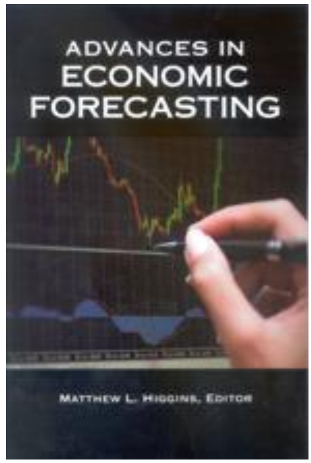
Upjohn Institute Press

Forecasting Regional and Industry-Level Variables: Challenges and Strategies



David E. Rapach
Saint Louis University

Chapter 4 (pp. 51-64) in:
Advances in Economic Forecasting
Matthew L. Higgins, ed.
Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2011
DOI: 10.17848/9780880993937.ch4

Advances in Economic Forecasting

Matthew L. Higgins
Editor

2011

W.E. Upjohn Institute for Employment Research
Kalamazoo, Michigan

Library of Congress Cataloging-in-Publication Data

Advances in economic forecasting / Matthew L. Higgins, editor.
p. cm.
The papers in this volume were presented at the 2009–2010 Werner Sichel Lecture-seminar Series held at Western Michigan University.
Includes bibliographical references and index.
ISBN-13: 978-0-88099-383-8 (pbk. : alk. paper)
ISBN-10: 0-88099-383-9 (pbk. : alk. paper)
ISBN-13: 978-0-88099-384-5 (hardcover : alk. paper)
ISBN-10: 0-88099-384-7 (hardcover : alk. paper)
1. Economic forecasting. I. Higgins, Matthew L.
HB3730.A484 2011
330.01'12—dc23
2011044001

W.E. Upjohn Institute for Employment Research
300 S. Westnedge Avenue
Kalamazoo, Michigan 49007-4686

Cover design by Alcorn Publication Design.
Index prepared by Diane Worden.
Printed in the United States of America.
Printed on recycled paper.

4
Forecasting Regional and Industry-Level Variables

Challenges and Strategies

David E. Rapach
Saint Louis University

Forecasting regional and industry-level (RIL) variables is an important task for a wide variety of economic agents. Policymakers at all levels of government utilize such forecasts, including local and state governments when planning budgets and the Federal Reserve when formulating U.S. monetary policy (e.g., the *Beige Book*). Businesses in the private sector also rely on such forecasts as inputs when taking employment, production, and investment decisions. The recent "Great Recession" highlights the relevance of forecasting RIL variables for policymakers and businesses: revenue reductions make accurate forecasts imperative for planning purposes and the efficient allocation of now-more-limited resources.

Forecasting almost any economic variable is, of course, extremely challenging. Nevertheless, forecasting RIL variables exacerbates typical forecasting difficulties. In particular, there are usually a plethora of potential predictors—global, national, regional, and industry variables—that are relevant for forecasting RIL variables. While theoretical models help to identify key determinants of a given RIL variable, such models are usually highly stylized and thus do not necessarily provide the most appropriate forecasting specifications, especially given the various idiosyncrasies surrounding individual RIL variables. A forecaster thus faces substantial *model uncertainty*. While the forecaster could include all potential predictors in a single forecasting model, such highly parameterized models usually fare very poorly in terms of out-of-sample forecasting, due in no small part to model uncertainty.[1]

Alternatively, the forecaster could preselect a relatively small number of predictors, but this ignores the potentially useful information available in the excluded variables. In this chapter, I outline some tractable approaches for incorporating information from a large number of potential predictors that avoid overly parameterized specifications. Recent research indicates that such approaches are quite beneficial for improving forecasts of RIL variables.

In addition to model uncertainty, *model instability* is a serious concern for forecasting RIL variables. Changes in institutions, public policy, and technology, among many other factors, can precipitate structural breaks that cause the predictive power of individual variables to vary significantly over time. Moreover, it is extremely difficult to predict the occurrence of structural breaks. Similar to model uncertainty, model instability causes highly parameterized models to break down in out-of-sample forecasting, so that a forecaster of RIL variables needs tractable approaches that are reasonably robust to structural breaks. Fortunately, approaches useful for dealing with model uncertainty also appear helpful for mitigating structural instability when forecasting RIL variables.

I outline three approaches—1) general-to-specific modeling with bagging (GETS-bagging), 2) forecast combination, and 3) factor models—for improving forecasts of RIL variables. GETS-bagging and forecast combination are methods for utilizing, in a tractable manner, information from a large set of potential predictors that are reasonably robust to model uncertainty and instability. Factor models focus on potentially strong relationships between RIL and national variables. I provide intuition and guidance on implementing these approaches. In addition, I discuss empirical results from recent research on forecasting RIL variables, highlighting examples pertaining to forecasting employment growth for Michigan and Missouri.

It is important to stress that the present chapter is relatively brief and is not meant as an exhaustive literature survey. Instead, it is intended to introduce the reader to strategies for improving forecasts of RIL variables from the recent literature—strategies designed to address the keen challenges posed by model uncertainty and instability.[2]

FORECASTING STRATEGIES

This section outlines forecasting approaches aimed at improving forecasts of RIL variables. I begin with a general ("kitchen sink") model that serves to illustrate some of the pitfalls that the GETS-bagging, combination, and factor model approaches are designed to avoid.

Kitchen-Sink Model

Consider the following general model specification:

$$(4.1)\qquad \Delta y^h_{k,t+h} = a_k + b_k \Delta y_{k,t} + \sum_{i=1}^{N} c_{k,i} x_{i,t} + e^h_{k,t+h}\ ,$$

where t denotes the time period, $\Delta y^h_{k,t+h} = (1/h)\sum_{j=1}^{h} \Delta y_{k,t+j}$, $\Delta y_{k,t} = y_{k,t} - y_{k,t-1}$, $y_{k,t}$ is in log-levels, and $e^h_{k,t+h}$ is a zero-mean disturbance term. The left-hand side of Equation (4.1) is a cumulative growth rate for the variable of interest that we wish to forecast.[3] The k subscript indicates that $y_{k,t}$ is an RIL variable, where k indexes the region or industry. The $x_{i,t}$ variables ($i = 1, \ldots, N$) on the right-hand side of Equation (4.1) represent N potential predictors of $\Delta y^h_{k,t+h}$, where N can be large. For expositional and notational simplicity, the right-hand side of Equation (4.1) includes only a single lag of $\Delta y_{k,t}$ and each $x_{i,t}$ variable; it is straightforward to allow for additional lags and thus a more general dynamic structure.

Consider forming a forecast of $\Delta y^h_{k,t^*+1}$ using information available through t^* based on the general model given by Equation (4.1):

$$(4.2)\qquad \Delta \hat{y}^h_{k,t^*+h} = \hat{a}_{k,t^*} + \hat{b}_{k,t^*} \Delta y_{k,t^*} + \sum_{i=1}^{N} \hat{c}_{k,i,t^*} x_{i,t^*}\ ,$$

where $\hat{a}_{k,t^*}$, $\hat{b}_{k,t^*}$, and $\hat{c}_{k,i,t^*}$ ($i = 1, \ldots, N$) are ordinary least square (OLS) estimates of the corresponding parameters in Equation (4.1) based on data from the beginning of the sample through t^*. When N is large, a serious drawback to this approach is that it can entail substantial in-sample overfitting, which translates into very poor out-of-sample forecasting performance. Intuitively, a highly parameterized model—a model with many $x_{i,t}$ variables—can deliver a substantial R^2 statistic for the in-sample period, but because of model uncertainty and structural instability, the good fit is specific to the sample and not robust.

Goyal and Welch (2008) and Rapach, Strauss, and Zhou (2010) provide recent examples of the poor forecasting performance of kitchen-sink models in the context of forecasting U.S. stock returns. Many potential predictors of aggregate market returns have been proposed in the finance literature, and different theoretical models emphasize different predictors. Goyal and Welch, as well as and Rapach, Strauss, and Zhou, find that general models with a large number of potential predictors from the literature substantially underperform when measured against the simple random-walk model with respect to U.S. stock returns. This type of result is common in the literature, so one can conclude that very simple models are almost always better than very general models for forecasting purposes. When forecasting RIL variables, one should thus avoid kitchen-sink models.[4]

GETS-Bagging

Pretesting provides a method for paring down Equation (4.1) into a more parsimonious model that includes only the important predictors of $\Delta y^h_{k,t+h}$. This is often referred to as general-to-specific (GETS) modeling. Consider again the problem of forming a forecast of $\Delta y^h_{k,t^*+h}$ using information available through t^*. Instead of including all N of the $x_{i,t}$ variables in the forecasting model, as in Equation (4.2), we first estimate Equation (4.1) and compute the t-statistic associated with each $x_{i,t}$. We then drop any variable from the forecasting model with a t-statistic whose absolute value is below a certain threshold, for example, 1.96 or 1.645. The forecasting model thus becomes a reduced version of Equation (4.2) that contains only the significant predictors. In this way, we attempt to identify a more parsimonious forecasting model that only includes what we deem to be important determinants of $\Delta y^h_{k,t+h}$.

While pretesting reduces the dimension of the forecasting model, the selection of the predictors to include in the forecasting model can be sample-specific, thereby representing in-sample overfitting in another guise. Breiman (1996) introduces the idea of bootstrap aggregating (bagging) as a procedure for stabilizing the pretesting decision rule. In essence, we harness the power of the computer to generate a large number of pseudo samples of observations for $\Delta y^h_{k,t+h}$ and $x_{i,t}$ ($i = 1, \ldots, N$)

using bootstrapping techniques. For each pseudo sample, we apply the decision rule and select the predictors to include in the forecasting model, forming a forecast based on the selected predictors under the pseudo sample. The GETS-bagging forecast is then a simple average of the forecasts corresponding to each of the pseudo samples. Intuitively, the pseudo samples provide new learning sets for the decision rule, thereby reducing the instability of the decision rule and its dependence on a specific sample and improving forecasting performance.[5]

Inoue and Kilian (2008) were the first to employ GETS-bagging in a macroeconomic forecasting context (the U.S. inflation rate). They find that GETS-bagging produces significant forecasting gains relative to a simple autoregressive (AR) time-series model and a general model similar to Equation (4.1), as well as relative to pretesting without bagging. More to the theme of this chapter, Rapach and Strauss (forthcoming) find that GETS-bagging produces consistent and significant out-of-sample gains for forecasting U.S. state-level employment growth. Results for forecasting Michigan and Missouri employment growth are discussed in more detail in the next section.

Forecast Combination

Instead of beginning with a general model, forecast combination takes a weighted average of forecasts generated by a large number of individual models. In the context of macroeconomic forecasting, Stock and Watson (1999, 2003, 2004) have popularized a combination approach that pools information from N individual autoregressive distributed lag (ARDL) models:

$$(4.3) \quad \Delta y^h_{k,t+h} = a_k + b_k \Delta y_{k,t} + c_{k,i} x_{i,t} + e^h_{k,t+h} \quad (i = 1, \ldots, N).$$

Analogous to Equation (4.2), we can form a forecast of $\Delta y^h_{k,t^*+h}$ at t^* for each ARDL based on estimates of the parameters in Equation (4.3) derived from data available through t^*.[6] A combination forecast of $\Delta y^h_{k,t^*+h}$ is then given by

$$(4.4) \quad \Delta \hat{y}^{h,C}_{k,t^*+h} = \sum_{i=1}^{N} \omega_{i,t^*} \Delta \hat{y}^{h,i}_{k,t^*+h},$$

where $\Delta\hat{y}^{h,i}_{k,t^*+h}$ (i = 1, . . . , N) is the forecast of $\Delta y^{h}_{k,t^*+h}$ based on the individual ARDL model with $x_{i,t}$, ω_{i,t^*} (i = 1, . . . , N) is the combining weight corresponding to $\Delta\hat{y}^{h,i}_{k,t^*+h}$, and $\sum_{i=1}^{N}\omega_{i,t^*}=1$. As stressed by Timmermann (2006), the intuition behind forecast combination is the same as that behind portfolio diversification: we reduce forecasting "risk" by averaging across a large number of individual forecasts, rather than by relying on a single forecasting model.

To implement the combination forecast, we need to determine the combining weights. There are a myriad of methods available for doing this, which are nicely surveyed by Timmermann (2006). An interesting result from the literature is that relatively simple schemes typically outperform more elaborate schemes, even though more elaborate schemes are theoretically optimal under certain assumptions. The problem is that model uncertainty and instability frequently render these assumptions inaccurate, limiting the usefulness of theoretically optimal weights in practice.

A simple combining scheme that often works well in practice is equal weighting: $\omega_{i,t^*}=1/N$ for all i. In the context of the general model, Equation (4.1), Rapach, Strauss, and Zhou (2010) show that equal weighting can be viewed as a type of "shrinkage" estimator. Intuitively, shrinkage limits the parameter space and prevents overfitting, thereby improving out-of-sample forecasting performance. While equal weighting often produces very consistent forecasting gains, additional gains can be realized by "tilting" the combining weights toward particular individual forecasts. For example, we could select the combining weights based on the performance of the individual forecasting models over a reasonably long holdout out-of-sample test period. The key, however, is not to overdo it. That is, it is typically best to hew fairly closely to equal weighting; otherwise, we have another manifestation of overfitting, and the forecasts become overly susceptible to model uncertainty and instability.[7]

Rapach and Strauss (forthcoming) find that combination forecasts outperform AR benchmark forecasts of U.S. state-level employment growth for 49 of the 50 individual states for a first-quarter 1990 to fourth-quarter 2010 forecast evaluation period, demonstrating the usefulness of the forecast combination approach for RIL variables. Specific results for Michigan and Missouri employment growth forecasts are presented on pp. 59–61.[8] In another recent application, Rapach and

Strauss (2009) show that combination forecasts improve upon AR benchmark forecasts of real housing price growth for a number of interior states for the period from first-quarter 1995 to fourth-quarter 2006. However, combination forecasts do not outperform the AR benchmark forecasts for a number of coastal states during this period, which could indicate that these coastal states experienced housing price bubbles.

Factor Models

Another potentially useful approach for forecasting RIL variables is factor modeling. If RIL variables have strong links to a national variable, factor models can exploit these links to generate improved forecasts. Consider the following simple factor model:

(4.5) $$\Delta y_{k,t} = \alpha_k + \beta_k f_t + \varepsilon_{k,t} \ ,$$

where f_t is an economy-wide or aggregate factor and $\varepsilon_{k,t}$ is a zero-mean disturbance that may be serially correlated. The coefficient on the factor (β_k) is referred to as the factor "loading" or "exposure." This coefficient captures the strength of the relationship between the RIL variable and the aggregate factor, with a larger β_k indicating a stronger response of $\Delta y_{k,t}$ to fluctuations in f_t. Perhaps the best-known example of a factor model in economics and finance is the canonical capital asset pricing model, where $\Delta y_{k,t}$ is the excess return on a particular stock and f_t represents the excess return on the market portfolio. The return on a stock with large β_k value, or "beta," responds more strongly to changes in the market return and thus has greater systemic risk exposure in the context of the capital asset pricing model.[9]

While f_t can be treated as an unobserved latent variable to be estimated (using, e.g., principal component analysis), f_t frequently coincides with an observable aggregate variable. It is then straightforward to construct a forecast of an RIL variable based on Equation (4.5). Consider, for example, forecasting U.S. state-level employment growth using the following factor model specification:

(4.6) $$\Delta \hat{y}^{h,f}_{k,t^*+h} = \hat{\alpha}_{k,t^*} + \hat{\beta}_{k,t^*} \Delta \hat{y}^{h}_{US,t^*+h} + \hat{\varepsilon}^{h}_{k,t^*+h} \ ,$$

where $\Delta\hat{y}^h_{US,t^*+h}$ is a forecast of aggregate U.S. employment growth; $\hat{\alpha}_{k,t^*}$ and $\hat{\beta}_{k,t^*}$ are OLS estimates of the intercept and slope coefficients, respectively, in a regression of state k employment growth on U.S. employment growth based on data through t^*; and $\hat{\varepsilon}^h_{k,t^*+h}$ is a forecast of the disturbance term in Equation (4.5) that takes into account the possible serial correlation in the disturbance term.[10] The forecast given by Equation (4.6) requires a forecast of U.S. employment growth to plug into the right-hand side. A GETS-bagging or combination forecast of U.S. employment growth is a natural choice.

Building on Owyang, Rapach, and Wall (2009), Rapach and Strauss (forthcoming) forecast U.S. state-level employment growth using Equation (4.6). They show that factor model forecasts outperform AR benchmark model forecasts for the vast majority of states. The forecasting gains are very sizable for a number of states (including Michigan and Missouri, as described in more detail in the next section). There are a few states, however, where the factor model performs much worse than the AR benchmark, so that factor model forecasts appear to offer gains on a somewhat less consistent basis than GETS-bagging and combination forecasts. Rapach et al. (2011) provide another application in the context of forecasting stock returns for industry-sorted portfolios. They find that a conditional version of the popular Fama-French three-factor model (Fama and French 1993) delivers statistically and economically significant out-of-sample gains for forecasting industry returns.

Estimation Window

The discussion thus far has assumed that the parameters of the forecasting model are estimated using data from the beginning of the available sample through the time of forecast formation. If we suspect the existence of substantial structural breaks, at first blush it may seem appropriate to use an estimation window that excludes prebreak data. As shown by Pesaran and Timmermann (2007) and Clark and McCracken (2009), however, it can be optimal to include prebreak data according to a mean-squared-error criterion; this is a manifestation of the classical bias-efficiency trade-off. Furthermore, Pesaran and Timmermann (2007) and Clark and McCracken (2009) show that the theoretically optimal estimation window is a complicated function of the timing and magnitude of structural breaks. A forecaster will not know these things a priori, so they must be estimated from the data. Estimating the timing of breaks is notoriously difficult. Moreover, by estimating these additional parameters, we again run the risk of having an overly parameterized forecasting model that performs poorly out-of-sample. In practice, it is thus often advisable to employ an expanding window (as assumed in the discussion above). Another strategy is to combine forecasts across models estimated using a variety of window sizes, since this approach recognizes that an expanding window is not necessarily optimal but still avoids the overfitting problem associated with trying to estimate the precise timing of structural breaks.

Amalgamating the Approaches

Finally, it is also worth considering amalgamating the GETS-bagging, forecast combination, and factor model approaches. We can straightforwardly accomplish this by taking an average of the GETS-bagging, combination, and factor model forecasts of an RIL variable. Indeed, Rapach and Strauss (forthcoming) find that such an amalgam forecast performs very well with respect to state-level employment growth: it outperforms the AR benchmark forecast for nearly every state, does not produce the outliers of the factor model approach, and delivers larger gains than the three individual approaches for the clear majority of states. Results for Michigan and Missouri are discussed in the next section.

FORECASTING MICHIGAN AND MISSOURI EMPLOYMENT GROWTH

This section reports more detailed results from Rapach and Strauss (forthcoming) on forecasting Michigan and Missouri state employment growth. The quarterly data composing the full sample span the first quarter of 1976 to the fourth quarter of 2010. Employment data are from the Bureau of Labor Statistics (BLS), and annualized employment growth is computed as 400 times the difference in the log levels of employment. As emphasized in this chapter, there are a host of potential predictors of RIL variables. Rapach and Strauss consider 11 potential predictors, which are given in Table 4.1. These predictors are representative of the types of national and regional determinants of state employment growth suggested by economic intuition and more formal models.[11]

Table 4.1 reports forecasting results for the first-quarter 1990–fourth-quarter 2010 forecast evaluation period and a forecast horizon of two quarters ($h = 2$ in the notation of the previous section). The "AR MSFE" row provides the mean squared forecast error (MSFE) for an AR benchmark model. This is a popular benchmark forecasting model that only relies on lagged values of the variable to be forecasted. While it is a seemingly naive time-series model, such simple time-series models are often difficult to beat in practice. The other rows in Table 4.1 report the ratio of the MSFE for the forecasting model specified in the row heading relative to the AR MSFE. A ratio below (above) unity thus indicates that the competing model outperforms (underperforms) the AR benchmark in terms of MSFE.

As seen in Table 4.1, the AR model produces an MSFE of 4.44 percent (2.57 percent) for Michigan (Missouri). The next rows report MSFE ratios for 11 ARDL models, each based on an individual predictor, as in Equation (4.3). Individual ARDL model results are reported to illustrate the difficulties in identifying a priori the most relevant predictors for a given RIL variable. While all 11 predictors appear plausible, they often vary significantly in their forecasting ability. For example, the ARDL model based on real housing price growth generates an MSFE that is 13 percent higher than the AR benchmark for Michigan, so the AR benchmark provides substantially more accurate forecasts. Housing permit growth, in contrast, reduces MSFE by 6 percent relative to the AR benchmark. In general, as emphasized throughout this chapter, model uncertainty and instability make it extremely difficult to determine a priori the most relevant variables for forecasting RIL variables.[12]

The "GETS-bagging," "Forecast combination," and "Factor model" rows in Table 4.1 report results for the forecasting strategies outlined on pages 54–58.[13] Finally, the "Amalgam" row reports results for an amalgam forecast that takes the form of a simple average of the GETS-bagging, combination, and factor-model forecasts (page 59). Table 4.1 shows that the suggested strategies produce MSFE ratios that are always

Table 4.1 Forecasting Results, State-Level Employment Growth, Two-Quarter Horizon, First-Quarter 1990 to Fourth-Quarter 2010 Evaluation Period

Forecasting model	Michigan	Missouri
AR MSFE	4.44	2.57
ARDL models		
State unemployment rate, differences	1.02	1.01
State real income growth	1.15	1.02
State real housing price growth	1.13	1.02
State housing building permit growth	0.94	1.07
U.S. manufacturing hours, differences	0.99	1.01
U.S. unemployment claims, log levels	0.96	0.79
U.S. new consumer good order growth	0.93	0.83
U.S. building permit growth	0.86	0.93
U.S. real stock price growth	0.82	0.88
U.S. real oil price growth	1.09	1.13
Average adjacent state employment growth	1.08	1.02
Suggested strategies		
GETS-bagging	0.91	0.80
Forecast combination	0.91	0.86
Factor model	0.76	0.65
Amalgam	0.78	0.68

NOTE: The AR (autoregressive) MSFE (mean squared forecast error) row reports the MSFE for the AR benchmark model. Other rows report the MSFE ratio for the forecasting model indicated in the row heading relative to the AR benchmark model.
SOURCE: Adapted from Rapach and Strauss (forthcoming).

below unity, so they consistently deliver forecasting gains relative to the AR benchmark.

Among the GETS-bagging, combination, and factor model forecasts, the factor model forecast performs the best for both Michigan and Missouri. The factor model forecast reduces MSFE by 24 percent (35 percent) relative to the AR benchmark for Michigan (Missouri).[14] For both states, the amalgam forecast also performs well: the MSFE reduction for the amalgam forecast relative to the AR benchmark is a very sizable 22 percent (32 percent) for Michigan (Missouri).

Overall, the results in Table 4.1, together with other results from recent research, illustrate the usefulness of the strategies suggested in this chapter for forecasting RIL variables. Of course, it is important not to read too much into these results and overgeneralize them. Forecasters of RIL variables should thus employ thorough back-testing of these strategies for a given application.[15] Nevertheless, the positive results in recent applications are very promising, so the suggested strategies should form an integral part of a forecaster's toolbox for dealing with the model uncertainty and instability inherent in forecasting RIL variables.

Notes

1. There also may simply be an inadequate number of time-series observations to feasibly estimate a model that includes a very large number of potential predictors.
2. For more extensive coverage of some of the topics covered in this chapter, see the volumes edited by Elliott, Granger, and Timmermann (2006) and Rapach and Wohar (2008), as well as the references at the end of this chapter.
3. The disturbance term will be serially correlated when $h > 1$.
4. Indeed, as mentioned in note 1, OLS estimation of the kitchen-sink model may not even be feasible if the timespan is limited relative to the large number of potential variables that exist for RIL variables.
5. See Inoue and Kilian (2008) and Rapach and Strauss (forthcoming) for more detailed expositions of the construction of bagging forecasts.
6. Again, we can include additional lags of the right-hand-side variables in Equation (4.3) to allow for a more general dynamic structure.
7. Hendry and Clements (2004) provide theoretical insight on how forecast combination can improve forecasting in the presence of structural breaks.
8. Also see Rapach and Strauss (2005), who investigate the performance of a large number of combining methods with respect to forecasting Missouri employment growth.
9. Under the capital asset pricing model, the intercept term should actually be zero in Equation (4.5), since it represents the abnormal, risk-adjusted return (or "alpha"), which will be zero in an efficient market.
10. See Rapach and Strauss (forthcoming) for details on the construction of the disturbance term forecast.
11. Rapach and Strauss (forthcoming) provide data sources for the predictors.
12. In addition, in unreported results, the kitchen-sink model performs very poorly for each state.
13. The combining weights in Equation (4.4) for the combination forecasts are selected based on the performance of the individual models over a relatively long holdout out-of-sample period, as discussed on page 56.
14. With respect to the factor model forecast for Michigan, the average estimate of β_k in Equation (4.5) used in the computation of the factor model forecasts is 1.45,

among the largest for individual U.S. states. Michigan employment thus has large "exposure" to national employment cycles, likely due in large part to the automobile industry's strong link to the national business cycle.

15. Even if back-testing provides positive results, as it says in the fine print, past performance is no guarantee of future success.

References

Breiman, Leo. 1996. "Bagging Predictors." *Machine Learning* 24(2): 123–140.

Clark, Todd E., and Michael W. McCracken. 2009. "Improving Forecast Accuracy by Combining Recursive and Rolling Forecasts." *International Economic Review* 50(2): 363–395.

Elliott, Graham, Clive W.J. Granger, and Allan Timmermann, eds. 2006. *Handbook of Economic Forecasting,* Vol. 1. Amsterdam: North-Holland.

Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33(1): 3–56.

Goyal, Amit, and Ivo Welch. 2008. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21(4): 1455–1508.

Hendry, David F., and Michael P. Clements. 2004. "Pooling of Forecasts." *Econometrics Journal* 7(1): 1–31.

Inoue, Atsushi, and Lutz Kilian. 2008. "How Useful Is Bagging in Forecasting Economic Time Series? A Case Study of U.S. CPI Inflation." *Journal of the American Statistical Association* 103(482): 511–522.

Owyang, Michael T., David E. Rapach, and Howard J. Wall. 2009. "States and the Business Cycle." *Journal of Urban Economics* 65(2): 181–194.

Pesaran, M. Hashem, and Allan Timmermann. 2007. "Selection of Estimation Window in the Presence of Breaks." *Journal of Econometrics* 137(1): 134–161.

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

———. 2009. "Differences in Housing Price Forecastability across U.S. States." *International Journal of Forecasting* 25(2): 351–372.

———. Forthcoming. "Forecasting U.S. State-Level Employment Growth: An Amalgamation Approach." *International Journal of Forecasting*.

Rapach, David E., Jack K. Strauss, Jun Tu, and Guofu Zhou. 2011. "Out-of-

Sample Industry Return Predictability: Evidence from a Large Number of Predictors." Working paper. St. Louis, MO: Saint Louis University, Department of Economics.

Rapach, David E., Jack K. Strauss, and Guofu Zhou. 2010. "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy." *Review of Financial Studies* 23(2): 821–862.

Rapach, David E., and Mark E. Wohar, eds. 2008. *Forecasting in the Presence of Structural Breaks and Model Uncertainty*. Frontiers of Economics and Globalization Series, Vol. 3. Bingley, West Yorkshire, UK: Emerald Group Publishing.

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

———. 2003. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature* 41(3): 788–829.

———. 2004. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23(6): 405–430.

Timmermann, Allan. 2006. "Forecast Combinations." In *Handbook of Economic Forecasting*, Vol. 1, Graham Elliott, Clive W.J. Granger, and Allan Timmermann, eds. Amsterdam: North-Holland, pp. 135–196.