain the fundamental factor.

Example 1 (continued)

Recall

$$\varphi_1(z) = \begin{bmatrix} 1-2z & 0 \\ \alpha(z) & 1 \end{bmatrix}.$$

At the zero of $\left| \varphi_1 \right|$ inside the unit circle, the singular value decomposition is

$$\varphi_1(½) = \begin{bmatrix} 0 & 0 \\ 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 0 \end{bmatrix} \left[\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \right] = U\,D\,V^*$$

Then

$$\varphi(z) = \varphi_1(z)\,V\,B_1(z)$$

$$= \begin{bmatrix} 1-2z & 0 \\ \alpha(z) & 1 \end{bmatrix} \left[\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \right] \begin{bmatrix} 1 & 0 \\ 0 & \dfrac{1-½z}{z-½} \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} 2(1-2z) & 2-z \\ 1+2\alpha(z) & (z-2)\dfrac{2-\alpha(z)}{1-2z} \end{bmatrix}$$

Since $\alpha(z) = 2$ at $z = ½$, $2-\alpha(z) = (1-2z)\,\beta(z)$ for some $\beta(z)$ analytic on the unit disk. Hence all elements of $\varphi(z)$ have one-sided convergent power series expansions in nonnegative powers of z on the unit disk. The determinant

$\left| \varphi(z) \right| = z-2$ vanishes nowhere on the unit disk.

∎

In summary, we collect the main results above as

Theorem (Rozanov) Let $S(\omega)$, ω in $(-\pi,\pi]$, be a given rational spectral density matrix with full rank almost everywhere. Let $C(z)$ be some factorization of $S(\omega)$ on the unit circle

$$S(\omega) = C(e^{-i\omega})C(e^{-i\omega})^*,$$

where $C(z)$ has the series expansion $C(z) = \sum_{k=-\infty}^{\infty} C_k z^k$ with all real coefficients. Then there exists a matrix rational function $\varphi(z)$ that is fundamental and generates the given spectral density matrix $S(\omega)$. Further $\varphi(z)$ may be constructed in finite steps from $C(z)$. The matrix function $\varphi(z)$ has a convergent one-sided power series expansion $\varphi(z) = \sum_{k=0}^{\infty} \psi_k z^k$ with real matrix coefficients $\{ \psi_k \}$.

∎

8. EMPIRICAL RESULTS

In this section, we report the results of applying the above methods to the Japanese-U.S. foreign exchange market. By arguments using iterated projections, tests of the null hypothesis of uncovered interest arbitrage remain valid when we restrict the information set to comprise only domestic and foreign interest rates and the exchange rate. The model therefore is a trivariate vector stochastic process comprising the Euroyen and Eurodollar 3-month (13-week) bid deposit rates in London and (the natural log of) the yen/dollar bid exchange rate. We performed tests on both monthly and weekly data for a number of different sample periods. The monthly data are sampled at the end of the month; the weekly data are Wednesday observations.[13]

Time domain results are displayed in Table 1. We estimated unrestricted vector autoregressions with lag lengths five (six) over sample periods June (July) 1975 to December 1984 and June (July) 1979 to December 1984. Wald tests of the cross-equation restrictions were then calculated in TSP (see Ito (1984)). Marginal confidence levels are low in all cases: they change noticeably however as we move across lag lengths. The nature of these regressions lead to design matrices that are close to singular. Their inverse will hence be imprecisely estimated and numerically unstable in finite sample.

For the frequency domain method, we obtained initial estimates as follows. We removed sample means and then estimated the matrix covariogram. The Levinson-Whittle algorithm was then used to find unrestricted projection coefficients. Under the identification conditions specified in Section III, an

---

[13]Monthly interest rates from 1975 to 1978 are taken from *World Financial Markets*, Morgan Guaranty. Monthly exchange rates from 1975 to 1978 are from the IFS database, IMF. Monthly and weekly data from 1979 are constructed from daily data kindly provided to us by DRI. The reader is referred to Ito (1983, 1984) for more careful description of the data and an account of developments in the market over the sample period.

unrestricted parametrization was extracted from the inverse of the matrix autoregressive operator. We chose rational functions of order four for the entries of $C(z)$ with the leading coefficient of the denominator polynomials normalized at unity.[14] Unrestricted estimates were obtained by leaving the third row of the matrix function $C(z)$ free in the maximization of the criterion function.

In Table 2, we display measures of the adequacy of the restricted representation in capturing dynamic properties of the monthly data. The sample period is January 1975 to December 1984. Since the estimated moving average representations were not fundamental (computing the zeroes of the determinantal function showed a number to be *inside* the unit circle), their characteristics are not necessarily identified with that informational content embodied in exchange and interest rate movements. Nevertheless, the results above assure us that we can still treat the likelihood ratio statistic in the usual manner. The estimated parametrization itself has no obvious economic interpretation. We present instead estimated innovation covariance matrices to let the reader judge the fit of the model relative to less restricted representations.

The labeled VAR innovation covariance matrix may be interpreted as follows. Estimate the spectral density by some unconstrained nonparametric method. One way is simply to calculate the periodogram of the data. Treat this as the counterpart of the restricted spectral density matrices computed in the paper. The vector autoregressive representation derived by this construction is essentially that computed using OLS equation by equation. The innovation covariance matrix should then be compared to the other innovation covariance

[14] We computed fourier transforms using Singleton's mixed radix fast fourier transform routines (1979). Davidon-Fletcher-Powell procedures in GQOPT were used for maximization of the criterion function. Other calculations used a collection of FORTRAN routines written by one of the authors designed for manipulating models of random processes with rational matrix spectral densities. The estimation was executed on a VAX 11-785 running VMS 4.1.

matrices presented that are obtained by restricting the spectral density matrix.

The variates are ordered in the sequence mentioned in the first paragraph of this section. Vector autoregressive representations used for the innovation covariance matrices here are obtained from an estimated covariogram derived by inverse fourier transforming the (un)restricted spectral density matrices. We also computed decompositions of the spectral density matrix using Laurent series expansion of the rational covariance generating function to recover the covariogram. To the extent that the determinantal poles of the fundamental moving average representation are well outside the unit circle, the estimates by Laurent series expansion and those by inverse fourier transforms differ little.

The marginal significance level of the χ^2 statistic is 63.4 percent. While small sample properties of this statistic are unknown, and the model that is estimated is liberally parametrized relative to the number of observations, it appears the data are not spectacularly inconsistent with the maintained hypotheses of uncovered interest arbitrage and rational expectations.

Table 3 collects similar results for weekly data. Here the sample period is the first week of 1981 to the last week of 1984. The variates are now ordered as the (log of) the spot exchange rate, the Eurodollar and Euroyen 13-week deposit rates in London. The marginal significance level for the likelihood ratio test of the restrictions is 7.4 percent. Individual variances of all three estimated innovation processes increase when the restrictions are imposed. The covariance matrices themselves do not change noticeably across computed six- and twelve-lag vector autoregressive representations within a given parametrization. They are also reasonably "close" across parametrizations.

Table 1 Monthly

Wald Tests (Time Domain) of Restricted Representation

Restricted vector autoregression

Number of Lags	Sample Period	χ^2	Marginal Confidence Level
5[a]	1975:6-1984:12	10.533	0.163
	1979:6-1984:12	18.059	0.680
6[b]	1975:7-1984:12	9.904	0.045
	1979:7-1984:12	15.583	0.315

[a]Degrees of freedom = 16
[b]Degrees of freedom = 19

Table 2 Monthly

Sample Period January 1975 to December 1984

Test of Restricted Representation, degrees of freedom =21
(Frequency Domain)

	χ^2	Marginal Confidence Level
Whittle Likelihood Ratio	18.237	0.366

Restricted Estimates: Innovation Covariance Matrix

$$6 \text{ lags } V_r = \begin{vmatrix} .12\times 10 & & \\ .14\times 10^{-2} & .76\times 10^{-2} & \\ .15\times 10^{-2} & .22\times 10^{-2} & .93\times 10^{-2} \end{vmatrix} \times 10^{-3},\ \ln\left| V_r \right| = -30.18$$

Unrestricted Estimates: Innovation Covariance Matrix

$$6 \text{ lags } V_u = \begin{vmatrix} .12\times 10 & & \\ .99\times 10^{-2} & .71\times 10^{-2} & \\ .82\times 10^{-2} & .12\times 10^{-2} & .63\times 10^{-2} \end{vmatrix} \times 10^{-3},\ \ln\left| V_u \right| = -30.56$$

VAR Estimates: Innovation Covariance Matrix

$$6 \text{ lags } V_v = \begin{vmatrix} .12\times 10 & & \\ .78\times 10^{-2} & .67\times 10^{-2} & \\ .26\times 10^{-2} & .18\times 10^{-2} & .62\times 10^{-2} \end{vmatrix} \times 10^{-3},\ \ln\left| V_v \right| = -30.68$$

Table 3 Weekly

Sample Period 1981:1 to 1984:52

Test of restricted representation, degrees of freedom=21
(Frequency Domain)

	χ^2	Marginal Confidence Level
Whittle Likelihood Ratio	30.984	0.926

Restricted Estimates: Innovation Covariance Matrix

$$6 \text{ lags } V_r = \begin{vmatrix} .58 & & \\ .82\times10^{-2} & .43\times10^{-2} & \\ .78\times10^{-2} & .83\times10^{-3} & .45\times10^{-3} \end{vmatrix} \times10^{-3},\ \ln\left| V_r \right| = -35.18$$

$$12 \text{ lags } V_r = \begin{vmatrix} .58 & & \\ .82\times10^{-2} & .43\times10^{-2} & \\ .78\times10^{-2} & .82\times10^{-3} & .43\times10^{-3} \end{vmatrix} \times10^{-3},\ \ln\left| V_r \right| = -35.25$$

Unrestricted Estimates: Innovation Covariance Matrix

$$6 \text{ lags } V_u = \begin{vmatrix} .41 & & \\ .10\times10^{-1} & .38\times10^{-2} & \\ .55\times10^{-2} & .67\times10^{-3} & .39\times10^{-3} \end{vmatrix} \times10^{-3},\ \ln\left| V_u \right| = -35.63$$

$$12 \text{ lags } V_u = \begin{vmatrix} .41 & & \\ .99\times10^{-2} & .37\times10^{-2} & \\ .55\times10^{-2} & .68\times10^{-3} & .39\times10^{-3} \end{vmatrix} \times10^{-3},\ \ln\left| V_u \right| = -35.67$$

VAR Estimates: Innovation Covariance Matrix

$$6 \text{ lags } V_v = \begin{vmatrix} .32 & & \\ .17\times10^{-1} & .26\times10^{-2} & \\ .40\times10^{-2} & .44\times10^{-3} & .31\times10^{-3} \end{vmatrix} \times10^{-3},\ \ln\left| V_v \right| = -36.43$$

$$12 \text{ lags } V_v = \begin{vmatrix} .30 & & \\ .25\times10^{-2} & .24\times10^{-2} & \\ .40\times10^{-2} & .44\times10^{-3} & .28\times10^{-3} \end{vmatrix} \times10^{-3},\ \ln\left| V_v \right| = -36.72$$

7. CONCLUSION

This paper has formulated and tested the restrictions implied by uncovered interest arbitrage. This was done within the context of a second order theory of vector random processes. The simplifying economic assumptions (such as risk neutrality) underlying such a tractable representation may be extreme. Nevertheless, we believe this to be a useful framework for examining explicit restrictions on the joint dynamics of economic variables.

The work here follows directly from Hansen and Sargent (1981) and Whiteman (1983). Hansen and Sargent explored the identification problem for exact linear models of the kind studied here. Whiteman examined the solution of linear rational expectations models by use of moving average representations. The estimation techniques (and machine software) discussed in this paper complement these in providing a unified set of tools for studying stochastic economic models where theory generates explicit predictions for dynamic properties. Hence this is yet another analytically tractable framework that allows the researcher to "keep in sight" all the second order properties of the data. At the same time the researcher can easily impose on the empirical analysis restrictions informed by a version of dynamic economic theory.

An earlier study had shown the validity of the linear representation for the market for U.S. - Japan foreign exchange. This work has verified that conclusion using an alternative parametrization. The ease with which these restrictions may be imposed directly on the spectral density matrix is encouraging for research on explicit dynamics.

REFERENCES

BAILLIE, R. , R.E. LIPPENS AND P.C. McMAHON: " Testing Rational Expectations and Efficiency in the Foreign Exchange Market, " *Econometrica*, 51(1983), 553-563.

CUMBY, R. AND M. OBSTFELD: " International Interest Rate and Price Level Linkages under Flexible Exchange Rates: A Review of Recent Evidence, " in BILSON, J.F.O. AND MARSTON, R.C. (eds.): *Exchange Rate Theory and Practice*. Chicago: University of Chicago, 1984.

GANTMACHER, F.R.: *Matrix Theory* Vol.I. New York: Chelsea, 1959.

HANNAN, E.J.: *Multiple Time Series*. New York: John Wiley, 1970.

HANSEN, L.P. AND R. HODRICK: " Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis, " *Journal of Political Economy*, 88(1980),829-53.

HANSEN, L.P. AND T.J. SARGENT: " Exact Linear Rational Expectations Models: Specification and Estimation. " Federal Reserve Bank of Minneapolis Research Department Staff Report 71, 1981.

ITO, T.: " Capital Controls and Covered Interest Parity, " NBER W.P. 1187, 1983.

———: " Use of (Time-Domain) Vector Autoregressions to Test Uncovered Interest Parity, " NBER W.P. 1493, 1984.

KOHN, R.: " Asymptotic Estimation and Hypothesis Testing Results for Vector Linear Time Series Models, " *Econometrica*, 47(1979),1005-1029.

KLEMA, V.C. AND A.J. LAUB: " The Singular Value Decomposition: Its Computation and Some Applications, " *IEEE Trans. Auto. Contr.*, AC-25, 2(1980).

LITTERMAN, R.B., D. QUAH AND T.J. SARGENT: " Fundamental Vector Moving Average Representations for Dynamic Index Models, " mimeo June 1985.

PHADKE, M.S. AND G. KEDEM: " Computation of the Exact Likelihood function of Multivariate Moving Average Models. " *Biometrika*, 65(1978),3,511-519.

ROZANOV, Y.A.: " Spectral Properties of Multivariate Stationary Processes and Boundary Properties of Analytic Matrices. " *Theory Probab. Appl.* 5(1960), 362-376. Reprinted in KAILATH, T. (ed.): *Linear Least-Squares Estimation*. Stroudsburg: Dowden, Hutchinson and Ross, 1977.

———: *Stationary Random Processes*. San Francisco: Holden-Day, 1967.

SARGENT, T.J.: " A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure, " *Journal of Monetary Economics*, 5(1979),133-143.

SARGENT, T.J. AND N. WALLACE: " Exploding Hyperinflation, " mimeo 1984.

SIMS, C.A.: " Macroeconomics and Reality, " *Econometrica*, 48(1980a),1-48.

__________: " Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered, " *American Economic Review*, 70(1980b),250-57.

__________: " Policy Analysis with Econometric Models, " *Brookings Papers on Economic Activity*, 1(1982),107-52.

SINGLETON, R.C.: " Mixed Radix Fast Fourier Transforms, " 1.4-1 in IEEE Acoustics, Speech, and Signal Processing Society, Digital Signal Processing Committee (ed.): *Programs for Digital Signal Processing*, New York: IEEE Press, 1979.

TAYLOR, J.B.: " Aggregate Dynamics and Staggered Contracts, " *Journal of Political Economy*, 88(1980),1-23. .

WHITEMAN, C.: *Linear Rational Expectations Models*, Minneapolis: University of Minnesota Press, 1983.