Forecasting Employment Growth in Missouri with Many Potentially Relevant Predictors: An Analysis of Forecast Combining Methods

David E. Rapach and Jack K. Strauss

In this paper, the authors examine different approaches to forecasting monthly Missouri employment growth in the presence of many potentially relevant predictors, including both regional and national economic variables. Following Stock and Watson (2003, 2004), they first generate simulated out-of-sample forecasts of Missouri employment growth at horizons of 3, 6, 12, and 24 months using individual autoregressive distributed lag (ARDL) models based on 22 potential predictors. They then consider 20 different methods from the extant literature for combining the forecasts generated by the individual ARDL models. At longer horizons of 12 and 24 months, combining methods based on Bayesian shrinkage techniques produce out-of-sample forecasts that are substantially more accurate than forecasts from an autoregressive (AR) benchmark model. Combining methods based on Bayesian shrinkage techniques also outperform simple combining methods (such as those that use the mean or median of the individual forecasts) at longer horizons. Nevertheless, simple combining methods consistently outperform the AR benchmark model at all horizons and appear to offer a low-cost way of generating reliable combination forecasts.

Federal Reserve Bank of St. Louis *Regional Economic Development*, 2005, *1*(1), pp. 97-112.

1 INTRODUCTION

Bates and Granger's (1969) seminal work showed that combinations of individual forecasts often outperform individual forecasts. Well over a quarter century later, Stock and Watson (2004) analyze more than a dozen different methods for combining forecasts of output growth in the G7 countries and find that combination forecasts are a useful way of incorporating information from a large number of potentially relevant predictors. They show that combination forecasts of output growth often outperform forecasts generated by a benchmark autoregressive (AR) model and that simple methods, such as simple averaging or trimmed averaging of a large number of individual autoregressive distributed (ARDL) model forecasts, typically outperform more complicated methods. Combination forecast methods can exploit the information in a large number of potential predictors. This is especially relevant when forecasting a variable like national output growth, subject to both supply and demand shocks and possible instabilities in the data. In this case, it is difficult to know a priori which particular variables are the most relevant and, moreover, it is unlikely that a forecaster can specify a single econometric model that closely corresponds to the actual—and perhaps unknowable—data-generating process.

Although most of the literature, including

David E. Rapach is an assistant professor of economics and Jack K. Strauss is a professor of economics in the Department of Economics at Saint Louis University. The authors thank Michael Owyang, Jeremy Piger, Howard Wall, and BERG seminar participants at the First Annual Conference of the Business and Economics Research Group (BERG) of the Federal Reserve Bank of St. Louis for helpful comments. The authors acknowledge research support from the Simon Center for Regional Forecasting at Saint Louis University. The views expressed are the authors and do not necessarily represent the official positions of Saint Louis University.

Stock and Watson (1999, 2003, 2004), examines combination forecasts of national economic variables such as output growth and inflation, we are interested in the usefulness of combining methods when forecasting a regional economic variable. As discussed above, combining methods are likely to be useful when predicting aggregate economic variables because they incorporate information from a large number of potential predictors. This is also likely to be the case when forecasting regional economic variables, as a large number of both national and regional variables may contain information useful for forecasting. In the present paper, we consider forecasting employment growth in Missouri in the presence of a large number of potentially relevant predictors. Generating accurate forecasts of regional variables is important for planning purposes for businesses and state and local governments. Evaluating forecast combining methods for predicting a regional economic variable represents a natural complement to the extant literature on national economic variables.

We analyze forecasts of Missouri employment growth over the 1995:01–2005:01 out-of-sample period. This period includes the expansion of the late 1990s, the 2001 recession, and the subsequent "jobless" recovery, so it should represent an informative laboratory for analyzing forecasts of Missouri employment growth. We consider 22 potential predictors of Missouri employment growth, including 9 regional and 13 national economic variables. Following Stock and Watson (2003, 2004), we first generate simulated out-of-sample forecasts of Missouri employment growth from individual ARDL models, with each ARDL model based on 1 of the 22 potential predictors. We then use 20 different methods from the extant literature to construct combination forecasts of the individual ARDL model forecasts. The combining methods are based on the following: simple averaging using the mean, median, or trimmed mean (Stock and Watson, 2003, 2004); ordinary least squares (OLS; Granger and Ramanathan, 1984); weighted least squares (WLS; Diebold and Pauly, 1987); discount mean squared forecast error (MSFE; Stock and Watson, 2004); Bayesian shrinkage techniques (Clemen and Winkler, 1986; Diebold and Pauly, 1990); clusters formed on the basis of MSFE (Aiolfi and Timmermann, 2005); model selection (Swanson and Zeng, 2001); principal components (Chan, Stock, and Watson, 1999; Stock and Watson, 2004); approximate Bayesian model averaging (Draper, 1995); and exponential reweighting (Yang, 2004).

Previewing our results, we find that forecast combining methods can improve the forecasting of employment growth in Missouri, especially at longer horizons of 1 and 2 years. In particular, combining methods based on Bayesian shrinkage techniques generate forecasts that are up to 29 percent and 49 percent more accurate, in terms of MSFE, than the forecasts produced by an AR benchmark model at horizons of 1 and 2 years, respectively. Combination forecasts based on Bayesian shrinkage techniques also have an MSFE that is close to or below that of the best individual ARDL model forecast at horizons of 1 and 2 years, and Bayesian shrinkage combination forecasts have a lower MSFE than simple combining methods at these horizons. It should be noted that a number of the forecast combining methods fail to outperform the AR benchmark model, implying that it is critical to carefully select combining methods when forecasting Missouri employment growth. Simple combining methods appear to offer a low-cost way of generating reliable combination forecasts, as they consistently outperform the AR benchmark model at all horizons, in agreement with the findings of Stock and Watson (2004).

The rest of the paper is organized as follows: Section 2 describes the econometric methodology, Section 3 reports the empirical results, and Section 4 concludes.

2 ECONOMETRIC METHODOLOGY

2.1 Individual Forecasts

Let $\Delta y_t = y_t - y_{t-1}$, where y_t is the log-level of Missouri employment at time t, and let

$$y_{t+h}^{h} = (1/h)\sum_{j=1}^{h} \Delta y_{t+j} ,$$

so that y_{t+h}^{h} is the growth rate of Missouri employment over the next h months expressed at a monthly rate. Consider the following ARDL model:

(1) $$y^h_{t+h} = \alpha + \sum_{j=0}^{q_1-1} \beta_j \Delta y_{t-j} + \sum_{j=0}^{q_2-1} \gamma_j x_{i,t-j} + \varepsilon^h_{t+h},$$

where $x_{i,t}$ is one of the potentially relevant predictors ($i = 1,\ldots,n$), h is the forecast horizon, and ε^h_{t+h} is an error term. We consider forecast horizons of 3, 6, 12, and 24 months ($h = 3,6,12,24$). In order to form recursive simulated out-of-sample forecasts of y^h_{t+h} using equation (1), we first divide the sample into in-sample and out-of-sample portions, where the first R observations comprise the in-sample period and the last P observations make up the out-of-sample period. We compute the initial out-of-sample forecast for y^h_{R+h} based on the predictor $x_{i,t}$ as

$$\hat{y}^h_{i,R+h|R} = \hat{\alpha}_R + \sum_{j=0}^{q_1-1} \hat{\beta}_{j,R} \Delta y_{R-j} + \sum_{j=0}^{q_2-1} \hat{\gamma}_{j,R} x_{i,R-j},$$

where $\hat{\alpha}_R$, $\hat{\beta}_{j,R}$, and $\hat{\gamma}_{j,R}$ are the OLS estimates of α, β_j, and γ_j, respectively, in equation (1) using data through period R. We select the lag lengths (q_1 and q_2) in equation (2) using the Akaike information criterion (AIC) and data through period R considering a minimum lag length of 0 for q_1 and 1 for q_2 (thus ensuring that the potential predictor $x_{i,t}$ appears in equation (1)) and a maximum lag length of 12 for q_1 and q_2. We form the second out-of-sample forecast by updating the above process using data through period R+1. Continuing in this manner, we end up with a series of $P - (h - 1)$ simulated out-of-sample forecasts corresponding to the predictor $x_{i,t}$, $\{y^h_{i,t+h|t}\}^{T-h}_{t=R}$. Note that q_1 and q_2 are selected anew when computing each recursive out-of-sample forecast, so that the ARDL lag lengths in the forecasting equations can vary over time. We consider 22 potential predictors that define the individual ARDL models ($n = 22$). Apart from data availability and revisions, these simulated out-of-sample forecasts mimic the situation of a forecaster in real time.[1]

An AR model, equation (1) with the restriction $\gamma_j = 0$ for all j imposed, serves as the benchmark model. This is a common benchmark model when forecasting time-series variables. The AR model forecasts are computed recursively in a manner similar to the ARDL model forecasts, with the lag length selected by the AIC given a minimum (maximum) lag length of 0 (12). This produces a series of $P - (h - 1)$ simulated out-of-sample forecasts corresponding to the AR benchmark model, $\{y^h_{AR,t+h|t}\}^{T-h}_{t=R}$.

2.2 Forecast Combining Methods

We consider 20 different methods for combining the individual forecasts generated by the $n = 22$ ARDL models, and the methods can be organized into 10 different classes. Most of the forecast combining methods require a holdout period to calculate the weights used to combine the individual ARDL model forecasts, and we use the first P_0 out-of-sample forecast observations as holdout observations. All of the combining methods take the form of a linear combination of the individual forecasts:

(2) $$\hat{y}^h_{c,t+h|t} = w_{0,t} + \sum_{i=1}^{n} w_{i,t} \hat{y}^h_{i,t+h|t},$$

where $\hat{y}^h_{c,t+h|t}$ is a given combination forecast whose weights, $\{w_{i,t}\}^n_{i=0}$, are typically calculated using the individual out-of-sample forecasts and y^h_{t+h} observations available from the start of the holdout out-of-sample period to time t. For each of the combining methods, we form combination forecasts over the post-holdout out-of-sample period, yielding $\{\hat{y}^h_{c,t+h|t}\}^{T-h}_{t=R+P_0}$, for a total of $T - (h - 1) - (R + P_0)$ forecasts available for evaluation. We compare the forecasts generated by each of the 20 combining methods, as well as the AR benchmark model, with the actual observations of employment growth over the post-holdout out-of-sample period, $\{y^h_{t+h}\}^{T-h}_{t=R+P_0}$.[2]

2.2.1 Simple Combining Methods. We consider three simple methods of combining individual forecasts: mean, median, and trimmed mean. Stock and Watson (2003, 2004) find that simple

[1] Although data availability is not an issue for financial variables, some nonfinancial variables are only available after a 1- to 2-month lag. Given this, it will generally be infeasible to use the procedure described in the text in real time at horizons of 1 to 2 months. At horizons beyond 2 months, say, 5 months, it is feasible to use the procedure to generate a forecast of cumulative employment growth over the previous 2 and subsequent 3 months.

[2] To be clear, out-of-sample forecasts are generated for the individual ARDL models over the entire out-of-sample period, which consists of both the holdout and post-holdout periods, using the recursive procedure described in Section 2.1.

combining methods work well in forecasting inflation and output growth using a large number of potential predictors in the G7 countries. The mean sets $w_{0,t} = 0$ and $w_{i,t} = (1/n)$ for all i in equation (2); the median uses the sample median of $\{\hat{y}^h_{i,t+h|t}\}^n_{i=1}$; the trimmed mean sets $w_{0,t} = 0$ and $w_{i,t} = 0$ for the individual models that produce the smallest and largest forecasts at time t, while $w_{i,t} = 1/(n-2)$ for the remaining individual models.[3]

2.2.2 OLS Combining Methods. Granger and Ramanathan (1984) recommend combining forecasts using unrestricted OLS. We consider OLS combination forecasts where the OLS coefficients are estimated using either a recursive or rolling window. To compute the initial OLS combination forecast (for $y^h_{R+P_0+h}$) using a recursive window, we regress $\{y^h_{s+h}\}^{R+(P_0-1)-(h-1)}_{s=R}$ on a constant and $\{\hat{y}^h_{i,s+h|s}\}^{R+(P_0-1)-(h-1)}_{s=R}$, $i = 1,\ldots,n$, and set the combining weights in equation (2) equal to the estimated OLS coefficients. To construct the second combination forecast (for $y^h_{R+(P_0+1)+h}$), the OLS coefficients are estimated by regressing $\{y^h_{s+h}\}^{R+(P_0-1)-(h-1)+1}_{s=R}$ on a constant and $\{\hat{y}^h_{i,s+h|s}\}^{R+(P_0-1)-(h-1)+1}_{s=R}$, $i = 1,\ldots,n$, and the fitted OLS coefficients again serve as the combining weights in equation (2). We proceed in this fashion through the end of the available out-of-sample period. The OLS combination forecasts based on a rolling window are computed in a similar manner, with the exception that in computing the second combination forecast, for example, the OLS coefficients that serve as the combining weights in equation (2) are estimated by regressing $\{y^h_{s+h}\}^{R+(P_0-1)-(h-1)+1}_{s=R+1}$ on a constant and $\{\hat{y}^h_{i,s+h|s}\}^{R+(P_0-1)-(h-1)+1}_{s=R+1}$, $i = 1,\ldots,n$.

2.2.3 WLS Combining Methods. Diebold and Pauly (1987) argue that combination forecasts based on time-varying weights can enhance forecasting performance in the presence of structural change. We use their "t-lambda" method. It follows the OLS combining method based on a recursive estimation window described in Section 2.2.2 above, with the exception that the combining weights are calculated using WLS instead of OLS. Diebold and Pauly (1987) recommend the weighting matrix $\Psi = diag[\psi_{tt}] = kt^{\lambda}$, where $k, \lambda > 0$, t = 1,…,T, and T is the number of observations used in the WLS regression. Under this approach, observations from the recent past receive more weight than observations from the distant past when computing the combining coefficients.[4] We consider $\lambda = 1$, which corresponds to weights that decrease at a constant rate as we move further into the past, and $\lambda = 3$, which corresponds to weights that decrease at an increasing rate.

2.2.4 Discount MSFE Combining Methods. Stock and Watson (2004) consider a combining method, where the weights in equation (2) depend inversely on the historical forecasting performance of the individual models. Their discount (or inverse) MSFE combining method employs the weights,

$$w_{i,t} = m_{it}^{-1} \Big/ \sum_{j=1}^{n} m_{j,t}^{-1}, \tag{3}$$

where

$$m_{i,t} = \sum_{s=R}^{t-h} \delta^{t-h-s} \left(y^h_{s+h} - \hat{y}^h_{i,s+h|s} \right)^2, \tag{4}$$

$w_{0,t} = 0$, and δ is a discount factor. Note that when $\delta = 1$, there is no discounting and equation (3) yields the optimal combination forecast derived by Bates and Granger (1969) for the case where the individual forecasts are uncorrelated; when $\delta < 1$, greater importance is attached to the recent forecasting performance of the individual models. We consider δ values of 1.0 and 0.9. Stock and Watson (2004) also consider a "most recently best" approach, where the "combination" forecast is the forecast corresponding to the individual model with the best forecasting performance over the previous year, and we include this approach in our analysis.

2.2.5 Bayesian Shrinkage Methods. In the presence of a relatively large number of individual forecasts, Bayesian shrinkage techniques may be helpful in forming combination forecasts, as suggested by Clemen and Winkler (1986) and

[3] The simple combining methods obviously do not require holdout out-of-sample observations.

[4] Using familiar notation, the WLS estimator can be expressed as $\hat{\beta}_{WLS} = (X'\Psi^{-1}X)^{-1}(X'\Psi^{-1}Y)$. Note that the value of k is arbitrary, because it disappears in the computation of the WLS estimator.

Diebold and Pauly (1990). We follow Stock and Watson (2004) and consider the following shrinkage combination forecast, which Diebold and Pauly (1990) show can be viewed as a Bayesian estimator:

$$w_{i,t} = \lambda \hat{\beta}_{i,t} + (1-\lambda)(1/n), \quad (5)$$

where $w_{0,t} = 0$, $\hat{\beta}_{i,t}$ is the OLS coefficient estimate corresponding to individual forecast i (the OLS coefficients are estimated using the recursive window scheme described in Section 2.2.2 above, with the exception that the intercept term is restricted to zero), $\lambda = \max\{0,1 - \kappa[n/(t-h-R-n)]\}$, and κ is a parameter that governs the degree of shrinkage toward equal weights. Larger values of κ correspond to smaller values of λ and thus more shrinkage toward equal weights. We consider κ values of 0.5 and 1.0.

2.2.6 Cluster Combining Methods. Aiolfi and Timmermann (2005) investigate persistence in forecasting performance and develop conditional combining methods. We use their $C(K, PB)$ algorithm, which proceeds as follows. To form the initial combination forecast, we first compute the MSFE for the individual forecasts $\{\hat{y}^h_{i,s+h|s}\}_{s=R}^{R+(P_0-1)-(h-1)}$, $i = 1,\ldots,n$, and group the individual models into K equal-sized clusters, where the first cluster contains the individual models with the lowest MSFE values, the second cluster contains the individual models with the next-lowest MSFE values, and so on. The first combination forecast, $\hat{y}^h_{c,R+P_0+h}$, is the average of the individual forecasts of $y^h_{R+P_0+h}$ generated by the models included in the first cluster. To form the second combination, we compute the MSFE for the individual forecasts $\{\hat{y}^h_{i,s+h|s}\}_{s=R+1}^{R+(P_0-1)-(h-1)+1}$, $i = 1,\ldots,n$, and group the individual models into clusters (so that the clusters are formed based on a rolling window), and the second combination forecast, $\hat{y}^h_{c,R+(P_0+1)+h}$, is the average of the individual forecasts of $y^h_{R+(P_0+1)+h}$ included in the first cluster. We proceed in this manner through the end of the available out-of-sample period. Following Aiolfi and Timmermann (2005), we consider $K = 2$ and $K = 3$ in our applications.

2.2.7 Model Selection Combining Methods. Swanson and Zeng (2001) consider combining methods based on model selection. We use their M-TST model selection approach, which uses a general-to-specific modeling procedure. We proceed as described in Section 2.2.2 above for the OLS combining method based on a rolling window, with the exception that we first examine the t-statistics corresponding to the estimated slope coefficients of the combining regression, where the t-statistics are calculated using heteroskedasticity and autocorrelation consistent (HAC) standard errors.[5] If any of the individual t-statistics are less then 1.645 in absolute value, we exclude these individual forecasts from the OLS regression used to estimate the combining weights. If all of the t-statistics are less than 1.645 in absolute value, we include all of the individual forecasts in the OLS regression.[6]

2.2.8 Principal Component Combining Methods. Chan, Stock, and Watson (1999) and Stock and Watson (2004) consider forming combination forecasts using the first m principal components of the individual forecasts. Let $\hat{F}^h_{1,s+h|s},\ldots,\hat{F}^h_{m,s+h|s}$, $s = R,\ldots,t$, represent the first m estimated principal components of the uncentered second-moment matrix of the individual forecasts, $\hat{y}^h_{i,s+h|s}$, $i = 1,\ldots,n$, $s = R,\ldots,t$. To form a combination forecast of y^h_{t+h} based on the fitted principal components, we estimate the regression model,

$$y^h_{s+h} = \phi_1 \hat{F}^h_{1,s+h|s} + \ldots + \phi_m \hat{F}^h_{m,s+h|s} + v^h_{s+h}, \quad (6)$$

where $s = R,\ldots,t-h$. The combination forecast is given by $\hat{y}^h_{c,t+h|t} = \hat{\phi}_1 \hat{F}^h_{1,t+h|t} + \ldots + \hat{\phi}_m \hat{F}^h_{m,t+h|t}$, where $\hat{\phi}_1,\ldots,\hat{\phi}_m$ are the OLS estimates of $\phi_1,\ldots,\phi_m$, respectively, in equation (6). We use $m = 1$ and $m = 2$ in computing forecasts using the principal component (PC) method.

2.2.9 Approximate Bayesian Model Averaging Combining Methods. Following Garratt et al. (2003), we compute combining weights using approximate Bayesian model averaging (ABMA),

[5] We use Newey and West (1987) HAC standard errors with a lag truncation of $h-1$.

[6] Swanson and Zeng (2001) also consider model selection based on the AIC or Schwarz information criterion (SIC). However, this involves computing the AIC or SIC for every possible combination of individual forecasts in the OLS regression model, which is impractical in our applications, as $n = 22$ so that there are $2^n - 1 = 4{,}194{,}303$ possible combinations of the individual forecasts.

in which functions of the SIC are used to approximate the posterior probabilities of the individual models (Draper, 1995). The combining weights can be expressed as

$$w_{i,t} = e^{\Delta_{i,t}} / \sum_{j=1}^{n} e^{\Delta_{j,t}}, \quad i = 1,\ldots,n, \tag{7}$$

where $\Delta_{i,t} = SIC^h_{i,t} - \max_j(SIC^h_{j,t})$, and $SIC^h_{i,t}$ is the SIC corresponding to the fitted ARDL model i given by equation (1) used to generate $\hat{y}^h_{i,t+h|t}$. Garratt et al. (2003) also follow Burnham and Anderson (1998) and compute weights using the AIC, so that $\Delta_{i,t} = AIC^h_{i,t} - \max_j(AIC^h_{j,t})$ in equation (7). We consider ABMA combining weights based on both the SIC and AIC.[7]

2.2.10 Exponential Reweighting Combining Methods. Yang (2004) develops what he labels the AFTER (aggregated forecast through exponential reweighting) algorithm to combine forecasts from individual models. Yang (2004) shows that the algorithm can be viewed as an optimal combination procedure under fairly general conditions. The weights for the AFTER algorithm are given by

$$w_{i,t} = \theta_{i,t} / \sum_{j=1}^{n} \theta_{j,t}, \tag{8}$$

where

$$\theta_{i,t} = \prod_{s=R}^{t-h} \hat{\upsilon}_{i,s}^{0.5} e^{-0.5 \sum_{s=R}^{t-h} [(y^h_{s+h} - \hat{y}^h_{i,s+h|s})^2 / \hat{\upsilon}_{i,s}]}, \tag{9}$$

and $\hat{\upsilon}_{i,t}$ is the OLS estimate of the variance of ε^h_{t+h} for the fitted ARDL forecasting model i (equation (1)) used to generate $\hat{y}^h_{i,t+h|t}$.

3 EMPIRICAL RESULTS

3.1 Data

Missouri employment growth is measured as the first difference in the log-levels of seasonally adjusted Missouri employment (multiplied by 100). The cumulative Missouri employment growth is divided by h, thereby expressing employment growth at the average monthly rate over the forecast horizon. We consider 9 regional and 13 national economic variables, for a total of 22 potential predictors ($x_{i,t}$ in equation (1)), which include labor market, production, and financial variables that are commonly used to forecast economic activity. The 9 regional variables are the Missouri unemployment rate and employment growth in the eight states that border Missouri (Arkansas, Illinois, Iowa, Kansas, Kentucky, Nebraska, Tennessee, and Oklahoma). The data are seasonally adjusted and, like the Missouri employment data, are from the Bureau of Labor Statistics (BLS). Based on availability, the data span from 1976:01 to 2005:01.[8] The 13 national economic variables are the following: U.S. employment; U.S. unemployment rate; capacity utilization rate; average weekly manufacturing hours; average weekly initial claims for unemployment insurance; manufacturers' new orders for consumer goods and materials; vendor performance (sales); manufacturers' new orders for nondefense capital goods; building permits; stock prices (S&P 500 index); interest rate spread (10-year Treasury bond yield minus the federal funds rate); national overtime; and industrial production. The U.S. employment and unemployment rate series are from the BLS; the remaining national variables are from the Conference Board. With three exceptions, all of the national variables are

[7] Like the simple combining methods, the ABMA combining methods do not require holdout out-of-sample observations.

[8] Supporting our specifications, the unit root tests of Ng and Perron (2001) clearly indicate that the log-levels of Missouri employment and employment in the eight bordering states are $I(1)$, while the Missouri unemployment rate is $I(0)$. We also tested for cointegration between the log-levels of Missouri employment and the log-levels of employment in each of the eight bordering states because equation (1) should potentially include an error-correction term in the event of cointegration. We only found evidence of cointegration between the log-levels of employment in Missouri and Kansas and Missouri and Oklahoma. However, Missouri employment appears weakly exogenous with respect to both of these variables, so the inclusion of an error-correction term in equation (1) is not necessary.

[9] With one exception, Ng and Perron (2001) unit root tests clearly support our specifications for the national variables. The exception is the U.S. unemployment rate, where the unit root null hypothesis cannot be rejected at the 10 percent level; however, the null hypothesis is very nearly rejected at the 10 percent level. We obtain similar results in our applications below if we the use first differences of the U.S. unemployment rate instead of the levels. Where relevant, we also found little evidence that an error-correction term needs to be included in equation (1) when any of the national variables serve as predictors.

Table 1

MSFE statistics for the Individual ARDL Models of Missouri Employment Growth, 1995:01–2005:01 Out-of-Sample Period

Variable	$h = 3$	$h = 6$	$h = 12$	$h = 24$
AR benchmark	1.92	1.21	0.88	0.84
Regional variables				
Missouri unemployment rate	0.97	0.94	1.01	0.98
Arkansas employment	1.00	1.00	0.92	0.93
Illinois employment	1.04	0.99	0.94	1.18
Iowa employment	0.98	0.92	0.95	0.91
Kansas employment	0.95	0.91	0.92	1.32
Kentucky employment	**0.90**	0.86	0.79	0.87
Nebraska employment	1.17	1.08	1.04	0.95
Tennessee employment	0.94	1.02	**0.70**	**0.76**
Oklahoma employment	0.97	1.00	1.06	1.22
National variables				
U.S. employment	1.11	0.93	0.98	1.16
U.S. unemployment rate	1.47	1.63	2.14	2.18
Capacity utilization	1.07	1.19	1.09	1.08
Average weekly hours, manufacturing	1.28	1.05	0.94	1.00
Unemployment claims	1.15	1.05	0.94	0.91
New manufacturing orders	1.06	1.03	0.96	0.98
Vendor sales	1.00	1.00	1.00	1.00
New manufacturing capital orders	1.08	1.15	1.26	1.30
Building permits	1.02	1.03	0.96	0.99
Stock market index	1.41	1.90	2.17	1.54
Interest rate spread	1.09	1.01	1.07	0.92
National overtime	0.95	**0.83**	0.87	0.94
Industrial production	1.30	0.98	1.00	1.00

NOTE: The first row reports the MSFE for the AR benchmark model; the remaining rows report the ratio of the MSFE for the individual ARDL model to the MSFE for the AR benchmark model. A bold entry signifies the ARDL model with the lowest MSFE at a given horizon.

measured in monthly growth rates (first differences of log-levels multiplied by 100); the three exceptions are the U.S. unemployment rate, capacity utilization rate, and interest rate spread, which are specified in levels.[9] While we do not claim that our list of 9 regional and 13 national variables constitutes an exhaustive list of potential predictors of Missouri employment growth, it does include a large number of potentially relevant predictors that are likely to be useful for our analysis.

3.2 Out-of-Sample Forecasting Results

We evaluate out-of-sample forecasts of Missouri employment growth over the 1995:01 to 2005:01 period. This period includes the late-1990s expansion, 2001 recession, and subsequent "jobless" recovery—an informative period in which to evaluate forecasts of Missouri employment growth. We consider "short" forecast horizons of 3 and 6 months and "long" forecast horizons of 12 and 24 months. As discussed in Section 2.2

Table 2

Forecast Combining Results for Missouri Employment Growth, 1995:01–2005:01 Out-of-Sample Period

	h = 3				*h* = 6			
Combination method	MSFE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	MSFE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2
AR benchmark	1.92	0.08	0.57*	0.08	1.21	–0.08	0.65	0.10
Mean	0.96	–0.02	0.67	0.10	0.92	–0.60	1.14	0.21
Median	0.94	–0.03	0.68	0.10	0.92	–0.54	1.07	0.22
Trimmed mean	**0.93**	–0.06	0.72	0.11	**0.90**	–0.57	1.12	0.23
OLS, recursive	1.44	0.53*	0.17**	0.02	1.56	0.33	0.35**	0.20
OLS, rolling	2.28	0.61**	0.07**	0.01	2.45	0.31	0.25**	0.21
WLS: t-lambda, $\lambda = 1$	1.65	0.58*	0.11**	0.01	1.90	0.36	0.28**	0.20
WLS: t-lambda, $\lambda = 3$	2.06	0.58*	0.10**	0.01	2.30	0.35	0.24**	0.20
Discount MSFE, $\delta = 1.0$	0.95	–0.02	0.67	0.10	0.92	–0.64	1.16	0.22
Discount MSFE, $\delta = 0.9$	0.95	0.01	0.65	0.10	0.91	–0.53	1.06	0.18
Most recently best	1.07	0.20	0.46**	0.10	1.03	0.55	0.12**	0.01
Shrinkage, $\kappa = 0.5$	1.18	0.47	0.25**	0.03	1.21	0.28	0.48**	0.24
Shrinkage, $\kappa = 1.0$	1.08	0.38	0.35**	0.04	1.02	0.19	0.59	0.24
Cluster: $C(2, PB)$	0.96	–0.02	0.67	0.09	0.94	–0.54	1.01	0.16
Cluster: $C(3, PB)$	0.96	0.07	0.61	0.08	0.94	–0.70	1.14	0.19
Model selection: M-TST	1.82	0.57*	0.12**	0.02	1.93	0.32	0.22**	0.13
PC, $m = 1$	0.96	0.00	0.65	0.10	0.97	–0.68	1.04	**0.30**
PC, $m = 2$	0.96	–0.07	0.71	0.09	0.97	–0.62	0.96	0.25
ABMA, SIC	1.07	0.21	0.45**	0.06	1.01	0.05	0.55	0.05
ABMA, AIC	1.05	0.13	0.50**	0.09	1.17	–0.26	0.77	0.23
AFTER	0.94	–0.05	0.69	**0.12**	0.99	–0.14	0.71	0.19

NOTE: The first row reports the MSFE for the AR benchmark model; the remaining rows report the ratio of the MSFE for the combining method to the MSFE for the AR benchmark model. $\hat{\alpha}_0$, $\hat{\alpha}_1$, and R^2 are the intercept estimate, slope estimate, and goodness-of-fit measure, respectively, for the MZ regression. A bold entry signifies the combining method with the lowest MSFE or the highest R^2 at a given horizon; * and ** indicate significance at the 5 percent and 1 percent levels, respectively, for a test of the null hypothesis that $\alpha_0 = 0$ ($\alpha_1 = 1$) for $\hat{\alpha}_0$ ($\hat{\alpha}_1$).

above, we need a holdout out-of-sample period in order to compute most of the combination forecasts, and we use the 60 observations preceding 1995:01 as the holdout out-of-sample period.

Table 1 reports out-of-sample forecasting results for the AR benchmark model and the individual ARDL models. The table reports the MSFE statistics for the AR benchmark model and the ratio of the MSFE for the individual ARDL models to the MSFE for the AR benchmark model. The first row of the table shows that the MSFE declines as the horizon increases for the AR benchmark model, suggesting that more accurate forecasts of average monthly Missouri employment growth are available at longer horizons. Observe that either five or six of the nine individual ARDL models based on the regional variables have lower MSFE statistics than those for the AR benchmark model at all reported horizons. The performance of national variables is poorer, as most individual ARDL models do not have lower MSFE statistics than those for the AR benchmark, particularly at shorter horizons. It would seem difficult to ascertain a priori which of the individual potential

h = 12				h = 24			
MSFE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	MSFE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2
0.88	–0.08	0.65	0.10	0.84	–0.18	0.63	0.04
0.84	–0.60	1.14	0.21	0.83	–0.39	0.88	0.05
0.85	–0.54	1.07	0.22	0.94	–0.49	0.90	0.07
0.82	–0.57	1.12	0.23	0.83	–0.36	0.85	0.07
1.27	0.33	0.35**	0.20	0.95	–0.12	0.58	0.52
2.81	0.31	0.25**	0.21	1.62	0.14	0.36	0.47
1.87	0.36	0.28**	0.20	1.30	–0.08	0.50	0.50
2.65	0.35	0.24**	0.20	1.94	–0.05	0.42**	**0.53**
0.85	–0.64	1.16	0.22	0.91	–0.72	1.10	0.07
0.87	–0.53	1.06	0.18	0.91	–0.68	1.08	0.05
1.32	0.55	0.12**	0.01	0.87	–0.09	0.60	0.15
0.85	0.28	0.48**	0.24	**0.51**	–0.16	0.80	0.33
0.71	0.19	0.59	0.24	0.53	–0.28	0.92	0.28
0.97	–0.54	1.01	0.16	1.06	0.54	0.07	0.00
0.96	–0.70	1.14	0.19	1.11	0.34	0.22	0.00
3.00	0.32	0.22**	0.13	1.73	0.13	0.36**	0.48
1.00	–0.68	1.04	**0.30**	1.36	–1.15	1.17	0.23
1.10	–0.62	0.96	0.25	1.27	–0.43	0.73	0.12
1.00	0.05	0.55	0.05	1.16	0.99**	–0.32**	0.04
0.93	–0.26	0.77	0.23	0.90	–0.95	1.28	0.13
0.92	–0.14	0.71	0.19	0.89	–0.89	1.23	0.15

predictors in the ARDL models are likely to display the best forecasting ability, and this provides a motivation for considering methods for combining the large number of individual ARDL forecasts.

Table 2 presents results for the 20 forecast combining methods over the 1995:01–2005:01 out-of-sample period. Similar to Table 1, Table 2 reports the MSFE for the AR benchmark model and the ratio of the MSFE for a given combining method to the MSFE for the AR benchmark model. Table 2 also reports the estimated intercept, estimated slope, and R^2 statistic for a Mincer and Zarnowitz (MZ, 1969) regression of the form

$$y^h_{t+h} = a_0 + a_1 \hat{y}^h_{c,t+h|h} + \eta^h_{t+h}, \qquad (10)$$

where $a_0 = 0$ and $a_1 = 1$ when the forecasts are unbiased. We indicate in Table 2 whether $\hat{a}_0$ ($\hat{a}_1$) is significantly different from 0 (1), where $\hat{a}_0$ ($\hat{a}_1$) is the OLS estimate of a_0 (a_1) in equation (10).[10]

[10] The *t*-statistics used to assess the statistical significance are based on Newey and West (1987) standard errors with a lag truncation of h–1.

Given that the forecasts are unbiased, the R^2 statistic provides a measure of the ability of the forecasts to explain movements in actual Missouri employment growth.

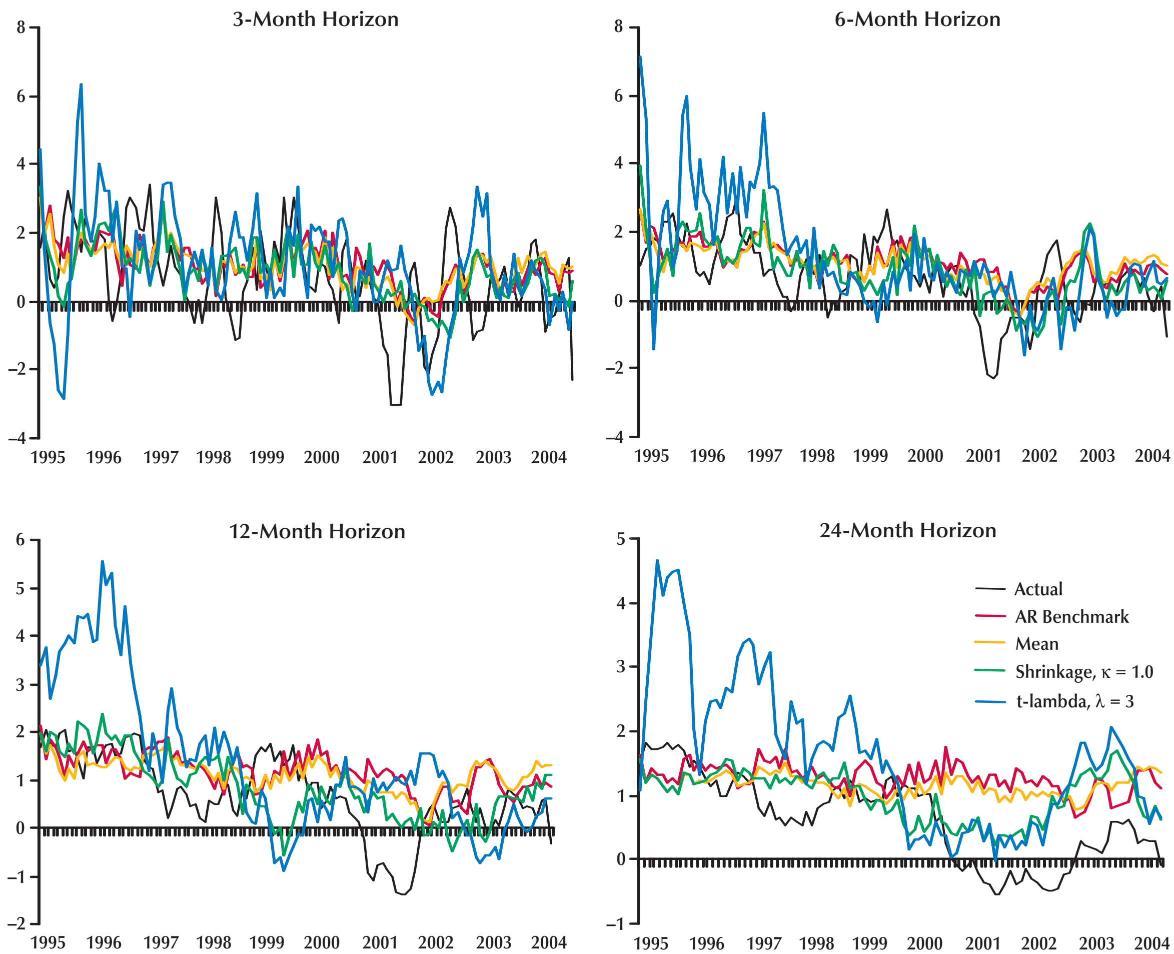
Among the forecast combining methods, there is considerable dispersion of results. The simple combining methods (mean, median, and trimmed mean) consistently outperform the AR benchmark, with reductions in MSFE of around 5 percent to 15 percent relative to the AR benchmark, and the R^2 statistic for the MZ regressions are greater than those for the AR benchmark. In addition, the estimated intercept and slope coefficients are not significantly different from 0 and 1, respectively, so that the simple combining methods appear to produce unbiased forecasts. At shorter horizons of 3 and 6 months, the trimmed mean combining method outperforms all of the other combining methods in terms of MSFE. In terms of the R^2 statistic of the MZ regression, the AFTER procedure performs marginally better than other simple combining methods at $h = 3$, and the PC ($m = 1$) method provides the best MZ fit at $h = 6$, with an R^2 statistic of 0.30. Both the AFTER and PC methods have estimated intercept and slope coefficients that are not significantly different from 0 and 1, respectively, in the MZ regression. Observe that many of the combining methods, especially the OLS and t-lambda methods, produce forecasts that are both substantially less accurate than the AR benchmark and biased according to the MZ regression results. Also note that the trimmed mean, the best performing combining method at horizons of 3 and 6 months, has an MSFE that is reasonably close to that of the best performing individual ARDL model at these horizons (Kentucky employment and national overtime, respectively). Given that it will be very difficult a priori for a researcher to select the individual variable that will perform the best, this helps to demonstrate the usefulness of the trimmed mean combining method at shorter horizons.

At longer horizons of 12 and 24 months, Table 2 shows that several of the combining methods lead to sizable reductions in MSFE and increases in the MZ R^2 statistics relative to the AR benchmark model. In particular, the shrinkage method with $\kappa = 1.0$ ($\kappa = 0.5$) leads to reductions in MSFE of 29 percent (15 percent) and 47 percent (49 percent) relative to the AR benchmark at horizons of 12 and 24 months, respectively. The coefficient estimates from the MZ regression indicate that the shrinkage forecasts are unbiased, with the exception of the slope coefficient when $\kappa = 0.5$ and $h = 12$. In addition, the shrinkage method with $\kappa = 1.0$ performs nearly as well as the best individual ARDL model (Tennessee employment) at the 12-month horizon and better than the best individual ARDL model (Tennessee employment) at the 24-month horizon, further demonstrating the usefulness of the shrinkage method with $\kappa = 1.0$ at longer horizons. The OLS and t-lambda procedures yield relatively high MZ R^2 statistics at $h = 12$ and $h = 24$, but the estimated slope coefficients in the MZ regressions are significantly less than 1 at the 1-year horizon. Moreover, these combining methods have MSFE statistics well above those for the AR benchmark, with the exception of the OLS recursive method when $h = 24$. The discount MSFE method appears to offer reasonably large reduction in MSFE relative to the AR benchmark model at horizons of 12 and 24 months. However, the reductions in MSFE associated with the discount MSFE methods are smaller than those for the simple combining methods. In fact, the simple combining methods offer fairly sizable reductions in MSFE relative to the AR benchmark and generate unbiased forecasts at horizons of 12 and 24 months. Overall, the simple combining methods perform consistently well at all reported horizons in Table 2 and seem to offer a low-cost way of generating reliable forecasts of Missouri employment growth.

To gain further insight into the relative forecasting performances of some of the combining methods, Figure 1 plots the realized observations of y_{t+h}^{h} and the forecasts generated by the AR benchmark model and the mean; shrinkage, $\kappa = 1.0$; and t-lambda, $\lambda = 3$ combining methods. The shrinkage, $\kappa = 1.0$ method is shown because it has the lowest MSFE at $h = 12$ and next-to-lowest MSFE at $h = 24$; the t-lambda, $\lambda = 3$ method is shown because it has the highest MZ R^2 statistic at $h = 24$. A problem with the t-lambda method,

Figure 1

Actual Missouri Employment Growth and Select Forecasts, 1995:01–2005:01 Out-of-Sample Period



especially at horizons of 6, 12, and 24 months, is that it is much more volatile than the actual realizations. Despite the fact that the t-lambda method has relatively high MZ R^2 statistics at horizons of 6, 12, and 24 months, the overly volatile nature of the t-lambda forecasts causes its MSFE to be substantially greater than that of the AR benchmark, and the estimated slope coefficient in the MZ regression is consistently significantly less than 1. At horizons of 12 and 24 months, Figure 1 shows that the shrinkage method tends to do the best job of tracking the decline in Missouri employment growth associated with the 2001 recession. Given that turning points are notoriously difficult to predict, this suggests that the shrinkage method forecasts are quite useful at longer horizons. We also see from Figure 1 that the mean generally does a better job than the AR benchmark model at tracking Missouri employment growth at all horizons. The mean forecasts are less volatile than the shrinkage and t-lambda forecasts, so they provide more reliable forecasts at shorter horizons.

Table 3

Shrinkage, κ = 1.0 Combining Weights for the Individual ARDL Model Forecasts of Missouri Employment Growth for Select Months, 1995:01–2005:01 Out-of-Sample Period, *h* = 12

Variable	1996:01	1997:01	1998:01	1999:01	2000:01	2001:01	2002:01	2003:01	2004:01
Regional variables									
Missouri unemployment rate	0.01	0.20	0.21	0.38	0.23	–0.18	–0.03	–0.02	–0.01
Arkansas employment	0.17	0.19	0.18	0.16	0.17	0.21	0.30	0.28	0.26
Illinois employment	0.02	–0.09	–0.10	–0.14	–0.05	0.02	0.06	0.01	–0.04
Iowa employment	0.28	0.25	0.28	0.25	0.26	0.31	0.28	0.26	0.25
Kansas employment	0.39	0.31	0.04	–0.16	–0.34	–0.17	–0.19	–0.06	0.00
Kentucky employment	0.00	0.27	0.23	0.11	0.23	0.37	**0.59**	**0.88**	**0.84**
Nebraska employment	**0.51**	**0.58**	**0.66**	**0.67**	**0.60**	0.45	0.43	0.40	0.35
Tennessee employment	0.01	–0.02	–0.08	–0.03	0.18	0.17	0.21	0.34	0.42
Oklahoma employment	0.17	0.20	0.21	0.38	0.23	–0.18	–0.03	–0.02	–0.01
National variables									
U.S. employment	–0.55	–0.68	–0.64	–0.68	–0.66	–0.58	–1.05	–1.27	–1.25
U.S. unemployment rate	–0.12	–0.29	–0.22	–0.23	–0.12	–0.12	–0.20	–0.26	–0.19
Capacity utilization	–0.22	–0.35	–0.45	–0.68	–0.60	–0.30	–0.47	–0.44	–0.35
Average weekly hours, manufacturing	–0.22	–0.14	0.06	0.08	0.12	0.02	0.02	0.00	–0.05
Unemployment claims	0.23	0.16	0.23	0.19	0.25	0.26	0.45	0.43	0.35
New manufacturing orders	0.02	–0.08	–0.09	–0.12	–0.11	–0.19	–0.22	–0.23	–0.26
Vendor sales	0.36	0.56	0.58	0.55	0.57	**0.58**	0.52	0.53	0.46
New manufacturing capital orders	0.03	0.06	0.02	0.05	–0.03	–0.07	–0.14	–0.22	–0.23
Building permits	0.13	0.07	0.06	0.10	0.05	0.01	0.01	–0.07	–0.12
Stock market index	–0.51	–0.78	–0.78	–0.62	–0.51	–0.42	–0.32	–0.28	–0.20
Interest rate spread	0.06	0.18	0.17	0.32	0.34	0.24	0.40	0.46	0.48
National overtime	0.12	0.21	0.12	0.09	–0.14	–0.15	–0.15	–0.17	–0.13
Industrial production	0.27	0.30	0.34	0.33	0.31	0.41	0.27	0.22	0.20

Note: A bold entry signifies the ARDL model that receives the highest weight at the given date; 0.00 indicates <0.005.

However, the mean forecasts are too smooth at longer horizons and thus are not as adept as the shrinkage forecasts in tracing swings in Missouri employment growth.

Given the relatively good forecasting performance of the shrinkage, $\kappa = 1.0$ method at longer horizons, it is interesting to examine the weights used to compute the combination forecasts for this method. Table 3 reports the weights associated with the individual forecasts using the shrinkage, $\kappa = 1.0$ method for the first month of most of the years of the forecast evaluation period. Each weight would be 0.045 under equal weighting. However, the weights on the individual forecasts often differ markedly from equal weighting in Table 3, and we also witness sizable changes in some of the weights from year to year. Several of the regional variables, such as Nebraska, Iowa, and Kentucky employment, are heavily positively weighted throughout most of the period, indicat-

Table 4

Forecast Combining Results for Missouri and Illinois Employment Growth, Alternative Out-of-Sample Periods, *h* = 12

	Missouri		Illinois	
Combination method	1995:01–1999:12	2000:01–2005:01	1995:01–2005:01	2000:01–2005:01
AR benchmark	0.47	1.41	1.72	3.22
Mean	0.88	0.83	0.81	0.81
Median	0.94	0.82	0.82	0.83
Trimmed mean	0.99	0.79	0.84	0.84
OLS, recursive	2.25	0.61	1.34	1.00
OLS, rolling	7.28	1.04	1.28	1.01
WLS: t-lambda, $\lambda = 1$	4.25	0.75	1.39	0.99
WLS: t-lambda, $\lambda = 3$	7.10	0.90	1.40	1.08
Discount MSFE, $\delta = 1.0$	0.89	0.84	0.82	0.80
Discount MSFE, $\delta = 0.9$	0.93	0.84	0.81	0.82
Most recently best	1.67	1.00	**0.59**	**0.53**
Shrinkage, $\kappa = 0.5$	1.25	**0.42**	0.90	0.64
Shrinkage, $\kappa = 1.0$	**0.88**	0.44	0.82	0.64
Cluster: $C(2, PB)$	0.97	0.90	0.89	0.78
Cluster: $C(3, PB)$	0.95	0.90	0.88	0.74
Model selection: M-TST	6.85	1.29	7.36	1.02
PC, $m = 1$	1.31	0.91	0.83	0.83
PC, $m = 2$	1.26	1.07	0.98	0.83
ABMA, SIC	1.04	1.00	1.41	1.39
ABMA, AIC	0.96	0.94	0.72	0.69
AFTER	0.98	0.84	1.06	1.03

NOTE: The first row reports the MSFE for the AR benchmark model; the remaining rows report the ratio of the MSFE for the combining method to the MSFE for the AR benchmark model. A bold entry signifies the combining method with the lowest MSFE.

ing the important role for some of the regional variables in forecasting Missouri employment growth.[11] Tennessee and Oklahoma employment are interesting in that the weights display significant increases and decreases, respectively, over the evaluation period. In terms of the national variables, vendor sales, industrial production, and the interest rate spread possess large positive weights throughout most of the period; other national variables such as U.S. employment, capacity utilization, and the stock market index have large negative weights throughout most of the period. Overall, the shrinkage method at the 1-year horizon appears able to identify the individual forecasts that are the most accurate and to accommodate changes in relative forecasting accuracy in computing the combining weights. It would be difficult a priori to identify the particular individual model or models with the best forecasting ability at different points in time, and the shrinkage method provides an a priori procedure to cull potentially useful information from a large number of individual models.

[11] For instance, if we were to exclude the eight neighboring state employment variables, the ratio of the MSFE for the shrinkage, $\kappa = 1.0$ procedure to the MSFE for the AR benchmark increases to 1.01 at the 12-month horizon, well above the 0.71 figure in Table 2.

3.3 Robustness Checks

Table 4 reports results for alternative forecast evaluation periods for Missouri employment growth and employment growth for a neighboring state, Illinois, at the 1-year horizon.[12] The results for Missouri show that the MSFE for the AR forecast procedure is more than three times lower for the 1995:01–1999:12 period than the 2000:01–2005:01 period. This reflects fairly sharp employment changes and potential data problems (for example, occasional contradictory reports between BLS monthly and quarterly reports on both the national and state levels) during the latter period.[13] Several of the combining methods in the earlier period produce very poor results, with MSFE ratios substantially above 1, while the shrinkage, $\kappa = 1.0$ method and mean combination forecast produce the lowest MSFE statistics. Both shrinkage forecasts produce the lowest MSFE during the latter period for Missouri. Overall, across our three evaluation periods (1995:01–2005:01, 1995:01–1999:01, 2000:01–2005:1), the shrinkage, $\kappa = 1.0$ method yields the lowest MSFE at the 1-year horizon for Missouri employment growth.

With respect to employment growth in Illinois, we consider the 1995:01–2005:01 and 2000:01–2005:01 forecast evaluation periods. The individual ARDL forecasts are based on the same set of national variables used for Missouri, as well as the Illinois unemployment rate and employment in the six states that border Illinois (Iowa, Indiana, Kentucky, Missouri, Michigan, and Wisconsin). The same time-series specifications given in Section 3.1 above are used for these variables. The most recently best method achieves the lowest MSFE for both evaluation periods, producing declines of 51 percent and 47 percent in MSFE relative to the AR benchmark. The simple combination forecasts outperform the AR benchmark over both evaluation periods, with declines in MSFE relative to the AR benchmark of 16 to 19 percent, and the shrinkage and ABMA, AIC methods also lead to sizable declines in MSFE relative to the AR benchmark, ranging from 10 to 36 percent. The results in the last two columns of Table 4 indicate that simple combining and shrinkage methods perform well with respect to forecasting employment growth at the 1-year horizon in both Missouri and Illinois.

4 CONCLUSION

There are a large number of potentially relevant predictors of regional economic variables such as Missouri employment growth. In this paper, we analyze 20 different methods from the extant literature for combining individual forecasts of Missouri employment growth from 1995:01 to 2005:01. The individual forecasts are generated by a large number of ARDL models based on potential predictors. Similar to Stock and Watson (2004), we find that simple combination methods work well, particularly at shorter forecast horizons of 3 and 6 months, and often outperform more complicated weighting procedures. At longer horizons of 12 and 24 months, we find that Bayesian shrinkage methods produce the most accurate forecasts, providing quite sizable reduction in MSFE relative to an AR benchmark model. The shrinkage combining methods also perform well at longer horizons over alternative evaluation periods and were able to track Missouri employment growth over the recent recession reasonably well. Examination of the combining weights used in the shrinkage methods indicates that a number of regional variables and a few national variables can play a significant role in improving forecasts of Missouri employment growth. Forecast combination procedures also lead to relatively large reductions in MSFE relative to an AR benchmark model when forecasting Illinois employment growth. Why do shrinkage combining methods work well in the present paper? Diebold and Lopez (1996, p. 256) offer helpful insight by observing that "the combining weights [of shrinkage combining methods] are coaxed toward the arithmetic mean, but the data are still allowed to speak, when (and if) the data have something to say." Shrinkage combining methods can thus take

[12] A 1-year horizon is typically important for state budgetary planning purposes.

[13] See Kliesen and Wall (2004) on reconciling the BLS jobless employment figures and Wall and Wheeler (2005) on large discrepancies in recent St. Louis employment.

advantage of the reliable performance of simple averaging across a diversity of variables while still allowing for particular variables to exert a stronger influence in certain situations. In future research, we plan to extend our analysis by developing combination forecasts of employment growth for each of the 50 individual U.S. states.

REFERENCES

Aiolfi, Marco and Timmermann, Allan. "Persistence in Forecasting Performance and Conditional Combination Strategies." *Journal of Econometrics*, 2005 (forthcoming).

Bates, J.M. and Granger C.W.J. "The Combination of Forecasts." *Operational Research Quarterly*, 1969, *20*, pp. 451-68.

Burnham, Kenneth P. and Anderson, David R. *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag, 1998.

Chan, Yeung Lewis; Stock, James H. and Watson, Mark W. "A Dynamic Factor Model Framework for Forecast Combination." *Spanish Economic Review*, July 1999, *1*(2), pp. 21-121.

Clemen, Robert T. and Winkler, Robert L. "Combining Economic Forecasts." *Journal of Economic and Business Statistics*, *4*(1), 1986, pp. 39-46.

Diebold, Francis X. and Lopez, Jose. "Forecast Evaluation and Combination," in G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics 14: Statistical Methods in Finance*. Amsterdam: Elsevier, 1996, pp. 241-68.

Diebold, Francis X. and Pauly, Peter. "Structural Change and the Combination of Forecasts." *Journal of Forecasting*, 1987, *6*, pp. 21-40.

Diebold, Francis X. and Pauly, Peter. "The Use of Prior Information in Forecast Combination." *International Journal of Forecasting*, December 1990, *6*(4), pp. 503-08.

Draper, David. "Assessment and Propagation of Model Uncertainty." *Journal of the Royal Statistical Society Series B*, 1995, *57*(1), pp. 45-97.

Garratt, Anthony; Lee, Kevin; Pesaran, Hashem M. and Shin, Yongcheol. "Forecast Uncertainties in Macroeconomic Modeling: An Application to the U.K. Economy." *Journal of the American Statistical Association*, December 2003, *98*(464), pp. 829-38.

Granger, Clive W.J. and Ramanathan, Ramu. "Improved Methods of Combining Forecasts." *Journal of Forecasting*, 1984, *3*, pp. 197-204.

Kliesen, Kevin L. and Wall, Howard J. "A Jobless Recovery with More People Working?" Federal Reserve Bank of St. Louis *Regional Economist*, April 2004, pp. 10-11.

Mincer, J. and Zarnowitz, Y. "The Evaluation of Economic Forecasts," in J. Mincer, ed., *Economic Forecasts and Expectations*. New York: National Bureau of Economic Research, 1969, pp. 3-46.

Newey, Whitney K. and West, Kenneth D. "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, May 1987, *55*(3), pp. 703-08.

Ng, Serena and Perron, Pierre. "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power." *Econometrica*, 2001, *69*(6), pp. 1519-54.

Stock, James H. and Watson, Mark W. "Forecasting Inflation." *Journal of Monetary Economics*, October 1999, *44*(2), pp. 293-335.

Stock, James H. and Watson, Mark W. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, September 2003, *41*(3), pp. 788-829.

Stock, James H. and Watson, Mark W. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting*, September 2004, *23*(6), pp. 405-30.

Swanson, Norman R. and Zeng, Tian. "Choosing Among Competing Econometric Forecasts:

Regression-Based Forecast Combination Using Model Selection." *Journal of Forecasting*, September, *20*(6), pp. 425-40.

Wall, Howard J. and Wheeler, Christopher H. "St. Louis Employment: A Tale of Two Surveys." Center for Regional Economics CRE8 Occasional Report No. 2005-01.

Yang, Yuhong. "Combining Forecasting Procedures: Some Theoretical Results." *Econometric Theory*, February 2004, *20*(1), pp. 176-222.