RSP 18(1): 100-116.

A Regional Payroll Forecasting Model That Uses Bayesian Shrinkage Techniques for Data Pooling

*James P. Lesage and Michael Magura**

Introduction

It has been long recognized that regional economies exhibit interregional linkages, as well as linkages with state and national economies. Despite these acknowledged interregional relations, many regional econometric models typically have been unlinked SMSA models or SMSA models that include only links with the state or national economies. This is not surprising, as econometric approaches to linking local economies to a multi-regional economy run into several problems. First, there is the issue of collinearity that arises from the multitude of explanatory variables that are typically used to achieve the links. Second, there is a degrees of freedom problem and overparameterization of these models. Finally, overfitting during estimation and difficulties in managing the data base for updating the model arise. All of these problems tend to degrade the ability of linked models to forecast more accurately than their unlinked counterparts.

This paper introduces a Bayesian approach to pooling data from across the SMSA models into a linked, regional forecasting model. We argue that this approach overcomes the problems mentioned above that are traditionally encountered in producing linked models. The focus of our investigation is whether this approach has an advantage over unlinked models in providing improved forecasts.

Our approach, like that of Garcia-Ferrer [2] represents a "bottom-up" attempt to build a simple model based on the time-series properties of the payroll data that will forecast well. This is in contrast to the work of Liu and Stocks [4] that we would classify as "top-down." They attempt to build elaborate models of the individual metropolitan area economies in order to approach the multi-regional forecasting problem. A set of simple unlinked models of payroll variation for each of the metropolitan areas provides a benchmark against that we assess the improvement attributable to the pooling of multi-regional information.

The paper proceeds as follows. Section 2 examines the time-series data for payroll in eight Ohio SMSAs demonstrating that there exists a fair amount of comovement over time. Simple OLS models using leading indicator variables provide benchmark forecasts that will be used to judge the improvement derived from our techniques. Section 3 describes and motivates the Bayesian prior proposed in the new technique. This estimator is shown to be a compact and efficient way to pool information from the multi-regional economy. Section 4 presents the results from forecasting experiments using three alternative estimation techniques. We compare the simple OLS model with a ridge regression model and the Bayesian pooling model. The ridge comparison is included in order to verify whether any reduction in forecast error truly arises from the Bayesian pooling information rather than shrinkage alone.

Analyses of the Data for Eight Ohio SMSAs

Our data represent quarterly total payroll data based on employment covered by unemployment insurance in each of the eight Ohio SMSAs: Akron, Canton, Cincinnati, Cleveland, Columbus, Dayton, Toledo, and Youngstown. These data are available from the Ohio Bureau of Employment Services, Labor Market Information Division. The data cover the period 1978 first quarter through 1985 fourth quarter and are in nominal value terms. These data were transformed to indices with a base of unity during the first quarter of 1982. This transformation was made so that the parameters of the various metropolitan payroll relations would take on magnitudes of comparable size. The requirement that the parameters be comparable in magnitude is necessitated by our shrinkage of the parameters from individual metropolitan relations toward the pooled value of the parameter estimates for all eight metropolitan

Figure 1 shows a plot of these indexed values of payroll for the eight SMSAs over the time period 1978 through 1985. It seems clear that these regional economies exhibit very similar patterns, often called comovements. It is this information concerning common influences from seasonality, business cycles, etc. that we will attempt to exploit in order to improve the forecasting performance of the individual metropolitan payroll relations.

The sample period covering 1978 first quarter through 1983 third quarter was used to estimate the models, and the data covering the 1984 and 1985 period were used for out-of-sample forecasting experiments. A vertical grid in the plots separates the forecast time period from the period used to fit the models. The fitted models were used to generate one through three-step-ahead forecasts during the out-of-sample period. In making these out-of-sample forecasts, the models were re-estimated using all past data prior to each forecast period. The fitted model that generated one-step-ahead forecasts was based on a single lagged dependent variable and a set of lagged leading indicator variables that were also lagged one period. The OLS version of this model is shown in equation (1):

$$P_{it} = \beta_0 + \beta_1 P_{it-1} + \beta_2 S_{it-1} + \beta_3 C_{it-1} + \beta_4 A_{t-1} + \beta_5 L_{t-1} + \varepsilon_{it}, \qquad (1)$$

where:
$i = 1,...,8$ denoting the eight metropolitan areas;
P_{it} = payroll in the ith SMSA at time t;
S_{it} = housing starts in the ith SMSA at time t;
C_{it} = unemployment insurance initial claims in the ith SMSA at time t;
A_t = national domestic auto sales at time t;
L_t = national index of 12 leading indicators at time t;
e_{it} = a Gaussian disturbance term for SMSA i at time t;

A separate model was used to generate two-period-ahead forecasts, that were based on the same variables as shown in equation (1) but a two period lag of these variables was imposed in order to achieve a two-step-ahead forecast. This is similar to the model that generated the three-step-ahead forecasts. The OLS versions of these two models are shown in equations (2) and (3).

$$P_{it} = \alpha_0 + \alpha_1 P_{it-1} + \alpha_2 P_{it-2} + \alpha_3 S_{it-2} + \alpha_4 C_{it-2} +$$
$$\alpha_5 A_{t-2} + \alpha_6 L_{t-2} + \varepsilon_{it} \qquad (2)$$

$$P_{it} = \tau_0 + \tau_1 P_{it-1} + \tau_2 P_{it-2} + \tau_3 P_{it-3} + \tau_4 S_{it-3} + \tau_5 C_{it-3} +$$
$$\tau_6 A_{t-3} + \tau_7 L_{t-3} + \varepsilon_{it} \qquad (3)$$

The models in equations (1) through (3) were estimated and forecast over the 1984 and 1985 period, along with some simpler variants of these models. The percentage root-mean-square error (RMSE) of one through three step-ahead forecasts are presented in Table 1. The simpler models that were estimated and forecast were (1) an autoregressive model (labeled AR1 through AR3 in Table 1) containing only a constant term and the lagged dependent variables, (2) an autoregressive model containing the national leading indicator variable and domestic auto sales, (labeled ARL1 through ARL3 in Table 1).

Equations (1) through (3) are labeled as OLS1 through OLS3 in Table 1. The simpler models represent an attempt to analyze by decomposition the information content contained in the sets of variables representing national influences, local influences, and autoregressive influences.

Table 1 has been organized in a way to facilitate comparison of the forecasting ability of the three models at one, two, and three-step-ahead out-of-sample forecasts. This comparison illustrates a number of points. First, none of the models perform in an unacceptable way, because all but one achieve an average forecast error under 6 percent, with most of them under 5 percent. Youngstown appears to be the most difficult to forecast, achieving the largest average percentage RMSE over this period. Second, no single model appears to dominate the others in terms of forecasting performance; that is, some of the models perform better at different forecasting horizons for some of the cities. In order to clearly illustrate this point in Table 1, an asterisk has been placed next to the minimum percentage RMSE forecast for each of the horizons when comparing across models. The results indicate that the autoregressive (AR) model was best in 16 of 48 cases, the autoregressive model with national leading indicators (ARL) was best in 13 of 48 cases, and the full model with both national and local leading indicators (OLS) was best in 19 of 48 cases. It should be noted that with regard to this enumeration of best models, some of the average percentage RMSEs are very slightly different from one model to the next, making this type of counting scheme somewhat deceptive. We will compare the out-of-sample forecast errors from the technique of Garcia-Ferrer [2] to the best forecasting models shown in Table 1 in order to assess the improvement in forecasting ability.

The Bayesian Pooling Technique

This section describes and motivates the Bayesian technique for pooling data information from all eight SMSAs and incorporating it into the model. The proposed Bayesian estimator could be classified as a variant on ridge regression that shrinks the vector of estimates toward values of zero. The Garcia-Ferrer [2] technique uses shrinkage but shrinks the estimates toward a vector of estimates obtained from a pooled regression using data from all of the SMSAs.

Let X_i denote the matrix of explanatory variables in the ith metropolitan area and Y_i be the dependent variable vector for $SMSA_i$. The estimator proposed by Garcia-Ferrer [2] is shown in equation (4):

$$\hat{\beta}_Z = (X_i'X_i + \lambda_i I)^{-1}(X_i'X_i\ \hat{\beta}_i + \lambda_i\tilde{\beta}). \tag{4}$$

where $\hat{\beta}_i$ is the OLS estimator for metropolitan area i derived from the usual formula $\hat{\beta}_i = (X_i'X_i)^{-1}X_i'Y_i$, and $\tilde{\beta}$ is a pooled estimator given by equation (5):

$$\tilde{\beta} = (X_1'X_1 + X_2'X_2 + ... + X_8'X_8)^{-1}(X_1'Y_2 + X_2'Y_2 + ... + X_8'Y_8). \tag{5}$$

The expression for $\tilde{\beta}$ is a matrix weighted average of the individual metropolitan data information that compactly summarizes the data variation occurring at the metropolitan area level. This average summary information about the rest of the region then is mixed with the data contained in the X_i data matrix and Y_i vector for the ith SMSA in order to achieve the Bayesian estimate for the ith model. Equation (4) shows that the mixing of this information is again a matrix weighted average of the information for the individual SMSA contained in X_i and Y_i with the pooled information contained in $\tilde{\beta}$. The λ_i parameter controls the relative weighting of the two types of information, that pertaining to the individual SMSA, and that representing all SMSAs pooled. This parameter can be given an interpretation as the relative confidence in the two types of information, such that when $\lambda_i \rightarrow \infty$ the estimator $\hat{\beta}_Z$ approaches the pooled estimator $\tilde{\beta}$. On the other hand, for very small values of λ_i the estimator $\hat{\beta}_Z$ for SMSA i approaches $\hat{\beta}_i$, the OLS estimator based on the individual SMSA information.

The value of λ_i was determined for each equation by "integrating out" the posterior mean of the distribution of the $\hat{\beta}_Z$ estimator. That is, we generated forecasts for a range of settings of the λ_i parameter for each SMSA and chose the value of λ_i that produced the minimum average percent RMSE over the eight quarter forecast period. Since our objective is to produce a forecasting model, this seems to be a reasonable procedure.

Garcia-Ferrer [2] note that the estimator $\hat{\beta}_Z$ is the posterior mean of β_Z given the following interpretation of our model. Based on the Lindley and Smith [3] pooling model, we can define $Y_i = X_i\beta_i + u_i$ ($i = 1, 2, ..., 8$) that is shown in (6):

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \\ . \\ . \\ Y_8 \end{pmatrix} = \begin{pmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ & & . & \\ . & . & . & \\ & . & . & \\ 0 & 0 & \dots & X_3 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ . \\ \\ . \\ \beta_8 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ . \\ \\ . \\ u_8 \end{pmatrix} \qquad (6)$$

We further assume that:

$$\beta_i = \theta + \delta_i, \quad i = 1,2,\dots,8 \qquad (7)$$

where θ is the mean vector for the individual β_i's. Assume the δ_i's are independently distributed, each with a $N(0,\Phi\sigma^2_{\delta i} I_k)$ distribution, where $0 < \Phi < \infty$, and $\sigma^2_{\delta i}$ is an individual variance for each δ_i. Further, the u_i's in (6) are assumed normally and independently distributed, with zero mean and common variance σ^2_u. The probability density function for β_Z, given the data, $\hat{\sigma}^2_u$, $\lambda_i = \sigma^2_u/\sigma^2_{\delta i}$, and conditioning on $\theta = \beta$, has a posterior mean given by $\hat{\beta}_Z$.

Some points must be noted about the estimator $\hat{\beta}_Z$ shown in (4). If the estimator were to shrink toward zero instead of the $\tilde{\beta}$ pooled estimate, we would have a traditional ridge estimator, that is shown in equation (8):

$$\tilde{\beta}_{ridge} = (X_i'X_i + \lambda_i I)^{-1}(X_i'X_i \ \hat{\beta}_i). \qquad (8)$$

This suggests that the improved forecasting ability of the estimator $\hat{\beta}_Z$ may derive from two sources: first, the ridge or shrinkage aspect of the estimator may contribute to better forecasts by overcoming the collinearity problems inherent in these linked models; second, the shrinkage toward the pooled estimate may contribute toward better forecasting as this pooled estimate contains important linkage information about the regional economy in which the SMSA economy is operating. In order to independently assess these two sources of forecasting improvement, we will compare the forecasting ability of both a ridge estimator and the $\hat{\beta}_Z$ estimator to that from our alternative simpler models shown in Table 1.

Another point to note about $\hat{\beta}_Z$ is that it differs from more traditional approaches to improve forecasting by incorporating the information provided by variables that reflect structural relationships among variables across SMSAs. Here the approach is to exploit a more limited aspect of interregional relationships, the information contained in the statistical commonality of payroll movements across regions found in the moment matrices from a stacked cross-section time-series model. This procedure has the potential to overcome the overfitting and collinearity problems inherent in more traditional approaches. It should also lead to a much easier model to maintain and update, because the individual metropolitan area models remain computationally separate. By this we mean that each of the metropolitan models incorporate only local explanatory variables in their explicit equational forms. The mixing of linkage information takes place at the matrix level during the estimation procedure requiring no individual computational treatment of the separate SMSA models that would ordinarily be required in a linked modeling effort.

A Comparison of the Forecasts

We turn our attention to a comparison of the forecasting performance of the models shown in equations (1) through (3) in Section 2 and the OLS model shown in equation (3) re-estimated using the Bayesian estimator. In addition, we estimate the OLS model using a ridge estimator in order to assess how much of the improved forecasting performance of the Bayesian estimator actually derives from the pooling information contained in the $\tilde{\beta}$ vector to which the Bayesian estimator shrinks. General shrinkage toward zero seems likely to provide some improvement as it is well known that ridge estimators perform well in the face of collinear data models. It seems reasonable that our SMSA models contain explanatory variables that are highly correlated because all are leading indicators of economic performance.

A collinearity diagnostic procedure suggested in Belsley, Kuh, and Welsch [1] was employed to determine whether near linear dependencies existed between the columns of the explanatory variables matrix in the SMSA models. This technique produces a variance decomposition proportions table based on a singular value decomposition of the data matrix X_i for each metropolitan area. A necessary condition for a severe collinearity problem is indicated by a condition index (that represents the ratio of the largest eigenvalue to each of the remaining eigenvalues of the data matrix X) in excess of 90. The condition indices for our eight metropolitan area data sets are reported in Table 2. The table shows that every SMSA except Akron exhibits at least four condition indices over 100. Akron exhibits only two of these large condition indices, still indicating problems. Each of the large condition indices is indicative of a near linear relationship among the columns of the explanatory variable matrix X, so that we have at least four such relations for each SMSA except Akron. Furthermore, the larger the condition index, the stronger is the near linear dependency, and every SMSA exhibits one such relation that is very strong, being associated with a condition index over 500. It should be noted that the diagnostics results presented in Table 2 were performed on the largest explanatory variables model, the OLS model with three lags of the dependent variable. There are then nine explanatory variables in this model including the constant term. This will produce nine eigenvalues for the X data matrix and eight condition indices, where the eight condition indices are calculated from the ratio of the largest eigenvalue to the eight remaining eigenvalues.

We now turn to a discussion of the forecasting experiments. These experiments involved estimating the three models shown in equations (1) to (3) that produce one-step-ahead, two-step-ahead and three-step-ahead forecasts with the Bayesian estimator. We varied the value of λ over a large range of values ranging from zero to unity. Analogous to ridge estimation techniques, we would expect to find a relatively small value for the optimal setting of λ because the larger the value of λ becomes, the more biased the coefficient estimates become. Fairly small values result in placing a relatively large weight on the non-sample information or, in the case of our Bayesian estimator, on the pooled information. The intuition for this large weight associated with a relatively small value of λ is that we are inflating a very small eigenvalue of the X'X matrix by the magnitude of λ that is small, thereby having a relatively large impact on the outcome of the estimator.

The average percentage RMSE results of our out-of-sample forecasting experiments are reported in Table 3. In order to focus on the relative performance of the Bayesian estimator with that of the alternative estimators shown in Table 1, we have replicated the best out-of-sample forecast from Table 1 in Table 3 and labeled it as "Best." In addition, we replicate the OLS model forecast RMSE in Table 3 to facilitate a comparison of how much improvement the Bayesian procedure yields relative to the OLS model. Table 3 also reports the forecast RMSE from an optimal ridge estimator. Here we followed the same procedure as with the Bayesian estimator to determine the setting for the ridge parameter λ. This parameter was varied over a range from zero to unity with the optimal setting chosen as that which produced the lowest average percentage RMSE over the forecast period.

In order to summarize the Table 3 results, we chose to define a clear-cut forecasting advantage as at least an average percent RMSE difference of 0.03, that is, of course, somewhat arbitrary. Table 3 shows that the Bayesian or ridge estimators surpassed or tied the forecasting performance of the best forecasts from Table 1 in 35 of 48 cases. That is, the "Best" forecast from Table 1 was a clear-cut winner in only 13 of the 48 cases. The ridge estimator appeared to outperform the Bayesian estimator as the ridge exhibited the lowest average percent RMSE in 13 of the cases whereas the Bayesian estimator did so in only four of the cases. The ridge and Bayesian estimator were tied or essentially tied for the lowest percent RMSE in four of the cases.

From the results of the experiments, it seems that there is a gain to be had from adopting some sort of shrinkage technique, although the ridge technique appears to be the preferred technique rather than the Bayesian. In this regard, it should be noted that shrinkage did not appear to be useful in 10 of the 48 cases because the results were essentially the same for the OLS and shrinkage estimators.

Conclusions

We find that the Bayesian technique doesn't provide as much of a dramatic improvement in forecasting performance for the local payroll models as found in Garcia-Ferrer [2]. An important finding was that shrinkage leads to improved forecasting ability in these models. The shrinkage aspect of the Bayesian estimator contributes significantly to improved forecasting performance for 17 of 48 models examined here. The 17 cases include 13 cases where ridge was superior plus four cases where the ridge and Bayesian were equal in performance. The ridge estimator is being interpreted as representative of the shrinkage aspect of the Bayesian estimator. On the other hand, the pooling aspect of the Bayesian estimator leads to improved forecasting in a smaller number of cases, four of 48. Here we interpret the four cases where the Bayesian estimator was a clear-cut winner as those cases where the pooling aspect of the Bayesian estimator contributed to the improved performance. It should be noted that Garcia-Ferrer [2] did not separately analyze the shrinkage and pooling aspects of their proposed Bayesian estimator. It could be the case that shrinkage accounted for their improved results as well.

A possible problem that would affect the ability of the pooling information to contribute to better forecasts is that some of the eight SMSAs exhibit significantly different variation than others. Pooling procedures based on statistical measures of comovement that would such relations for each SMSA except Akron. Furthermore, the larger the condition index, the stronger is the near linear dependency, and every SMSA exhibits one such relation that is very strong, being associated with a condition index over 500. It should be noted that the diagnostics results presented in Table 2 were performed on the largest explanatory variables model, the OLS model with three lags of the dependent variable. There are then nine explanatory variables in this model including the constant term. This will produce nine eigenvalues for the X data matrix and eight condition indices, where the eight condition indices are calculated from the ratio of the largest eigenvalue to the eight remaining eigenvalues.

We now turn to a discussion of the forecasting experiments. These experiments involved estimating the three models shown in equations (1) to (3) that produce one-step-ahead, two-step-ahead and three-step-ahead forecasts with the Bayesian estimator. We varied the value of λ over a large range of values ranging from zero to unity. Analogous to ridge estimation techniques, we would expect to find a relatively small value for the optimal setting of λ because the larger the value of λ becomes, the more biased the coefficient estimates become. Fairly small values result in placing a relatively large weight on the non-sample information or, in the case of our Bayesian estimator, on the pooled information. The intuition for this large weight associated with a relatively small value of λ is that we are inflating a very small eigenvalue of the X'X matrix by the magnitude of λ that is small, thereby having a relatively large impact on the outcome of the estimator.

The average percentage RMSE results of our out-of-sample forecasting experiments are reported in Table 3. In order to focus on the relative performance of the Bayesian estimator with that of the alternative estimators shown in Table 1, we have replicated the best out-of-sample forecast from Table 1 in Table 3 and labeled it as "Best." In addition, we replicate the OLS model forecast RMSE in Table 3 to facilitate a comparison of how much improvement the Bayesian procedure yields relative to the OLS model. Table 3 also reports the forecast RMSE from an optimal ridge estimator. Here we followed the same procedure as with the Bayesian estimator to determine the setting for the ridge parameter λ. This parameter was varied over a range from zero to unity with the optimal setting chosen as that which produced the lowest average percentage RMSE over the forecast period.

In order to summarize the Table 3 results, we chose to define a clear-cut forecasting advantage as at least an average percent RMSE difference of 0.03, that is, of course, somewhat arbitrary. Table 3 shows that the Bayesian or ridge estimators surpassed or tied the forecasting performance of the best forecasts from Table 1 in 35 of 48 cases. That is, the "Best" forecast from Table 1 was a clear-cut winner in only 13 of the 48 cases. The ridge estimator appeared to outperform the Bayesian estimator as the ridge exhibited the lowest average percent RMSE in 13 of the cases whereas the Bayesian estimator did so in only four of the cases. The ridge and Bayesian estimator were tied or essentially tied for the lowest percent RMSE in four of the cases.

From the results of the experiments, it seems that there is a gain to be had from adopting some sort of shrinkage technique, although the ridge technique appears to be the preferred technique rather than the Bayesian. In this regard, it should be noted that shrinkage did not appear to be useful in 10 of the 48 cases because the results were essentially the same for the OLS and shrinkage estimators.

Conclusions

We find that the Bayesian technique doesn't provide as much of a dramatic improvement in forecasting performance for the local payroll models as found in Garcia-Ferrer [2]. An important finding was that shrinkage leads to improved forecasting ability in these models. The shrinkage aspect of the Bayesian estimator contributes significantly to improved forecasting performance for 17 of 48 models examined here. The 17 cases include 13 cases where ridge was superior plus four cases where the ridge and Bayesian were equal in performance. The ridge estimator is being interpreted as representative of the shrinkage aspect of the Bayesian estimator. On the other hand, the pooling aspect of the Bayesian estimator leads to improved forecasting in a smaller number of cases, four of 48. Here we interpret the four cases where the Bayesian estimator was a clear-cut winner as those cases where the pooling aspect of the Bayesian estimator contributed to the improved performance. It should be noted that Garcia-Ferrer [2] did not separately analyze the shrinkage and pooling aspects of their proposed Bayesian estimator. It could be the case that shrinkage accounted for their improved results as well.

A possible problem that would affect the ability of the pooling information to contribute to better forecasts is that some of the eight SMSAs exhibit significantly different variation than others. Pooling procedures based on statistical measures of comovement that would only pool more similar data samples to achieve the pooled estimates might be helpful.

In conclusion we find that the Bayesian pooling estimator proposed by Garcia-Ferrer [2] equals or outperforms the forecasting performance of competing models in 35 of 48 cases considered here. However, most of the improved performance emanates from the shrinkage aspect of this estimation procedure and not the pooling of information from other regions. Garcia-Ferrer [2] did not analyze these two aspects of their proposed estimator and may have achieved improved forecasts for the same reasons. If this is true, ridge estimation would provide the same improvement in forecasting ability, yet be much simpler from a computational standpoint.

Endnote

*The authors are with the University of Toledo. This research was supported by a grant from the Ohio Board of Regents Research Excellence Program.

References

1. Belsley, David A., Edwin Kuh, and Roy E. Welsch, *Regression Diagnostics* (John Wiley and Sons, Inc. 1980).

2. Garcia-Ferrer, A., R.A. Highfield, F. Palm, and A. Zellner, "Macroeconomic Forecasting Using Pooled International Data," *Journal of Business & Economic Statistics* (1987), pp. 53-68.

3. Lindley, D.V. and A.F.M. Smith, "Bayes Estimates for the Linear Model," *Journal of the Royal Statistical Society* (1972), Ser. B, 34, pp. 1-41.

4. Liu, Y. and A.H. Stocks, "A Labor Oriented Quarterly Econometric Forecasting Model for the Youngstown-Warren SMSA," *Regional Science and Urban Economics*, 13 (1983), pp. 317-340.

Table 1
Average Percent RMSE of Forecast Over 1984 and 1985

	Akron	Canton	Cincinnati	Cleveland	Columbus	Dayton	Toledo	Youngstown
1-step								
AR1	2.84*	3.43	2.35*	2.52	1.76	2.72	2.87*	3.67
ARL1	3.08	2.04	2.56	2.49*	1.45*	2.62	3.02	2.94*
OLS1	2.89	2.02*	2.70	2.80	3.75	2.56*	3.45	3.19
1-step								
AR2	1.77*	2.83	1.29*	2.30	1.42	2.39*	2.80	3.18
ARL2	2.06	1.76	1.83	1.68	1.34*	2.49	2.06	2.68*
OLS2	1.98	1.73*	2.21	1.57*	1.37	2.53	1.97*	3.26
2-step								
AR2	2.71*	3.95	2.34*	3.38	1.81	3.69*	3.19	4.86
ARL2	3.19	1.67*	3.03	3.20	2.61	4.72	3.17	2.99*
OLS2	3.54	2.71	3.04	2.69*	1.54*	4.23	3.05*	4.15
1-step								
AR3	1.53*	2.62*	1.66*	2.37	1.60	2.87	3.04	3.92
ARL3	1.81	2.64	1.92	2.58	1.92	2.64*	3.54	3.45*
OLS3	1.93	2.84	2.31	1.86*	1.43*	2.66	2.98*	3.74

Table 1 (continued)
Average Percent RMSE of Forecast over 1984 and 1985

	Akron	Canton	Cincinnati	Cleveland	Columbus	Dayton	Toledo	Youngstown
2-step								
AR3	2.75	4.01	2.18*	3.57	1.60*	4.30	3.28*	5.67
ARL3	3.09	3.97	2.28	3.63	2.42	4.49	5.11	3.84*
OLS3	2.44*	2.82*	2.57	3.18*	2.00	3.04*	4.52	5.93
3-step								
AR3	4.04	6.11	3.33	4.53	2.50	5.65	2.04*	8.83
ARL3	2.59*	5.43	2.85*	2.35*	2.71	5.23	4.16	5.08
OLS3	3.35	2.39*	3.30	3.80	2.12*	3.42*	4.98	4.22*

Table 2
Condition Indices for the Data Matrices

Akron	Canton	Cincinnati	Cleveland	Columbus	Dayton	Toledo	Youngstown
4.3	4.2	5.4	5.2	5.6	4.3	5.5	5.6
14.6	19.6	18.3	34.7	29.2	24.5	21.1	32.1
22.1	36.9	22.6	38.7	39.7	31.2	37.9	60.7
38.8	86.2	57.5	96.0	86.3	74.4	85.1	119.9
65.5	120.2	128.1	184.6	203.7	142.9	177.6	182.4
72.7	149.5	170.7	215.2	266.8	177.5	195.8	234.8
149.7	187.0	234.4	338.8	452.3	221.5	215.5	254.6
671.0	650.2	532.8	845.8	969.7	469.1	550.1	831.8

Table 3

Forecast Error Comparison of the Techniques

	Akron	Canton	Cincinnati	Cleveland	Columbus	Dayton	Toledo	Youngstown
1-step								
Best	2.84	2.02	2.35	2.49*	1.45*	2.56	2.87*	2.94
OLS1	2.89	2.02	2.70	2.80	3.75	2.56	3.45	3.1
Bayesian	2.89	1.64*	2.20*	2.80	1.72	2.56	3.45	3.19
ridge	2.76*	1.64*	2.20*	2.56	1.44*	2.43*	3.17	2.79*
1-step								
Best	1.77	1.73	1.29*	1.57	1.34	2.39*	1.97*	2.68
OLS2	1.98	1.73	2.21	1.57	1.37	2.53	1.97*	3.26
Bayesian	1.49*	1.73	2.07	1.57	0.95*	2.53	1.97*	2.49
ridge	1.47*	1.47*	1.58	1.47*	0.97*	2.53	1.96*	2.41*
2-step								
Best	2.71*	1.67*	2.34*	2.69*	1.54*	3.69*	3.05*	2.99
OLS2	3.54	2.71	3.04	2.69*	1.54*	4.23	3.05*	4.15
Bayesian	3.03	1.67*	3.05	2.70*	1.53*	4.24	3.05*	2.94
ridge	2.77	1.99	2.93	2.72*	1.59	4.24	3.03*	2.79*

Table 3 (continued)
Forecast Error Comparison of the Techniques

	Akron	Canton	Cincinnati	Cleveland	Columbus	Dayton	Toledo	Youngstown
1-step								
Best	1.53*	2.62	1.66*	1.86*	1.43*	2.64*	2.98*	3.45*
OLS3	1.93	2.84	2.31	1.86*	1.43*	2.66*	2.98*	3.74
Bayesian	1.67	2.70	2.06	1.86*	1.42*	2.66*	2.98*	3.71
ridge	1.85	2.58*	2.17	1.89*	1.44*	2.66*	2.99*	3.68
2-step								
Best	2.44	2.82*	2.18	3.18	1.60*	3.04	3.28*	3.84*
OLS3	2.44	2.82*	2.57	3.18	2.00	3.04	4.52	5.93
Bayesian	2.13*	2.82*	2.14	3.00*	2.00	3.04	4.52	5.67
ridge	2.27	2.85*	2.03*	3.14	1.97	2.99*	3.77	4.49
3-step								
Best	2.59	2.39	2.85	2.35	2.12	3.42*	2.04*	4.22
OLS3	3.35	2.39	3.30	3.80	2.12	3.42*	4.98	4.22
Bayesian	2.26*	2.38	2.43*	2.06	1.64	3.43*	3.48	4.06
ridge	2.44	2.32*	2.54	1.60*	1.51*	3.42*	2.04*	3.85*