Using Large Data Sets to Forecast Sectoral Employment

Rangan Gupta
Department of Economics
University of Pretoria
Pretoria, 0002, SOUTH AFRICA

Alain Kabundi
Department of Economics and Econometrics
University of Johannesburg
Johannesburg, 2006, SOUTH AFRICA

Stephen M. Miller
Department of Economics,
University of Nevada, Las Vegas
Las Vegas, Nevada, 89154-6005 USA

Josine Uwilingiye
Department of Economics and Econometrics
University of Johannesburg
Johannesburg, 2006, SOUTH AFRICA

Abstract

We implement several Bayesian and classical models to forecast employment for eight sectors of the US economy. In addition to standard vector-autoregressive and Bayesian vector autoregressive models, we also include the information content of 143 additional monthly series in some models. Several approaches exist for incorporating information from a large number of series. We consider two multivariate approaches – extracting common factors (principle components) in a factor-augmented vector autoregressive or vector error-correction, Bayesian factor-augmented vector autoregressive or vector error-correction models, or Bayesian shrinkage in a large-scale Bayesian vector autoregressive models. Using the period of January 1972 to December 1989 as the in-sample period and January 1990 to March 2009 as the out-of-sample horizon, we compare the forecast performance of the alternative models. Finally, we forecast out-of sample from April 2009 through March 2010, using the best forecasting model for each employment series. We find that factor augmented models, especially error-correction versions, generally prove the best in out-of-sample forecast performance, implying that in addition to macroeconomic variables, incorporating long-run relationships along with short-run dynamics play an important role in forecasting employment.

Keywords: Sectoral Employment, Forecasting, Factor Augmented Models, Large-Scale BVAR models

JEL classification: C32, R31

1. Introduction

Unlike the standard post-WWII recession, analysts called the recoveries from recession in the early 1990s and 2000s "jobless" recoveries. Most analysts also predict a jobless recovery from the recent Great Recession. Pundits argue that the midterm election results of 2010 depended in great measure on the state of the national and local economies, the lack of employment growth, and the stubbornly high unemployment rate. Macroeconomists debate whether the Great Recession largely reflects insufficient aggregate demand or structural issues. As such, forecasting employment should receive more attention in the literature. Rapach and Strauss (2008) state "forecasting employment growth has received little attention ... relative to such macroeconomic stalwarts as inflation, GDP growth, and the unemployment rate." (p. 75).

The most recent Great Recession affected employment in sectors differently. More specifically, the largest percentage employment loss from peak to trough occurred in the construction sector with a loss of 27.7 percent or 2.14 million jobs and the smallest percentage loss occurred in the leisure and hospitality sector with a loss of 4.1 percent or 551 thousand jobs. Thus, forecasting total employment, which Rapach and Strauss (2008, 2010a) do, hides much more potentially important volatility between sub-sectors. When forecasting macroeconomic variables such as employment, researchers must decide which other macroeconomic variables may help improve forecast performance. One approach uses economy theory and the intuitive judgment of the researcher to select the other variables used in the forecasting exercise. A second approach, an agnostic view, collects a large set of variables that can potentially improver the forecasting performance and lets the data speak for themselves. We adopt this second approach and gather a large data set of 143 variables plus the eight sectoral employment series. This paper considers the dynamics of employment and the ability of different pure time-series models to forecast sectoral employment.[1] The main focus considers how the researcher can incorporate large data sets into forecasting equations, using dynamic factor analysis or shrinking large-scale BVAR models. We illustrate the process using employment from eight sectors -- mining and logging; construction; manufacturing; trade, transportation, and utilities; financial activities; professional and business services; leisure and hospitality; and other services.

More specifically, we compare the out-of-sample forecasting performance of various time-series models – vector autoregressive (VAR) vector error-correction (VEC), factor augmented VAR (FAVAR), factor augmented VEC (FAVEC), and various Bayesian time-series models. For the Bayesian models, we estimate Bayesian VAR (BVAR), Bayesian VEC (BVEC), Bayesian factor augmented VAR (BFAVAR), Bayesian factor augmented VEC (BFAVEC), and large-scale BVAR (LBVAR) models. A factor-augmented model generally performs the best across the eight employment series, using the average root-mean-squared-error (RMSE) criteria. The LBVAR models come in a close second to the factor-augmented models on several occasions, and actually outperform the factor-augmented models for an extremely small number of forecast horizons. Finally, the models that exclude the information from the large set of data generally come in a distant third in forecast performance and only prove the best forecasting models on a few occasions, implying that the macroeconomic fundamentals partly drive employment.

We organize the rest of the paper as follows. Section 2 provides a brief review of the literature on using large data sets in forecasting models. Section 3 discusses the literature on forecasting employment. Section 4 specifies the various time-series models estimated and used for forecasting. Section 5 discusses the data and the results. Section 6 concludes.

[1] Focusing on the employment numbers, however, obscures a large part of employment dynamics. That is, much job churning occurs in the labor markets. New businesses open and hire thousands of workers each month, while other businesses close and thousands of other workers find themselves without employment.

2. Forecasting with Large Data Sets

Zellner and Palm (1974) theoretically rationalize why time-series models generally perform as well as or better than dynamic simultaneous equation models (SEM) in structural form.[2] An important issue involves determining how additional information can or cannot improve the forecasting performance over a simple univariate autoregressive or autoregressive-moving-average representation.

A bivariate approach uses an autoregressive distributed lag (ARDL) model (Stock and Watson 1999, 2003, 2004), a transfer function model (Enders 2004, Ch. 5). That is, the researcher runs a transfer function model, where the variable to forecast enters as an autoregressive process and one driver variable enters as a distributed lag. The researcher compares the baseline model, the pure autoregressive specification forecasts with the forecasts for the transfer function or ARDL specification. Researchers extend this further and repeat the process for a whole series of potential driver variables. Now, one can aggregate across the individual forecasts to generate a combined forecast. Combination forecasts range from simple means or medians to more complicated principal-components- or mean-square-forecast-error-weighted forecasts.

Other multivariate methods adopt "atheoretical" VAR or VEC models to generate forecasts. These models do not impose exogeneity assumptions on the included predictor variables. Unlike the single-equation bivariate ARDL or transfer function model, the VAR or VEC approaches assume that lagged values of each variable may provide valuable information in

[2] Any dynamic SEM in structural form or dynamic structural model implicitly generates a series of univariate time-series models for each endogenous variable. The dynamic structural model, however, imposes restrictions on the parameters in the reduced-form time-series specification. Dynamic structural models prove most effective in performing policy analysis, albeit subject to the Lucas critique. Time-series models prove most effective at forecasting. That is, in both cases errors creep in whenever the researcher makes a decision about the specification. Clearly, more researcher decisions relate to a dynamic structural model than a univariate time-series model, suggesting that fewer errors enter the time-series model and allowing the model to produce better forecasts.

forecasting each endogenous variable. VAR and VEC models, however, come with their own issues such as over-parameterization, since the estimated number of parameters increases dramatically with additional variables or additional lags in the system. One solution to the over-parameterization problem extracts common factors from a large data set, which then get added to the VAR or VEC specifications (Bernanke, Boivin, and Eliazs 2005, Stock and Watson 2002a, 2005). Adding a few common factors from the large dataset to VAR and VEC systems economizes on the number of new parameters to estimate.

Bayesian VAR (BVAR) or VEC (BVEC) models overcome the over-parameterization problem by defining a small number of hyper-parameters in the specification that determines all parameters in the system. Linking all parameters of the system to these hyper-parameters, however, also introduces errors in the forecasting exercise. Since the Bayesian approach already addresses the over-parameterization problem through Bayesian shrinkage, researchers can estimate BVAR or BVEC systems that include a large number of additional explanatory variables, obviating the need to extract common factors. Nothing prevents, however, the extraction of common factors from the large set of macroeconomic variables to include in factor-augmented VAR (FAVAR) or VEC (FAVEC) systems, which we also do.

The bivariate ARDL method, in contrast to the VAR, BVAR, VEC, or BVEC modeling approaches, uses information from a large dataset one variable at a time and then aggregates across all forecasts. As a result, this approach does not differentiate between common factors and non-common factors in the large dataset. Each exhibits the same effect on the forecast, over and above the autoregressive part of the model. In the factor-augmented approach, the researcher potentially leaves information on the table by only extracting the common factor information and leaving the remaining information out of the analysis. On the other hand, the Bayesian approach, includes all the information from the large set of data, but restricts the estimation by imposing conditions on the parameters of the estimating equation through the hyper-parameters. In sum, all methods introduce restrictions on the way information from the large dataset affects the estimation process. Thus, any of the individual approaches may lead to better forecasts *a priori*.

In this paper, we consider the multivariate factor-augmented and large-scale Bayesian methods for incorporating the information from a large dataset. These methods provide the natural extension of the VAR, VEC, BVAR, and BVEC models. The bivariate ARDL model involves a single-equation, whereas the VAR, VEC, BVAR, and BVEC models involve multiple equations. Thus, we exclude the ARDL approach from our analysis.

3. Forecasting Employment

As noted in the introduction, little work exists on forecasting national employment trends. Much forecasting of employment does exist, however, at the regional level. Regional economists use employment, since other macroeconomic indicators such as GDP or industrial production either do not exist at the regional level, do not provide sufficient disaggregation, or appear too infrequently. As a result, regional economists use employment trends by sector to help understand the growth of the regional economy.

Regional economists developed the ideas of economic base and shift-share analysis to track and predict regional growth, using employment data. The popularity of these analyses comes from the simplicity of execution and the easily accessible data to execute the analysis. Lane (1966) and Williamson (1975) provide some history and background on economic base analysis; whereas Stevens and Moore (1980) provide a critical review of shift-share analysis as a forecasting tool. Since these analyses do not consider structural issues, but instead rely on simple constructs from the employment data itself, we can consider the approaches as a rudimentary time-series forecasting technique.

In another related line of research, regional economists consider the relative advantages and disadvantages of forecasting regional economic activity, including employment, using time-series and structural models. Early efforts compare the forecasting performance of structural and autoregressive integrated moving average (ARIMA) models (Taylor 1982, Glennon, Lane and Johnson 1987).

More recently, a few economists consider the performance of different models in forecasting employment at the national level. For example, Stock and Watson (2002b) forecast eight monthly macroeconomic time-series variables, including nonagricultural employment, from 1970 through 1998. They use a larger data set of 215 additional potential predictors, extracting principle components using dynamic factor modeling, to see if forecasting accuracy improves over simpler time-series models. They conclude that these new forecasts outperform univariate ARs, small VARs, and leading indicator models.

Rapach and Strauss (2008) forecast employment growth, using monthly seasonally adjusted civilian employment from the Conference Board data set and an autoregressive distributed lag (ARDL) model framework, containing 30 determinants, to forecast national employment growth. Given the difficulty in determining *a priori* the particular variables that prove the most important in forecasting employment growth, the authors also use various methods to combine the individual ARDL model forecasts, which result in better forecasts of employment growth. The combining method based on principle components does the best, while those methods that rely on simple averaging, clusters, and discounted mean square forecast error also produce forecasts better than the individual ARDL without combining. In an earlier paper, Rapach and Strauss (2005) obtain similar results when forecasting the employment growth in

Missouri, using an ARDL approach based on 22 regional and national predictors. They observe that combining methods based on Bayesian shrinkage techniques produce substantially more accurate out-of-sample forecasts than those from a benchmark AR model.

Rapach and Strauss (2010a) forecast national employment growth, using the same data set in Rapach and Strauss (2008), by applying bootstrap aggregating (bagging) to a general-to-specific procedure based on a general dynamic linear regression model. When they compared bagging to the forecast combination approaches, the authors find bagging forecasts often deliver the lowest forecast errors. Further, the authors note that incorporating information from both bagging and combination forecasts (based on principal components) often leads to further gains in forecast accuracy.

More recently, Rapach and Strauss (2010b) forecast state employment growth using several distinct econometric approaches, such as combinations of individual ARDL models, general-to-specific modeling coupled with bagging, and factor models. As in their earlier studies, the results show that these forecasting approaches consistently deliver sizable reductions in forecast errors relative to the benchmark AR model across states. Further, they observe forecasting improvements on amalgamating these approaches, especially during national business-cycle recessions.

Banbura *et al.*, (2010) show that a VAR model with Bayesian shrinkage, incorporating a large number of explanatory variables, often produces better forecasts for non-farm employment than those from small-scale VAR and FAVAR models.

Against this backdrop, our paper extends the above mentioned studies, in the sense that we use a variety of large-scale models that facilitate the role of a wider possible set of fundamentals to affect the dynamic movement of employment. Note that the motivation to use a large data set (143 explanatory variables), rather than 20 to 30 variables used as predictors in the ARDL model, received support since the models based on the large data set tend to outperform medium-scale models that used 20 variables in forecasting employment.[3]

4. VAR, VEC, BVAR, BVEC, FAVAR, FAVEC, BFAVAR, BFAVEC, and LBVAR Specifications and Estimation[4]

4.1 VAR, VEC, BVAR, BVEC, and LBVAR:

Following Sims (1980), we can write an unrestricted VAR model as follows:

$$Y_t = A_0 + A(L)Y_t + \varepsilon_t \, , \qquad (1)$$

where Y equals a $(n \times 1)$ vector of variables to forecast; A_0 equals an $(n \times 1)$ vector of constant terms; $A(L)$ equals an $(n \times n)$ polynomial matrix in the backshift operator L with lag length *p*,[5] and ε equals an $(n \times 1)$ vector of error terms. In our case, we assume that $\varepsilon \sim N(0, \sigma^2 I_n)$, where I_n equals an $(n \times n)$ identity matrix.

The VAR method typically use equal lag lengths for all variables, which implies that the researcher must estimate many parameters, including many that prove statistically insignificant. This over-parameterization problem can create multicollinearity and a loss of degrees of freedom, leading to inefficient estimates, and possibly large out-of-sample forecasting errors. Some researchers exclude lags with statistically insignificant coefficients. Alternatively, researchers use near VAR models, which specify unequal lag lengths for the variables and equations.

Imposing additional restrictions on a standard VAR model generates a VEC model that

---

[3] See footnote 15 for further details.

[4] The discussion in this section relies heavily on LeSage (1999), Gupta and Miller (forthcoming a, forthcoming b), and Das *et al*., (2009).

[5] That is, $A(L) = A_1 L + A_2 L^2 + ... + A_p L^p$;

uses cointegrated non-stationary series. While including short-run dynamic adjustment, the VEC model also incorporates the cointegration relationship so that it restricts the movement of endogenous variables to converge to their long-run relationships. The cointegration term, called the error correction term, gradually corrects through a series of partial short-run adjustments.

More explicitly, assume that Y_t includes n time-series variables integrated of order one, (i.e., *I(1)*).[6] The error-correction counterpart of the VAR model in equation (1) converts into a VEC model as follows:[7]

$$\Delta Y_t = \pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-1} + \varepsilon_t \qquad (2)$$

where $\pi = -[I - \sum_{i=1}^{p} A_i]$ and $\Gamma_i = -\sum_{j=i+1}^{p} A_j$.

Litterman (1981), Doan *et al.*, (1984), Todd (1984), Litterman (1986), and Spencer (1993) use the BVAR model to overcome the over-parameterization problem. Rather than eliminating lags, the Bayesian method imposes restrictions on the coefficients across different lag lengths, assuming that the coefficients of longer lags may more closely approach zero than the coefficients on shorter lags. If, however, stronger effects come from longer lags, the data can override this initial restriction. Researchers impose the constraints by specifying normal prior distributions with zero means and small standard deviations for most coefficients, where the standard deviation decreases as the lag length increases and implies that the zero-mean prior holds with more certainty. The first own-lag coefficient in each equation proves the exception with a unitary mean. Finally, Litterman (1981) imposes a diffuse prior for the constant. We employ this "Minnesota prior" in our analysis, where we implement Bayesian variants of the

---

[6] See Lesage (1999) and references cited therein for further details regarding the non-stationary of most macroeconomic time series.

[7] See, Dickey *et al.* (1991) and Johansen (1995) for further technical details.

classical VAR models.

Formally, the means of the Minnesota prior take the following form:

$$\beta_i \sim N(1, \sigma^2_{\beta_i}) \text{ and } \beta_j \sim N(0, \sigma^2_{\beta_j}) \tag{3}$$

where β_i equals the coefficients associated with the lagged dependent variables in each equation of the VAR model (i.e., the first own-lag coefficient), while β_j equals any other coefficient. In sum, the prior specification reduces to a random-walk with drift model for each variable, if we set all variances to zero. The prior variances, $\sigma^2_{\beta_i}$ and $\sigma^2_{\beta_j}$, specify uncertainty about the prior means, $\overline{\beta}_i$ = 1, and $\overline{\beta}_j$ = 0.

Doan *et al.,* (1984) propose a formula to generate standard deviations that depend on a small numbers of hyper-parameters: *w*, *d*, and a weighting matrix *f*(*i*, *j*) to reduce the over-parameterization in the VAR models. This approach specifies individual prior variances for a large number of coefficients, using only a few hyper-parameters. The specification of the standard deviation of the distribution of the prior imposed on variable *j* in equation *i* at lag *m*, for all *i*, *j* and *m*, equals $S_1(i, j, m)$, defined as follows:

$$S_1(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j}, \tag{4}$$

where $f(i, j) = 1$, if $i = j$ and k_{ij} otherwise, with ($0 \leq k_{ij} \leq 1$), and $g(m) = m^{-d}$, with $d > 0$. The estimated standard error of the univariate autoregression for variable *i* equals $\hat{\sigma}_i$. The ratio $\hat{\sigma}_i / \hat{\sigma}_j$ scales the variables to account for differences in the units of measurement and, hence, causes the specification of the prior without consideration of the magnitudes of the variables. The term *w* indicates the overall tightness, with the prior getting tighter as the value falls. The parameter *g(m)* measures the tightness on lag *m* with respect to lag 1, and equals a harmonic shape with decay factor *d*, which tightens the prior at longer lags. The parameter *f*(*i*, *j*) equals the tightness of variable *j* in equation *i* relative to variable *i*, and by increasing the interaction (i.e., the value of k_{ij}), we loosen the prior.[8] Note that, following LeSage (1990), in the Bayesian versions of the VEC models, we impose no priors on the error-correction terms.

We also follow Banbura, Giannone, and Reichlin (2010) and set the value of the overall tightness parameter as an alternative to obtain a desired average fit for the eight employment variables of interest in the in-sample period (1972:1 to 1989:12). For example, the optimal value of *w(Fit)* (= 0.0230), with *d* =2.0, for the LBVAR model with 151 variables obtained in this fashion is then retained for the entire evaluation period. The values of *w* for the eight variable BVAR and one variable Bayesian Autoregressive (BAR) models, given *d*=2, equal 0.2366 and 1.8250, respectively. While, the corresponding value of *w* for the BVEC with eight employment series equals 0.2419. Specifically, for a desired *Fit* of 0.50, we choose *w* as follows:

$$w(Fit) = \arg\min_{w} \left| Fit - \frac{1}{8}\sum_{i=1}^{8} \frac{MSE_i^w}{MSE_i^0} \right|, \tag{5}$$

where $MSE_i^w = \sqrt{\frac{1}{T_0 - p - 1}\sum_{t=p}^{T_0-2} y_{i,t+1|t}^w - y_{i,t+1})^2}$. That is, we evaluate the one-step-ahead mean squared error (*MSE*) using the training sample $t = 1,..... T_0 -1$, where T_0 is the beginning of the sample period and *p* is the order of the VAR. The value MSE_i^0 is the *MSE* of variable *i* with the prior restriction imposed exactly (*w*=0).[9]

---

[8] For an illustration, see Dua and Ray (1995).

[9] In addition to using a *Fit* of 0.50, we also experiment with a *Fit* as the average relative MSE from an OLS-estimated VAR containing the eight sectoral employment variables, i.e., $Fit = \frac{1}{8}\sum_{i=1}^{8}\frac{MSE_i^\infty}{MSE_i^0}$, as well as a *Fit* value of 0.25. In both cases, the forecasting performances of the alternative Bayesian models deteriorates. These results are available upon request from the authors.

We estimate the alternative BVARs using Theil's (1971) mixed estimation technique. Essentially then, the method involves supplementing the data with prior information on the distribution of the coefficients. The number of observations and degrees of freedom increase artificially by one for each restriction imposed on the parameter estimates. Thus, the loss of degrees of freedom from over-parameterization in the classical VAR models does not emerge as a concern in the alternative BVAR specifications.

4.2 FAVAR and BFAVAR:

We use the dynamic factor (DF) model to extract common components between stationary macroeconomic series and then use these common components to forecast employment, adding three extracted factors to the 8-variable VAR model to create a factor-augmented VAR (FAVAR) model in the process.[10] We choose the three factors by the cumulative variance share, under which, the fourth eigenvalue fell below the threshold of 5 percent. Furthermore, we estimate idiosyncratic component (see below) with AR(*p*) processes as suggested by Boivin and Ng (2005).

The DF model expresses individual times series as the sum of two unobserved components: a common component driven by a small number of common factors and an idiosyncratic component for each variable. The DF model extracts the few factors that explain the co-movement of the US economy. Forni *et al.* (2005) demonstrate that for a small number of factors relative to the number of variables and a heterogeneous panel, we can recover the factors from present and past observations.

Suppose that Z_t equals a $n \times 1$ covariance stationary vector standardized to possess a mean zero and a variance equal to one, obtained from the original $n \times 1$ vector of *I(1)* variables

[10] That is, we first transform all data to induce stationarity. Then, using the transformed data, we extract the common components.

Y_t. Under DF models, we write Z_t as the sum of two orthogonal components as follows:

$$Z_t = \lambda f_t + \xi_t \tag{6}$$

where f_t equals a $r \times 1$ vector of static factors, λ equals an $n \times r$ matrix of factor loadings, and ξ_t equals a $n \times 1$ vector of idiosyncratic components. In a DF model, f_t and ξ_t are mutually orthogonal stationary processes, while, $\chi_t = \lambda f_t$ equals the common component.

Since dynamic common factors are latent, we must estimate them. We note that the estimation technique used matters for factor forecasts. This paper adopts the Stock and Watson (2002b) method, which employs the static principal component (PC) approach on Z_t. The factor estimates, therefore, equal the first principal components of Z_t, (i.e., $\hat{f_t} = \hat{\Lambda}' Z_t$, where $\hat{\Lambda}$ equals the $n \times r$ matrix of the eigenvectors corresponding to the r largest eigenvalues of the sample covariance matrix $\hat{\Sigma}$).

For forecasting purposes, we use an 8-variable VAR augmented by extracted common factors using the Stock and Watson (2002a) approach. This approach is similar to the univariate Static and Unrestricted (SU) approach of Bovin and Ng (2005). Therefore, the forecasting equation to predict Y_t is given by

$$\begin{bmatrix} \hat{Y}_{t+h} \\ \hat{f}_{t+h} \end{bmatrix} = \hat{\Phi}(L) \begin{bmatrix} Y_t \\ f_t \end{bmatrix} \tag{7}$$

where h equals the forecasting horizon, $\hat{\Phi}(L)$ equal lag polynomials, which we estimate with and without restrictions. As Boivin and Ng (2005) clearly note, VAR models are special cases of equation (8). With known factors and the parameters, the FAVAR approach should produce smaller mean squared errors. In practice, however, one does not observe the factors and we must estimate them. Moreover, the forecasting equation should reflect a correct specification. We consider the following DF model specifications:

- FA(V)AR: includes the employment in 1 (8) sector(s) and the three common static factors; and
- BFA(V)AR: the FA(V)AR specification with Bayesian restrictions on lags of the employment in 1 (8) sector(s) and the three factors, based on the priors outlined above. The values of *w* obtained for the BFAAR and BFAVAR, given *d*=2, were 0.4673 and 0.1699, respectively.

4.3 FAVEC and BFAVEC:

For the FAVEC models, we follow the procedure proposed by Banerjee and Marcellino (2009) and Banerjee, Marcellino, and Masten (2010).[11] We begin with a common trend representation, given below, for a set of *n* $I(1)$ variables (Y_t), whose standardized version is defined as X_t:[12]

$$\Delta X_t = \lambda \Delta F_t + \Delta \xi_t . \tag{8}$$

Bai and Ng (2004) and Bai (2004) allow for the possibility that ξ_t or some elements of ξ_t are $I(1)$.

We can rewrite equation (8) as follows:

$$\Delta X_t = \alpha \beta' \Delta F_t + \varepsilon_t \tag{9}$$

where $\beta' = \Lambda'_{\perp}$ and hence $\beta' X_t$ is $I(0)$ and an over-time correlation can exist between the errors $\Delta \xi_t$ and ε_t.

[11] See these papers for more details on the model and the estimation.

[12] When we extract the common factors for the FAVAR and BFAVAR models, we transform all variables to induce stationarity. Now, we transform all variables to induce non-stationarity. That is, for stationary variables, we accumulate to make them I(1). We also extract three common factors from the non-stationary variables, excluding the stationary variables. The findings prove similar to the three factors extracted when we accumulate the I(0) variables to make them I(1).

The literature on cointegration focuses mainly on equation (9), also known as the VEC model, while Banerjee and Marcellino (2009) reconcile the factor analysis in equation (8) and the cointegration concept in equation (9). The new hybrid model addresses the problem associated with large number of data sets that the simple VEC model (equation 2) does not consider. Hence, if important information does not enter the VEC model, then the model results in biased coefficients caused by omitted variables. In this case, the FAVEC model improves on the standard VEC model. Banerjee, Marcellino, and Masten (2010) demonstrate that the information set in the FAVEC model improves the forecasting performance of models, especially at the longer horizon.

By including the error-correction terms in the DF model, the FAVEC model enhances the former model, especially in the presence of cointegration. Thus, the factors extracted from a large panel of economic variables in levels jointly associate with the limited set of economic variables of main interest while allowing for cointegration. The FAVEC model naturally generalizes the FAVAR model developed by Bernanke, Boivin, and Eliaz (2005) and Stock and Watson (2005).

Assume that we only want to forecast a few variables in the entire economy. We, therefore, divide our panel into two parts, N^A including the variables of interest, X_t^A and $N^B = n - N^A$ containing the remaining variables, X_t^B. Equation (7) becomes:

$$\begin{pmatrix} X_t^A \\ X_t^B \end{pmatrix} = \begin{pmatrix} \Lambda^A \\ \Lambda^B \end{pmatrix} F_t + \begin{pmatrix} \xi_t^A \\ \xi_t^B \end{pmatrix} \qquad (10)$$

where Λ^A is $N^A \times r$ matrix and Λ^B is $N^B \times r$. The dimension of Λ^A does not change as n increases while the dimension of Λ^B increases with n. The theory requires that the rank of Λ^B, $r^B = r$, whereas the rank of Λ^A, $r^A \leq r$. That is, a smaller number of trends drives X_t^A. From equation (11), we see that X_t^A and F_t are cointegrated, while F_t are uncorrelated random walks.

From the Granger representation theorem, there exists an error correction specification as follows:

$$\begin{pmatrix} \Delta X_t^A \\ \Delta F_t \end{pmatrix} = \begin{pmatrix} \gamma^A \\ \gamma^B \end{pmatrix} \delta' \begin{pmatrix} X_{t-1}^A \\ F_{t-1} \end{pmatrix} + \begin{pmatrix} \nu_t^A \\ \nu_t \end{pmatrix} \tag{11}$$

We can extend equation (11) by adding additional lags to account for correlation in the errors as follows:

$$\begin{pmatrix} \Delta X_{t\ At}^A \\ \Delta F_t \end{pmatrix} = \begin{pmatrix} \gamma^A \\ \gamma^B \end{pmatrix} \delta' \begin{pmatrix} X_{t-1}^A \\ F_{t-1} \end{pmatrix} + A_1 \begin{pmatrix} \Delta X_{t-1}^A \\ \Delta F_{t-1} \end{pmatrix} + \cdots + A_q \begin{pmatrix} \Delta X_{t-q}^A \\ \Delta F_{t-q} \end{pmatrix} + \begin{pmatrix} u_t^A \\ u_t \end{pmatrix} \tag{12}$$

where the errors $\left(u_t^{A}{}', u_t'\right)'$ are *i.i.d.* Equation (12) is known as a FAVEC model.

Banerjee and Marcellino (2009) show that there must be N^A cointegrating relationships in equation (12), given that equation (12) includes $N^A + r$ dependent variables and that X_t^A is driven by F_t or a subset of F_t, and that elements of F_t are uncorrelated random walks.

Since Λ^A is $N^A \times r$, but can have a reduced rank of r^A, $N^A - r^A$ cointegrating relationships exist, including X_t^A variables only. Banerjee and Marcellino (2009) demonstrate that this emerges from a standard VEC model. The remaining r^A cointegrating relationships involve X_t^A and F_t. Therefore, potentially $n - N^A$ omitted cointegrating relationships exist in the standard VEC model.

Similarly, equation (12) improves on DF model and FAVAR models, given that the error-correction terms do not appear. That is, the FAVAR does not account for the long-run information and, hence, $\gamma^A = \gamma^B = 0$. Like the DF model, the FAVAR model does not account for cointegration and, therefore, it is misspecified in the presence of long-run relationships. It follows that the FAVEC model nests the VEC, FAVAR, and VAR models and, hence, it should outperform these models in forecasting.

- FAVEC: includes the employment in eight sectors, the three common static factors, and the error-correction terms; and
- BFAVEC: the FAVEC specification with Bayesian restrictions on lags of the FAVEC model based on the priors outlined above. The values of *w* for the BFAVEC model, given *d*=2, equals 01782.

Note that even though the Bai (2004) approach suggests four static factors, we decide to use 3 factors, just as in the case of the FAVAR models, since using the cumulative variance share of common component, we find that that the fourth factor explains only 3 percent of the variation, which is less than our pre-specified cut-off limit of 5 percent.

4.4 Comparing Forecasts:

For each of one- to twelve-months-ahead forecasts, we test whether the gain (loss) in the RMSE from the alternative "optimal" models relative to the random-walk model is significant. The optimal models minimize the average RMSE across all twelve forecast horizons. We use the *ENC-T* test of Clark and McCracken (2001). This test applies to nested models, given that the "optimal" models nest the random-walk model.

The test statistic is defined as follows:

$$ENC-T=(P-1)^{1/2}\frac{\bar{c}}{(P^{-1}\sum_{t=R}^{T-1}(c_{t+h}-\bar{c}))^{1/2}}, \qquad (13)$$

where, $c_{t+h}=\hat{v}_{0,t+h}(\hat{v}_{0,t+h}-\hat{v}_{1,t+h})$ and $\bar{c}=\sum_{t=R}^{T-1}c_{t+1}$, R denotes the estimation period, P is the prediction period, f is some generic loss function ($f(v_{0,t+h})=v_{0,t+h}^2$, in our case), $h\geq 1$ is the forecast horizon, $\hat{\upsilon}_{0,t+h}$ and $\hat{\upsilon}_{1,t+h}$ are h-step ahead prediction errors for models 0 and 1 (where model 0 is the "optimal" model), constructed using Newey and West (1987) type consistent estimators.

The hypotheses of interest are:

$$H_0 : E\left(f\left(\upsilon_{0,t+h}\right) - f\left(\upsilon_{1,t+h}\right)\right) = 0 \text{, and} \tag{14}$$

$$H_A : E\left(f\left(\upsilon_{0,t+h}\right) - f\left(\upsilon_{1,t+h}\right)\right) > 0 . \tag{15}$$

The limiting distribution is $N(0, 1)$ for $h = 1$. The limiting distribution for $h > 1$ is non-standard, as discussed in Clark and McCraken (2001). As long as a Newey and West (1987) type estimator is used when $h > 1$, however, then the tabulated critical values closely approximate the $N(0, 1)$ values (Bhardwaj and Swanson, 2006).

5. Data Description, Model Estimation, and Results

5.1 Data

While the small-scale VARs, both the classical and Bayesian variants, only include employment data for the eight sectors, the large-scale BVARs and the DF model also include the 143 monthly national and regional series. Seasonally adjusted employment data come from the Bureau of Labor Statistics. For the remaining 143 seasonally adjusted national and regional variables, we collected the data from various sources such as the Conference Board, the Global Insight database, the FREDII database of the St. Louis Federal Reserve Bank, the US Census Bureau, and the National Association of Realtors.

We transformed all data to induce stationarity for the FAVAR-type models before extracting the three factors. We can use non-stationary data, however, with the BVAR. Sims *et al.* (1990) indicate that with the Bayesian approach entirely based on the likelihood function, the associated inferences do not require special treatment for non-stationarity, since the likelihood function exhibits the same Gaussian shape regardless of the presence of non-stationarity. Following Banbura, Giannone, and Reichlin (2010) for the variables in the panel that are characterized by mean-reversion, however, we set a white-noise prior (i.e., $\bar{\beta}_i$ = 0); otherwise, we impose the random walk prior (i.e., $\bar{\beta}_i$ = 1). Note that when considering the large-scale BVAR model based on 151 variables, given that the system defined by equation (1) contains both *I(1)* and *I(0)* variables, we use the random-walk prior or white-noise prior accordingly. As for the FAVEC models, we begin with 115 I(1) variables, not counting the eight employment series, and we then cumulate the remaining 28 I(0) variables to transform them into non-stationary variables, before extracting the three factors. Appendix A lists these variables as well as the transformations used prior to analyzing the data.

The real activity group consists of variables such as industrial production, capacity utilization, retail sales, real personal consumption, real personal income, new orders, inventories, new housing starts (national and regional), housing sales (national and regional), employment, average working hours, and so on. The price and inflation group consists of variables such as the consumer price index, the producer price index, real housing prices (national and regional), the personal consumption expenditure deflator, average hourly earnings, exchange rates, and so on. The monetary sector group consists of variables such as monetary aggregates, various interest rates, credit outstanding, and so on.

5.2 Estimation and Results

In this section, we first select the optimal model for forecasting each sector's employment, using the minimum average root mean squared error (RMSE) across the one-, two-, … , and twelve-month-ahead out-of-sample forecasts. Then second, we consider ex ante out-of-sample forecasts.

The data sample for all eight employment series runs from January 1972 (1972:1) through March 2009 (2009:3). First, the cointegration tests amongst the eight employment series for the (B)VEC models as well as amongst the eight employment series and the three common static factors for the (B)FAVEC models, use data from 1972:1 through 1989:12. Further, this sample provides the base for estimating all of the various specifications considered for possible out-of-sample forecasting experiments. Second, the out-of-sample forecasting experiments cover 1990:1 through 2009:3. Third, we keep the number of factors extracted for the FAVAR and FAVEC models fixed over the forecasting period, but recursively update their estimates. Fourth, as each forecasting recursion also includes model selection, we choose the number of cointegrating vectors for the (B)VEC and (B)FAVEC models by using the trace test proposed by Johansen (1991). Fifth, we base the leg-length for the various models at each recursive estimation on the unanimity of at least two of the following five lag length selection criteria, namely, the sequential modified likelihood ratio (LR) test statistic (each test at the 5-percent level), the final prediction error (FPE), the Akaike information criterion (AIC), the Schwarz information criterion (SIC), and the Hannan-Quinn information criterion (HQIC).[13] Finally, for the large-scale BVAR, we use the lag-length chosen for the eight variable small-scale VAR containing only the eight sectoral employment series.

5.2.1 One- to Twelve-Month-Ahead Forecast Accuracy

Given the different forecasting models specified in Section 4, we estimate these alternative small- and large-scale models for the eight employment series in our sample over the period

---

[13] After determining the in-sample lag length for the VEC- and FAVEC-type models, we apply the trace test of cointegration to the eight employment series, and the eight employment series and the three factors for the FAVEC models, The tests suggest 5 and 10 cointegrating vectors, respectively, implying that the system contains 3 and 1 common trend(s), respectively. According to these results, the number of common factors (r) equals 3) and common trends (r^A) equal 3, which is the eight employment series (N^A) less five, the number of cointegrating vectors in the VEC. Note that, at each recursion, we choose the number of cointegrating vectors for the (B)VEC and (B)FAVEC models by using the trace test, hence the number of cointegrating relations are updated over the out-of-sample. Interestingly, at the end of the out-of-sample, we find that the number of cointegrating vectors falls to 3 in the VEC, while the number stays at 10 for the FAVEC. These results are available upon request from the authors.

1972:1 to 1989:12 using monthly data. We then compute out-of-sample one-, two-, …, and twelve-month-ahead forecasts for the period of 1990:1 to 2009:3, and compare the forecast accuracy relative to the forecasts generated by the benchmark random-walk (RW) with drift and estimated in levels. Note that the choice of the in-sample period, especially the starting date, depends on data availability. The starting point of the out-of-sample period precedes by a few months the recession in the 1990 and the jobless recovery that followed that recession as well as the recession in the 2001.

We estimate the multivariate versions of the classical AR, VAR, and VEC, the small-scale BVARs and BVECs, the large-scale BVARs, and the classical and Bayesian FAVARs and FAVECs over the period 1972:1 to 1989:12, and then forecast from 1990:1 through 2009:3. Depending on the number of lags selected, specific initial months feed the lags. We re-estimate the models each month over the out-of-sample forecast horizon in order to update the estimate of the coefficients, before producing the one-, two-, …, and twelve-month-ahead forecasts. We implemented this iterative estimation and the forecast procedure for 219 months, with the first forecast beginning in 1990:1. This produced a total of 219 one-, 219 two-, …, and 219 twelve-month-ahead forecasts. We calculate the root mean squared errors (RMSE)[14] for the 219 one-, two-, …, and twelve-month-ahead forecasts for the eight employment series across all of the different specifications. We then examine the average of the RMSE statistic for one-, two-, …, and twelve-month-ahead forecasts over 1990:1 to 2009:3. We select the model that produces the lowest average RMSE values as the 'optimal' specification for a specific state.

Tables 1 to 8 report the average of the one-, two-, …, and twelve-month-ahead RMSEs

---

[14] Note that if A_{t+n} denotes the actual value of a specific variable in period $t + n$ and ${}_tF_{t+n}$ equals the forecast made in period t for $t + n$, the RMSE statistic equals the following: $\sqrt{\left[\sum_1^N \left({}_tF_{t+n} - A_{t+n}\right)^2 / N\right]}$ where N equals the number of forecasts.

across the eight employment series, respectively. The benchmark for all forecast evaluations is the random-walk (RW) with drift model forecast RMSEs. Thus, the 0.3068 entry for the BFAVEC model in Table 1 means that the BFAVEC model experienced a forecast RMSE of only 30.68 percent of the forecast RMSE for the RW model. First, we consider the best performing model based on the average RMSE across the one-, two-, …, and twelve-month-ahead forecasts. Three different specifications prove optimal across the eight employment series. One, the BFAVEC models with w=0.1782 and d=2 prove optimal for mining and logging; manufacturing; financial activities; leisure and hospitality; and other service employment. Two, the LBVAR models with w=0.0230 and d=2 prove optimal for construction; and professional and business services, and come in a close second to the BFAAR models with w= 0.4672 and d=2, which proves optimal for forecasting trade, transportation and utilities employment. These results appear as the bold numbers in the Average column in Tables 1 to 8.[15]

Table 9 also tests whether the difference in forecasting performance proves significant relative to the RW forecasts, using the *ENC-T* test statistic. The BFAVEC model provide significantly better forecasts at the 1-percent level. The BFAAR models provide significantly better forecasts at only the 10-percent level, as does the LBVAR model for the construction employment. The LBVAR model for professional and business services employment provide better forecast relative to the RW model at the 5-percent level.

The forecasting results for the one-, two-, …, and twelve-month-ahead forecasts

---

[15] We also analyze the forecasting ability of a medium-scale BVAR model based on 28 variables, which includes 20 macroeconomic variables, besides the 8 employment series. Note that the 20 variables chosen are the same ones used in Banbura et al.,'s (2010) medium-scale BVAR, and are a subset of the 151 variables used in our large-scale models. We find that, barring the cases of financial activities employment; and professional and business services employment, the large-scale models tends to outperform the medium-scale model – a result similar to that obtained by Banbura et al., (2010). This result, in turn, strengthens our motivation to use large models that include 143 variables, over and above the 8 employment series. As with the Bayesian models for the large data set, the tightness of the prior is based on an in-sample fit of 50 percent. In addition, for the factor models, we use one factor. These results are available upon request from the authors.

generally follow a similar pattern. In most cases, a VEC (i.e., VEC, BVEC, FAVEC, or BFAVEC) model provides the best forecasting performance. This conclusion holds no matter whether the optimal model based on the average of the one-, two-, ..., and twelve-month-ahead forecasts yields the BFAVEC, LBVAR, or BFAAR models. That is, even when the optimal models for the average across all forecast horizons are the LBVAR or the BFAAR models, the VEC (i.e., VEC, BVEC, FAVEC, and BFAVEC) models frequently still provide the best forecasts in many instances.[16] .[17]

In sum, different specifications yield the best forecast performance based on RMSEs for different employment series and at different forecast horizons. One common pattern does emerge, nevertheless. No matter the forecast horizon, the VEC (i.e., VEC, BVEC, FAVEC, and BFAVEC) models generally provide the best forecast performance.

5.2.2 Comparing One- to Twelve-Month-Ahead Forecasts with the Actual Series

Figures 1 to 8 plot the out-of-sample forecasts and actual values from April 2009 through March 2010, using the best forecasting model for each employment series (see Table 9 for models). We used the average RMSEs reported in Tables 1 to 8 to select the best models.[18] Note that since the

---

[16] Note that the RMSE from the VEC-type models relative to the RW model is quite volatile, when compared to the (V)AR-type models. This mainly reflects the fact that the VECs are first estimated for the growth rates of the employment series, and hence, produce forecasts for the growth rates of employment. These growth rate forecasts are then converted back into their corresponding level-form using the actual data. Given that growth rate forecasts often tend to experience outliers, as is our case as well. This, in turn, tends to blow-up the RMSE ststitic. Similar results have also been observed by Gupta and Ziramba (2011) and Balcilar et al., (forthcoming) when using VECs for forecasting. The forecast results from the alternative models based on the growth rates of the employment series are available upon request from the authors.

[17] Following Carriero et al., (2011), we also analyze the forecasting performance of the LBVAR and the factor models in predicting the growth rate of the employment. As with the Bayesian models, forecasting the employment in levels, the tightness of the prior is based on an in-sample fit of 50 percent. We observe that the FAAR model performs the best for mining and logging; construction; and trade, transportation and utilities employment growth, while the BFAAR model stands out in the remaining five cases. Recall that there are four cases when the LBVAR model outperforms the factor models, when compared to forecasting employment in level-form, These results are available upon request from the authors.

[18] In addition to the ex ante out-of-sample forecasting exercise over 2009:4 to 2010:3, we also analyze the in-sample (1972:1-1989:12) and out-of-sample (1990:1-2009:3) forecasts obtained from the optimal models for each of the

BFAVEC performs the best in five of the eight employment series, we plot the ex ante forecasts from this model, even for the cases where the LBVAR and BFAAR models prove optimal. In addition, we also plot the combination forecasts obtained based on the average forecasts from the 14 (including the RW) different models estimated.

The forecast period captures the preliminary turn around in employment for all series except financial activities. Of course, whether the employment series actually bottom during this period or continue to fall with future releases remains an unanswered question. The worst forecast performance occurs in mining and logging employment, where the actual employment series bottomed in October 2009 while the forecast series continues on a downward trend throughout the forecast period.

The best forecast performance occurs for construction employment, where the actual and forecast series track each other closely. But, construction employment appears to bottom only in February 2010. The forecast series for manufacturing, financial activities, and leisure and hospitality employment each show a turnaround in employment over this period. But the forecast values recover too rapidly as compared to the actual series. For the remaining series – trade, transportation, and utilities; business services; and other services employment, the actual series show a more rapid turnaround over this period than the forecast values.

6. Conclusion

We forecast employment in eight sectors, using the AR, VAR, VEC, and their Bayesian counterparts, both with and without the information content of 143 additional monthly economic series. We examine two approaches for incorporating information from a large number of data series – extracting common factors (principle components) in a FAVAR, FAVEC, and their

---

eight employment series. The difference between the actual data and the predicted data for the in-sample is virtually inseparable, while the out-of-sample forecasts from the optimal models tend to predict the turning points quite well. We suppress these results to save space, but are available upon request from the authors.

Bayesian counterparts or Bayesian shrinkage in a LBVAR models.

Using the period of 1972:1 to 1989:12 as the in-sample period and 1990:1 to 2009:3 as the out-of-sample horizon, we compare the forecast performance of the alternative models for one- to twelve-month-ahead forecasts. Based on the average root mean squared error (RMSE) for the one-, two-, …, and twelve–month-ahead forecasts, we find that the factor-augmented models, albeit with different values for *w* and *d*, generally outperform the large-scale models for the eight employment series examined. A LBVAR model only provides the best forecasting performance for two employment series – construction employment at one-step ahead forecast horizon and professional and business services employment at one-, two-, and three-step ahead forecast horizons. In addition, amongst the factor augmented models, generally the VEC (i.e., FAVEC and BFAVEC) generally perform the best, highlighting the importance of modeling the long-run equilibrium relationship over and above the short-run dynamics.

We also compare the forecast and actual values of the employment series over April 2009 through March 2010 when all employment series, save one, show preliminary evidence of bottoming and starting to increase. The worst performing model forecasts mining and logging employment while the best performing model forecasts construction employment.

In sum, the utilization of a large dataset of economic variables, as well as long-run relationship with the short-run dynamics, improve the forecasting performance over models that do not use this data. In other words, macroeconomic fundamentals do matter when forecasting the eight employment series.

References:

Bai, J. (2004). Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data, *Journal of Econometrics* 122(1), 137-183.

Bai, J., and Ng, S., 2004. A PANIC Attack on Unit Roots and Cointegration. *Econometrica* 72(4), 1127-1177.

Balcilar, M., Gupta, R., Majumdar, A., and Miller, S. M. (forthcoming). Forecasting Nevada Gross Gaming Revenue and Taxable Sales Using Coincident and Leading Employment Indexes. *Empirical Economics.*

Bańbura, M., Giannone, D., and Reichlin, L., 2010. Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics* 25(1), 71–92.

Banerjee, A., and Marcellino, M. G., 2009. Factor-Augmented Error Correction Models. In Castle , J. L., and Shephard, N. (eds.), *The Methodology and Practice of Econometrics - A Festschrift for David Hendry*. Oxford: Oxford University Press, 227-254.

Banerjee, A., Marcellino, M. G., and Maston, I., 2010. Forecasting with Factor-Augmented Error Correction Models. CEPR Discussion Paper No. DP7677.

Bhardwaj, G., and Swanson, N. R., 2006. An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. *Journal of Econometrics* 131(1-2), 539-578.

Bernanke, B. S., Boivin, J., and Eliazs, P. (2005) Measuring the Effects of Monetary Policy: A Factor-Augmented Vector autoregressive (FAVAR) Approach, *The Quarterly Journal of Economics*, 120, 387–422.

Boivin, J., and Ng, S. (2005). Understanding and comparing factor based forecasts. *International Journal of Central Banking*, 1(3), 117-152.

Carriero, A., Kapetanios, G., and Marcellino, M. (2011). Forecasting large datasets with Bayesian reduced rank multivariate models. *Journal of Applied Econometrics 26*, 735–761.

Clark, T. E., and McCracken, M. W., 2001. Tests of Equal Forecast Accuracy and Encompassing for Nested Models. *Journal of Econometrics* 105(1), 85–110.

Das, S., Gupta, R., and Kabundi, A. (2009). Could We Have Forecasted the Recent Downturn in the South African Housing Market? *Journal of Housing Economics* 18(4), 325-335..

Doan, T. A., Litterman, R. B., and Sims, C. A. (1984). Forecasting and Conditional Projections Using Realistic Prior Distributions. *Econometric Reviews*, 3(1), 1-100.

Dua, P., and Ray, S. C. (1995). A BVAR Model for the Connecticut Economy. *Journal of Forecasting*, 14(3), 167-180.

Enders, W., 2004. *Applied Econometric Time Series*, 2nd Ed. John Wiley & Sons.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The Generalized Dynamic Factor Model, One Sided Estimation and Forecasting. *Journal of the American Statistical Association*, 100(471), 830–840.

Glennon, D., Lane, J., and Johnson, S., 1987. Regional Econometric Models that Reflect Labor Market Relations. *International Journal of Forecasting* 3(2), 299-312.

Gupta, R., and Miller, S. M. (forthcoming a) "'Ripple Effects" and Forecasting Home Prices in Los Angeles, Las Vegas, and Phoenix." *Annals of Regional Science*.

Gupta, R., and Miller, S. M. (forthcoming b) "The Time-Series Properties on Housing Prices: A Case Study of the Southern California Market." *Journal of Real Estate Finance and Economics*.

Gupta, R., and Ziramba, E. (2011) "Is the Permanent Income Hypothesis Really Well-Suited for Forecasting?" *Eastern Economic Journal*,37(2), 165-177.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59, 1551-1580.

Lane, T., 1966. The Urban Base Multiplier: An Evaluation of the State of the Art. *Land Economics* 42(3), 339-347.

LeSage, J. P., (1990). A Comparison of the Forecasting Ability of ECM and VAR Models, *Review of Economics and Statistics*, 72(4), 664-671.

LeSage, J. P. (1999). *Applied Econometrics Using MATLAB*, www.spatial-econometrics.com.

Litterman, R. B. (1981). A Bayesian Procedure for Forecasting with Vector Autoregressions. *Working Paper*, Federal Reserve Bank of Minneapolis.

Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions – Five Years of Experience. *Journal of Business and Economic Statistics*, 4(1), 25-38.

Newey, W. K., and West, K. D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3), 703-708.

Rapach, D. E., and Strauss, J. K. (2005). Forecasting Employment Growth in Missouri with Many Potentially Relevant Predictors: An Analysis of Forecast Combining Methods. Federal Reserve Bank of Saint Louis *Regional Economic Development* 1(1), 97-112.

Rapach, D. E., and Strauss, J. K. (2008). Forecasting US Employment Growth using Forecast Combining Methods. *Journal of Forecasting* 27(1), 75–93.

Rapach, D. E., and Strauss, J. K. (2010a). Bagging or Combining (or Both)? An Analysis Based on Forecasting U.S. Employment Growth. *Econometric Reviews* 29(5-6), 511–533.

Rapach, D. E., and Strauss, J. K. (2010b). Forecasting U.S. State-Level Employment Growth: An Amalgamation Approach. Saint Louis University, *Mimeo*.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1), 1-48.

Sims, C. A., Stock, J. H., and Watson, M. W., 1990. Inference in Linear Time Series Models with Some Unit Roots. *Econometrica* 58(1), 113–144.

Spencer, D. E. (1993). Developing a Bayesian Vector Autoregression Model. *International Journal of Forecasting*, 9(3), 407-421.

Stevens, B. H., and Moore, C. L., 1980. A Critical Review of the Literature on Shift-Share as a Forecasting Technique. *Journal of Regional Science* 20(4), 419–437.

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

Stock, J. H., and Watson, M.W. (2002a) Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, 97(460), 147–162.

Stock, J. H., and Watson, M. W., (2002b) Macroeconomics forecasting using diffusion indexes, *Journal of Business and Economic Statistics*, 20(2), 147–162.

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

Stock, J. H., and Watson, M. W., (2004). Combination Forecasts of Output Growth in a Seven-Country Data Set. *Journal of Forecasting*, *23*(6), 405-430.

Stock, J. H., and Watson, M. W., (2005). Implications of dynamic factor models for VAR analysis. NBER Working Paper No. 11467.

Taylor, C. A., 1982. Econometric Modeling of Urban and Other Substate Areas: An Analysis of Alternative Methodologies. *Regional Science and Urban Economics* 12(3), 425-448.

Theil, H. (1971). *Principles of Econometrics*. New York: John Wiley.

Todd, R. M. (1984). Improving Economic Forecasting with Bayesian Vector Autoregression. *Quarterly Review*, Federal Reserve Bank of Minneapolis, Fall, 18-29.

Williamson, R. B., 2006. Predictive Power of the Export Base Theory. *Growth and Change* 6(1), 3–10.

Zellner, A., and Palm, F. (1974). Time Series Analysis and Simultaneous Equation Econometric Models. *Journal of Econometrics*, 2(1), 17-54.

Table 1: One- to twelve-months-ahead forecast for Mining & Logging Employment (1990:1-2009:3)

	Models	1	2	3	4	5	6	7	8	9	10	11	12	Average
	AR	0.9853	1.0092	1.0137	1.0123	1.0013	0.9969	0.9979	1.0021	1.0073	1.0123	1.0160	1.0193	1.0061
	VAR	3.7004	4.0382	3.9870	4.1106	4.1408	4.2559	4.3870	4.5101	4.6408	4.8198	4.9835	5.1304	4.3920
	FAAR	4.2751	4.2552	4.2043	4.2717	4.2057	4.1661	4.1154	4.0807	4.0174	4.0637	4.0564	4.0396	4.1459
	FAVAR	1.4424	1.3539	1.2048	1.1112	1.0096	0.9496	0.8922	0.8610	0.8300	0.7941	0.7680	0.7454	0.9968
	VECM	2.7565	**0.1835**	**0.2627**	3.6758	3.7609	1.8321	1.9592	0.6140	5.3147	9.5587	21.4553	22.8251	6.1832
	FAVECM	**0.4043**	1.7191	1.5952	2.2532	2.2682	2.3753	2.3648	2.5441	2.2523	2.1070	1.8466	1.6341	1.9470
w=1.8250, d=2	**BAR**	0.9732	0.9879	0.9902	0.9883	0.9726	0.9692	0.9723	0.9806	0.9892	0.9979	1.0038	1.0071	0.9860
w=0.2336, d=2	**BVAR**	1.3440	1.4630	1.5297	1.5742	1.6090	1.6385	1.6677	1.6931	1.7141	1.7350	1.7550	1.7744	1.6248
w=0.4673, d=2	**BFAAR**	0.9722	0.9378	0.9110	0.8708	0.8534	0.8448	0.8360	0.8273	0.8188	0.8106	0.8045	0.8017	0.8574
w=0.1699, d=2	**BFAVAR**	1.3228	1.4151	1.4749	1.5265	1.5713	1.6022	1.6336	1.6593	1.6795	1.6996	1.7192	1.7390	1.5869
w=0.2419, d=2	**BVECM**	1.2087	3.2360	3.2332	2.6029	2.0898	1.8864	1.7423	1.6014	1.3586	1.1620	0.9201	0.7085	1.8125
w=0.1782, d=2	**BFAVECM**	0.5087	0.3333	0.3405	**0.4444**	**0.5178**	**0.4802**	**0.3967**	**0.2699**	**0.2483**	**0.0981**	**0.0288**	**0.0146**	**0.3068**
w=0.0230, d=2	**LBVAR**	0.8531	0.8323	0.8547	0.8804	0.9249	0.9852	1.0304	1.0720	1.1013	1.1197	1.1232	1.1314	0.9924

Note: AR, VAR, FAAR, FAVAR, VEC, and FAVEC refer to autoregressive, vector autoregressive, factor-augmented vector autoregressive, factor-augmented vector autoregressive, vector error-correction, and factor-augmented error-correction models. BAR, BVAR, BFAAR, BFAVAR, BVEC, BFAVEC, and LBVAR refer to Bayesian AR, VAR, FAAR, FAVAR, VEC, and FAVEC models. The text identifies various priors and parameterizations. RMSE means root mean square error. The entries measure the average RMSE across all forecasts at each horizon – one-, two-, …, and twelve-month-ahead forecasts as well as the average RMSE across the individual forecasts. Bold numbers represent the minimum value in each column.

Table 2: One- to twelve-months-ahead forecast for Construction Employment (1990:1-2009:3)

	Models	1	2	3	4	5	6	7	8	9	10	11	12	Average
	AR	0.7659	0.6775	0.6622	0.6694	0.6927	0.7127	0.7334	0.7536	0.7753	0.7958	0.8150	0.8342	0.7406
	VAR	1.2265	1.1850	1.2043	1.2825	1.3994	1.5316	1.6489	1.7582	1.8682	1.9590	2.0251	2.0771	1.5972
	FAAR	1.3792	1.4090	1.4248	1.5056	1.5906	1.6826	1.7573	1.8346	1.9152	1.9870	2.0276	2.0618	1.7146
	FAVAR	0.8272	0.7222	0.6682	0.6637	0.6817	0.7030	0.7296	0.7502	0.7751	0.7980	0.8125	0.8281	0.7466
	VECM	2.4944	3.9569	6.5679	46.5333	3.7338	2.2800	2.3069	0.9573	1.0346	0.4862	0.8052	0.9031	6.0050
	FAVECM	1.0519	1.5828	2.8189	19.2000	1.3959	0.5832	0.4014	0.0388	**0.1023**	**0.2305**	**0.2900**	**0.3192**	2.3346
w=1.8250, d=2	**BAR**	0.7533	0.6596	0.6389	0.6409	0.6600	0.6766	0.6947	0.7130	0.7331	0.7523	0.7708	0.7901	0.7069
w=0.2336, d=2	**BVAR**	0.7461	0.6591	0.6420	0.6506	0.6773	0.7056	0.7349	0.7622	0.7891	0.8182	0.8442	0.8669	0.7413
w=0.4673, d=2	**BFAAR**	0.7320	0.6352	0.6115	0.6139	0.6312	0.6458	0.6696	0.6903	0.7134	0.7353	0.7536	0.7688	0.6834
w=0.1699, d=2	**BFAVAR**	0.7333	0.6427	0.6223	0.6328	0.6624	0.6912	0.7225	0.7498	0.7767	0.8054	0.8313	0.8541	0.7270
w=0.2419, d=2	**BVECM**	1.9593	3.0522	4.5309	14.5667	**0.0239**	0.6653	0.4486	0.5849	0.4534	0.5660	0.4352	0.4185	2.3087
w=0.1782, d=2	**BFAVECM**	1.1611	1.4580	2.3663	14.6000	1.0444	**0.5495**	**0.2208**	**0.0068**	0.1203	0.2586	0.3005	0.3450	1.8693
w=0.0230, d=2	**LBVAR**	**0.6843**	**0.5665**	**0.5361**	**0.5375**	0.5630	0.5933	0.6214	0.6531	0.6841	0.7116	0.7328	0.7610	**0.6371**

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 3: **One- to twelve-months-ahead forecast for Manufacturing Employment (1990:1-2009:3)**

	Models	1	2	3	4	5	6	7	8	9	10	11	12	Average
	AR	0.6953	0.6663	0.7107	0.7766	0.8424	0.9079	0.9609	1.0119	1.0564	1.0924	1.1250	1.1585	0.9170
	VAR	1.0242	1.1526	1.3530	1.5924	1.8180	2.0323	2.2204	2.3976	2.5559	2.6869	2.7974	2.8891	2.0433
	FAAR	1.1576	1.3119	1.5674	1.8819	2.1660	2.4231	2.6308	2.8040	2.9325	3.0269	3.1066	3.1769	2.3488
	FAVAR	0.7027	0.6302	0.6298	0.6483	0.6600	0.6911	0.7298	0.7668	0.7996	0.8259	0.8426	0.8555	0.7319
	VECM	0.6800	2.1061	1.9359	1.4385	1.4723	0.9680	0.4596	0.7809	1.5969	1.9798	1.8783	1.7495	1.4205
	FAVECM	0.8400	0.6970	0.2692	**0.0738**	**0.0686**	**0.0064**	**0.1553**	**0.1977**	**0.2211**	0.2375	0.3457	0.3422	0.2879
w=1.8250, d=2	**BAR**	0.6859	0.6579	0.7035	0.7706	0.8381	0.9048	0.9583	1.0087	1.0519	1.0872	1.1207	1.1555	0.9119
w=0.2336, d=2	**BVAR**	0.6164	0.5616	0.5815	0.6259	0.6764	0.7223	0.7587	0.7932	0.8226	0.8481	0.8717	0.8941	0.7310
w=0.4673, d=2	**BFAAR**	0.6373	0.5791	0.5942	0.6274	0.6564	0.6947	0.7386	0.7803	0.8194	0.8527	0.8774	0.8991	0.7297
w=0.1699, d=2	**BFAVAR**	0.5984	0.5417	0.5613	0.6026	0.6489	0.6891	0.7298	0.7652	0.7958	0.8218	0.8458	0.8680	0.7057
w=0.2419, d=2	**BVECM**	**0.3429**	2.6364	1.3718	1.0533	0.6675	0.6247	0.3851	0.2846	0.2495	**0.2260**	**0.0408**	**0.0387**	0.6601
w=0.1782, d=2	**BFAVECM**	0.8400	0.6515	**0.2308**	0.0820	0.0844	0.0107	0.1677	0.2065	0.2277	0.2404	0.3472	0.3428	**0.2860**
w=0.0230, d=2	**LBVAR**	0.4883	**0.4076**	0.4000	0.4396	0.4903	0.5372	0.5711	0.5957	0.6171	0.6204	0.6137	0.6187	0.5333

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 4: **One- to twelve-months-ahead forecast for Trade, Transportation and Utilities Employment (1990:1-2009:3)**

	Models	1	2	3	4	5	6	7	8	9	10	11	12	Average
	AR	0.7121	0.6151	0.6034	0.6174	0.6362	0.6508	0.6671	0.6785	0.6937	0.7098	0.7255	0.7399	0.6708
	VAR	1.0339	0.9817	0.9906	1.0421	1.1235	1.2288	1.3329	1.4258	1.5168	1.5983	1.6587	1.7066	1.3033
	FAAR	1.1844	1.0964	1.0895	1.1363	1.2161	1.3031	1.3750	1.4349	1.4976	1.5404	1.5545	1.5613	1.3325
	FAVAR	0.7127	0.5824	0.5507	0.5603	0.5719	0.5870	0.6099	0.6291	0.6561	0.6782	0.6945	0.7110	0.6287
	VECM	2.5405	8.9744	12.6923	5.8182	4.6164	4.0094	3.1185	2.4518	2.3433	2.0924	1.8485	1.7739	4.3566
	FAVECM	1.3986	4.2564	3.8205	1.5091	0.6119	0.3125	0.1754	0.0214	0.0851	0.1582	0.1899	0.2290	1.0640
w=1.8250, d=2	**BAR**	0.7119	0.6198	0.6109	0.6268	0.6487	0.6655	0.6834	0.6967	0.7139	0.7313	0.7489	0.7650	0.6852
w=0.2336, d=2	**BVAR**	0.7010	0.6212	0.6158	0.6334	0.6562	0.6791	0.7033	0.7239	0.7468	0.7699	0.7908	0.8090	0.7042
w=0.4673, d=2	**BFAAR**	0.6598	0.5545	0.5386	0.5566	0.5739	0.5889	0.6161	0.6373	0.6624	0.6849	0.7034	0.7197	**0.6247**
w=0.1699, d=2	**BFAVAR**	0.6840	0.5950	0.5809	0.5995	0.6248	0.6476	0.6763	0.7003	0.7274	0.7542	0.7781	0.7992	0.6806
w=0.2419, d=2	**BVECM**	1.4054	3.6154	3.2564	1.1818	0.4658	0.2531	0.1422	**0.0071**	**0.0358**	**0.0785**	**0.1120**	**0.1204**	0.8895
w=0.1782, d=2	**BFAVECM**	1.2568	2.5641	1.8718	0.5818	**0.1005**	**0.0344**	**0.0995**	0.1929	0.2149	0.2367	0.2525	0.2498	0.6380
w=0.0230, d=2	**LBVAR**	**0.6522**	**0.5533**	**0.5343**	**0.5397**	0.5633	0.5884	0.6153	0.6386	0.6690	0.6957	0.7175	0.7427	0.6258

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 5: **One- to twelve-months-ahead forecast for Financial Activities Employment (1990:1-2009:3)**

	Models	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Average**
	AR	**0.5940**	**0.5559**	0.5604	0.5878	0.6195	0.6548	0.6861	0.7134	0.7400	0.7643	0.7846	0.8022	0.6719
	VAR	0.7615	0.7362	0.7624	0.8210	0.8830	0.9627	1.0351	1.1120	1.1940	1.2782	1.3578	1.4410	1.0288
	FAAR	0.8136	0.7673	0.8012	0.8622	0.9147	0.9861	1.0472	1.1094	1.1690	1.2331	1.2948	1.3599	1.0299
	FAVAR	0.6268	0.5771	0.5776	0.6064	0.6403	0.6866	0.7292	0.7690	0.8092	0.8511	0.8885	0.9218	0.7236
	VECM	0.8696	3.7619	2.7143	3.6486	2.7458	5.7813	6.9231	4.8396	2.2470	1.1867	0.9511	1.2754	3.0787
	FAVECM	0.8261	1.2381	**0.2143**	1.1081	0.7627	0.9688	0.9231	0.6509	0.1687	**0.0444**	**0.2117**	0.3449	0.6218
w=1.8250, d=2	**BAR**	0.5950	0.5560	0.5609	**0.5862**	0.6162	0.6492	0.6777	0.7017	0.7247	0.7449	0.7612	0.7752	0.6624
w=0.2336, d=2	**BVAR**	0.6352	0.6169	0.6429	0.6905	0.7430	0.7989	0.8533	0.9056	0.9587	1.0106	1.0598	1.1064	0.8351
w=0.4673, d=2	**BFAAR**	0.6025	0.5636	0.5679	0.5892	0.6153	0.6460	0.6737	0.6960	0.7199	0.7423	0.7610	0.7761	0.6628
w=0.1699, d=2	**BFAVAR**	0.6353	0.6161	0.6416	0.6880	0.7394	0.7945	0.8486	0.9001	0.9527	1.0043	1.0532	1.0995	0.8311
w=0.2419, d=2	**BVECM**	1.2174	1.4762	1.0952	0.8649	0.6949	0.5156	0.4487	0.4717	0.5783	0.6356	0.6906	0.7295	0.7849
w=0.1782, d=2	**BFAVECM**	1.2609	1.4286	1.0714	0.7027	**0.4746**	**0.1094**	**0.1026**	**0.1321**	**0.0422**	0.1200	0.2150	**0.2878**	**0.4956**
w=0.0230, d=2	**LBVAR**	0.6424	0.6308	0.6515	0.6821	0.7218	0.7682	0.8146	0.8534	0.8871	0.9193	0.9517	0.9889	0.7927

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 6: **One- to twelve-months-ahead forecast for Professional and Business Services Employment (1990:1-2009:3)**

	Models	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Average**
	AR	0.6495	0.5882	0.5821	0.5951	0.6103	0.6335	0.6582	0.6782	0.6987	0.7192	0.7370	0.7539	0.6587
	VAR	0.7868	0.7398	0.7656	0.8175	0.8657	0.9266	0.9902	1.0477	1.0982	1.1398	1.1696	1.1877	0.9613
	FAAR	0.8324	0.7664	0.7792	0.8301	0.8734	0.9295	0.9816	1.0280	1.0618	1.0870	1.1028	1.1113	0.9486
	FAVAR	0.6732	0.5995	0.5849	0.5986	0.6129	0.6410	0.6724	0.6993	0.7283	0.7561	0.7782	0.7978	0.6785
	VECM	6.2353	1.9263	3.6421	14.9394	3.6012	1.3808	0.6113	0.3212	0.7858	0.7650	0.9017	0.5929	2.9752
	FAVECM	3.5294	1.8947	2.0737	3.7879	**0.1429**	**0.0042**	**0.0475**	0.3078	0.3840	0.4755	0.4954	0.5292	1.1393
w=1.8250, d=2	**BAR**	0.6503	0.5886	0.5811	0.5924	0.6071	0.6308	0.6560	0.6758	0.6957	0.7156	0.7324	0.7482	0.6562
w=0.2336, d=2	**BVAR**	0.6613	0.6104	0.6100	0.6280	0.6490	0.6757	0.7023	0.7246	0.7468	0.7685	0.7883	0.8060	0.6976
w=0.4673, d=2	**BFAAR**	0.6413	0.5739	0.5652	0.5784	0.5913	0.6111	0.6355	0.6556	0.6771	0.6974	0.7129	0.7264	0.6388
w=0.1699, d=2	**BFAVAR**	0.6586	0.6038	0.6020	0.6208	0.6422	0.6686	0.6964	0.7191	0.7419	0.7640	0.7840	0.8022	0.6920
w=0.2419, d=2	**BVECM**	2.5882	1.6211	2.1474	4.0606	0.3631	0.2259	0.0593	0.1931	**0.2762**	**0.3928**	**0.4193**	**0.4654**	1.0677
w=0.1782, d=2	**BFAVECM**	1.2941	0.9158	1.0947	**0.1818**	0.5536	0.5146	0.5252	0.6023	0.6145	0.6540	0.6484	0.6590	0.6882
w=0.0230, d=2	**LBVAR**	**0.5409**	**0.4670**	**0.4528**	0.4544	0.4665	0.4919	0.5178	0.5389	0.5634	0.5850	0.6072	0.6306	**0.5264**

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 7: **One- to twelve-months-ahead forecast for Leisure and Hospitality Employment (1990:1-2009:3)**

	Models	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Average**
	AR	0.9533	0.9236	0.9021	0.9084	0.9271	0.9390	0.9425	0.9422	0.9433	0.9456	0.9483	0.9512	0.9356
	VAR	1.5839	1.7317	1.7335	1.6661	1.6182	1.6311	1.6621	1.7261	1.7963	1.8488	1.8845	1.9015	1.7320
	FAAR	1.8260	2.0116	1.9889	1.8602	1.7806	1.7553	1.6964	1.6442	1.5873	1.5389	1.5056	1.4687	1.7220
	FAVAR	1.1834	1.1795	1.1478	1.1082	1.0563	1.0229	0.9828	0.9455	0.9229	0.9209	0.9255	0.9248	1.0267
	VECM	**0.1143**	4.2023	4.6432	2.7578	1.2642	1.1381	1.1932	1.1863	1.1091	1.3201	1.0572	1.1373	1.7603
	FAVECM	2.3429	0.7688	1.9437	0.8022	**0.1169**	0.0844	0.4022	0.5213	0.3900	0.4418	0.1627	**0.1061**	0.6736
w=1.8250, d=2	**BAR**	0.9499	0.9237	0.9086	0.9162	0.9329	0.9429	0.9457	0.9448	0.9447	0.9453	0.9464	0.9478	0.9374
w=0.2336, d=2	**BVAR**	0.9331	0.8938	0.8681	0.8508	0.8408	0.8388	0.8423	0.8447	0.8586	0.8704	0.8881	0.9020	0.8693
w=0.4673, d=2	**BFAAR**	0.9163	0.8665	0.8310	0.8178	0.8156	0.8107	0.8140	0.8137	0.8225	0.8350	0.8426	0.8471	0.8361
w=0.1699, d=2	**BFAVAR**	0.9381	0.8988	0.8704	0.8516	0.8411	0.8356	0.8378	0.8383	0.8515	0.8638	0.8814	0.8950	0.8669
w=0.2419, d=2	**BVECM**	0.8000	1.1387	1.5399	0.5156	1.3051	0.1867	0.2807	0.2258	0.1996	**0.0567**	0.1345	0.1108	0.5412
w=0.1782, d=2	**BFAVECM**	0.1714	**0.7341**	1.1174	**0.2244**	0.5057	**0.0818**	**0.0243**	**0.0718**	**0.0906**	0.1955	**0.1278**	0.1346	**0.2900**
w=0.0230, d=2	**LBVAR**	0.8288	0.7583	**0.7171**	0.7036	0.7027	0.6904	0.6847	0.6839	0.6891	0.6987	0.7036	0.7119	0.7144

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 8: **One- to twelve-months-ahead forecast for Other Services Employment (1990:1-2009:3)**

	Models	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Average**
	AR	0.7668	0.7284	**0.7255**	**0.7440**	0.7601	0.7699	0.7861	0.8024	0.8182	0.8358	0.8498	0.8604	0.7873
	VAR	1.1039	1.0204	1.0366	1.1111	1.1882	1.2731	1.3637	1.4380	1.5027	1.5643	1.6178	1.6517	1.3226
	FAAR	1.2979	1.1704	1.1704	1.2407	1.3013	1.3563	1.4144	1.4556	1.4886	1.5261	1.5545	1.5636	1.3783
	FAVAR	0.8161	0.7571	0.7478	0.7647	0.7813	0.8025	0.8304	0.8514	0.8760	0.9034	0.9269	0.9465	0.8337
	VECM	2.2048	2.5195	6.8551	18.2857	7.3548	1.3196	0.5487	0.5084	1.1493	1.4086	1.5349	1.1271	3.7347
	FAVECM	1.5422	1.6623	2.2174	2.4000	0.7849	**0.0365**	0.2813	0.4051	0.4577	0.4400	0.4736	0.4672	0.9307
w=1.8250, d=2	**BAR**	0.7704	0.7310	0.7296	0.7502	0.7680	0.7793	0.7961	0.8113	0.8257	0.8425	0.8558	0.8653	0.7938
w=0.2336, d=2	**BVAR**	0.7803	0.7484	0.7508	0.7673	0.7851	0.7996	0.8175	0.8321	0.8446	0.8582	0.8706	0.8796	0.8112
w=0.4673, d=2	**BFAAR**	0.7897	0.7533	0.7527	0.7682	0.7765	0.7779	0.7945	0.8096	0.8258	0.8428	0.8554	0.8636	0.8008
w=0.1699, d=2	**BFAVAR**	0.7841	0.7500	0.7510	0.7668	0.7828	0.7938	0.8114	0.8253	0.8379	0.8513	0.8631	0.8714	0.8074
w=0.2419, d=2	**BVECM**	1.3855	2.0130	2.7971	3.8571	1.3226	0.1963	**0.1393**	**0.2468**	0.3267	0.3486	0.3894	0.4078	1.1192
w=0.1782, d=2	**BFAVECM**	1.5301	1.4416	1.5797	1.0571	**0.3226**	0.1553	0.2869	0.2932	**0.3002**	**0.2643**	**0.2608**	**0.2394**	**0.6443**
w=0.0230, d=2	**LBVAR**	**0.7414**	**0.7237**	0.7380	0.7541	0.7877	0.8130	0.8480	0.8822	0.9176	0.9533	0.9824	1.0070	0.8457

Note: See Table 1. Bold numbers represent the minimum value in each **column.**

Table 9: ENC-T Test of Differences between Optimal and Random-Walk Models

		QA											
Employment Series	**Optimal Model**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Mining & Logging	**BFAVECM (w=0.1782,d=2)**	-49.13*	-66.67*	-65.95*	-55.56*	-48.22*	-51.98*	-60.33*	-73.01*	-75.17*	-90.19*	-97.12*	-98.54*
Construction	**LBVAR (w=0.0230,d=2)**	-31.57†	-43.35**	-46.39**	-46.25**	-43.70**	-40.67**	-37.86**	-34.69†	-31.59†	-28.84†	-26.72†	-23.90†
Manufacturing	**BFAVECM (w=0.1782,d=2)**	-16.00	-34.85†	-76.92*	-91.80*	-91.56*	-98.93*	-83.23*	-79.35*	-77.23*	-75.96*	-65.28*	-65.72*
Trade, Transport. & Utilities	**BFAAR (w=0.2,d=1)**	-34.02†	-44.55**	-46.14**	-44.34**	-42.61**	-41.11**	-38.39**	-36.27**	-33.76†	-31.51†	-29.66†	-28.03†
Financial Activities	**BFAVECM (w=0.1782,d=2)**	26.09†	42.86**	7.14	-29.73†	-52.54*	-89.06*	-89.74*	-86.79*	-95.78*	-88.00*	-78.50*	-71.22*
Professional & Business Services	**LBVAR (w=0.0230,d=2)**	-45.91**	-53.30*	-54.72*	-54.56*	-53.35*	-50.81*	-48.22**	-46.12**	-43.66**	-41.50**	-39.28**	-36.94**
Leisure & Hospitality	**BFAVECM (w=0.1782,d=2)**	-82.86*	-26.59†	11.74	-77.56*	-49.43*	-91.82*	-97.57*	-92.82*	-90.94*	-80.45*	-87.22*	-86.54*
Other Services	**BFAVECM (w=0.1782,d=2)**	53.01*	44.16**	57.97*	5.71	-67.74*	-84.47*	-71.31*	-70.68*	-69.98*	-73.57*	-73.92*	-76.06*

Note: The ENC-T statistics test the difference in RMSEs betweenthe optimal model relative to the random-walk model. Negative signs mean that the optimal model forecasts better than the random-walk model.

* means significant at the 1-percent level.

** means significant at the 5-percent level.

† means significant at the 10-percent level.

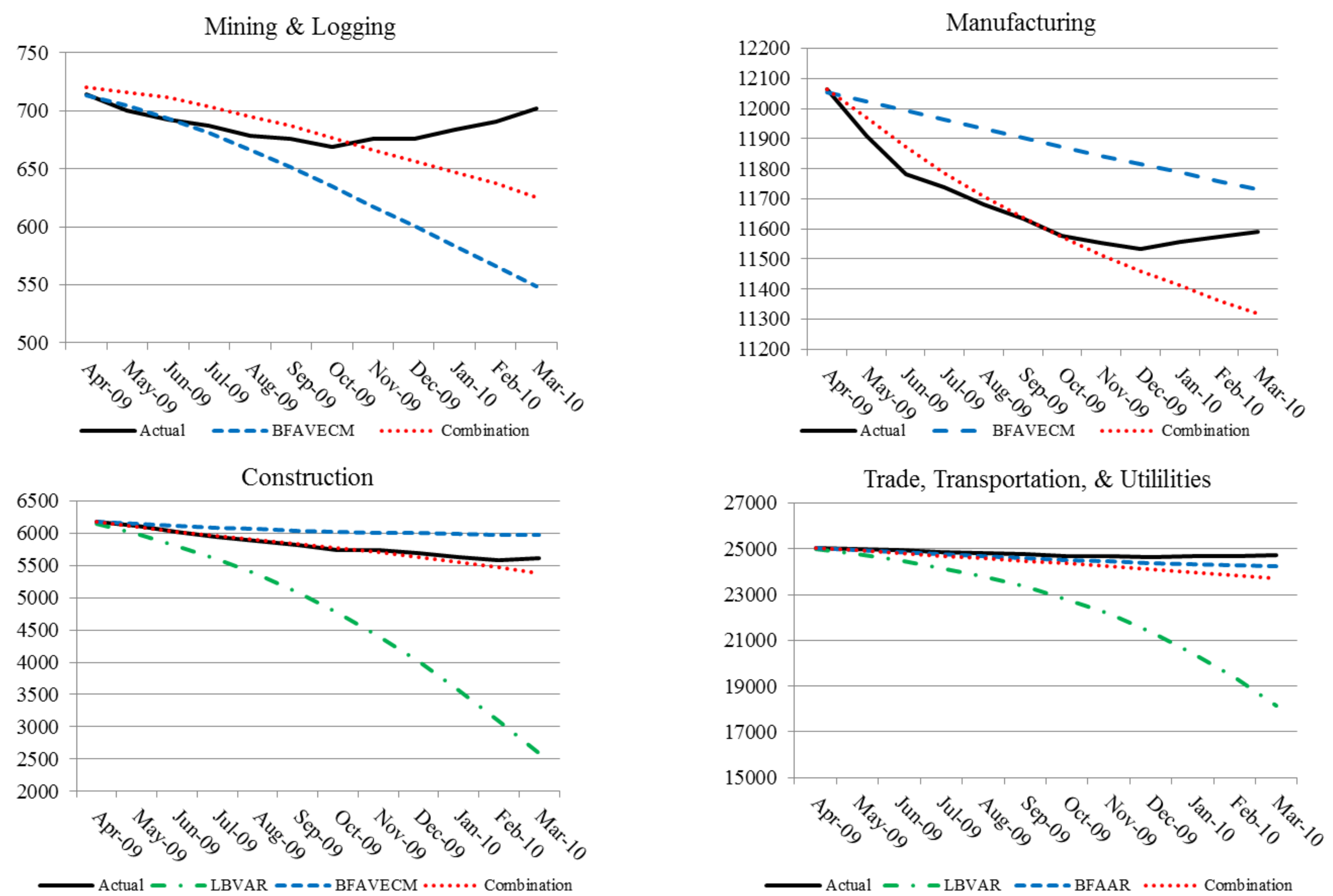

Figure 1: Actual and Forecast Values of Eight Employment Series

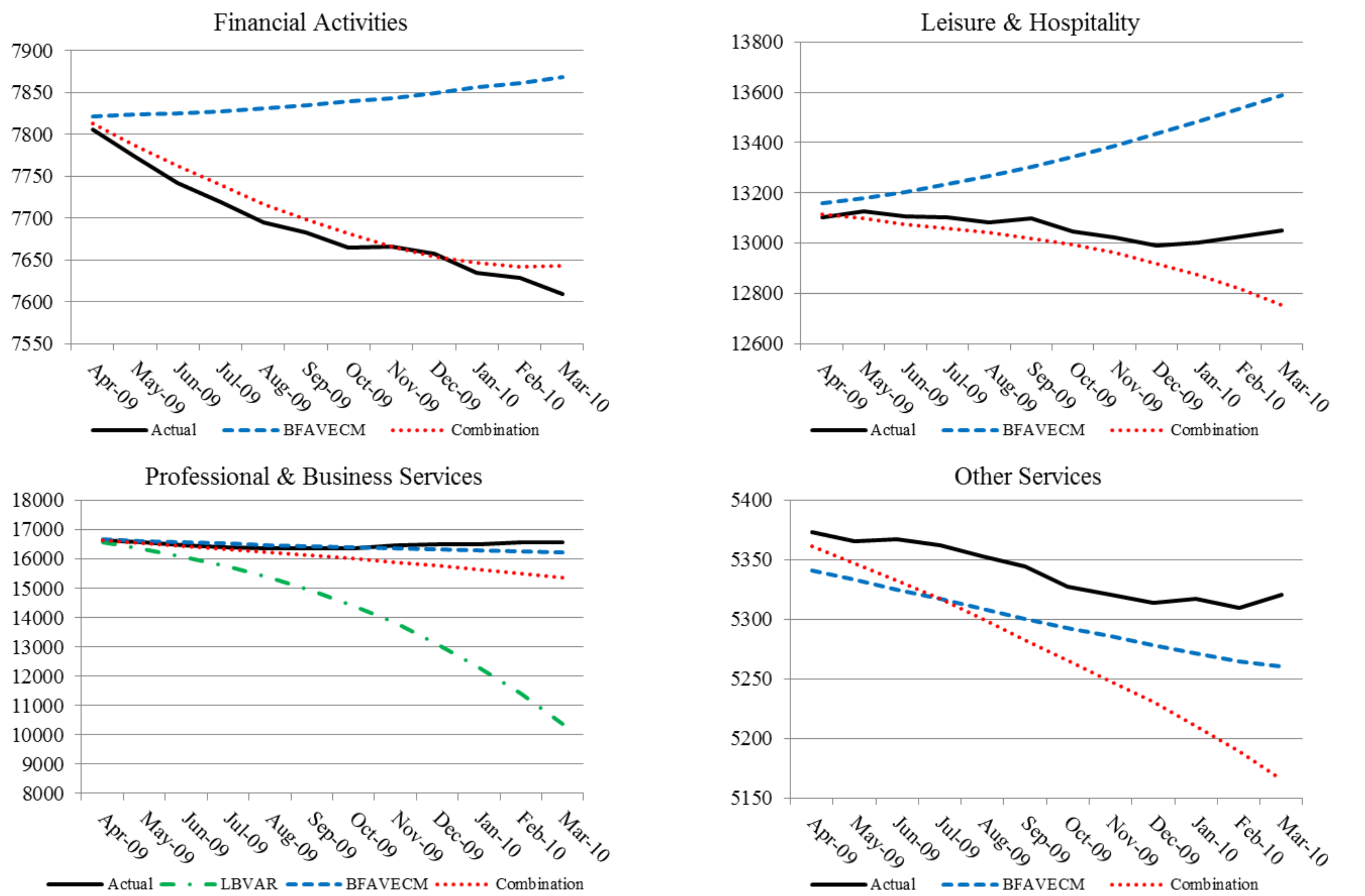

Figure 1: Actual and Forecast Values of Eight Employment Series (continued)

Appendix A:

Table A1: Variables

Data Code	Variable Name	Format
a0m052	PERSONAL INCOME (AR, BILL. CHAIN 2000 $)	5
A0M051	PERSONAL INCOME LESS TRANSFER PAYMENTS (AR, BILL. CHAIN 2000 $)	5
A0M224_R	REAL CONSUMPTION (AC) A0M224/GMDC	5
A0M057	MANUFACTURING AND TRADE SALES (MIL. CHAIN 1996 $)	5
A0M059	SALES OF RETAIL STORES (MIL. CHAIN 2000 $)	5
IPS10	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX	5
IPS11	INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL	5
IPS299	INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS	5
IPS12	INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS	5
IPS13	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS	5
IPS18	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS	5
IPS25	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT	5
IPS32	INDUSTRIAL PRODUCTION INDEX - MATERIALS	5
IPS34	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS	5
IPS38	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS	5
IPS43	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)	5
IPS307	INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES	5
IPS306	INDUSTRIAL PRODUCTION INDEX - FUELS	5
IPDM	INDUSTRIAL PRODUCTION: DURABLE MANUFACTURING (NAICS)	5
IPNDM	INDUSTRIAL PRODUCTION: NONDURABLE MANUFACTURING (NAICS)	5
IPM	INDUSTRIAL PRODUCTION: MINING	5
IPGEU	INDUSTRIAL PRODUCTION: ELECTRIC AND GAS UTILITIES	5
PMP	NAPM PRODUCTION INDEX (PERCENT)	1
A0m082	CAPACITY UTILIZATION (MFG)	2
LHEL	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)	2
LHELX	EMPLOYMENT: RATIO; HELP-WANTED ADS: NO. UNEMPLOYED CLF	2
LHEM	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)	5
LHNAG	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)	5
LHUR	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA)	2
LHU680	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)	2
LHU5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)	5
LHU14	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)	5
LHU15	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS., SA)	5
LHU26	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)	5
LHU27	UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS, SA)	5
A0M005	AVERAGE WEEKLY INITIAL CLAIMS, UNEMPLOYMENT INSURANCE (THOUS.)	5
CES002	EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE	5
CES003	EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING	5
CES006	EMPLOYEES ON NONFARM PAYROLLS - MINING	5
CES017	EMPLOYEES ON NONFARM PAYROLLS - DURABLE GOODS	5
CES033	EMPLOYEES ON NONFARM PAYROLLS - NONDURABLE GOODS	5
CES046	EMPLOYEES ON NONFARM PAYROLLS - SERVICE-PROVIDING	5
CES049	EMPLOYEES ON NONFARM PAYROLLS - WHOLESALE TRADE	5
CES053	EMPLOYEES ON NONFARM PAYROLLS - RETAIL TRADE	5
CES140	EMPLOYEES ON NONFARM PAYROLLS - GOVERNMENT	5
CESNRM	ALL EMPLOYEES: NATURAL RESOURCES & MINING	5
CEML	MINING & LOGGING EMPLOYMENT	5
CEC	CONSTRUCTION EMPLOYMENT	5
CEM	MANUFACTURING EMPLOYMENT	5
CETTU	TRADE, TRANS. & UTIL. EMPLOYMENT	5
CEFA	FINANCIAL ACTIVITIES EMPLOYMENT	5
CEPBS	PROF & BUS. SERV. EMPLOYMENT	5

Data Code	Variable Name	Format
CELH	LEISURE & HOSPITALITY EMPLOYMENT	5
CEOS	OTHER SERVICES EMPLOYMENT	5
CES151	AVERAGE WEEKLY HOURS: MANUFACTURING	1
CES155	AVERAGE WEEKLY HOURS: OVERTIME: MANUFACTURING	2
PMEMP	NAPM EMPLOYMENT INDEX (PERCENT)	1
HSFR	HOUSING STARTS:TOTAL (THOUS.U)S.A.	4
HSNE	HOUSING STARTS: NORTHEAST (THOUS.U.)S.A.	4
HSMW	HOUSING STARTS: MIDWEST (THOUS.U.)S.A.	4
HSSOU	HOUSING STARTS: SOUTH (THOUS.U.)S.A.	4
HSWST	HOUSING STARTS: WEST (THOUS.U.)S.A.	4
HSBR	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)	4
HSBNE	HOUSES AUTHORIZED BY BUILD. PERMITS: NORTHEAST (THOU.U.)S.A	4
HSBMW	HOUSES AUTHORIZED BY BUILD. PERMITS: MIDWEST (THOU.U.)S.A.	4
HSBSOU	HOUSES AUTHORIZED BY BUILD. PERMITS: SOUTH (THOU.U.)S.A.	4
HSBWST	HOUSES AUTHORIZED BY BUILD. PERMITS: WEST (THOU.U.)S.A.	4
HPNE	REAL HOUSE PRICE NORTHEAST	6
HPMW	REAL HOUSE PRICE MIDWEST	6
HPS	REAL HOUSE PRICE SOUTH	6
HPW	REAL HOUSE PRICE WEST	6
HPUS	REAL HOUSE PRICE US	6
SNE	HOME SALES NORTHEAST	6
SMW	HOME SALES MIDWEST	6
SS	HOME SALES SOUTH	6
SW	HOME SALES WEST	6
SUS	HOME SALES US	6
HMOB	MOBILE HOMES: MANUFACTURERS' SHIPMENTS (THOUS.OF UNITS,SAAR)	4
PMI	PURCHASING MANAGERS' INDEX (SA)	1
PMNO	NAPM NEW ORDERS INDEX (PERCENT)	1
PMDEL	NAPM VENDOR DELIVERIES INDEX (PERCENT)	1
PMNV	NAPM INVENTORIES INDEX (PERCENT)	1
A0M008	MFRS' NEW ORDERS, CONSUMER GOODS AND MATERIALS (BILL. CHAIN 1982 $)	5
A0M007	MFRS' NEW ORDERS, DURABLE GOODS INDUSTRIES (BILL. CHAIN 2000 $)	5
A0M027	MFRS' NEW ORDERS, NONDEFENSE CAPITAL GOODS (MIL. CHAIN 1982 $)	5
A1M092	MFRS' UNFILLED ORDERS, DURABLE GOODS INDUS. (BILL. CHAIN 2000 $)	5
A0M070	MANUFACTURING AND TRADE INVENTORIES (BILL. CHAIN 2000 $)	5
A0M077	RATIO, MFG. AND TRADE INVENTORIES TO SALES (BASED ON CHAIN 2000 $)	2
FM1	MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)	6
FM2	MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP(BIL$,	6
FM3	MONEY STOCK: MZM(BIL$,SA)	6
FM2DQ	MONEY SUPPLY - M2 IN 2005 DOLLARS (BCI)	5
FMFBA	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)	6
FMRRA	DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA)	6
FMRNBA	DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL$,SA)	6
FCLNQ	COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1996 DOLLARS (BCI)	6
FCLBMC	NET CHANGE IN BUSINESS LOANS	1
CCINRV	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)	6
A0M095	RATIO, CONSUMER INSTALLMENT CREDIT TO PERSONAL INCOME (PCT.)	2
FSPCOM	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)	5
FSPIN	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)	5
FSDXP	S&P'S COMPOSITE COMMON STOCK: PRICE-DIVIDEND RATIO (%NSA)	5
FSPXE	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA)	5
FYFF	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)	2
CP90	COMMERCIAL PAPER RATE (AC)	2
FYGM3	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)	2
FYGM6	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)	2
FYGT1	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)	2
FYGT5	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)	2
FYGT10	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)	2

Data Code	Variable Name	Format
FYAAAC	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)	2
FYBAAC	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)	2
scp90	CP90-FYFF	1
sfygm3	FYGM3-FYFF	1
sFYGM6	FYGM6-FYFF	1
sFYGT1	FYGT1-FYFF	1
sFYGT5	FYGT5-FYFF	1
sFYGT10	FYGT10-FYFF	1
sFYAAAC	FYAAAC-FYFF	1
sFYBAAC	FYBAAC-FYFF	1
EXRUS	UNITED STATES; EFFECTIVE EXCHANGE RATE (MERM) (INDEX NO.)	5
EXRSW	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)	5
EXRJAN	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)	5
EXRUK	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)	5
EXRCAN	FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)	5
PWFSA	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)	6
PWFCSA	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)	6
PWIMSA	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)	6
PWCMSA	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)	6
PSCCOM	SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100)	6
NFS	NON-FERROUS SCRAP (1982=100)	6
PMCP	NAPM COMMODITY PRICES INDEX (PERCENT)	1
PUNEW	CPI-U: ALL ITEMS (82-84=100,SA)	6
PU83	CPI-U: APPAREL & UPKEEP (82-84=100,SA)	6
PU84	CPI-U: TRANSPORTATION (82-84=100,SA)	6
PU85	CPI-U: MEDICAL CARE (82-84=100,SA)	6
PUC	CPI-U: COMMODITIES (82-84=100,SA)	6
PUCD	CPI-U: DURABLES (82-84=100,SA)	6
PUS	CPI-U: SERVICES (82-84=100,SA)	6
PUXF	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)	6
PUXHS	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)	6
PUXM	CPI-U: ALL ITEMS LESS MIDICAL CARE (82-84=100,SA)	6
PUE	CPI-U: ALL ITEMS LESS ENERGY (82-84=100,SA)	6
GMDC	PCE, IMPL PR DEFL:PCE (1987=100)	6
GMDCD	PCE, IMPL PR DEFL:PCE; DURABLES (1987=100)	6
GMDCN	PCE, IMPL PR DEFL:PCE; NONDURABLES (1996=100)	6
GMDCS	PCE, IMPL PR DEFL:PCE; SERVICES (1987=100)	6
CES275	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NO	6
CES277	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NO	6
CES278	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NO	6
HHSNTN	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)	2

Note: **For BVAR models: 1, 2 = No transformation; 4, 5 and 6 = Log(data) × 100; For FAVAR models: 1 = No transformation; 2 = First-difference of data; 4 = Log(data) × 100; 5. 6: Growth rate of data in percentage.**