A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates

ROBERT ENGLE and MARK WATSON*

The paper formulates and estimates a single-factor multivariate time series model. The model is a dynamic generalization of the multiple indicator (or factor analysis) model. It is shown to be a special case of the general state space model and can be estimated by maximum likelihood methods using the Kalman filter algorithm. The model is used to obtain estimates of the unobserved metropolitan wage rate for Los Angeles, based on observations of sectoral wages within the Standard Metropolitan Statistical Area. Hypothesis tests, model diagnostics, and out-of-sample forecasts are used to evaluate the model.

KEY WORDS: State space model; Dynamic factor analysis; Kalman filter; Method of scoring; Unobserved component estimation.

1. INTRODUCTION

Much of the growth and decline of regional economies can be attributed to changes in comparative advantage, and the single most important component of this comparative advantage is probably wage rates. Therefore, considerable interest centers on the measurement of local wage rates and on the determinants of their movements. Because a region within a national economy can be thought of as a very open economy, there are strong economic pressures for wages to equalize between regions, both through commodity trade which tends to equate factor prices and through regional migration of labor and capital. For further discussion of these issues, see Engle (1974).

The measurement of a regional wage rate and its determinants is complicated by the differing wage in different industries and by differing skill mixes in different industries. In this article a statistical technique will be employed to separate movements in a metropolitan wage rate into a national industrial component, a metropolitan area-wide component, and a local industry specific component. For example, the wage rate in contract construction in Los Angeles will be decomposed into one component determined by the wage rate in contract construction in the United States, a second determined by the overall wage rate in Los Angeles, and a third resulting from factors particular to Los Angeles contract construction.

There are good economic reasons for expecting each of these components to be important. The national component measures not only changes in the U.S. economy as a whole through inflation and business cycles, but also measures changes in technology, changes in preferences, changes in the supply or demand for the output of the industry nationally, and collective bargaining outcomes that may affect industrial wages for a broad geographical region. The metropolitan component reflects the demand and supply of labor in the metropolitan labor market. Presumably, no industry can avoid the effect of the local labor market entirely, but some may be more strongly influenced than others. This component would reflect migration patterns of capital and labor, the cost of living in the region, and the tightness of the local labor market. The specific effect is the remainder which measures situations peculiar to this industry and region. By definition, the three effects are independent.

To illustrate the problem, consider the least squares regression of the log of the wage rate in industry i in Los Angeles, w_{it}, on the log of the national wage rate in this industry, n_{it}, using annual data. The residuals from this regression are composed of metropolitan effects and local industry specific effects. The metropolitan effects are common to each industry and therefore produce correlation across industries while the specific effects are by definition independent of other industries. In Table 1, these regressions and residual correlations are presented; the large cross-sectional correlations suggest the importance of the metropolitan effect. A factor analysis of these residual correlations indicates that one factor could explain 70 percent of the variance.

Because the data are a time series of cross-sections, the dynamic effects must also be considered and standard factor analysis is not appropriate. The first-order lagged correlation matrix, also presented in Table 1, shows the importance of the dynamics in the data set. Cross-correlations between sectors must result from serial correlation in the metropolitan component, while autocorrelations could arise from serial correlation in the specific effect. The frequency domain version of factor analysis of Geweke (1977) and Geweke and Singleton (1981) can

Table 1. Ordinary Least Squares Results:
$$W_t = a_0 + a_2 n_t + e_t$$

Sector	a_0	a_2	DW	R^2
Contract constructions	.1371	1.0685	.24	.99
	(.026)	(.019)		
Durable manufactures	.1396	.9282	.20	.99
	(.018)	(.017)		
Nondurable manufactures	.2497	.8826	.17	.99
	(.013)	(.014)		
Wholesale trade	.1315	1.0231	.73	.99
	(.011)	(.010)		
Retail trade	.3493	.9640	.38	.99
	(.011)	(.016)		

Residual Correlation:

Sector	CC	DM	NM	WT	RT
Contract constructions	1.000				
Durable manufactures	.728[a]	1.000			
Nondurable manufactures	.717[a]	.936[a]	1.000		
Wholesale trade	.713[a]	.304	.350	1.000	
Retail trade	.872[a]	.711[a]	.778[a]	.515[a]	1.000

Correlation of Residuals with Lagged Residuals:

Sector	CC-1	DM-1	NM-1	WT-1	RT-1
Contract construction	.878[a]	.641[a]	.645[a]	.621[a]	.768[a]
Durable manufactures	.796[a]	.891[a]	.893[a]	.542[a]	.703[a]
Nondurable manufactures	.799[a]	.901[a]	.901[a]	.567[a]	.732[a]
Wholesale trade	.700[a]	.173	.297	.614[a]	.544[a]
Retail trade	.844[a]	.710[a]	.725[a]	.524[a]	.810[a]

[a] Significant at the 95% level.

be applied to this problem. Their test of a one-factor model against an unrestricted vector stationary process accepts the one-factor model with a statistic of 11.97, which is distributed as chi-squared with 11 degrees of freedom under the null hypothesis. Thus, it appears appropriate to seek a dynamic one-factor model of these data where estimated values of the factor would be interpreted as the movements in the metropolitan component of the wage rate.

One would presumably like to estimate the coefficients on the national wage rates jointly with the factor and bring as much economics to bear on the separation of the components as possible. Consequently, several models are proposed for the metropolitan component that include varying amounts of economic information.

In the second section we introduce the general state space model, and the third section discusses the estimation and diagnostic problems of the general model. Section 4 describes the economics of the formulation, while Section 5 presents the estimation results and tests the forecasting ability of the model.

2. GENERAL FORMULATION OF THE MODEL

The models considered in this paper are special cases of the "state space" model used in engineering to represent a variety of physical processes. In fact, a wide range of models used in econometrics can be viewed as

special cases of state space models, as will be shown. An introduction and comparison between econometric and engineering applications is given in Mehra (1974). The advantage of viewing the models in this way is that general maximum likelihood estimates are available based upon the Kalman filter recursive algorithm.

The state space model consists of two sets of equations: transition or process equations and measurement equations. The transition equations describe the evolution of a $J \times 1$ vector x_t of characteristics of a physical process in response to a $K \times 1$ vector z_t of exogenous or lagged dependent variables and a $J \times 1$ vector v_t of disturbances. The state vector x_t is unobservable and hence corresponds to the unobserved components which are to be isolated in this paper. The measurement equations describe the relation between the unobserved state x_t and a $P \times 1$ vector of measurements y_t. The predetermined variables z_t and another vector of disturbances e_t may also enter the measurement equation.

The model can be specified as

$$\underset{J \times 1}{x_t} = \underset{J \times J}{\phi} \underset{J \times 1}{x_{t-1}} + \underset{J \times K}{\gamma} \underset{K \times 1}{z_t} + \underset{J \times 1}{v_t} \qquad (2.1)$$

$$\underset{P \times 1}{y_t} = \underset{P \times J}{\alpha} \underset{J \times 1}{x_t} + \underset{P \times K}{\beta} \underset{K \times 1}{z_t} + \underset{P \times 1}{e_t} \qquad (2.2)$$

and

$$\begin{pmatrix} v_t \\ e_t \end{pmatrix} \sim \text{N.I.D} \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix}. \qquad (2.3)$$

Regression models, time series models, and regression models with ARIMA disturbances are special cases of the state space model when $P = 1$. In some cases such as in Harvey and Phillips (1979) and Hannan (1976) this formulation has computational or theoretical advantages. More interestingly, unobserved components models such as Nerlove (1971), Pagan (1975), and Engle (1979a) can be formulated and estimated in this framework. If x_t is interpreted as a vector of regression coefficients and α is given a time subscript, (2.1) and (2.2) become a time varying or random coefficients model, as discussed by Cooley and Prescott (1973) (1976), Harvey and Collier (1977), Garbade (1977), Brown, Durbin and Evans (1975), Pagan (1980), and Rosenberg (1973).

In contrast to this wide range of applications of the state space model with one measurement equation, there appear to be no time series applications that fully utilize the model when $P > 1$. When $\phi = 0$ a rich collection of cross-sectional models immediately appear as special cases. Of particular interest is the standard factor analysis model for which $\beta = \gamma = 0$ and R and Q are diagonal and for which α gives the factor loadings. The model is called the multiple indicator model (MI) by Goldberger (1972) because the observable variables, y, are the effects or indicators of the latent variable x_t. Closely related is the multiple-indicator, multiple-cause model (MIMIC) originally formulated by Zellner (1970) and then extended by Goldberger (1972, 1977), and Jöreskog and Goldberger (1975). Here $\gamma \neq 0$, so the z variables cause the latent

variable while the y's indicate the effect. A slight generalization also allows $\beta \neq 0$. If we further relax the model to allow R nondiagonal, the factor structure of the disturbances is no longer present and the model becomes a seemingly unrelated regression set-up where there may be coefficient restrictions across equations.

When $\phi \neq 0$, the state space model provides dynamic generalizations of these cross-sectional models. For example, a multivariate stationary time series can be written in state space form with β and $\gamma = 0$. Thus, such a model could be considered a dynamic factor analysis or dynamic multiple indicator (DYMI) model; the interest in such models occurs when there are a small number of independent factors which make up the full multivariate time series. In this case, forecasting would be done by forecasting the small number of factors and then combining these forecasts according to the factor loadings. This formulation could be thought of as a parsimonious version of a vector ARIMA model.

Similarly, when β is nonzero, the model is a generalization of vector transfer function models or reduced form systems with vector ARIMA disturbances. If in addition γ is nonzero, then we have a dynamic MIMIC, or DYMIMIC model that can again be considered as a parsimonious version of a vector transfer function model.

Although several frequency domain procedures have been proposed for identifying unobserved components in multivariate time series by Geweke (1977), Brillinger (1975), and Sargent and Sims (1977), only one time domain approach is known to the authors. Sargent and Sims (1977), Sims (1980), and Neftci (1979) introduce an "observable index" model that defines the unobserved component to be a nonstochastic linear combination of lagged dependent variables. Thus the z vector includes lagged y's and the associated disturbances have zero variances so that $Q = 0$.

In this article we use economic theory to specify and estimate a time domain dynamic factor analysis model that restricts $\gamma = 0$ but allows the factors to be serially correlated.

3. ESTIMATION AND DIAGNOSIS IN THE STATE SPACE MODEL

Maximum likelihood estimation of the model (2.1), (2.2), (2.3) can be undertaken using brute force by successively substituting (2.1) in (2.2) to eliminate the unobserved variable x. This yields a "reduced form" set of equations in y and z with many lags and a disturbance covariance matrix that includes a variety of the same parameters found in the lag structures. Although such a problem can be solved by a generalized least squares approach (GLS), each case will be different. The Kalman filter algorithm provides a general solution to this problem and also calculates estimates of the x_t. The algorithm does not require inverting a $T \times T$ matrix, where T is the length of the time series, as would a brute force, GLS approach.

The state space model has two sets of unknowns. There

are unknown parameters α, β, ϕ, γ, Q, R, and there are unobserved states, x. The estimation of these two classes of unknowns is done in two related steps. Conditional on these parameter values, best linear unbiased estimates of the states x_t, based on all information through time t, are obtained with the Kalman filter. As a by-product of the state estimation, the value of the likelihood function is obtained. Then, using a grid search or some form of nonlinear routine, the likelihood can be maximized with respect to the unknown parameters. A final pass through the Kalman filter produces the estimates of the state vector based upon the maximum likelihood parameters. At this point it is also possible to get "smoothed" estimates of the states. These estimates are the best linear unbiased estimates based on all the information, not just that up through t, conditional on the parameter estimates.

The Kalman filtering algorithm is becoming well known in economics through work by Chow (1975), Taylor (1970), Pagan (1975), Engle (1979a), Garbade (1977), and others, and will not be discussed in detail here. Basically, it constructs an estimate $\hat{x}_t$ of x_t based on combining information newly available in time t, (y_t, z_t), with $\hat{x}_{t-1}$. If the x_t process is stationary, the recursion is initialized by setting the mean and variance of x_0 equal to their steady state values. Otherwise the method discussed in Rosenberg (1973) can be used. We define the innovations η_t as the difference between y_t and the best estimate of y_t based on information up to $t - 1$ plus any exogenous variables in time t.

$$\eta_t = y_t - \hat{y}_t \tag{3.1}$$

$$= y_t - \alpha\phi\hat{x}_{t-1} - (\alpha\gamma + \beta)z_t$$

The innovations for a correctly specified model must be white noise and must also have no lagged cross correlations. If either of these features is not present, then the forecast of y_t would not have used all the information through $t - 1$. The contemporaneous variance covariance matrix of the innovations, H_t, is a function only of the parameters of the model and is therefore known. Schweppe (1965) has shown that the log-likelihood of the sample can be written very simply in terms of the innovations as

$$L = \sum_t L_t = \sum_t -\frac{1}{2}(\log |H_t| + \eta_t'H_t^{-1}\eta_t). \tag{3.2}$$

Therefore it is in principle a simple task to maximize the likelihood function with respect to the unknown parameters.

Unfortunately, this maximization is computationally demanding as there are generally large numbers of parameters and each evaluation of the likelihood function requires an appreciable number of calculations. Although several nonlinear optimization procedures have been used in this project, the most successful procedure was the generalization of Pagan's (1980) approach to the method of scoring and only this will be discussed.

Let all the parameters be gathered into a vector θ. The

iterative procedure involves finding $\mathscr{I}_k$, the information matrix evaluated at θ^k; and λ_k, a scalar step length, to obtain new estimates θ^{k+1} based upon estimates from the kth iteration.

$$\theta^{k+1} = \theta^k + \lambda_k \mathscr{I}_k^{-1} (\partial L/\partial \theta) \mid \theta^k. \quad (3.3)$$

Pagan (1980) has derived an expression for the information matrix in the case $P = 1$. Here we derive the equivalent expression for P greater than one and correct an error in Gupta and Mehra (1974). We use the well-known expressions for derivatives of a symmetric matrix B

$$\frac{\partial \mid B \mid}{\partial x} = \mid B \mid \text{tr}\left(B^{-1} \frac{\partial B}{\partial x}\right). \quad (3.4)$$

$$\frac{\partial B^{-1}}{\partial x} = -B^{-1} \frac{\partial B}{\partial x} B^{-1}. \quad (3.5)$$

Differentiating L_t in (3.2) with respect to one parameter θ_i using (3.4) and (3.5) gives

$$\partial L_t/\partial \theta_i = -\frac{1}{2} \text{tr}(H_t^{-1} \partial H_t/\partial \theta_i)$$

$$- (\partial \eta_t/\partial \theta_i)' H_t^{-1} \eta_t \quad (3.6)$$

$$+ \frac{1}{2} \eta_t' H_t^{-1} \partial H_t/\partial \theta_i H_t^{-1} \eta_t,$$

which can be rewritten by taking the trace of the last term as

$$\partial L_t/\partial \theta_i = -\frac{1}{2} \text{tr}(H_t^{-1} \partial H_t/\partial \theta_i)(I - H_t^{-1} \eta_t \eta_t')$$

$$- (\partial \eta_t/\partial \theta_i)' H_t^{-1} \eta_t \quad (3.7)$$

$$= L_{1_t} + L_{2_t}.$$

To find the second derivative matrix of the log-likelihood, write

$$\partial L_{1_t}/\partial \theta_j = -\frac{1}{2} \text{tr}[\partial(H_t^{-1} \partial H_t/\partial \theta_i)/\partial \theta_j]$$

$$\times [I - H_t^{-1} \eta_t \eta_t']$$

$$- \frac{1}{2} \text{tr}[(H_t^{-1} \partial H_t/\partial \theta_i)H_t^{-1} \partial H_t/\partial \theta_j H_t^{-1} \eta_t \eta_t']$$

$$+ \frac{1}{2} \text{tr}\{H_t^{-1}(\partial H_t/\partial \theta_i)H_t^{-1}$$

$$\times [(\partial \eta_t/\partial \theta_j) \eta_t' + \eta_t(\partial \eta_t/\partial \theta_j)']\}. \quad (3.8)$$

The only random variables in this expression are the η_t, hence taking the expected value of (3.8) the first term vanishes and the last two matrices in the second cancel. From (3.1) it can be seen that $\partial \eta_t/\partial \theta_i$ depends only on past innovations (through $\hat{x}_{t-1}$) and current exogenous variables. Since current and past innovations are independent by the normality assumption, the third term in (3.8) vanishes in expected value as well. Thus we have

$$E(\partial L_{1_t}/\partial \theta_j) = -\frac{1}{2} \text{tr}[H_t^{-1} \partial H_t/\partial \theta_i H_t^{-1} \partial H_t/\partial \theta_j]. \quad (3.9)$$

Now differentiate L_{2_t} with respect to θ_j to obtain

$$\partial L_{2_t}/\partial \theta_j = -\partial^2 \eta_t/\partial \theta_i \partial \theta_j H_t^{-1} \eta_t$$

$$- (\partial \eta_t/\partial \theta_i)' \partial H_t^{-1}/\partial \theta_j \eta_t \quad (3.10)$$

$$- (\partial \eta_t/\partial \theta_i)' H_t^{-1} \partial \eta_t/\partial \theta_j.$$

Taking expected values, the first two terms vanish. The third depends only on past innovations and thus is equal to its expected value conditional on the past. This gives

$$E(\partial L_{2_t}/\partial \theta_j) = -(\partial \eta_t/\partial \theta_i)' H_t^{-1} \partial \eta_t/\partial \theta_j. \quad (3.11)$$

The ijth element of the information matrix is the negative of the sum of (3.9) and (3.11) summed over all time periods. Thus

$$\mathscr{I}_{ij} = \sum_t \frac{1}{2} \text{tr}[H_t^{-1} \partial H_t/\partial \theta_i H_t^{-1} \partial H_t/\partial \theta_j]$$

$$+ \sum_t (\partial \eta_t/\partial \theta_i)' H_t^{-1} \partial \eta_t/\partial \theta_j. \quad (3.12)$$

The advantage of the scoring algorithm is now apparent. The expression (3.12) requires only first derivatives, which can be calculated numerically by K passes of the Kalman filter where K is the dimension of θ. Starting from consistent estimates asymptotic efficiency will be reached in one step. Computational gains are expected from using the derivatives of H, the steady state matrix to which H_t converges, rather than H_t itself in (3.12).

Three general classes of model diagnostics are available for evaluating the performance of the model. First, from the information matrix, a set of standard errors are available for all the parameters of the model. These may be used to test the significance of parameters in an overfitted model. Second, the innovations are available and can be tested for whiteness and lack of cross-correlation. Finally, Lagrange Multiplier or Score tests can be derived for testing the adequacy of the model.

Significant serial correlation of any of the innovations, or lagged cross-correlations between innovations for different observation vectors each indicate faults with the model specification. Test statistics designed for this procedure have not yet been proposed, hence in this article standard serial correlation coefficient tests have been applied. The Brown, Durbin, and Evans (1975) approach to recursive residuals might suggest using the cusum or cusum squared tests on each innovation series.

If significant serial correlation is observed, then the model should be augmented in some fashion. If the serial correlation is in only one innovation, then presumably there should be an autocorrelated error in this measurement equation. If there are lagged cross-correlations, then the common factor is not sufficiently well specified. Either it must be redefined, allowed to enter the equations differently, or perhaps a second factor is necessary.

Lagrange Multiplier or Score tests are asymptotically equivalent to Wald and Likelihood Ratio tests for local alternatives and require that the model be estimated only under the null. This is an attractive feature for the models considered in this article, since numerous calculations are required to overfit the model. A discussion of Lagrange Multiplier tests can be found in Engle (1979c). In Engle and Watson (1980) a numerical method for calculating the test statistic is presented, and an example of the testing procedure is given in Section 5.

4. ECONOMIC FORMULATION FOR THE UNOBSERVED COMPONENT

The economic model of wage determination can be simply formulated for each sector i as

$$w_{it} = \alpha_i m_t + \beta_i n_{it} + \beta_{0i} + e_{it}, \qquad (4.1)$$

where w and n are the logs of the Los Angeles and national wage rates in industry i and year t, and $Ee_{it}e_{jt} = 0$, $Ee_{it}m_t = 0$, for all t and $i \neq j$. The unobserved component m is common to all sectors and defined to have mean zero. The parameters α_i and β_i indicate the sensitivity of wage determination in industry i to metropolitan and national wage rates. The assumption that n_{it} is exogenous to Los Angeles is universally incorporated in regional models and is plausible because Los Angeles provides no more than 6 percent of the employment in any of the industries.

The data are analyzed for five sectors: Contract Construction (CC), Durable Manufacturing (DM), Non-Durable Manufacturing (NDM), Wholesale Trade (WT), and Retail Trade (RT). Simple regional economic theory (see, for example, Engle 1979b) suggests that export industries facing elastic demands would be particularly sensitive to national wage rates because excessive local wage pressure would lead to decreased local production in favor of alternative sites, thereby keeping the observed local industry wage in line with the national. On the other extreme, local serving industries would be more sensitive to the wage rate in the metropolitan economy because there is no possibility of relocating the production. Hence we anticipate large values of α for contract construction and retail trade, and smaller but still positive values for manufacturing. If wholesale trade in Los Angeles primarily provides services as a transshipment point, then this sector would behave like an export industry.

To complete the specification of the model, it is necessary to formulate an equation for the unobserved metropolitan component. An initially appealing equation made the wage a function of the tightness of the local labor market, as measured by the local unemployment rate or by the difference between the local and the national unemployment rates. This can be generally specified as

$$m_t = \phi m_{t-1} + \gamma u_t + v_t. \qquad (4.2)$$

With this equation, the model becomes a dynamic multiple-indicator, multiple-cause (DYMIMIC) model.

Estimates of variants of (4.1) with (4.2) produced plausible values of α_i and β_i, but invariably had the wrong sign for γ. In retrospect, this is to be expected and has been found in the literature in other contexts, by Metcalf (1971) and Hall (1975) for example. For a regional economy, unemployment is determined jointly with wages; high wages lead to high unemployment through migration both of capital and of labor. Thus the simultaneous equations bias is presumably responsible for the unexpected sign of γ. For further details see Engle and Watson (1978).

To improve the specification requires formulation of the system of equations for the endogenous local variables. Equation (4.2) could then be replaced by its reduced form. For a regional economy, there are few local exogenous variables, hence this reduced form would consist primarily of lagged dependent variables. Following Wallis (1977), and Zellner and Palm (1974), the exogenous variables could be represented by a multivariate time series and the full system solved so that m satisfies an ARIMA model.

A rather extravagant approach to the estimation of such a model is to allow each data point to be given by an unknown fixed constant. Letting D_t be one in period t and zero otherwise, we can define a fixed effects model as

$$m_t = \sum_t D_t \gamma_t. \qquad (4.3)$$

This model imposes no structure on the γ_t but uses a large number of parameters to estimate a sequence of very highly correlated points. It is also useless for forecasting. When this model was estimated, it appeared that the γ_t followed a smooth time pattern that could most likely be explained by a low order ARMA model. Again, for further details, see Engle and Watson (1978).

These preliminary results suggest the desirability of simple dynamic factor-analysis models, which turn out to be most successful for this data set. In the simplest case, each e_i is assumed to be serially uncorrelated and m is assumed to be a first-order autoregression

$$m_t = \phi m_{t-1} + v_t. \qquad (4.4)$$

Together, (4.4) and (4.1) describe a state space system which can be estimated and tested using the techniques of Section 3. This model is called Model A.

As will be shown, Model A fails a variety of specification tests and hence a generalization called Model B was formulated. This model allows m to follow a second-order autoregression while each e_i is a first-order autoregression with its own coefficient. Letting e_t be the 5×1 column vector of disturbances, this can be written and estimated in state space form as

$$\begin{pmatrix} m_t \\ m_{t-1} \\ e_t \end{pmatrix} = \left(\begin{array}{cc|c} \phi_1 & \phi_2 & 0 \\ 1 & 0 & \\ \hline & & \rho_1 \quad 0 \\ 0 & & 0 \quad \ddots \quad \rho_5 \end{array} \right) \begin{pmatrix} m_{t-1} \\ m_{t-2} \\ e_{t-1} \end{pmatrix} + v_t$$

$$w_t = \left(\begin{array}{cc|c} 1 & 0 & \\ & & I \\ \alpha & 0 & \end{array} \right) \begin{pmatrix} m_t \\ m_{t-1} \\ e_t \end{pmatrix} + \begin{pmatrix} \beta_1 & 0 \\ & \ddots \\ 0 & \beta_5 \end{pmatrix} n_t$$

$$Q = \begin{pmatrix} \sigma^2_m & & & 0 \\ & 0 & & \\ & & \sigma^2_1 & \\ & & & \ddots \\ 0 & & & \sigma^2_5 \end{pmatrix}, \quad R = 0. \qquad (4.5)$$

5. DYNAMIC FACTOR ANALYSIS-ESTIMATION

Model A was estimated using a variant of the technique described in Section 3. The results were not particularly encouraging. There appeared to be substantial serial correlation in the residuals of several sectors and lagged cross-correlations between sectors were sometimes severe. For first- and second-order lags these are shown in Table 2. The finite sample properties of the serial correlation coefficients are not known, but the asymptotic 5 percent tests of correlation coefficients of white noise random variables are rejected for 10 of the 50 coefficients. Presumably these tests are very conservative for the same reasons given by Durbin (1970), and Davies and Newbold (1979).

Further evidence of misspecification comes from overfitting Model A. The estimates of Model B are given in Table 3. For the second-order coefficient in the m autoregression, ϕ_2, the t statistic is -4.3 and for the individual serial correlation coefficients, the ρ_i's, they range from 1.2 to 8.4. These all lead to a rejection of the specification in A; they also suggest that in CC and RT, the metropolitan component may account for all the dynamics.

The results for Model B are rather encouraging. The largest factor loadings are on contract construction and

Table 2. Auto- and Cross-Correlations of Innovations from Dynamic Factor Analysis Models

		Model A			
Sector	CC_{-1}	DM_{-1}	NM_{-1}	WT_{-1}	RT_{-1}
Contract construction	.70[a]	$-.03$	.11	.53[a]	.38
Durable manufactures	.31	.28	.21	.27	.28
Nondurable manufactures	.28	.35	.25	.24	.33
Wholesale trade	.49[a]	.00	.12	.50[a]	.15
Retail trade	.60[a]	$-.01$	.12	.53[a]	.46[a]
Sector	CC_{-2}	DM_{-2}	NM_{-2}	WT_{-2}	RT_{-2}
Contract construction	.43[a]	.18	.20	.33	.18
Durable manufactures	.26	.03	.10	$-.05$	.31
Nondurable manufactures	.22	$-.05$	$-.03$	.11	.31
Wholesale trade	.39	.15	.16	.20	.09
Retail trade	.47[a]	.04	.23	.48[a]	.15
		Model B			
Sector	CC_{-1}	DM_{-1}	NM_{-1}	WT_{-1}	RT_{-1}
Contract construction	.05	$-.01$	$-.10$	.02	$-.09$
Durable manufactures	.07	.09	.21	.11	.06
Nondurable manufactures	$-.36$	$-.10$	$-.25$	$-.40$	$-.23$
Wholesale trade	$-.13$	$-.06$	$-.09$	$-.10$	$-.23$
Retail trade	.29	$-.02$	.15	.27	.23
Sector	CC_{-2}	DM_{-2}	NM_{-2}	WT_{-2}	RT_{-2}
Contract construction	.19	.11	.05	.07	.18
Durable manufactures	$-.28$	$-.16$	$-.20$	$-.40$	.03
Nondurable manufactures	$-.19$	$-.29$	$-.11$	$-.18$	$-.10$
Wholesale trade	.24	.11	.11	.13	$-.02$
Retail trade	$-.07$	$-.16$	$-.25$	$-.08$	$-.15$

[a] Significant at 95% level. Asymptotic critical value is .43.

Table 3. Dynamic Factor Analysis (Model B)[a]

Where $m_t = \phi_1 m_{t-1} + \phi_2 m_{t-2} + v_{1t}$

$w_{it} = \alpha_i m_t + \beta_i n_{it} + e_{it}$ For sectors $i = 1, \ldots, 5$

$e_{it} = \rho_i e_{it-1} + v_{i+it}$

Sector	α	β	ρ	$\sigma^2 \times 10^4$	SE
Contract construction	1.	.874	.628	.598	.008
		(.078)	(.389)	(.329)	
Durable manufactures	.549	.786	.742	.835	.009
	(.090)	(.053)	(.155)	(.266)	
Nondurable manufactures	.380	.786	.898	.466	.007
	(.091)	(.040)	(.107)	(.149)	
Wholesale trade	.302	.959	.519	1.191	.011
	(.075)	(.032)	(.227)	(.352)	
Retail trade	.663	.810	.340	.941	.010
	(.070)	(.059)	(.289)	(.343)	

	ϕ_1	ϕ_2	$\sigma^2 \times 10^4$	σ_{v_i}
Metropolitan component	1.606	$-.619$	1.229	.011
	(.125)	(.145)	(.585)	

[a] Standard errors are in parentheses.

retail trade, with durable manufacturing next. All the factor loadings are positive and, judging by the standard errors of the loadings, all are highly significantly different from zero. The standard errors are generally 1 percent or less. It appears that there may be a unit root in the process for the metropolitan component, which merely implies that over this sample period there does not appear to be an equilibrating process operating to bring Los Angeles wage rates into a par with other locations. If a unit root is found in this model, there is no change in the estimation procedure or the inference required except that the initial conditions must be estimated as nuisance parameters, and asymptotic approximations used for inference may be poor (see Evans and Savin 1981). Models formulated in terms of first differences and an intercept might be more attractive in some cases.

It should also be noted that a large number of parameters have been estimated using a sample of moderate size. Maximum likelihood estimates may therefore not be very good, as only asymptotic optimality is guaranteed.

Diagnostic checks based on the innovations were also carried out for Model B. Auto- and cross-correlations presented in Table 2 show that in the first- and second-order auto- and cross-correlation matrix no coefficients are significant. A further test of the adequacy of Model B was carried out using the Lagrange Multiplier test described in Engle and Watson (1980). Under the null, Model B is the correct specification, while under the alternative each e_i is allowed to follow an AR(2) process and m_t is allowed to be AR(3). The resulting test statistic, which is asymptotically distributed as a χ_6^2 random variable under H_0, had a value of 9.1. This can be compared with the 5 percent critical value of 12.6.

The estimated model also produces estimates of the

Table 4. Estimates of the Metropolitan Wage Index

Year	Dynamic Factor Analysis
1952	.90
53	.91
54	.91
55	.93
56	.93
57	.94
58	.97
59	.99
1960	1.01
61	1.03
62	1.05
63	1.07
64	1.08
65	1.09
66	1.10
67	1.10
68	1.09
69	1.08
1970	1.08
71	1.10
72	1.11
73	1.12
74	1.13
75	1.17

metropolitan component of wage rates. These estimates, when converted to index form, measure the extent to which the metropolitan wage rate varies from its normal relation with the national average after adjusting for industrial mix. The index is simply the exponential of the smoothed estimate of the unobserved component, that is, the best linear unbiased estimate based on all the data, assuming all the parameters are correct. The series are given in Table 4. Examination of the series shows that wage rates were 10 percent below normal in the early 1950's and rose steadily to 17 percent above normal in 1975. There is a slowdown in the late 60's and early 70's, which would correspond to the slowdown in aerospace and military contracts.

6. FORECASTING PERFORMANCE

The forecasting performance of the factor analysis model was compared with a simple regression model for the two post-sample years 1976 and 1977. For the second year, a two-period ahead forecast was used. The regression model included the national wage rate and a lagged value of the Los Angeles industry wage rate, and allowed a first order autoregressive error process. The model was, therefore, slightly more parsimonious than the dynamic factor model. The criterion for performance was the sum of squared errors for the four forecastable sectors (no new data has been reported for contract construction).

The sum of squared errors ($\times 10^4$) for the factor analysis and the regression models in 1976 were 10.68 and 13.62, respectively. In 1977, the corresponding figures were 28.33 and 29.37. For a few sectors the regression model forecast better, while for others the factor analysis model forecast better. However, the sum of squared errors favored the factor model in both years, and the larg-

est error was always in the regression forecast. This is consistent with our anticipation that a joint forecast can be more accurate when a joint estimation procedure is employed. It is also encouraging evidence for the usefulness of these procedures not only for economic analysis but also for forecasting.

APPENDIX: SOURCES OF DATA

Average hourly earnings for the five industries in Los Angeles for 1952–1975 were from U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings, States and Areas 1939–1975*, Washington, D.C.: Government Printing Office, 1977.

Average hourly earnings for the four industries in Los Angeles for 1976 were from *California Statistical Abstract 1977*.

Average hourly earnings for the five industries in the U.S. for 1952–1975 were from U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings, United States, 1909–1975*, Washington, D.C.: Government Printing Office, 1977.

Average hourly earnings for the four industries in the U.S. for 1976 were from U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings*, Washington, D.C.: Government Printing Office, March 1977.

National unemployment figures were from U.S. Department of Labor, Bureau of Labor Statistics, *Handbook of Labor Statistics* (various dates), Washington, D.C.: Government Printing Office.

Los Angeles unemployment figures for 1970–1976 were from the *California Statistical Abstract, 1976*, for 1952–1970 from the California Employment Development Department, Los Angeles. Due to the definitional change, these data were not comparable with the 1970–1976 data. The overlapping year (1970) was used to adjust the old figures to the new definition.

REFERENCES

BRILLINGER, D.R. (1975), *Time Series Data Analysis and Theory*, New York: Holt, Rinehart and Winston.

BROWN, R.L., DURBIN, J., and EVANS, J.M. (1975), "Techniques for Testing the Constancy of Regression Relationships Over Time, With Comments," *Journal of the Royal Statistical Society*, Ser. B, 37, 149–192.

CHOW, G.C. (1975), *Analysis and Control of Dynamic Economic Systems*, New York: John Wiley.

COOLEY, T.F., and PRESCOTT, E.C. (1973), "Tests of an Adaptive Regression Model," *Review of Economics and Statistics*, 55, 248–256.

——— (1976), "Estimation in the Presence of Stochastic Parameter Variation," *Econometrica*, 44, 167–184.

DAVIES, W., and NEWBOLD, P. (1979), "Some Power Studies of the Portmanteau Test of Time Series Model Specification," *Biometrika*, 66, 153–155.

DURBIN, J. (1970), "Testing for Serial Correlation in Least Squares Regression When Some of the Regressions are Lagged Dependent Variables," *Econometrica*, 38, 410–421.

ENGLE, R.F. (1974), "Issues in the Specification of an Econometric Model of Metropolitan Growth," *Journal of Urban Economics*, 1, 250–267.

———— (1979a), "Estimating Structural Models of Seasonality," in *Seasonal Analysis of Economic Time Series*, Economic Research Report, ER-1, U.S. Department of Commerce, Bureau of the Census.

———— (1979b), "Estimation of the Price Elasticity of Demand Facing Metropolitan Producers," *Journal of Urban Economics*, 6, 42–64.

———— (1979c), "A General Approach to the Construction of Model Diagnostics Based on the Lagrange Multiplier Principle," Discussion Paper 79-43, Department of Economics, University of California, San Diego.

ENGLE, R.F., and WATSON, M. (1978), "A One Factor Multivariate Time Series Model of Metropolitan Wage Rates," Discussion Paper 78-15, Department of Economics, University of California, San Diego.

———— (1980), "A Time Domain Approach to Dynamic Factor and MIMIC Models," *Les Cahiers du Deminaire d'Econometrie*, 22, 109–125.

EVANS, G.B.A., and SAVIN, N.E. (1981), "Testing for Unit Roots: 1," *Econometrica*, 49, 753–782.

GARBADE, K. (1977), "Two Methods for Examining the Stability of Regression Coefficients," *Journal of the American Statistical Association*, 72, 54–63.

GEWEKE, J. (1977), "The Dynamic Factor Analysis of Economic Time-Series Models," in *Latent Variables in Socio-Economic Models*, eds. D.J. Aigner and A.S. Goldberger, New York: North-Holland.

GEWEKE, J., and SINGLETON, K. (1981), "Maximum Likelihood Confirmatory Factor Analysis of Economic Time Series," *International Economic Review*, 22, 37–54.

GOLDBERGER, A.S. (1972), "Structural Equation Methods in the Social Sciences," *Econometrica*, 40, 979–1001.

———— (1977), "Maximum-Likelihood Estimation of Regressions Containing Unobservable Independent Variables," in *Latent Variables in Socio-Economic Models*, eds. D.J. Aigner and A.S. Goldberger, New York: North-Holland.

JÖRESKOG, K., and GOLDBERGER, A.S. (1975), "Estimation of a Model With Multiple Indicators and Multiple Causes of a Single Latent Variable," *Journal of the American Statistical Association*, 70, 631–639.

GUPTA, N.K., and MEHRA, R.K. (1974), "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations," *IEEE Transactions on Automatic Control*, AC-19, 774–783.

HALL, R. (1975), "The Rigidity of Wages and the Persistence of Unemployment," *Brookings Papers on Economic Activity*, 2, 301–335.

HANNAN, E.J. (1976), "The Identification and Parameterization of Armax and State Space Forms," *Econometrica*, 44, 713–724.

HARVEY, A., and COLLIER, P. (1977), "Testing for Functional Misspecification in Regression Analysis," *Journal of Econometrics*, 6, 103–120.

HARVEY, A., and PHILLIPS, G.D.A. (1979), "The Estimation of Regression Models With Autoregressive-Moving Average Disturbances," *Biometrika*, 66, 49–58.

MEHRA, R.K. (1974), "Identification in Control and Economics: Similarities and Differences," *Annals of Economic and Social Measurement*, 3, 21–47.

METCALF, D. (1971), "The Determinants of Earnings Changes: A Regional Analysis for the U.K., 1960–68," *International Economic Review*, 12, 273–282.

NEFTCI, S.N. (1979), "Modeling the Price Side of Econometric Models: An Analysis of the Underlying Hypothesis," *Journal of Econometrics*, 10, 71–84.

NERLOVE, M. (1971), "Analysis of Economic Time-Series by Box-Jenkins and Related Techniques," Report 7156, Center for Mathematical Studies in Business and Economics, University of Chicago.

PAGAN, A. (1975), "A Note on the Extraction of Components From Time Series," *Econometrica*, 43, 163–168.

———— (1980), "Some Identification and Estimation Results for Regression Models With Stochastically Varying Coefficients," *Journal of Econometrics*, 13, 341–363.

ROSENBERG, B. (1973), "The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression," *Annals of Economic and Social Measurement*, 2, 399–428.

SARGENT, T.J., and SIMS, C.A. (1977), "Business Cycle Modeling Without Pretending to Have Too Much *A Priori* Economic Theory," in *New Methods in Business Cycle Research: Proceedings from a Conference*, ed, C.A. Sims, Minneapolis, Minn.: Federal Reserve Bank of Minneapolis.

SCHWEPPE, F.C. (1965), "Evaluation of Likelihood Functions for Gaussian Signals," *IEEE Transactions on Information Theory*, 11, 61–70.

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

TAYLOR, L. (1970), "The Existence of Optimal Distributed Lags," *Review of Economic Studies*, 37, 95–106.

WALLIS, K.F. (1977), "Multiple Time Series Analysis and the Final Form of Econometric Models," *Econometrica*, 45, 1481–1498.

ZELLNER, A. (1970), "Estimation of Regression Relationships Containing Unobservable Variables," *International Economic Review*, 11, 441–454.

ZELLNER, A., and PALM, F. (1974), "Time Series Analysis and Simultaneous Equation Models," *Journal of Econometrics*, 2, 17–54.