Federal Reserve Bank of Minneapolis
Research Department Working Paper

USING UNOBSERVABLE INDEX MODELS TO ESTIMATE UNOBSERVABLES AND FORECAST OBSERVABLES

Robert B. Litterman and Thomas J. Sargent

Working Paper 125
PACS File 2700

Revised April 1984

NOT FOR DISTRIBUTION
WITHOUT AUTHORS' APPROVAL

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. The material contained is of a preliminary nature, is circulated to stimulate discussion, and is not to be quoted without permission of the authors.

Several people made generous contributions to the ideas and computations contained in this paper. John Geweke and Kenneth Singleton graciously supplied us with their computer program for maximum likelihood confirmatory and exploratory dynamic factor analysis. That program is a modification of a program written earlier at the University of Minnesota by John Geweke and Paul Anderson. Kenneth Singleton also kindly answered numerous questions and made suggestions about how to modify his program for some of the calculations performed here. Robert Litterman modified the Geweke-Singleton program to estimate confirmatory models with orthogonal indexes and to calculate the projections and predictions described in this paper. Litterman carried out all of the calculations in this paper. Finally, our thinking about index models was initially stimulated by R. E. Lucas's "Equilibrium Model of the Business Cycle," [13]. Robert E. Lucas, Jr., and Preston Miller made useful comments on an earlier draft.

This paper describes how the unobservable index model of Sargent and Sims [16] can be used to estimate the effects of the unobservable indices on observables. The paper also shows how the unobservable index model can be used to constrain the parameters of a vector autoregression for the purpose of forecasting. These methods are illustrated in the context of a two-unobservable index model that was motivated by Lucas's model of business cycles. This particular index model is designed to represent the natural rate hypothesis, with one index being potentially interpretable as "unexpected aggregate demand," while the second index is supposed to represent purely neutral price fluctuations.

Given an (nx1) vector stochastic process x_t, a k unobservable index model is given by

$$(1.1) \qquad x_{it} = \sum_{j=1}^{k} \lambda_{ij}(L) f_{jt} + u_{it}, \qquad i = 1, \ldots, n$$

where $Eu_{it}u_{js} = 0$ for all t and s when $j \neq i$; $Eu_{it}f_{js} = 0$ for all t, s, i, j; $Ef_{jt}f_{is} = 0$ for $i \neq j$ and all s, t; $Ef_{it}f_{it-s} = Ef_{jt}f_{jt-s} = 0$ for $t \neq s$. That is, f_{jt}, $j = 1, \ldots, k$ and u_{it}, $i = 1, \ldots, n$ are a set of mutually orthogonal stochastic processes, with the f_{jt}'s being serially uncorrelated. The u_{it}'s are permitted to be serially correlated. The $\lambda_{ij}(L)$ are square summable polynomials in the lag operator L. The x_{it}'s are observable while the f_{jt} and u_{it}'s are unobservable.

For k sufficiently small relative to n, equation (1.1) restricts the covariances of the process x_t, namely, $C(\tau) = Ex_t x_{t-\tau}^T$. Via the Yule-Walker equations, the covariogram $C(\tau)$ determines the p^{th} order vector autoregression

$$(1.2) \qquad x_t = \sum_{h=1}^{p} A_j^{(p)} x_{t-h} + \varepsilon_t^p$$

where $E\varepsilon_t^p x_{t-h}^T = 0$ for $h = 1, \ldots, p$. One purpose of this paper is to describe and to illustrate methods for imposing the constraints on (1.2) that are implied by the model (1.1). Motivations for wanting some source of prior restrictions to impose on the vector autoregression (1.2) are described by Sims [20] and Litterman.

The effects of index j on observable variable i, $\lambda_{ij}(L)f_{jt}$, are unobservable. However, given sufficient restrictions, the second moments (i.e., the cross covariogram) of $\lambda_{ij}(L)f_{jt}$ with the entire x_t process can be identified and estimated, as can the covariogram of the $\lambda_{ij}(L)f_{jt}$ process. These second moments contain all of the information that is needed to construct the linear least squares estimator of $\lambda_{ij}(L)f_{jt}$ in the space spanned by $\{x_{t-s}\}_{s=-m_1}^{m_2}$:

$$(1.3) \qquad \hat{E}\left[\lambda_{ij}(L)f_{jt} \,\middle|\, \{x_{t-s}\}_{s=-m_1}^{m_2}\right] = \sum_{s=-m_1}^{m_2} B_s x_{t-s}.$$

Another goal of this paper is to describe procedures by which the projection (1.3) can be calculated. We illustrate (1.3) by presenting time series of $\lambda_{ij}(L)f_{jt}$ for a two-index model of the aggregate time series, where one index is interpretable as representing purely neutral nominal disturbances.

When $k > 1$, identifying restrictions must be imposed on the (nxk) matrix of $[\lambda_{ij}(L)]$'s in order to compute (1.3), although (1.2) can be computed in the absence of such restrictions. Below we shall appeal to the Geweke and Singleton characterization of restrictions that are sufficient for our purposes.

2. Identification and Estimation

We use the unobservable index model

$$x_t = \sum_{j=-\infty}^{\infty} \Lambda_j f_{t-j} + u_t$$

or

$$(2.1) \qquad x_t = \Lambda(L) f_t + u_t$$

where x_t is an (nx1) vector stochastic process, $\{\Lambda_j\}$ is a square summable sequence of (nxk) matrics, f_t is a (kx1) vector of unobservable indexes, and u_t is an (nx1) vector of disturbances. The spectral density matrix of u_t is assumed diagonal. The spectral density matrix of f_t is the identity matrix. It is assumed that $Ef_t u'_{t-s} = 0$ for all t and s, and that both f_t and u_t are second-order stationary stochastic processes.

For representing aspects of business cycles, we are interested in a two-index model of the restricted form

$$(2.2) \qquad x_t = \begin{pmatrix} \Lambda_{11}(L) & 0 \\ \Lambda_{21}(L) & \Lambda_{22}(L) \end{pmatrix} \begin{pmatrix} f_{1t} \\ f_{2t} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

where $\Lambda_{11}(L)$ is (rx1), $\Lambda_{22}(L)$ is (n-r) x 1, and $\Lambda_{21}(L)$ is (n-r) x 1 in dimension. In (2.2),

$$x_t = \begin{pmatrix} y_t \\ p_t \end{pmatrix}$$

where y_t is an r x 1 vector of real economic variables, and p_t is an (n-r) x 1 vector of nominal variables. Note the (n-r) x 1 vector of zeroes in the upper-right corner of $\Lambda(L)$, which generally give additional restrictions on the covariance structure of

the x_t process beyond those contained in the two-index hypothesis. The zero restrictions contained in (2.2) are designed to represent a version of the natural rate hypothesis: the second index f_{2t} impinges on nominal magnitudes, but not on real magnitudes. In Appendix C, we motivate and qualify the sense in which (2.1) represents the natural rate hypothesis.

Expressing (2.1) in terms of the restrictions that it places on cross-spectral density matrices gives

$$(2.3) \qquad S_x(w) = \Lambda(w)S_f(w)\Lambda(w)' + S_u(w), \quad w\varepsilon[-\pi,\pi]$$

where $S_x(w)$ is the nxn cross-spectral density matrix of x, $\Lambda(w)$ the Fourier transform of the sequence $\{\Lambda_j\}$, $S_f(w)$ is the (kxk) spectral density matrix of f_t, and $S_u(w)$ is the diagonal cross-spectral density matrix of u_t.[1/] Here the prime denotes transposition and complex conjugation. Equation (2.3) is an analogue in the frequency domain of the ordinary factor analysis model.

A. Identification

There are $n(n+1)/2$ parameters in $S_x(w)$, nk parameters in $\Lambda(w)$, and n parameters in $S_u(w)$. If $n(n+1)/2 > nk + n$, or equivalently, if $n/2 - 1/2 > k$, there are fewer parameters in $\Lambda(w)$ and $S_u(w)$ than there are parameters in $S_x(w)$ which we would like to match. If n is large relative to k, the model places restrictions on $S_x(w)$.

There are two types of unobservable index models, which differ with regard to whether they leave $\Lambda(w)$ an unrestricted nxk matrix, or impose linear restrictions on the elements of $\Lambda(w)$ as in the example of equation (2.2) In an exploratory unobservable

index model, $\Lambda(w)$ is left unrestricted. In a confirmatory unobservable index model, linear restrictions, typically zero-one exclusion restrictions, are imposed on elements of $\Lambda(w)$. By imposing more restrictions, confirmatory models permit more aspects of $\Lambda(w)$ to be identified and estimated.

Without a priori restrictions on $\Lambda(w)$ and $S_f(w)$, the individual elements of $\Lambda(w)$ and $S_f(w)$ are not identifiable for $k > 2$, although $\Lambda(w)S_f(w)\Lambda(w)'$ and $S_u(w)$ are identified even without any such restrictions.[2/] For taking $M(w)$ to be any (kxk) nonsingular, complex-valued matrix, define the new "factor loadings" $\Lambda^*(w) = \Lambda(w)M(w)^{-1}$ and define the new spectral density $S_f^*(w) = M(w)S_f(w)M(w)'$. Then notice that if (2.3) is satisfied originally, it remains satisfied when we replace $\Lambda(w)$ by $\Lambda^*(w)$ and $S_f(w)$ by $S_f^*(w)$. Without some a priori information, any nonsingular matrix $M(w)$ defines an admissible transformation or "rotation" of a given initial solution of (2.3). The elements of $\Lambda(w)$ and $S_f(w)$ are said to be identified if enough a priori information is imposed to make choosing $M(w)$ equal to the (kxk) identity matrix the only admissible transformation. Heuristically, since $M(w)$ is a (kxk) matrix, k^2 a priori restrictions are necessary to identify the model. If we impose that $S_f(w) = S_f^*(w) = I$, this implies that $M(w)M(w)' = I$, supplying $k(k+1)/2$ restrictions on admissible transformation matrices $M(w)$. Further, suppose that we require that the j^{th} column of $\Lambda(w)$ (or any $\Lambda^*(w)$ obtained from $\Lambda(w)$ by postmultiplying by an admissible $M(w)^{-1}$) have $(k-j)$ zeroes for $j = 1, 2, \ldots, k$. This leads to a set of equations of the form

$$\sum_i \Lambda_{hi}(w)M(w)^{-1}_{ik} = 0$$

for the elements of $\Lambda(w)_{hk}$ set to zero a priori. These equations supply an additional $k(k-1)/2$ restrictions on $M(w)$. This means that the restrictions on $S_f(w)$ and $\Lambda(w)$ comprise a total of $k^2 = k(k+1)/2 + k(k-1)/2$ restrictions on admissible $M(w)$ matrices. These restrictions are, in general, necessary (though not sufficient) to imply that $\Lambda^*(w)\Lambda^*(w)^{-1} = \Lambda^*(w)\Lambda^*(w)^{-1}$, or equivalently, that $M(w)^{-1}M(w) = I$. However, for a complex matrix, $M(w)^{-1}M(w) = I$ does not imply that $M(w) = I$. Therefore, to identify $\Lambda(w)$ once $\Lambda(w)\Lambda(w)'$ has been identified, k additional normalization[5] on $M(w)$ are required so that $M(w)^{-1}M(w) = I$ implies that $M(w) = I$. For example, Geweke and Singleton [7] impose the condition that one row of $\Lambda(w)$ be real. With this or an equivalent normalization, $\Lambda(w)$ is identified. If more zero restrictions than described above are available, the model becomes overidentified.

The preceeding "order conditions" for identification are necessary, but not sufficient, for the model with orthogonal factors to be identified. For a statement and proof of a set of sufficient conditions, see Geweke and Singleton [7].

In our model (2.2) with $S_f(w) = I$, the second column of $\Lambda(w)$ has r zeroes, where r is the dimension of y_t, i.e., the number of real variables in the system. So for $r > 1$, the model is identified according to the preceding counting rules. For the model (2.2), Geweke and Singleton's sufficient condition requires that the rank of the (rx1) matrix $\Lambda_{11}(w) = \sum_{j=0}^{\infty} \Lambda_{11,j} e^{-iwj}$ equals one. This condition is evidently satisfied so long as $r > 1$. If $r > 2$, the model with orthogonal indexes is overidentified.

B. Estimating the Confirmatory Model

To estimate a confirmatory model, an unobservable k-index model is first estimated with no prior constraints placed on Λ or S_f. This is the "exploratory dynamic factor analysis" model in Geweke's terminology. In the course of estimating the unconstrained k-index model, the following normalizations are arbitrarily imposed in order to determine versions of Λ and S_f: (i) the spectral density of f_t, $S_f(w)$, is set equal to the identity matrix, supplying $k(k+1)/2$ constraints; (ii) the matrix $\Lambda'(w)S_u^{-1}(w)\Lambda(w)$ is required to be diagonal, supplying an additional $k(k-1)/2$ constraints required to determine $\Lambda(w)\Lambda(w)'$. These are the usual arbitrary restrictions used in estimating the unconstrained model (see Lawley and Maxwell [9, pages 7-8 and Chapter 4] for the details).

These estimates from the unconstrained or exploratory model are used to construct starting values for estimating the constrained or "confirmatory" model. Given the estimate $\Lambda(w)$ from the exploratory model, one constructs a k x k unitary matrix $U(w)$ which defines a new set of loadings $\Lambda^*(w) = \Lambda(w)U(w)$ which have zero elements where the theory indicates they should be. Geweke and Singleton [7] construct $U(w)$ by using the factor rotation scheme of Lawley and Maxwell [9, pages 79-84], appropriately modified to account for the fact that $\Lambda(w)$ and $U(w)$ are matrices of complex variables in the present case. As mentioned above, the zero identifying restrictions described above are sufficient to restrict the set of matrices $M(w)$ defining admissible transformations to satisfy $M(w)M(w)^{-1} = I$; but since $M(w)$ is complex, this

does not imply that $M(w) = I$. An additional k normalizations are required to restrict the admissible transformation matrix $M(w)$ to equal I. Geweke and Singleton impose the condition that one row of $\Lambda(w)$ be real. For example, where the last row of $\Lambda^*(w)$ is to be normalized, this is accomplished by taking $\Lambda^*(w) = \Lambda(w)U(w)$ and postmultiplying by the diagonal (kxk) matrix

$$A(w) = \begin{pmatrix} \frac{\Lambda^*_{n1}}{|\Lambda^*_{n1}|} & 0 & \cdots & 0 \\ 0 & \frac{\Lambda^*_{n2}}{|\Lambda^*_{n2}|} & & \\ \vdots & & & \\ 0 & 0 & & \frac{\Lambda^*_{nk}}{|\Lambda^*_{nk}|} \end{pmatrix} .$$

The initial estimates of the factor loadings from which to start the constrained optimization are then $\Lambda^{**}(w) = \Lambda(w)U(w)A(w)$. From here on, estimation of the confirmatory model proceeds by using a hill-climbing procedure to maximize (2.8) given below.

C. Estimating Exploratory and Confirmatory Models

The model is estimated using the following procedures, which are described in more detail in Geweke [6], Geweke and Singleton [7], and Sargent and Sims [16]. First, the original data are regressed on a constant, seasonal dummies, and a linear trend. The residuals from these regressions are then treated as the basic observations, which are to be modeled as linearly indeterministic time series. These observations constitute the components of the vector x_t, in terms of which the model to be estimated is

$$(2.4) \qquad x_t = \Lambda(L)f_t + u_t,$$

or in terms of spectral density matrices

$$(2.5) \qquad S_x(w) = \Lambda(w)\Lambda(w)' + S_u(w).$$

The first step in estimation is to quasi-difference the data to form a set of filtered series x_t^a defined by

$$(2.6) \qquad x_t^a = D(L)x_t$$

where $D(L) = D_0 - D_1L - D_2L^2$, and where $D_0 = I$ and D_1 and D_2 are diagonal matrices whose components are determined from univariate second-order autoregressions for each component of x_t. So x_t^a consists of residuals of x_t from second-order univariate autoregressions. The purpose of filtering x_t in this way is approximately to flatten the spectral density matrix of x_t^a, which is necessary for reliable estimation. Operating on (2.4) with D(L) gives

$$x_t^a = D(L)\Lambda(L)f_t + D(L)u_t.$$

The theoretical spectral density matrix of x_t^a under the k-index hypothesis is

$$S_{x^a}(w) = D(w)\Lambda(w)\Lambda(w)'D(w)' + D(w)S_u(w)D(w)'$$

where $D(w) = \sum_{j=0}^{2} D_j e^{-iwj}$ is the Fourier transform of the D_j sequence, and where the prime again denotes transposition and complex conjugation. The model is parameterized as

$$(2.7) \qquad S_{x^a}(w) = \Lambda_a(w)\Lambda_a(w)' + S_{u^a}(w)$$

where it is understood that $\Lambda_a(w)$ is the Fourier transform of $D(L)\Lambda(L)$ and $S_{u^a}(w)$ is the spectral density of $D(L)u_t$. In saying that the model is parameterized as (2.7), we intend to indicate that it is the parameters of $\Lambda_a(w)$ and $S_{u^a}(w)$ themselves that are free at each frequency. Once the parameters of $\Lambda_a(w)$ and $S_{u^a}(w)$ have been estimated, $\Lambda(w)$ and $S_u(w)$ are estimated from the inverse relationships

$$\Lambda(w) = D(w)^{-1}\Lambda_a(w)$$

$$S_u(w) = D(w)^{-1}S_{u^a}(w)D(w)^{-1},$$

where estimated quantities are substituted in the right-hand sides of these equations.

Estimation proceeds by first calculating the Fourier transform of x_t^a for $t = 1, \ldots, T$,

$$x_a(w_j) = \frac{1}{T}\sum_{t=0}^{T-1} x_t^a e^{iw_j(t-1)}$$

$$w_j = \frac{2\pi j}{T} \qquad -\left[\frac{T-1}{2}\right] \leq j \leq \left[\frac{T}{2}\right]$$

where $[s]$ means the greatest integer less than or equal to s. For the purposes of calculating test statistics and coherences (though <u>not</u> for purposes of projection and prediction), the cross-spectral density matrix of x_t^a in a given frequency band is estimated as a simple average across m frequencies of the cross-periodogram ordinates

$$\hat{S}_{x^a}(w_k) = \frac{T}{m}\sum_{j\varepsilon J_k} x_a(w_j)x_a(w_j)'$$

where the prime denotes conjugation and transposition, and where J_k indexes the frequencies included in the band centered at w_k. At the band in question, the free parameters in $\Lambda_a(w_k)$ and $S_{ua}(w_k)$ are chosen by minimizing the criterion

$$(2.8) \qquad \zeta(w_k) = \ln\left|\Lambda_a(w_k)\Lambda_a(w_k)' + S_{ua}(w_k)\right| + \mathrm{tr}\left[\hat{S}_{xa}(w_k)\left(\Lambda_a(w_k)\Lambda_a(w_k)' + S_{ua}(w_k)\right)^{-1}\right],$$

which is equivalent with maximizing the likelihood function, as Lawley and Maxwell [9] and Geweke [6] note. In computing test statistics and coherences, the model is estimated across a number of nonoverlapping bands of frequencies.

For forming projections of $\Lambda(L)f_t$ on observables and also for estimating vector autoregressions, a tent-like moving average is used to estimate the spectral density matrix $S_{xa}(w)$ at overlapping bands that may be larger in number than the nonoverlapping bands used for hypothesis testing and computing coherences. This procedure sacrifices the asymptotic independence of estimated spectral statistics at different bands in exchange for a finer resolution of the spectral densities that is useful in prediction. The precise nature of the tent moving average is described in Section 7 below.

3. Measures of Fit

Under the specification of orthogonal factors with $S_f(w) = I$, (3.2) becomes

$$S_x(w) = \Lambda(w)\Lambda(w)' + S_u(w).$$

The $(i,h)^{th}$ element of $\Lambda(w)\Lambda(w)'$ is given by

$$[\Lambda(w)\Lambda(w)']_{ih} = \sum_{j=1}^{k} \Lambda_{ij}(w)\overline{\Lambda}_{hj}(w)$$

where the bar denotes complex conjugation. The term $\Lambda_{ij}\overline{\Lambda}_{hj}$ is the part of $[S_x(w)]_{ih}$, the cross spectrum between variables i and h, that is accounted for by their common dependence on the j^{th} index. It is useful to calculate the coherence

$$\frac{|\Lambda_{ij}(w)|^2}{S_x(w)_{ii}},$$

which is the proportion of variance in the i^{th} variable that is explained by the j^{th} index at frequency w. The proportion of the variance at frequency w in the i^{th} variable that is explained by all k indexes is simply

$$\frac{[\Lambda(w)\Lambda(w)']_{ii}}{S_x(w)_{ii}} = \frac{\sum_{j=1}^{k} |\Lambda_{ij}(w)|^2}{S_x(w)_{ii}}.$$

We obtain a measure of the percentage of the variance in x_i over all frequencies accounted for by the j^{th} index by taking

$$\frac{\sum_w |\Lambda_{ij}(w)|^2}{\sum_w S_x(w)_{ii}}.$$

Similarly, we obtain a measure of the percentage of the variance in x_i over all frequencies accounted for by all k indexes by

$$\frac{\sum_w \sum_{j=1}^{k} |\Lambda_{ij}(w)|^2}{\sum_w S_x(w)_{ii}}.$$

4. Projecting Unobservables on Observables

Although the indexes f_{bt}, b = 1, ..., k are unobservable, we are able to estimate the cross-second-moment matrices between $\Lambda_b(L)f_{bt}$ and x_t, where $\Lambda_b(L)$ is the b^{th} column of $\Lambda(L)$ so that $\Lambda_b(L)f_{bt}$ is the vector of "cumulative effects of the b^{th} index on x_t at time t." These moment matrices can be used to form estimates of the projections of components of $\Lambda_b(L)f_{bt}$ on past, present, and maybe future values of the observable variables x_t. For example, consider the projection equation

$$\Lambda_b(L)f_{bt} = \sum_{j=0}^{m} B_j^b x_{t-j} + \xi_t^b$$

where the B_j^b are (nxn) matrices and ξ_t^b is an (nx1) vector of least squares residuals that satisfy the orthogonality conditions $E\xi_t^b x'_{t-s} = 0$ for s = 0, 1, ..., m. The random vector $\sum_{j=0}^{m} B_j^b x_{t-j}$ is the estimator of $\Lambda_b(L)f_{bt}$ which, among estimators in the linear space spanned by $\{x_t, x_{t-1}, \ldots, x_{t-m}\}$, is closest to $\Lambda_b(L)f_{bt}$ in the mean square norm. We can construct an estimate $\sum_{j=0}^{m} B_j^b x_{t-j}$ of $\Lambda_b(L)f_{bt}$ by using the estimates of $S_x(w)$ and $\Lambda(w)\Lambda(w)'$. The idea is simply that $S_u(w)$ and $\Lambda(w)\Lambda(w)'$ contain all of the information about the covariances required to construct the projection $\sum_{j=0}^{m} B_j^b x_{t-j}$. This will permit us to calculate actual time series that estimate, for example, "the cumulative effect of the index unexpected aggregate demand on variable x_{it} at time t." The details of the projection calculation are relegated to Appendix A. Section ___ below describes time series of estimates $\Lambda_b(L)f_{bt}$.

5. Using the Unobservable Index Model to Forecast

The k-index model restricts the dimensionality of the parameter space used to model an n-variable covariance stationary vector stochastic process. As such, the model can be viewed as a device for estimating vector autoregressions via parameterizations that are parsimonious, but that still permit substantial dynamic interactions among variables.[15/] In Section ____, we report the results of using a two-index "exploratory" model to form vector autoregressions which are used to forecast outside the sample period over which they were estimated. The procedure we use was initially suggested by Sargent and Sims [16, page 63]. Here we fill in a few details not explicitly described by Sargent and Sims [16].

It is possible to construct vector autoregressions from unobservable index models in two ways which are asymptotically equivalent under the appropriate limiting procedures, but which give different results in small samples. The procedures differ in how one "recolors," that is, moves from representations for the filtered variables x_t^a to those for the unfiltered data x_t. The first procedure, which might be termed "recoloring in the frequency domain," works as follows. First, by maximum likelihood, we estimate the cross-spectral density matrix of x_t^a under the k-index hypothesis as

$$\hat{S}^k_{x^a}(w_j) = \hat{\Lambda}_a(w_j)\hat{\Lambda}_a(w_j)' + \hat{S}_{u^a}(w_j)$$

where "hats" denote maximum likelihood estimates, and $\hat{S}^{k}_{x^a}(w_j)$ is the maximum likelihood estimate of $S_{x^a}(w_k)$ under the k-unobservable-index hypothesis.[16/] The estimates $\hat{S}^{k}_{x^a}(w_j)$ are obtained by using maximum likelihood to estimate $\Lambda_a(w_j)$ and $S_{u^a}(w_j)$ from $S_{x^a}(w)$ formed from averages averaging over the periodogram ordinates at a number of overlapping bands centered at frequencies w_j, using a tent-like moving average, as described in Section 2 and in more detail in Section ____ below. Next, we obtain an estimate of the spectral density matrix of the unfiltered data under the k-unobservable-index hypothesis from

$$\hat{S}^{k}_{x}(w_j) = \hat{D}(w_j)^{-1}\hat{S}^{k}_{x^a}(w_j)\hat{D}(w_j)^{-1'}$$

where $\hat{D}(w)$ is the Fourier transform of the estimated diagonal matrix of filters $D(L)$, and the prime denotes transposition and complex conjugation. Letting $C_{x,x}(\tau) = Ex_t x'_{t-\tau}$, we then estimate the matrix cross covariogram of the x process from the inversion formula

$$\hat{C}^{k}_{x,x}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{S}^{k}_{x}(w)e^{+iw\tau}dw$$

or

$$\hat{C}^{k}_{x,x}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{D}(w)^{-1}[\hat{\Lambda}_a(w)\hat{\Lambda}_a(w)' + S_{u^a}(w)]\hat{D}(w)^{-1}e^{+iw\tau}dw.$$

These estimates of $Ex_t x'_{t-\tau} = C_{x,x}(\tau)$ are then used to compute the coefficients in the m^{th}-order vector autoregression for x_t

$$(5.1) \qquad x_t = \sum_{j=1}^{m} A^{k}_{j} x_{t-j} + \xi_t.$$

The A_j^k's are computed by substituting $\hat{C}^k_{x,x}(\tau)$ for the population $C_{x,x}(\tau)$ in the population orthogonality conditions (or Yule-Walker equations)[3/]

$$(5.2) \qquad C_{xx}(s) = \sum_{j=1}^{m} A_j^k C_{xx}(s-j) \qquad s = 1, 2, \ldots, m.$$

The estimates of the A_j^k's are then used to generate forecasts of $x_{t+1}, x_{t+2}, \ldots$, given $\{x_t, x_{t-1}, \ldots\}$, by using the chain rule of forecasting.[4/] Once the forecasts of future x_t's are generated, the deterministic parts, i.e., constant, trend, and seasonal dummy terms are added in to form forecasts of the original data.

The alternative method of forming the vector autoregression, which could be termed "recoloring in the time domain," proceeds as follows. From the spectral density of x_t^a estimated under the k-index hypothesis, $\hat{S}^k_{x^a}(w)$, we use the inversion formula to obtain an estimate of the cross covariogram $Ex_t^a x_{t-\tau}^{a'} = C_{x^a,x^a}(\tau)$,

$$\hat{C}^k_{x^a x^a}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{S}^k_{x^a}(w) e^{+iw\tau} dw$$

or

$$(5.3) \qquad \hat{C}^k_{x^a x^a}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} [\hat{\Lambda}_a(w)\hat{\Lambda}_a(w)' + \hat{S}_{u^a}(w)] e^{+iw\tau} dw.$$

Next, we use these estimates of $C_{x^a x^a}(\tau)$ to calculate the p^{th}-order vector autoregression for x_t^a,

$$(5.4) \qquad x_t^a = \sum_{j=1}^{p} \hat{F}^k_j x^a_{t-j} + \varepsilon_t^a$$

where the ε_t^a's satisfy the orthogonality conditions $Ex_{t-j}^a \varepsilon_t^{a'} = 0$, $j = 1, \ldots, p$. The $\hat{F}^k_j$'s are calculated from the normal equations

$$C_{x^a x^a}(s) = \sum_{j=1}^{p} \hat{F}_j^k C_{x^a x^a}(s-j), \qquad s = 1, \ldots, p,$$

where the $C_{x^a x^a}(\tau)$'s are estimated subject to the k-index model from (5.3). To obtain the implied vector autoregression for x_t, we write (5.4) as

$$\hat{F}^k(L)x_t^a = \varepsilon_t^a$$

where $\hat{F}^k(L) = I - \hat{F}_1^k L - \ldots - \hat{F}_p^k L^p$. Then substituting $x_t^a = \hat{D}(L)x_t$ into the above equation gives

$$\hat{F}^k(L)\hat{D}(L)x_t = \varepsilon_t^a$$

or

$$\hat{A}(L)x_t = \varepsilon_t \tag{5.5}$$

where

$$\hat{A}_s = \sum_{j=0}^{p} \hat{F}_j^k \hat{D}_{s-j}, \qquad s = 0, \ldots, p+2, \tag{5.6}$$

and where $\hat{F}_j^k$'s and $\hat{D}_j$'s not previously defined (for example, corresponding to negative subscripts) are defined as zero. It is an implication of (5.6) that $\hat{A}_0 = I$. Once the $\hat{A}_s$'s have been formed, the rest of the forecasting procedure is as described above for the "frequency domain recoloring procedure."

For systems estimated over a small number of frequency bands, our preference is for the procedure that recolors in the time domain. The reason is that with a small number of bands, the frequency domain procedure gives a crude approximation to the spectral density matrix of an autoregressive process. We have found that when only a small number of bands are used with the

frequency domain recoloring procedure, forecasts are generated that exhibit high first-order serial correlation of residuals, even within the estimation period, indicating that the model is not even accounting for the low-order, own-serial correlation of each series. This problem can be avoided either by using the time domain recoloring procedure, or by retaining frequency domain recoloring but increasing the number of bands over which the model is estimated, or perhaps by interpolating to generate more points in $\hat{S}_x^k(w)$ before inverse Fourier transforming to obtain $C_{xx}(s)$.

Having obtained estimates of the A_j's in (6.5), we can obtain a vector moving average for x_t of the form

$$x_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$$

where C_j is an (nxn) matrix and $\{\varepsilon_t\}$ is the (nx1) vector of innovations in the x_t process. The C_j's are related to the A_j's by

$$(5.7) \qquad (I-A_1 e^{-iw}-\ldots-A_m e^{-iwm})^{-1} = \sum_{j=0}^{\infty} C_j e^{-iwj} = C(w).$$

The C_j's can be obtained either by inverse Fourier transforming $C(w)$, or else by recursively solving the difference equations (5.5) in response to inputs ε_t that are identically zero except for ε_t being a unit vector at time 0. Such vector moving average representations from unconstrained vector autoregressions are used by Sims [20] as a vehicle for making some economic interpretations. Below we present the vector moving average representation for one of our index models.

Appendix A

Projecting Unobservables on Observables

In this appendix we describe how we calculate the projection of $\Lambda_b(L)f_{bt}$ on $\{x_t, x_{t-1}, \ldots, x_{t-m}\}$. That is, we describe how we calculate the B_j's in the projection equations

$$\Lambda_b(L)f_{bt} = \sum_{j=0}^{m} B_j^b x_{t-j} + \xi_t^b$$

where

$$Ex_{t-j}\xi_t^{b'} = 0 \qquad \text{for } j = 0, \ldots, m.$$

By estimating model (3.1), and identifying $\Lambda_b(w)$ for $b = 1, \ldots, k$, we recover an estimate of the cross spectrum between x_t and $\Lambda_b(L)f_{bt}$. In particular,

$$S_{x,\Lambda_b f_b}(w) = \Lambda_b(w)\Lambda_b(w)'$$

$$= \sum_{=-\infty}^{\infty} C_{x,\Lambda_b f_b}(\tau)e^{-iw\tau}$$

where $C_{x,\Lambda_b f_b}(\tau) = Ex_t \cdot (\Lambda(L)f_{bt-\tau})'$. By computing the inverse Fourier transform of the nxn matrix $\Lambda_b(w)\Lambda_b(w)'$, we can recover the (nxn) matrix cross covariogram $C_{x,\Lambda_b f_b}(\tau)$:

$$C_{x,\Lambda_b f_b}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Lambda_b(w)\Lambda_b(w)'e^{+iw\tau}dw.$$

The spectral matrix of the (nx1) vector process $\Lambda_b(L)f_{bt}$ is given by

$$S_{\Lambda_b f,\Lambda_b f}(w) = \Lambda_b(w)\Lambda_b(w)'.$$

It follows that the spectral matrix of the (nx1) vector process $\Lambda_b(L)f_{bt}$ equals the cross-spectral matrix between x_t and $\Lambda_b(L)f_{bt}$. It also follows that the matrix covariogram of the process $\Lambda_b(L)f_{bt}$ equals the matrix cross covariogram between x_t and $\Lambda_b(L)f_{bt}$.

For example, let

$$\Lambda_b(L) = \begin{pmatrix} \sum_{i=-\infty}^{\infty} \lambda_{1i} L^i \\ \sum_{i=-\infty}^{\infty} \lambda_{2i} L^i \\ \vdots \\ \sum_{i=-\infty}^{\infty} \lambda_{ni} L^i \end{pmatrix} .$$

Then the matrix $C_{x,\Lambda_b f_b}(\tau) = C_{\Lambda_b f_b, \Lambda_b f_b}(\tau)$ is given by

$$\begin{pmatrix} \sum_{i=-\infty}^{\infty} \lambda_{1i}\lambda_{1\tau+i} & \sum_{i=-\infty}^{\infty} \lambda_{2i}\lambda_{1\tau+i} & \cdots & \sum_{i=-\infty}^{\infty} \lambda_{ni}\lambda_{1\tau+i} \\ \sum_{i=-\infty}^{\infty} \lambda_{1i}\lambda_{2\tau+i} & \sum_{i=-\infty}^{\infty} \lambda_{2i}\lambda_{2\tau+i} & \cdots & \sum_{i=-\infty}^{\infty} \lambda_{ni}\lambda_{2\tau+i} \\ \vdots & & & \\ \sum_{i=-\infty}^{\infty} \lambda_{1i}\lambda_{n\tau+i} & \sum_{i=-\infty}^{\infty} \lambda_{2i}\lambda_{n\tau+i} & & \sum_{i=-\infty}^{\infty} \lambda_{ni}\lambda_{n\tau+i} \end{pmatrix} .$$

Notice that this matrix is not symmetric, unless $\tau = 0$.

The cross-spectral matrix of x_t with $\Lambda(L)f_t$ equals the spectral matrix of $\Lambda(L)f_t$ with itself and is given by

$$S_{x,\Lambda f} = S_{\Lambda f, \Lambda f} = \Lambda(w)\Lambda(w)'$$

$$= \sum_{b=1}^{k} \Lambda_b(w)\Lambda_b(w)' .$$

The cross covariogram between x_t and $\Lambda(L)f_t$ can be obtained by inverse Fourier transforming $\Lambda(w)\Lambda(w)'$, or by simply summing over b the cross covariogram of x_t with $\Lambda_b(L)f_{bt}$:

$$C_{x,\Lambda_f}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Lambda(w)\Lambda(w)'e^{+iw\tau}dw$$
$$= \sum_{b=1}^{k} C_{x,\Lambda_b f_b}(\tau)$$

where

$$C_{x,\Lambda f}(\tau) = Ex_t \cdot (\Lambda(L)f_{t-\tau})'.$$

Further, the spectral density matrix of x_t, as constrained by the model, is given by $\Lambda(w)\Lambda(w)' + S_u(w)$. By inverse Fourier transforming this, we estimate $C_{x,x}(\tau) = Ex_t x'_{t-\tau}$.

The estimates of the cross covariograms $C_{x,x}(\tau)$ and $C_{x,\Lambda_b f_b}(\tau)$ can be used to estimate the projection of $\Lambda_b(L)f_{bt}$ on x_t. For example, we can project the (nx1) process $\Lambda_b(L)f_{bt}$ on current and m lagged x's,

$$(A.1) \qquad \Lambda_b(L)f_{bt} = \sum_{j=0}^{m} B_j^b x_{t-j} + \xi_t^b$$

where the B_j^b are (nxn) matrices and ξ_t^b is an (nx1) vector of least squares residuals that satisfy the orthogonality conditions $E\xi_t^b x'_{t-s} = 0$ for $s = 0, \ldots, m$. Postmultiplying (A.1) by x'_{t-s} and taking expectations gives the least squares normal equations

$$E(\Lambda_b(L)f_{bt})x'_{t-s} = \sum_{j=0}^{m} B_j^b \, Ex_{t-j}x'_{t-s}$$

or

$$C_{\Lambda_b f_b,x}(s) = \sum_{j=0}^{m} B_j^b C_{x,x}(s-j), \qquad s = 0, \ldots, m.$$

Noting that $C_{\Lambda_b f_b, x}(s)' = C_{x, \Lambda_b f_b}(-s)$ and transposing both sides, we obtain

$$C_{x,\Lambda_b f_b}(-s) = \sum_{j=0}^{m} C_{xx}(s-j)'B_j^{b'}, \qquad s = 0, \ldots, m.$$

Notice that $C_{x,x}(\tau)' = Ex_{t-\tau}x_t' = C_{xx}(-\tau)$, so we can write

$$C_{x,\Lambda_b f_b}(-s) = \sum_{j=0}^{\infty} C_{xx}(j-s)B_j^{b'}, \qquad s = 0, \ldots, m.$$

Stacking the above equations for s = 0, ..., m gives

$$\text{(A.2)} \qquad \begin{bmatrix} C_{x,\Lambda_b f_b}(0) \\ C_{x,\Lambda_b f_b}(-1) \\ \vdots \\ C_{x,\Lambda_b f_b}(-m) \end{bmatrix} = \begin{bmatrix} C_{xx}(0) & C_{xx}(1) & \ldots & C_{xx}(m) \\ C_{xx}(-1) & C_{xx}(0) & \ldots & C_{xx}(m-1) \\ \vdots & & & \\ C_{xx}(-m) & C_{xx}(-m+1) & \ldots & C_{xx}(0) \end{bmatrix} \begin{bmatrix} B_0^{b'} \\ B_1^{b'} \\ \vdots \\ B_m^{b'} \end{bmatrix}.$$

Assuming that the n(m+1) x n(m+1) matrix on the right is of full rank, the normal equations (A.2) uniquely determine the B_j^b's, j = 0, ..., m.

Consider the i^{th} element of equation (A.1),

$$\Lambda_{i,b}(L)f_{bt} = \sum_{j=0}^{m} B_{ij}^b x_{t-j} + \xi_{it}^b$$

where $\Lambda_{i,b}(L)$ is the i^{th} row of $\Lambda_b(L)$ and B_{ij}^b is the i^{th} row of B_j^b. Squaring this equation and taking expected values gives

$$E(\Lambda_{ib}(L)f_{bt})^2 = \sum_{j=0}^{m} B_{ij}^b \sum_{s=0}^{m} Ex_{t-j}x_{t-s}' B_{is}^{b'} + E\xi_{it}^{b\,2}$$

or

$$\text{(A.3)} \qquad E(\Lambda_{ij}(L)f_t)^2 = \sum_{j=0}^{m} B_{ij}^b \sum_{s=0}^{m} C_{xx}(s-j)B_{is}^{b'} + E\xi_{it}^{b\,2}.$$

Substituting from (A.2), we can write the double sum in matrix terms as

$$\begin{bmatrix} B^b_{i,0} B^b_{i,1} \cdots B^b_{i,m} \end{bmatrix} \begin{bmatrix} C_{xx}(0) & C_{xx}(1) \cdots \\ C_{xx}(-1) & C_{xx}(0) \cdots \\ \vdots & \end{bmatrix} \begin{bmatrix} C_{xx}(0) & C_{xx}(1) \cdots \\ C_{xx}(-1) & C_{xx}(0) \\ \vdots & \end{bmatrix}^{-1} \begin{bmatrix} C_{x,\Lambda_b f}(0) \\ C_{x,\Lambda_b f}(-1) \\ \\ C_{x,\Lambda_b f}(-m) \end{bmatrix} \leftarrow i^{th} \text{ column}$$

or

$$\begin{bmatrix} B^b_{i,0} B^b_{i,1} \cdots B^b_{i,m} \end{bmatrix} \qquad \begin{bmatrix} C_{x,\Lambda_b f_b}(0) \\ C_{x,\Lambda_b f_b}(-1) \\ C_{x,\Lambda_b f_b}(-m) \end{bmatrix} \quad i^{th} \text{ column.}$$

We estimate $E(\Lambda_{ib}(L)f_t)^2$ from the Fourier transform of the i^{th} diagonal element of $\Lambda_b(w)\Lambda_b(w)'$,

$$E(\Lambda_{ib}(L)f_t)^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} [\Lambda_b(w)\Lambda_b(w)']_{ii}\, dw.$$

The variance decomposition (A.3) allows us to compute the percentage of the variance in the effect of the b^{th} index on the i^{th} variable x_{it}, that is explained by $\{x_t, \ldots, x_{t-m}\}$.

By summing equation (A.1) over b, we obtain the projection of $\Lambda(L)f_t$ on x_{t-j}, $j = 0, \ldots, m$:

$$\Lambda(L)f_t = \sum_{b=1}^{k} \Lambda_b(L)f_{bt}$$

$$= \sum_{j=0}^{m} \left(\sum_{b=1}^{k} B_j^b \right) x_{t-j} + \sum_{b=1}^{k} \xi_t^b.$$

$$= \sum_{j=0}^{m} B_j x_{t-j} + \xi_t$$

where $\xi_t = \sum_{b=1}^{k} \xi_t^b$ and $B_j = \sum_{b=1}^{k} B_j^b$, and ξ_t by construction satisfies the least squares orthogonality conditions $E\xi_t x'_{t-j} = 0$, $j = 0, \ldots, m$.

Appendix B

Here we describe the methods used to generate the forecasts, the Theil U statistics for which are reported in Section 8. Let x_{it} be the original data for the series listed in Table 1. The following steps were employed:

1. Deterministic components are removed from each x_{it} by regressing on a constant trend and the seasonal dummies over the period 1948I-1971II. The residuals from those regressions are taken as the series x_{it}. Next, each x_{it} is regressed on two lagged values to create estimates of the diagonal matrix of filters $D(L)$. The filtered data $x^a_{it} = D_i(L)x_{it}$ are formed.

2. The finite Fourier transform of the x^a_{it}, call it $\tilde{x}^a_i(w_j)$, is calculated for $w_j = 2\pi j/T$, $j = 0, 1, \ldots, T/2$. To use the fast Fourier transform, the x^a_{it}'s are filled out with zeroes to create a series of length 128, so that $T = 128$. The $\tilde{x}^a_i(w_j)$'s are used to generate the cross periodogram.

3. The cross periodogram is smoothed to produce an estimate of the cross-spectral density matrix at frequencies centered at each of m bands. A "tent" filter is used, which is 29 frequencies (w_j's) wide. The center of the tent is 1 1/2 times the height at the ends. The first band is always centered at zero angular frequency, the last at π.

4. A two-index exploratory model is estimated by maximum likelihood at each of the frequencies for the x^a_{it}. That is, we estimate

$$\hat{S}^k_{x^a}(w_j) = \hat{\Lambda}_a(w_j)\Lambda_a(w_j)' + \hat{S}_{u^a}(w_j).$$

5. The matrix covariogram of x_t^a is calculated from the inversion formula

$$\hat{C}^k_{x^a x^a}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} [\hat{\Lambda}_a(w)\hat{\Lambda}_a(w)' + \hat{S}_{u^a}(w)]e^{+iw\tau}dw.$$

These estimates of the covariogram are used in the least squares normal equations

$$\hat{C}_{x^a x^a}(s) = \sum_{j=1}^{p} \hat{F}^k_j \hat{C}_{x^a x^a}(s-j), \qquad s = 1, \ldots, p$$

to estimate the coefficient $\hat{F}^k_j$ of a p^{th}-order autoregression for x_t^a.

6. An autoregressive representation for x_t is formed, as described in the text, from

$$\hat{A}(L) = \hat{F}^k(L)D(L)$$

where $\hat{A}(L) = \sum_{j=0}^{p+2} \hat{A}^k_j L^j$, and $\hat{A}^k_j$ is the matrix of coefficients on the j^{th} lag of x_t in the matrix autoregressive representation of x_t.

7. The chain rule of forecasting is used to form the forecast $E[x_{s+j}|x_s, x_{s-1}, \ldots, x_{s-p-2}]$ for s = 1971II, 1971III, 1971IV, and 1972I, and for j = 1, ..., 8.

8. Given the predicted x_{it}'s, the deterministic components are added back in to form forecasts of the original x_{it}. These forecasts are used together with actual outcomes to produce sets of forecast errors for forecasts made in 1971II, 1971III, 1971IV, and 1972I. Notice that the model is not reestimated every quarter (to save money).

9. The model is reestimated every year. Data from an additional year is added, and all of steps 1 through 8 are re-

peated. After the forecasts and forecast errors are computed, another year of data is added and steps 1 through 8 are repeated again. This continues through data for 1975II, when the last reestimation occurs. The last forecast was made in 1976I. Notice that since data through only 1976IV were used in this study, forecast errors were not calculated for the longer horizons for the last few dates on which forecasts were made. This accounts for why the Theil U statistics are based on fewer observations for the longer horizons.

The univariate autoregressions used as benchmarks were fourth order, each included a constant term. These autoregressions were reestimated at the same dates that the index models were reestimated.

The Theil U statistics reported in Section 8 were obtained from a system with a number of frequency bands between 0 and π, equal to 8, and with the order of the autoregression p equal to 9.

Appendix C

Index Models with Neutral Price Variations

We focus on a class of models in which the behavior of private economic agents makes the force "nominal aggregate demand" impinge on real quantities and nominal prices in ways sufficiently different that it requires two distinct indexes to account for the effects of nominal aggregate demand. In particular, we want to admit the possibility that agents respond differently to expected and to unexpected movements in nominal aggregate demand.

A statistical model that admits this possibility can be written

$$y_t = a(L)(n_t - \hat{n}_t) + b(L)\hat{n}_t + u_{yt}$$

$$p_t = c(L)(n_t - \hat{n}_t) + d(L)\hat{n}_t + u_{pt}. \qquad \text{(C.1)}$$

Here y_t is an (rx1) vector of real aggregate economic variables, and p_t is an (hx1) vector of nominal economic variables, such as nominal prices; a, b, c, and d are vectors of polynomials in the lag operator L, each being square summable and one-sided in non-negative powers of L. The (rx1) vector u_{yt} and the (hx1) vector u_{pt} consist of random variables whose variances are small relative to the variances of the left-hand-side variables to which they correspond. Thus, (C.1) is asserted to "fit well." Below, the u_t's will be required to satisfy some additional conditions which we must impose to proceed with estimation. In (C.1), n_t is "nominal aggregate demand" and $\hat{n}_t$ is the level of nominal aggregate demand that the public had expected to prevail as of some date

possibly, but not necessarily, earlier than t. Equation (C.1) permits unexpected movements in demand to affect prices and quantities differently than expected movements in demand.

A special case of the model occurs when the distinction between expected and unexpected aggregate demand makes no difference, so that $a(L) = b(L)$ and $c(L) = d(L)$. Here the idea is that agents respond to unexpected components in aggregate demand in the same way that they respond to expected components. In this case, (C.1) collapses to the one-index model

$$y_t = a(L)n_t + u_{yt}$$

$$p_t = c(L)n_t + u_{pt}.$$

This model permits nominal aggregate demand to influence y_t and p_t differently to the extent that this difference can be captured by the different polynomials $a(L)$ and $c(L)$.

A second case in which (C.1) collapses to a one-index model is the case in which $n_t = \hat{n}_t$, or more generally, when the variance of $n_t - \hat{n}_t$ is sufficiently small that little goodness of fit is lost by lumping the terms in $(n_t - \hat{n}_t)$ with the disturbance u_t. This is another way of representing the notion that the distinction between unexpected and expected aggregate demand is unimportant, in this particular case because agents presumably have full enough information sets and short enough information lags to prevent large squared forecasting errors on average.

A third case in which (C.1) collapses to a one-index model is one in which expected aggregate demand $\hat{n}_t$ is a function only of lagged values of n_t, say

$$\hat{n}_t = f(L)n_t$$

where $f(L) = \sum_{j=1}^{\infty} f_j L^j$. In this case, (C.1) can be written in the one-index form

$$y_t = \{a(L)(I-f(L))+b(L)f(L)\}n_t + u_{yt}$$

$$p_t = \{c(L)(I-f(L))+d(L)f(L)\}n_t + u_{pt}.$$

This case is of practical importance, especially if nominal aggregate demand is thought to be observable to private agents with a short lag, thereby making it likely that $\hat{n}_t = f(L)n_t$ provides a good approximation in the mean squared error sense to the actual law by which $\hat{n}_t$ is formed.

We assume that none of these degenerate cases obtains, and that it requires two indexes to account for the covariation of y_t and p_t and to fit well. To avoid the last-mentioned degeneracy, we follow Lucas and posit that nominal aggregate demand $\{n_t\}$ is unobservable to agents: agents don't even see lagged values of n_t. At time t, agents have observations on a set of variables Ω_t. Agents are assumed to form their expectation of n_t according to

$$\text{(C.2)} \qquad \hat{n}_t = \sum_{j=0}^{\infty} h_j \Omega_{t-j}$$

where $\{h_j\}$ is a vector of distributed lag weights. At this point, we have imposed no restrictions on the expectation generating process (C.2) except to require that its outcome $\hat{n}_t$ cannot be well approximated by a process lying in the space spanned by past n_t's. In particular, we have not imposed that expectations be

"rational," as we could be requiring that $\hat{n}_t$ be the projection of n_t on the space spanned by $\{\Omega_1, \Omega_{t-1}, \ldots\}$.

A version of (C.1) which represents the "natural rate" or "neutrality" hypothesis is one in which $b(L) = 0$ and $d(L) \neq 0$. Here the model becomes

$$y_t = a(L)(n_t - \hat{n}_t) + u_{yt}$$

(C.3)

$$p_t = c(L)(n_t - \hat{n}_t) + d(L)\hat{n}_t + u_{pt}.$$

We intend that $d(L)\hat{n}_t$ account for a substantial proportion of the variance of p_t. The force of (C.3) is that expected changes in aggregate demand leave real quantities unaltered, but do affect nominal quantities. In effect, (C.3) asserts that a one-index model is adequate to account for most of the covariation of real quantities, but that a two-index model is required when two or more nominal variables are added to the system.

If the information set Ω_t in (C.2) includes current and/or lagged observations on (y_t, p_t), then substituting (C.2) into (C.1) or (C.3) gives rise to a mixed observable index, unobservable index model. An unobservable index arises from the fact that n_t itself is not directly observed either by private agents or the econometrician, making $n_t - \hat{n}_t$ unobservable. An observable index emerges if current and/or lagged (y_t, p_t)'s are important components of Ω_t in (C.2), making $\hat{n}_t$ observable. While the mixed observable-unobservable index model is probably the most faithful statistical representation of the theoretical ideas described above, practical methods are not presently available for estimat-

ing such models with large dimension. For this reason, we shall somewhat compromise the theoretical ideas by fitting unobservable index models. As aruged by Sargent and Sims [16], this will do little harm to the extent that the index model fits the data well.

To begin preparing the model for empirical implementation, we shall renormalize the model formed by (C.2) and (C.3) as follows. Form the projection of $\hat{n}_t$ on past, present, and future values of the $(n_t-\hat{n}_t)$ process,

$$\hat{n}_t = \left(\sum_{j=-\infty}^{\infty} g_j L^j\right)(n_t-\hat{n}_t) + \eta_t \tag{C.4}$$

where η_t obeys the least squares orthogonality conditions $E(n_t-\hat{n}_t)\cdot \eta_s = 0$ for all t and s. Substituting (C.4) into (C.3) gives

$$\begin{aligned} y_t &= a(L)(n_t-\hat{n}_t) + u_{yt} \\ p_t &= \{c(L)+d(L)g(L)\}(n_t-\hat{n}_t) + d(L)\eta_t + u_{pt}. \end{aligned} \tag{C.5}$$

The system (C.5) is in the form of a two-index model in which the two indexes are $(n_t-\hat{n}_t)$ and η_t, "unexpected aggregate demand" and "that part of expected aggregate demand that is orthogonal to the entire $n - \hat{n}$ process," respectively. The cross covariogram between $n_t - \hat{n}_t$ and η_t is identically zero, as is the cross spectrum. Notice that the polynomial in the lag operator g(L) is in general two-sided.

The preceding theory imposes no restrictions on the vector $u_t = (u_{yt},u_{pt})'$ other than that Eu_tu_t' be small relative to the variances of y and p. To proceed with estimation, we impose

the conditions on u_t required to convert (C.5) into the unobservable index model of Sargent and Sims [16]. In particular, we require (i) that the spectral density matrix of the vector u_t be diagonal, so that each component of u is orthogonal to the past, present, and future of every other component of u; (ii) that u_t be orthogonal to the past, present, and future of both $(n_t - \hat{n}_t)$ and η_t. With these conditions on $\{u_t\}$, (C.5) is an "unobservable index" or "dynamic factor" model with a set of zero identifying restrictions. As emphasized by Sargent and Sims [16], the side conditions that have been imposed on the u's have no foundation in the theoretical ideas motivating the model. On the contrary, if the observable series (y_t, p_t) are elements of the information set Ω_t in (C.2), one would, in general, expect the orthogonality condition (ii) to be violated. The conditions are imposed not because they are believed to be true, but because some such conditions must be imposed to proceed with estimation.

To transform the model (C.5) into a form matching (2.1), let $n_t - \hat{n}_t$ and η_t have univariate Wold moving average representations

$$n_t - \hat{n}_t = \sum_{j=0}^{\infty} v_{1j} \varepsilon_{1t-j} = v_1(L) f_{1t}$$

$$\eta_t = \sum_{j=0}^{\infty} v_{2j} \varepsilon_{2t-j} = v_2(L) f_{2t}$$

where $\{v_{1j}\}$ and $\{v_{2j}\}$ are each square summable sequences, and where f_{1t} is a fundamental white noise for $n_t - \hat{n}_t$ and f_{2t} is a fundamental white noise for η_t. We are free to normalize so that $Ef_{1t}^2 = Ef_{2t}^2 = 1$. Since $E(n_t - \hat{n}_t) \cdot \eta_s = 0$ for all t and s, t follows that $Ef_{1t}f_{2s} = 0$ for all t and s. System (C.5) can now be written in the form of (2.1), with the definitions

$$x_t = \begin{pmatrix} y_t \\ p_t \end{pmatrix}, \quad \Lambda(L) = \begin{pmatrix} a(L)v_1(L) & 0 \\ [c(L)+d(L)g(L)]v_1(L) & d(L)v_2(L) \end{pmatrix}$$

$$u_t = \begin{pmatrix} u_{yt} \\ u_{pt} \end{pmatrix}, \quad f_t = \begin{pmatrix} f_{1t} \\ f_{2t} \end{pmatrix} .$$

REFERENCES

1. Anderson, T. W., and H. Rubin. 1956, "Statistical Inference in Factor Analysis," in J. Neyman, ed., Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability 5, (University of California Press, Berkeley) (A); 111-150.

2. Barro, R. J. 1976, "Rational Expectations and the Role of Monetary Policy," Journal of Monetary Economics 2 (January); 1-31.

3. Brillinger, David. 1975, Time Series Data Analysis and Theory, Holt, Rinehart, and Winston; Chapters 9 and 10.

4. Chow, Gregory C. 1973, "Multiperiod Predictions from Stochastic Difference Equations by Bayesian Methods," Econometrica 41; 109-118.

5. Fair, Ray C. 1978b, "Estimating the Expected Prediction Accuracy of Econometric Models," Cowles Foundation Discussion Paper 480, forthcoming in International Economic Review.

6. Geweke, John F. 1977, "The Dynamic Factor Analysis of Economic Time Series," in D. Aigner and A. Goldberger, eds., Latent Variables in Socio-economic Models (North Holland, Amsterdam); 365-383.

7. Geweke, John F., and K. J. Singleton. 1978, "Confirmatory Maximum Likelihood Factor Analysis of Economic Time Series," manuscript, revised 1978.

8. Koopmans, T. C., and O. Reirsøl. 1950, "The Identification of Structural Characteristics," Annals of Mathematical Statistics, Vol 21; 165-181.

9. Lawley, D. N., and A. E. Maxwell. 1971, Factor Analysis as a Statistical Method (Butterworths, London), second edition.

10. Litterman, Robert B. 1979, "Techniques of Forecasting Using Vector Autoregressions," unpublished Ph.D. thesis, University of Minnesota.

11. Lucas, R. E., Jr. 1973, "Some International Evidence on Output-Inflation Trade-offs," American Economic Review 63 (June); 326-334.

12. _____. 1972, "Econometric Testing of the Natural Rate Hypothesis." In The Econometrics of Price Determination Conference, ed. by O. Eckstein. Washington: Board of Governors of the Federal Reserve System; 50-59.

13. _____. 1975, "An Equilibrium Model of the Business Cycle," Journal of Political Economy (December).

14. McCallum, Bennett T. 1977, "Price-Level Stickiness and the Feasibility of Monetary Stabilization Policy with Rational Expectations," Journal of Political Economy 85 (June).

15. Priestly, M. B., T. S. Rao, and H. Tong. 1974, "Applications of Principal Components and Factor Analysis in the Identification of Multivariable Systems," IEEE Transactions on Automatic Control, Vol. AC-19, 6 (December); 730-734.

16. Sargent, Thomas J., and Christopher A. Sims. 1977, "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory," in C. A. Sims, ed., New Methods in Business Cycle Research, Federal Reserve Bank of Minneapolis.

17. Sargent, Thomas J. 1973, "Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment," Brookings Papers on Economic Activity 2; 429-472.

18. _____. 1979, Macroeconomic Theory, Academic Press; Chapters X and XI.

19. Sims, Chirstopher A. 1975, "Strategies for Controlling Dimensionality in Models of Multivariate Time Series." Paper presented to Third World Congress of the Econometrics Society (August), Toronto, manuscript (revision in preparation).

20. _____. "Macroeconomics and Reality," University of Minnesota discussion paper, forthcoming in Econometrica.

21. _____. 1972, "The Role of Approximate Prior Restrictions in Distributed Lag Estimation," Journal of the American Statistical Association, Vol 67, 337 (March); 169-75.