# Forecasting national recessions using state-level data

Owyang, Michael T. and Piger, Jeremy and Wall, Howard J.

Federal Reserve Bank of St. Louis, University of Oregon, ISEE, Lindenwood University

10 April 2012

# Forecasting National Recessions Using State Level Data*

Michael T. Owyang† Jeremy M. Piger‡ Howard J. Wall §

first draft: September 19, 2011
this draft: April 10, 2012

**Abstract:** A large literature studies the information contained in national-level economic indicators, such as financial and aggregate economic activity variables, for forecasting U.S. business cycle phases (expansions and recessions.) In this paper, we investigate whether there is additional information regarding business cycle phases contained in subnational measures of economic activity. Using a probit model to predict the NBER expansion and recession classification, we assess the forecasting benefits of adding state-level employment growth to a common list of national-level predictors. As state-level data adds a large number of variables to the model, we employ a Bayesian model averaging procedure to construct forecasts. Based on a variety of forecast evaluation metrics, we find that including state-level employment growth substantially improves short-horizon forecasts of the business cycle phase. The gains in forecast accuracy are concentrated during months of national recession. Posterior inclusion probabilities indicate substantial uncertainty regarding which states belong in the model, highlighting the importance of the Bayesian model averaging approach.

**Keywords:** turning points, probit, covariate selection

**JEL Classification Numbers:** C52, C53, E32, E37

*This paper benefitted from conversations with Graham Elliott and Oscar Jordá. Kristie M. Engemann and Kate Vermann provided research assistance. The views expressed herein do not reflect the official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Research Division, Federal Reserve Bank of St. Louis, (Michael.T.Owyang@stls.frb.org)

‡Department of Economics, University of Oregon, (jpiger@uoregon.edu)

§Department of Economics, Lindenwood University, (hwall@lindenwood.edu)

# 1 Introduction

A traditional view of the U.S. business cycle is that of alternating phases of expansion and recession, where expansions correspond to widespread, persistent growth in economic activity, and recessions consist of widespread, relatively rapid, decline in economic activity.[1] A large literature investigates different aspects of these business cycle phases and documents asymmetries across them. Such work experienced a resurgence following Hamilton (1989), who built a modern statistical model of the alternating phases characterization of the business cycle by describing the latent business cycle phase as following a first-order Markov process that influences the mean growth rate of output.

Of particular interest to academics, policymakers, and practitioners is the prediction of business cycle phases. Hamilton's Markov-switching model characterized phase transitions as random events with fixed probabilities, and, therefore, was not particularly advantageous for forecasting using conditioning variables. More recent work has investigated the notion that business cycle phases can be predicted using a variety of economic and financial time series. These studies typically take the NBER's chronology of the dates of U.S. business cycle phases as data and use discrete choice models (e.g., probit, logit, etc.) to attach probabilities of expansion and recession to current and future periods. The broad conclusion is that economic variables measuring aggregate real activity, such as employment or output growth, provide valuable information about the business cycle phase at short horizons, while only measures of the interest rate term structure are informative at longer horizons.[2] For a recent summary of this literature, see Katayama (2008).

The existing literature has focused primarily on the use of predictors measured at the national level. However, there is reason to believe that variables measured at the subnational level might be useful for predicting the national-level business cycle phase. In a recent paper studying the propagation of state-level recessions, Hamilton and Owyang (2011) find

[1]See Mitchell (1927) and Burns and Mitchell (1946).

[2]See e.g., Estrella (1997), Estrella and Mishkin (1998), and Kauppi and Saikkonen (2008).

that some states occasionally lead the nation into recession. Further, Hernandez-Murillo and Owyang (2006) showed that adding regional employment data can assist in predicting aggregate employment. This suggests that geographically-disaggregated economic indicators may have some predictive power for aggregate business cycle phases. State-level data may be a particularly useful indicator if we believe, as suggested by Temin (1998), that recessions can have a number of root causes. In this case, different regions would enter and exit recessions at different times relative to the average (or the nation).

In this paper, we investigate the predictive ability of state-level economic indicators for business cycle phases. Using the NBER chronology to define business cycle phase dates, we construct a monthly probit model of future business cycle phases. We include both national- and state-level variables to generate forecasts of the national business cycle phase. The national variables we use are those found (or expected) to be good predictors of national business cycle phases by the existing literature. In particular, we focus on interest rates, asset prices, aggregate employment, and aggregate industrial production. To these, we add state-level employment growth to capture the predictive ability of subnational economic activity measures.

For the purposes of forecasting, simply adding state-level data to a model is problematic. It is likely that many states will not be informative about future national business cycle phases at all, or perhaps any, forecast horizons. Further, there is significant collinearity in employment growth across U.S. states. Put together, the naive use of all state-level data will likely lead to an overparameterized model with a high level of estimation uncertainty, which will not bode well for improved forecasting performance. One may reduce, though not solve, these problems by aggregating across states to the regional level. However, this aggregation would potentially average states that contain very different forecasting information.

In this paper, we take a Bayesian model averaging (BMA) approach to incorporate state-level predictors in a forecasting model. In particular, we explicitly incorporate the selection of predictors into the estimation of the model, and average forecasts across different models by constructing the posterior predictive distribution for the future business cycle phase. This approach allows individual states with predictive content for the business cycle phase at a particular horizon to be highlighted in producing forecasts, while pushing out those states that are not informative. Notably, the Bayesian approach to constructing forecasts also incorporates uncertainty regarding model parameters.

Based on a variety of forecast evaluation metrics, we find that including state-level employment growth significantly improves short-horizon forecasts of the NBER business cycle phase over those produced by a model using only national-level data. We document the incremental information content of the state-level data based on the model's in-sample fit over the past 50 years, and also on out-of-sample forecast performance over the past 30 years. We also show that the forecasting improvement comes primarily from improved classification of recession months. To give two examples, for one-month ahead forecasts, 88% of recession months are correctly classified using the model that includes state-level data, as compared to only 64% for a model based on national-level data only. Also, again based on one-month ahead forecasts, the 2008-2009 recession would have been identified by late February 2008 when using state-level data, as opposed to late July 2008 when using only national-level data. Posterior inclusion probabilities indicate significant uncertainty about whether or not some states should be included in the model, which argues for the BMA approach we take to construct forecasts.

The balance of the paper is as follows: Section 2 outlines the empirical model used for forecasting recessions and describes the Bayesian approach to estimation and construction of forecasts. Section 3 describes the national- and state-level data used to estimate the model and evaluate forecasts. Section 4 presents the results for both in-sample model fit and the performance of out-of-sample forecasts. Section 5 summarizes and concludes.

# 2 The Empirical Approach

## 2.1 Model

Define $S_t \in \{0, 1\}$ as a binary random variable that indicates whether month $t$ belongs to an expansion (0) or recession (1) phase. Our objective is to forecast the business cycle phase, $S_t$, based on information available to a forecaster at the end of month $t - h$. This information may include national-, state-, or regional-level variables in any combination and is collected in the $n \times 1$ vector $X_{t-h}$.

Following the bulk of the existing literature, we use a probit model to link $S_t$ to $X_{t-h}$. Here, the probability that $S_t = 1$ is given by:

$$\Pr\left[S_t = 1|\rho_\gamma\right] = \Phi\left(\alpha + X'_{t-h}\beta\right), \tag{1}$$

where the link function, $\Phi(.)$, is the standard normal cumulative density function, $\beta$ is an $n \times 1$ vector of coefficients, and $\rho = [\alpha, \beta']'$ is the $(n+1) \times 1$ vector collecting the model parameters.

The number of potentially relevant forecasting variables available in $X_{t-h}$ may be large. This is especially true with the inclusion of subnational data, as variables are measured repeatedly across regions or states. From a forecasting perspective, this is problematic as it is well established that highly parameterized models tend to have poor out-of-sample forecasting performance. Moreover, because the probit is nonlinear, the marginal change in the predicitve probabilities are functions of the values of all of the included variables. This means that including irrelevant variables could bias the forecasted probabilities. Here, we focus on a modified version of (1), in which not all variables in $X_{t-h}$ need be included in the model. In particular, define $\gamma$ as an $n \times 1$ vector of zeros and ones, with a one indicating that the corresponding variable in $X_{t-h}$ should be included in the model. We rewrite (1) to incorporate this variable selection as follows:

$$\Pr\left[S_t = 1 | \rho_\gamma, \gamma\right] = \Phi\left(\alpha + X'_{\gamma,t-h}\beta_\gamma\right), \tag{2}$$

where $X_{\gamma,t-h}$, $\rho_\gamma$, and $\beta_\gamma$ contain the elements of $X_{t-h}$, $\rho$, and $\beta$ relevant for the variables selected by $\gamma$. As is described in the next subsection, we treat $\gamma$ as unknown, and estimate its value along with the parameters of the model using Bayesian techniques.

## 2.2 Estimation

To estimate the model in (2), we take a Bayesian approach, which has some key advantages for our purposes. For one, uncertainty about which variables should be included in the model – that is uncertainty about $\gamma$ – can be formally incorporated into Bayesian estimation in a straightforward manner. Related to this, the Bayesian framework provides a mechanism, through the posterior predictive density, to obtain forecasts that average over different choices for variable inclusion and the values of unknown parameters.

Bayesian estimation requires priors be placed on the model parameters, $\rho_\gamma$, as well as the covariate selection vector, $\gamma$. We specify diffuse, i.i.d., mean-zero normal distributions for the individual parameters collected in $\rho_\gamma$:

$$p\left(\rho_\gamma\right) = N\left(\mathbf{0}_{k_\gamma+1}, \sigma^2\mathbf{I}_{k_\gamma+1}\right); \quad \sigma^2 = 10, \tag{3}$$

where $k_\gamma = \gamma'\gamma$ is the number of covariates selected by $\gamma$, $\mathbf{0}_{k_\gamma+1}$ is a $k_\gamma + 1 \times 1$ vector of zeros, and $\mathbf{I}_{k_\gamma+1}$ is the $k_\gamma + 1 \times k_\gamma + 1$ identity matrix. For $\gamma$, we specify a multinomial distribution defined across the $2^n$ different possible choices of $\gamma$. Let $N_i = \binom{n}{i}$ be the number of choices of $\gamma$ for which $k_\gamma = i$. The prior probability over $\gamma$ is then:

$$\Pr\left(\gamma\right) \propto \frac{1}{N_{k_\gamma}}. \tag{4}$$

This distribution is flat in two dimensions. First, it assigns equal probability to all choices of $\gamma$ that have the same $k_\gamma$. In other words, versions of (2) with the same number of covariates will receive equal prior probability. Second, the prior assigns equal cumulative probability to groups of choices for $\gamma$ that imply different numbers of covariates. That is, $\Pr(k_\gamma = i) = \Pr(k_\gamma = j)$, $i, j = 0, 1, \cdots, n$.[3]

To implement Bayesian estimation, we employ the Gibbs sampler to obtain draws from the joint posterior distribution, $\pi(\rho_\gamma, \gamma|\mathbf{S})$, where $\mathbf{S} = [S_1, \cdots, S_T]'$ represents the observed data.[4] The Gibbs sampler is facilitated by augmenting the system with a continuous variable $y_t$ that is deterministically related to the observed state variable $S_t$ (Tanner and Wong (1987)). Define $y_t$ as:

$$y_t = \alpha + X'_{\gamma,t-h}\beta_\gamma + u_t, \tag{5}$$

where $u_t \sim$ i.i.d.$N(0,1)$. Given (2), the relationship between $y_t$ and $S_t$ is:

$$S_t = 1 \quad \text{if} \quad y_t \geq 0.$$

The Gibbs sampler is then implemented in two blocks. In the first, $\rho_\gamma$ and $\gamma$ are sampled conditional on $\mathbf{S}$ and the augmented data $\mathbf{y} = [y_1, \cdots, y_T]'$, as a draw from the conditional posterior distribution $\pi(\rho_\gamma, \gamma|\mathbf{y}, \mathbf{S})$. As $\mathbf{y}$ and $\mathbf{S}$ are deterministically related, this distribution simplifies to $\pi(\rho_\gamma, \gamma|\mathbf{y})$. In the second, the augmented data $\mathbf{y}$ is sampled conditional on $\rho_\gamma$, $\gamma$, and $\mathbf{S}$, from the conditional posterior distribution $\pi(\mathbf{y}|\rho_\gamma, \gamma, \mathbf{S})$. We now describe each of these blocks in detail:

[3]Note that this prior does not assign equal probability to all possible choices of $\gamma$. While seemingly attractive as a "flat" prior, an equal weights prior would give substantially different prior weight to the number of variables included. For example, if there are 50 possible variables, the cumulative prior probability of all models with 3 variables would be 16 times the cumulative prior probability of all models with 2 variables.

[4]See, for example, Albert and Chib (1993), Gelfand and Smith (1990), Casella and George (1992), and Carter and Kohn (1994).

**Sampling** $\pi\left(\rho_\gamma, \gamma|\mathbf{y}\right)$

As suggested by Holmes and Held (2006), we jointly sample $\rho_\gamma$ and $\gamma$ from:

$$\pi\left(\rho_\gamma, \gamma|\mathbf{y}\right) = \pi_\rho\left(\rho_\gamma|\gamma, \mathbf{y}\right)\Pr\left(\gamma|\mathbf{y}\right),$$

by employing a Metropolis step. Given a previous draw of $\rho_\gamma$ and $\gamma$, denoted $\left[\rho_\gamma^{[g]}, \gamma^{[g]}\right]$, we obtain a candidate for the covariate selection vector, denoted $\gamma^*$, by sampling a proposal distribution $q\left(\gamma^*|\gamma^{[g]}\right)$. Conditional on $\gamma^*$, we then obtain a candidate for $\rho_\gamma$, denoted $\rho_\gamma^*$, by sampling from the full conditional posterior density $\pi_\rho\left(\rho_\gamma|\gamma^*, \mathbf{y}\right)$.

The proposal distribution $q\left(\gamma^*|\gamma^{[g]}\right)$ is set as follows. Conditional on $\gamma^{[g]}$, the candidate covariate selection vector $\gamma^*$ is drawn with equal probability from the set of vectors that includes $\gamma^{[g]}$ and all other vectors that alter a single element of $\gamma^{[g]}$ (either from 0 to 1 or 1 to 0.) In other words, the candidate covariate selection vector will either select the same covariates as $\gamma^{[g]}$, take away one covariate from $\gamma^{[g]}$, or add one covariate to $\gamma^{[g]}$. One notable property of this proposal distribution is that $q(\gamma^*|\gamma^{[g]})$ will equal $q(\gamma^{[g]}|\gamma^*)$.

The full conditional distribution for $\rho$, $\pi_\rho\left(\rho_\gamma|\gamma, \mathbf{y}\right)$, is derived as follows. Define $\mathbf{X}_{\gamma,t-h} = \left[1, X'_{\gamma,t-h}\right]$, $t = 1, \cdot\cdot\cdot, T$ and let $\mathbf{X}_\gamma$ represent the $T \times k_\gamma$ matrix of stacked $\mathbf{X}_{\gamma,t-h}$. Then, given the prior distribution in (3), the conditional posterior for $\rho_\gamma$ is:

$$\pi_\rho\left(\rho_\gamma|\gamma, \mathbf{y}\right) \sim N\left(\mathbf{m}_\gamma, \mathbf{M}_\gamma\right),$$

where

$$\mathbf{M}_\gamma = \left(\frac{1}{\sigma^2}\mathbf{I}_{k_\gamma} + \mathbf{X}'_\gamma\mathbf{X}_\gamma\right)^{-1},$$

$$\mathbf{m}_\gamma = \mathbf{M}_\gamma\left(\mathbf{X}'_\gamma\mathbf{y}\right).$$

The candidate, $\rho_\gamma^*$ can then be sampled from $N\left(\mathbf{m}_{\gamma^*}, \mathbf{M}_{\gamma^*}\right)$.

The Metropolis step assigns an acceptance probability $A$ to determine whether or not the candidate will be accepted. Given the $g^{th}$ draw $\left[\rho_\gamma^{[g]}, \gamma^{[g]}\right]$, the $(g+1)^{th}$ draw is determined by:

$$\left[\rho_\gamma^{[g+1]}, \gamma^{[g+1]}\right] = \begin{cases} \left[\rho_\gamma^{*}, \gamma^{*}\right] & \text{with probability} \quad A \\ \left[\rho_\gamma^{[g]}, \gamma^{[g]}\right] & \text{with probability} \quad 1-A \end{cases},$$

where,

$$A = \min\left\{1, \frac{\Pr(\gamma^*|\mathbf{y})\pi_\rho(\rho_\gamma^*|\gamma^*, \mathbf{y})q(\gamma^{[g]}|\gamma^*)\pi_\rho(\rho_\gamma^{[g]}|\gamma^{[g]}, \mathbf{y})}{\Pr(\gamma^{[g]}|\mathbf{y})\pi_\rho(\rho_\gamma^{[g]}|\gamma^{[g]}, \mathbf{y})q(\gamma^*|\gamma^{[g]})\pi_\rho(\rho_\gamma^*|\gamma^*, \mathbf{y})}\right\},$$

which, given the symmetry of $q(\cdot|\cdot)$, simplifies to

$$A = \min\left\{1, \frac{\Pr(\gamma^*|\mathbf{y})}{\Pr(\gamma^{[g]}|\mathbf{y})}\right\}.$$

From Bayes' Rule:

$$\Pr(\gamma|\mathbf{y}) \propto f(\mathbf{y}|\gamma)\Pr(\gamma),$$

where $f(\mathbf{y}|\gamma)$ is the marginal likelihood for the augmented data, $\mathbf{y}$, conditional on the choice of variables $\gamma$, and $\Pr(\gamma)$ is the prior distribution over $\gamma$. We can then rewrite the acceptance probability as:

$$A = \min\left\{1, \frac{f(\mathbf{y}|\gamma^*)\Pr(\gamma^*)}{f(\mathbf{y}|\gamma^{[g]})\Pr(\gamma^{[g]})}\right\}.$$

To compute $A$, we must compute $f(\mathbf{y}|\gamma)$. Given the prior distribution in (3), this is available analytically, as:

$$f(\mathbf{y}|\gamma) \sim N(0, \Sigma_\gamma)$$

where $\Sigma_\gamma = I_T + \sigma^2 \mathbf{X}_\gamma \mathbf{X}_\gamma'$. Using the equation for the multivariate normal we then have:

$$A = \min\left\{1, \frac{\left|\Sigma_{\gamma^{[g]}}\right|^{0.5} \exp(-0.5\mathbf{y}'\left(\Sigma_{\gamma^*}\right)^{-1}\mathbf{y})\Pr\left(\gamma^*\right)}{\left|\Sigma_{\gamma^*}\right|^{0.5} \exp(-0.5\mathbf{y}'\left(\Sigma_{\gamma^{[g]}}\right)^{-1}\mathbf{y})\Pr\left(\gamma^{[g]}\right)}\right\}. \tag{6}$$

**Sampling** $\pi\left(\mathbf{y}|\rho_\gamma, \gamma, \mathbf{S}\right)$

Conditional on $\rho_\gamma$ and $\gamma$, we can draw $y_t$ from a normal distribution with mean $\delta_{\gamma,t}$ and unit variance, where

$$\delta_{\gamma,t} = \alpha + X'_{\gamma,t-h}\beta_\gamma.$$

However, the target distribution also conditions on $\mathbf{S}$, which adds the requirement that the sign of $y_t$ must match the realization of $\mathbf{S}$. In this case, $y_t$ can be drawn from the truncated normal density:

$$y_t \sim \begin{cases} N\left(\delta_{\gamma,t}, 1\right) I_{[y_t \geq 0]} & \text{if} \quad S_t = 1 \\ N\left(\delta_{\gamma,t}, 1\right) I_{[y_t < 0]} & \text{if} \quad S_t = 0 \end{cases},$$

where the indicator $I_{[.]}$ reflects the direction of the truncation. This can be repeated for $t = 1, \cdots, T$ to obtain a draw from $\mathbf{y}$.

Given arbitrary starting values for $\rho_\gamma$ and $\gamma$, the above two sampling steps can be iterated to obtain draws $\left[\rho_\gamma^{[g]}, \gamma^{[g]}\right]$, for $g = 1, \cdots, G$. Following a suitably large number of initialization samples, these draws will be from the joint posterior distribution of interest, $\pi\left(\rho_\gamma, \gamma|\mathbf{S}\right)$. In all estimations reported below, we sample 20,000 initialization draws to achieve convergence. Results are then based on an additional set of $G = 20,000$ draws. For all estimation results presented, we verified the adequacy of the initialization period by running the Gibbs Sampler for two dispersed sets of starting values.

## 2.3 Construction of Forecasts and Forecast Evaluation

Denote an alternative realization of the phase indicator variable as $S_t^*$. This alternative realization may be for a time period for which we already have an observation, that is from $t = 1, \cdots, T$, or for a future period. To predict an alternative realization, we require the posterior predictive distribution:

$$\Pr\left[S_t^* = 1|\mathbf{S}\right]. \tag{7}$$

We can simulate from (7) as follows. The posterior predictive distribution is factored as:

$$\Pr\left[S_t^* = 1|\mathbf{S}\right] = \int_{\rho_\gamma,\gamma} \Pr\left[S_t^* = 1|\rho_\gamma, \gamma, \mathbf{S}\right] \pi\left(\rho_\gamma, \gamma|\mathbf{S}\right). \tag{8}$$

Equation (8) suggests a Monte Carlo integration approach to calculate the posterior predictive distribution. Specifically, given a draw $\left[\rho_\gamma^{[g]}, \gamma^{[g]}\right]$ from $\pi\left(\rho_\gamma, \gamma|\mathbf{S}\right)$ we simulate a value of $S_t^*$, denoted $S_t^{*[g]}$, from:

$$\Pr\left[S_t^* = 1|\rho_\gamma, \gamma, \mathbf{S}\right] = \Pr\left[S_t^* = 1|\rho_\gamma, \gamma, \right] = \Phi\left(\alpha + X'_{\gamma,t-h}\beta_\gamma\right),$$

where the validity of the first equality sign comes from the fact that, given the model in (2) and the model parameters, the observed data $\mathbf{S}$ are not informative for the distribution of $S_t^*$. This simulated value of $S_t^{*[g]}$ is a draw from the posterior predictive distribution, $\Pr\left[S_t^*|\mathbf{S}\right]$, and it follows that:

$$\lim_{G\to\infty} \frac{1}{G}\sum_{g=1}^{G}\left(S_t^{*[g]}\right) = \Pr[S_t^* = 1|\mathbf{S}], \quad t = 1, \cdots, T. \tag{9}$$

Thus, we can construct a simulation consistent estimate of the posterior predictive density $\Pr[S_t^* = 1|\mathbf{S}]$, which will serve as our (point) forecast of $S_t$. In the following, we refer to this forecast as $\widehat{S}_t$. It is worth emphasizing that $\widehat{S}_t$ is not conditional on model parameters or the choice of which variables to include in the model. These sources of uncertainty have been integrated, over their respective posterior distributions, out of the prediction. Note that the integration over the posterior distribution for $\gamma$ gives $\widehat{S}_t$ the interpretation of a Bayesian model averaged prediction.

To evaluate $\widehat{S}_t$, we consider several forecast evaluation metrics that are standard in the existing literature for the evaluation of probability forecasts. The first is the correspondence, defined as the proportion of months for which $\widehat{S}_t$ correctly indicates the NBER business cycle phase. We use 0.5 as the threshold between expansion and recession, and define $I_{\left[\widehat{S}_t \geq 0.5\right]}$ as an indicator function denoting the predicted business cycle phase. The correspondence (CSP) is then given by:

$$CSP = \frac{1}{\tau_2 - \tau_1} \sum_{t=\tau_1+1}^{\tau_2} \left( I_{\left[\widehat{S}_t \geq 0.5\right]} S_t + \left(1 - I_{\left[\widehat{S}_t \geq 0.5\right]}\right)(1 - S_t) \right),$$

where $\tau_1$ and $\tau_2$ are chosen to cover the period over which $\widehat{S}_t$ is being evaluated. Lower values of the CSP indicate worse forecast performance. The second is the Brier (1950) quadratic probability score (QPS), which is a probability analog of mean squared error:

$$QPS = \frac{2}{\tau_2 - \tau_1} \sum_{t=\tau_1+1}^{\tau_2} \left( \widehat{S}_t - S_t \right)^2.$$

The QPS ranges from 0 to 2, with 0 indicating perfect forecast accuracy. Lastly, we compute the log probability score (LPS):

$$LPS = -\frac{1}{\tau_2 - \tau_1} \sum_{t=\tau_1+1}^{\tau_2} \left( S_t \ln\left(\widehat{S}_t\right) - (1 - S_t) \ln\left(1 - \widehat{S}_t\right) \right).$$

The LPS ranges from 0 to $\infty$, with 0 indicating perfect forecast accuracy. The LPS penalizes large errors more heavily than does the QPS.

Each of the above metrics assumes a symmetric loss function across errors made in categorizing recessions as expansions (false negatives) vs. expansions as recessions (false positives). However, it is reasonable to think that certain agents, such as policymakers, may have a loss function that is asymmetric across these different types of errors. To evaluate the performance of $\widehat{S}_t$ in recessions vs. expansions, we also compute the CSP, QPS, and LPS separately for the expansion and recession months in an evaluation period.

# 3 Data

Our predictor variables consist of both national- and state-level variables, all of which are sampled at the monthly frequency. For the national-level variables, we include a measure of the term spread, the federal funds rate, and the return on the S&P 500 stock market index. Each of these variables have been shown to help predict recessions at various horizons in the existing literature.[5] We also include two direct measure of aggregate economic activity, namely payroll employment growth and industrial production growth. Existing studies, such as Estrella and Mishkin (1998), have shown that economic activity measures have some power to forecast recessions at short horizons.

In addition to this standard set of national-level variables, we also include state-level payroll employment growth to capture state-level economic activity. We choose payroll employment growth as the measure of state-level economic activity for two reasons. First, we are interested in relatively high frequency monitoring of business cycle phases. Payroll employment is the broadest measure of state-level economic activity that is sampled at a monthly frequency.[6] Second, as compared to other monthly measures of state-level activity, such as retail sales, payroll employment is released quickly, roughly three weeks following the end of the month. This timeliness makes payroll employment attractive for forecasting.[7]

[5]See e.g., Estrella and Mishkin (1998), Wright (2006), and King et al. (2007).

[6]The most comprehensive measure of state-level economic activity, Gross State Product, is only released annually.

[7]The term spread is measured as the difference between the monthly averages of the 10-year Treasury bond and the 3-month Treasury bill. The Federal Funds rate is measured as its monthly average value. The S&P 500 return is the three month growth rate of the S&P 500 index. Industrial production and national- and state-level payroll employment growth are the three month growth rates of the underlying levels of these variables. In each case, the data were seasonally-adjusted.

For the dependent variable in our estimation, $S_t$, we require a monthly measure of the national business cycle phase. We follow the standard practice in the recession forecasting literature of using the chronology of recessions and expansion dates provided by the NBER's Business Cycle Dating Committee.

All data series were collected over the period from January 1960 to June 2011. After constructing growth rates and adjusting for the maximum forecast horizon considered, the sample period for $S_t$ covers from August 1960 to June 2011. Over this period there are eight NBER defined recessions, and 15% of the monthly observations are recession months.

In our primary analysis, we use variables collected as of the July 2011 vintage. Thus, for variables that are revised, which is the case for the economic activity measures in our sample, we use ex-post revised data in our out-of-sample forecasting experiments rather than the vintage of data that would have been available to a forecaster in real time. We make this choice due to difficulties with obtaining long histories for state-level payroll employment at a substantial number of vintages over our out-of-sample forecasting period. However, as a robustness check, we additionally report results of an out-of-sample forecasting experiment over a shorter time period for which we were able to obtain real-time data.

# 4 Results

## 4.1 In-Sample Predictions

We begin by presenting results for in-sample predictions over the full sample period, where $\widehat{S}_t$ is computed over the same sample period on which we perform estimation, August 1960 to June 2011. We consider four alternative forecast horizons, consisting of $h = 0, 1, 2, 3$ months ahead. In all cases, we assume that the information used to predict $S_t$ consists of the information available at the end of month $t - h$. For the financial variables in our data set, this includes the values of these variables measured for month $t - h$. For each of the real-activity measures in the data set, both at the national- and state-level, this includes the values of these variables measured for month $t-h-1$. As an example, a one-month-ahead forecast in our context refers to a prediction of $S_t$ formed using financial variables measured for month $t-1$ and economic activity variables measured for month $t-2$. It is worth highlighting that the case of $h=0$ corresponds to a prediction of the business cycle phase in month $t$, formed using data available at the end of month $t$, some of which is data measured for month $t$. Such "nowcasts" are of substantial interest, since definitive classification of business cycle phases, particularly around turning points, are generally only available with a substantial lag.[8]

Our primary interest is on forecasting improvements achieved through the addition of state-level predictors. To assess these improvements, we set a baseline model in which $X_{t-h}$ includes only the national-level variables in our predictor set. We then compare the performance of this model to an alternative model in which $X_{t-h}$ contains both national- and state-level data.

Table 1 reports the in-sample forecast evaluation metrics for the baseline model and the model including state-level data for each of the various forecast horizons. These metrics tell a consistent story: including state-level data substantially improves the in-sample fit of the probit recession prediction model. The inclusion of state-level data lowers the QPS and LPS by 25%-70% compared to the model with only national-level data. The CSP improvements are from 0.02-0.05 percentage points, indicating that between 12-30 more months over the sample were correctly classified by the model that includes state-level data. Looking across forecast horizons, the improvements are largest at the $h=0$ and $h=1$ horizons, although they remain substantial at longer horizons as well.

We next investigate whether the improvements generated by the inclusion of state-level data are symmetric across expansion and recession months. Table 2 presents the in-sample forecast evaluation metrics computed separately for expansion and recession months. Com-

[8] The NBER has historically announced turning points with a lag of between 5-19 months. Statistical models using only coincident data improve on the NBER's timeliness considerably, but still identify turning points only after several months have passed, in part due to data reporting lags (see, e.g. Chauvet and Piger (2008).)

paring across phases, it is apparent from both the baseline and extended model that the QPS and LPS are substantially lower, and the CSP substantially higher, for expansion months vs. recession months. This suggests that correct classification using the probit model is less difficult for expansion months than recession months.[9] For expansion months, the inclusion of state-level data lowers the QPS and LPS by 20%-70%. The improvements in the CSP statistic are less impressive, ranging from 0 ($h = 3$) to 2 ($h = 0$) percentage points, which correspond to between 0-10 more months, out of the 521 expansion months in the sample, being correctly classified through the use of state-level data. For recession months, the QPS and LPS reductions are again large, ranging between 30% and 70%. The CSP improvements are more striking for recession months, ranging from 12 ($h = 2$) to 22 ($h = 1$) percentage points. This corresponds to 11-20 of the 90 recession months in the sample.

Figure 1 shows plots of $\widehat{S}_t$ over the estimation sample period, along with recession shading determined by the NBER recession chronology. These figures visually confirm the results of the evaluation metrics. For all forecast horizons, and particularly for $h = 0$ and $h = 1$, conditioning on both national- and state-level data improves the delineation of the NBER expansion and recession phases over the model that conditions on national-level data alone.

Given the improvements in the in-sample fit generated by including state-level data, we next ask which states provide this improvement. Table 3 reports the posterior inclusion probabilities for the model that includes both national- and state-level data. These inclusion probabilities measure the posterior probability that a particular variable is included in the true model, and are computed as the proportion of the $G$ samples from the Gibbs sampler for which $\gamma^{[g]}$ includes that variable. The table reports these inclusion probabilities for each of the national-level variables, and for all state-level variables that achieved at least 50%

[9] This is not surprising given that expansion months make up a much larger portion of the sample then recession months. For example, suppose a sample period is made up of 90% expansion months and 10% expansion months. Then a simple probit model with no covariates would predict expansion with probability 0.9 and recession with probability 0.1. Conditional on being in an expansion month, the probit model would then have a 90% successful prediction rate, but only a 10% successful prediction rate conditional on being in a recession month. Expansion months are also more persistent than recession months, which generates additional predictability during periods of expansion.

inclusion probability for one or more forecast horizons.

Beginning with the national-level variables, both the federal funds rate and S&P 500 return have inclusion probabilities of 100% for all forecast horizons considered. Consistent with the existing literature, the term spread variable becomes generally more important as the forecast horizon increases, obtaining 100% inclusion probability by the $h = 3$ month horizon. Also consistent with the existing literature, industrial production has some predictive ability at short horizons, but has a very low inclusion probability by $h = 3$ months. Finally, aggregate employment growth has a posterior inclusion probability that is above 50% in only one case, that of $h = 0$. For the baseline model with only national-level predictors, both aggregate industrial production and employment growth are assigned very high inclusion probabilities across all horizons considered (not reported). Thus, the inclusion of state-level variables in the information set diminishes the importance of these aggregate variables.

Turning to the state-level variables, there are six states with high inclusion probabilities across a majority of forecast horizons, namely California, Florida, Illinois, Nebraska, Pennsylvania and Washington. These states tend to cluster by population, with four being in the top six most populous states as measured by the 2010 U.S. Census. However, there are also a number of high population states, notably Texas, New York, and Michigan that never receive high inclusion probabilities at any forecast horizon. This demonstrates the value added of considering subnational data along with variable selection, in that it allows those portions of the national employment data that are less predictive of the future business cycle phase to be be downweighted in a formal statistical way.

Variation in the importance of states for forecasting national recessions is not surprising. Owyang et al. (2005) and Hamilton and Owyang (2011) show that state-level business cycles are often out of phase with each other and with the national-level cycle. Also, as argued by Temin (1998), different national recession experiences likely arose from different root causes. For example, some recessions may begin because of weakness in the manufacturing sector while others may begin due to uncertainty in financial markets. If true, it is then reasonable to assume that which states will be leading indicators of a national recession will depend on the type of recession. Further, the timing of recessions with different causes may vary depending on how long the shocks take to propagate across the country. In this sense, we find it intuitive that different states would have more explanatory power at different horizons.

The bottom panel of Table 3 shows the mean of the posterior distribution for $k_\gamma$, the number of variables selected by covariate selection vector $\gamma$. This is equivalent to the average number of variables included in the model in (2) across the $G$ draws from the Gibbs sampler. Note that this average is substantially higher than would be obtained by simply summing up the number of individual states with high inclusion probabilities for a particular forecast horizon. This suggests the presence of a substantial amount of uncertainty about which state-level employment growth variables belong in the model, and argues for the BMA approach used here vs. simply conditioning on a particular model.

## 4.2 Out-of-Sample Predictions

The previous section presented evidence that including state-level data in addition to national-level data substantially improves the in-sample fit of a probit model for predicting U.S. business cycle phases, particularly during recession months. Of course, it is well appreciated that improved in-sample fit does not guarantee improved out-of-sample forecast performance. Thus, in this section we evaluate the out-of-sample forecast performance of the probit model augmented with state-level data.

To assess out-of-sample performance, we construct a series of out-of-sample forecasts computed recursively. Beginning with an initial estimation period of August 1960 to December 1978, we form forecasts at horizons $h = 0, 1, 2, 3$ over the period January 1979 to June 2011.[10] After each out-of-sample forecast is produced, the estimation sample is extended by one month, and the model re-estimated. In terms of business cycle episodes, the out-of-sample period includes 5 NBER-defined recessions, accounting for approximately 15% of the

[10] The first out-of-sample forecast for $h = 0$ and $h = 1$ is January 1979, for $h = 2$ is February 1979, and for $h = 3$ is March 1979. The last out-of-sample forecast for all horizons is June 2011.

390 months over this period.

Table 4 presents the forecast evaluation metrics for both the baseline model that uses only national-level predictors, and the extended model that uses both national- and state-level predictors. These metrics suggest that the out-of-sample forecast performance of the model that includes state-level data is better than the baseline model for short horizon forecasts, namely $h = 0$ and $h = 1$. At these horizons, the CSP is higher, and the QPS lower for the model including state-level data for all forecast horizons, while the LPS is lower for $h = 0$. The forecast improvements as measured by several metrics are substantial. As one example, the CSP is 4 percentage points higher for the model including state-level data when $h = 1$, meaning that approximately 16 more months were correctly classified by the model that includes state-level data. For longer horizons, the inclusion of state-level data appears less helpful. For $h = 2$ and $h = 3$, the forecast metrics generally show a deterioration for the model that includes state-level data.

Table 5 presents the forecast evaluation metrics computed separately for expansion vs. recession months in the out-of-sample period. These results demonstrate that the forecast improvements generated by the inclusion of state-level data in the out-of-sample period are concentrated in recession months. In particular, the forecast evaluation metrics computed for expansions are generally similar for the baseline model and the model that includes state-level data, with which model has better performance differing across horizon and forecast evaluation metric. However, when we focus on recession months, there is a clear benefit from incorporating state-level data for short-horizon forecasts. For $h = 0$ and $h = 1$, the QPS are reduced by 50% to 60% and the LPS by 12% to 60% during recession months. The CSP improvements are approximately 25 percentage points at these horizons, meaning that a quarter of the recession months over the sample period are correctly classified by the model that includes state-level data, but not by the model that includes only national-level data. In terms of absolute performance, the CSP suggests that the model including state-level data correctly classifies recession months over the out-of-sample period around 90% of the time at the $h = 0$ and $h = 1$ horizons. There are again no clear improvements from the addition of state-level data at longer horizons.

To provide an example from a specific recession, Table 6 presents the out-of-sample forecasts, $\widehat{S}_t$, for the $h = 1$ case around the 2008-2009 recession. Beginning with the model that includes only national-level variables, $\widehat{S}_t$ does not cross 50% probability of recession until August 2008, eight months following the beginning of the NBER-defined recession. For the $h = 1$ horizon, this forecast would have been available at the end of July 2008. This is consistent with the considerable uncertainty that persisted well into 2008 about whether the economy had entered a recession phase. For example, the NBER did not announce the December 2007 peak until December 1, 2008. Also, as discussed in Hamilton (2011), statistical models designed to track business cycle turning points using national-level data did not send a definitive signal that the recession had begun until mid-to-late 2008. However, Table 6 also reveals that incorporating state-level data would have provided a much quicker signal of the beginning of this recession. Specifically, $\widehat{S}_t$ moved above 50% probability of recession for March 2008, where this forecast would have been available as of the end of February 2008, an impressive five month improvement over the model using only national-level data. Notably, both models produce accurate one-month ahead forecasts of the end of the 2008-2009 recession.

We next evaluate which state-level variables are providing the out-of-sample forecast improvements. Table 7 reports the posterior inclusion probabilities for the model that includes both national- and state-level data, averaged over the recursive estimations conducted to construct the out-of-sample forecasts. The table provides these inclusion probabilities for all of the national-level variables, and for all state-level variables that achieved greater than 50% inclusion probability for at least one forecast horizon. For the national-level variables, the S&P 500 return is a robust predictor across all forecast horizons, with average inclusion probabilities close to 100%. The Federal Funds rate also has average inclusion probabilities above 50% for several forecast horizons, but these inclusion probabilities are significantly lower than they were over the full sample period, and are not above 50% for the $h = 3$ horizon. As was the case in the full sample period, the term spread becomes a more robust predictor as the forecast horizon lengthens. Finally, neither aggregate employment or aggregate industrial production growth have average inclusion probabilities above 50% for most forecast horizons, the single exception being industrial production growth when $h = 0$. These variables have very high inclusion probabilities when state-level data is not included (not reported), implying the importance of the aggregate level variables is substantially diminished by the inclusion of state-level data.

For the state-level variables, we focus on the $h = 0$ and $h = 1$ forecast horizons, where the forecast improvements from the addition of state level data were concentrated. There are five states with average inclusion probabilities above 50% for the $h = 0$ horizon, namely California, Florida, Illinois, Iowa and Pennsylvania. For the $h = 1$ horizon there are also five such states, in this case California, Connecticut, Illinois, Nebraska and Pennsylvania. There are also a large number of state-level variables that, while not breaking the 50% inclusion probability barrier, have inclusion probabilities substantially greater than 0%. This is demonstrated by the average number of variables included in the forecasting model, which at roughly 17 and 16 for the $h = 0$ and $h = 1$ horizons, is significantly above the sum of only those variables with inclusion probabilities above 50%. This can also be seen in the maps provided in Figure 2, which groups states by ranges of inclusion probability. The maps show that while there are few states with very high average inclusion probabilities (darker shading), there are many states with average inclusion probabilities in the 20%-60% range, meaning a large number of states influence the BMA forecast. These probabilities also indicate significant uncertainty regarding exactly which state-level variables should be included in the model. This highlights the potential importance of the BMA approach we take to select predictors and incorporate uncertainty about this selection.

Given this potential importance, we next present results meant to evaluate whether the BMA predictor selection algorithm is a significant factor for the out-of-sample forecast improvements generated with the addition of state-level data. Specifically, Table 8 reports the forecast evaluation metrics for out-of-sample forecasts produced from a model in which all national- and state-level variables are always included. As the forecasting improvements from adding state-level data were concentrated in short-horizon forecasts of recession months, we focus on the forecast evaluation metrics computed for nowcasts ($h = 0$) and one-month-ahead forecasts ($h = 1$) of recession months over the out-of-sample period. By comparing Table 8 to the bottom panel of Table 5, we can gauge the value added of using the BMA predictor selection algorithm to construct forecasts, vs. simply including all possible variables. Indeed, this comparison shows a deterioration in the out-of-sample forecast performance from conditioning on a model that includes all possible variables rather than using BMA. However, with the exception of the LPS for $h = 1$, it is also the case that the model with all variables included is still preferred to the model that doesn't include state-level data.

Finally, as was discussed above, our out-of-sample forecasts are constructed using ex-post revised data for the predictors taken from the July 2011 vintage for each series. In Table 9 we evaluate the robustness of the out-of-sample forecasting results when we instead use "real-time" data of vintages that would have been available to a forecaster in real time. Due to difficulties with obtaining long histories for state-level payroll employment at a substantial number of vintages, we are able to construct out-of-sample forecasts with real-time data over a shorter out-of-sample period running from July 2007 to June 2011, a period that includes the most recent NBER-defined recession. As forecast improvements from the addition of state-level data were primarily at short horizons, we focus on one month ahead forecasts.

Table 9 demonstrates that our primary conclusions from the longer out-of-sample period using ex-post data are confirmed for the shorter out-of-sample period using real-time data. In particular, there is a general improvement in the forecast evaluation metrics computed for recession months from the addition of state-level data.[11] As an example, the CSP is

[11] The log probability score does indicate a deterioration in the one-month ahead forecast performance of recession months from adding state-level data. This is due to a single recessionary month, January 2008, where the model including state-level data assigns a probability of recession very close to zero. The log probability score severely penalizes such extreme forecast misses, as it will include the logarithm of an

17 percentage points higher when state-level data is included, which corresponds to roughly 3 more recession months during the 2008-2009 recession being correctly classified. Also, as before, there is no apparent improvement from the addition of state-level data for one-month ahead predictions of expansion phases.

# 5 Conclusion

A large literature has investigated the predictive content of variables measured at the national level, such as aggregate employment and output growth, for forecasting U.S. business cycle phases (expansions and recessions.) Motivated by recent studies showing differences in the timing of business cycle phases in nationally aggregated data from those for geographically disaggregated data, we investigate the information contained in state-level employment growth for forecasting national business cycle phases. We use as a baseline a probit model to explain NBER-defined business cycle phases, where the conditioning information consists of national-level economic activity and financial variables. We then add to this model state-level employment growth. To avoid issues associated with overparameterization of forecasting models, we use a Bayesian model averaging procedure to construct forecasts.

Using a variety of forecast evaluation metrics, we find that adding state-level employment growth improves short-horizon forecasts of the NBER business cycle phase over a model that uses data measured at the national-level only. The gains in forecasting accuracy are concentrated during months of recession, and are substantial. Posterior inclusion probabilities indicate substantial uncertainty regarding which states belong in the model, highlighting the importance of the Bayesian model averaging approach.

argument close to zero. However, as the baseline model only assigns a probability of recession for January 2008 of 8%, it is questionable whether this severe relative penalty is reflective of significant differences in forecast performance.

# References


Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association 88*(422), 669–679.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review 75*(January), 1–3.

Burns, A. F. and W. C. Mitchell (1946). *Measuring business cycles.* New York: National Bureau of Economic Research. ID: 169122.

Carter, C. K. and R. Kohn (1994). On gibbs sampling for state space models. *Biometrika 81*(3), 541–553. ID: 476895942.

Casella, G. and E. I. George (1992). Explaining the gibbs sampler. *American Statistician 46*(3), 167–174. ID: 481283991.

Chauvet, M. and J. Piger (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics 26*(1), 42–49.

Estrella, A. (1997). A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics 16*(2), 198–205.

Estrella, A. and F. S. Mishkin (1998). Predicting u.s. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics 80*(1), 45–61.

Gelfand, A. E. and A. F. M. Smith (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association 85*(410), 398–409.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*(2), 357–384. ID: 480697643.

Hamilton, J. D. (2011). Calling recessions in real time. *International Journal of Forecasting 27*(4), 1006–1026.

Hamilton, J. D. and M. T. Owyang (2011). The propagation of regional recessions. NBER working paper no. 16657.

Hernandez-Murillo, R. and M. T. Owyang (2006). The information content of regional employment data for forecasting aggregate conditions. *Economic Letters 90*(3), 335–339.

Holmes, C. C. and L. Held (2006). Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis 1*(1), 145–168.

Katayama, M. (2008). Improving recession probability forecasts in the u.s. economy. Working Paper.

Kauppi, H. and P. Saikkonen (2008). Predicting u.s. recessions with dynamic binary response models. *The Review of Economics and Statistics 90*(4), 777–791.

King, T. B., A. T. Levin, and R. Perli (2007). Financial market perceptions of recession risk. Federal Reserve Board Finance and Economics Discussion Series working paper no. 2007-57.

Mitchell, W. C. (1927). *Business Cycles: The Problem and its Setting*. New York: National Bureau of Economic Research.

Owyang, M. T., J. Piger, and H. J. Wall (2005). Business cycle phases in u.s. states. *Review of Economics and Statistics 87*(4), 604–616.

Tanner, M. A. and W. H. Wong (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association 82*(398), 528–540.

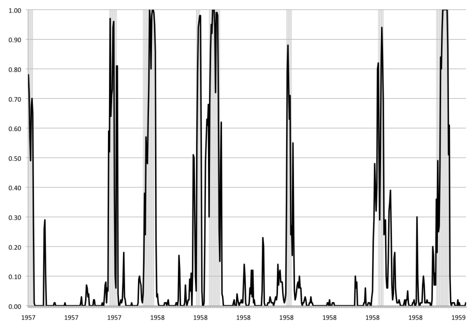
Temin, P. (1998). The causes of american business cycles: An essay in economic historiography. In J. C. Fuhrer and S. Schuh (Eds.), *Beyond Shocks: What Causes Business Cycles?*, Conference Series No. 42, Boston, pp. 37–59. Federal Reserve Bank of Boston.

Wright, J. H. (2006). The yield curve and predicting recessions. Federal Reserve Board Finance and Economics Discussion Series working paper no. 2006-07.

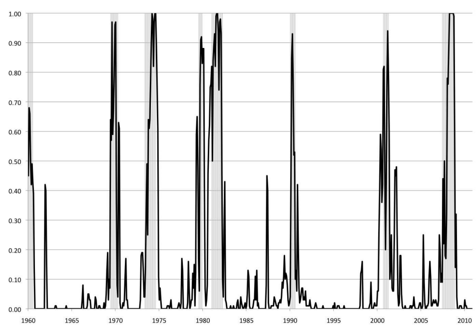
## Figure 1
## In-Sample Recession Predictions

### National-Level Predictors

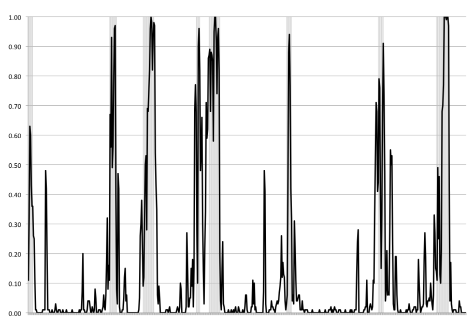
*h*=0



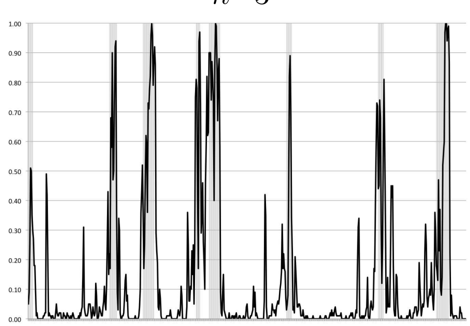
*h*=1



*h*=2

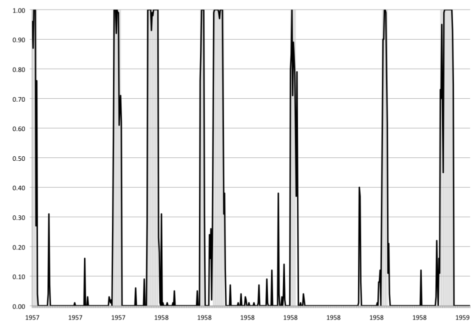


*h*=3

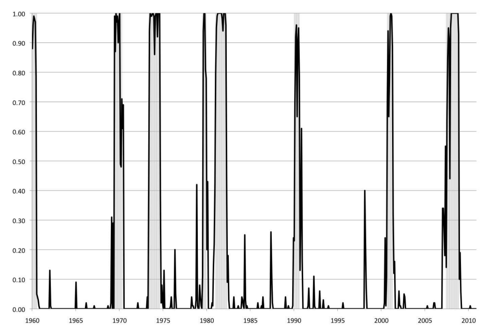


### National- and State-Level Predictors

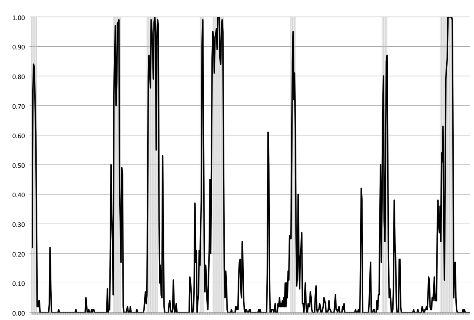
*h*=0



*h*=1



*h*=2



*h*=3

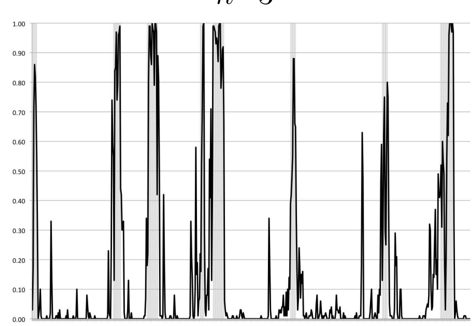
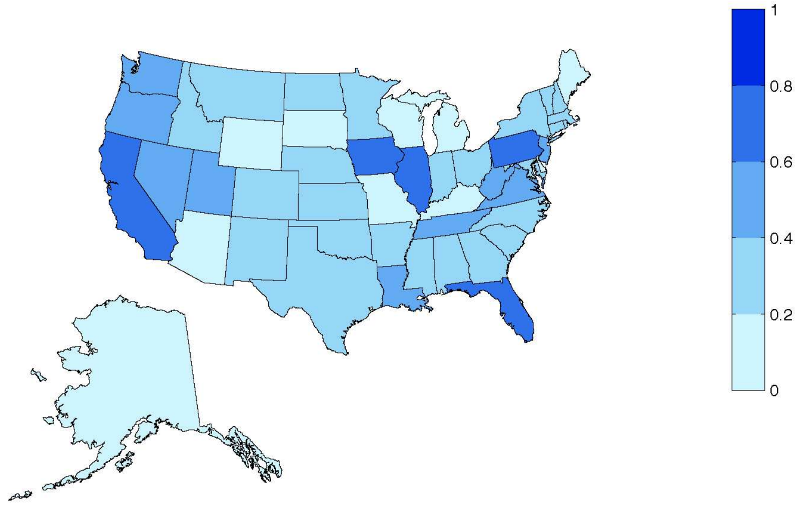
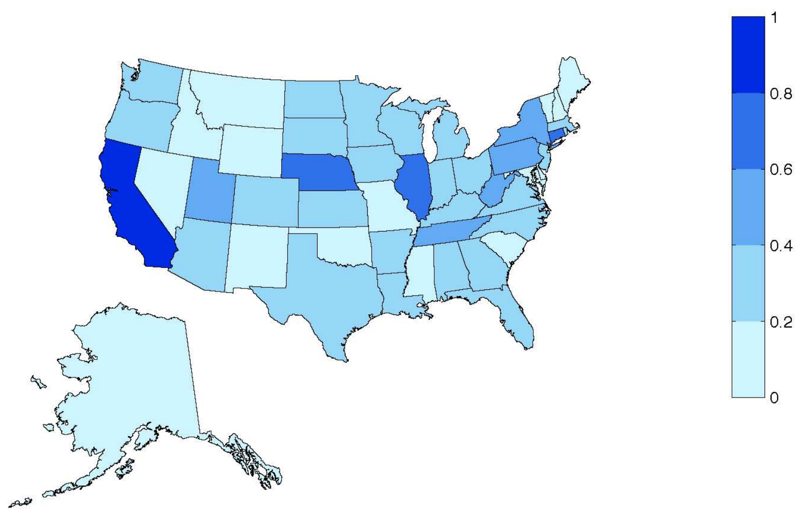


**Figure 2**
**Average Predictor Inclusion Probabilities for Recursive Estimations**

$h$=0




$h$=1




**Notes:** These maps indicate the average posterior probability that state-level employment variables are included in the model given by (2), where averaging is across the multiple recursive estimations beginning over the period August 1960-December 1978, and ending with the period August 1960-mid 2011, with the exact ending month dependent on the forecasting horizon.

**Table 1**
**In-Sample Forecast Evaluation Metrics**

| | **National-Level Predictors** | | | **National and State-Level Predictors** | | |
|---|---|---|---|---|---|---|
| Forecast Horizon | $CSP$ | $QPS$ | $LPS$ | $CSP$ | $QPS$ | $LPS$ |
| $h = 0$ | 0.94 | 0.09 | 0.15 | 0.98 | 0.03 | 0.05 |
| $h = 1$ | 0.92 | 0.11 | 0.17 | 0.97 | 0.04 | 0.07 |
| $h = 2$ | 0.92 | 0.12 | 0.18 | 0.94 | 0.08 | 0.14 |
| $h = 3$ | 0.91 | 0.13 | 0.20 | 0.94 | 0.09 | 0.15 |

**Notes:** This table holds the forecast evaluation metrics defined in Section 2.3 constructed for in-sample predictions of the business cycle phase (expansion or recession) produced over the period August 1960 to June 2011. The in-sample predictions are constructed from the posterior predictive density of two versions of the model in (2), which differ on the inclusion of state-level predictors.

**Table 2**
**In-Sample Forecast Evaluation Metrics - Expansion vs. Recession Months**

| | National-Level Predictors | | | National and State-Level Predictors | | |
|---|---|---|---|---|---|---|
| Forecast Horizon | $CSP$ | $QPS$ | $LPS$ | $CSP$ | $QPS$ | $LPS$ |
| *Expansion Months* | | | | | | |
| $h = 0$ | 0.97 | 0.04 | 0.08 | 0.99 | 0.02 | 0.03 |
| $h = 1$ | 0.97 | 0.06 | 0.12 | 0.98 | 0.02 | 0.04 |
| $h = 2$ | 0.97 | 0.05 | 0.09 | 0.98 | 0.04 | 0.07 |
| $h = 3$ | 0.97 | 0.06 | 0.10 | 0.97 | 0.04 | 0.08 |
| *Recession Months* | | | | | | |
| $h = 0$ | 0.73 | 0.37 | 0.55 | 0.91 | 0.12 | 0.19 |
| $h = 1$ | 0.68 | 0.44 | 0.66 | 0.90 | 0.15 | 0.28 |
| $h = 2$ | 0.63 | 0.49 | 0.72 | 0.75 | 0.34 | 0.52 |
| $h = 3$ | 0.59 | 0.54 | 0.79 | 0.73 | 0.37 | 0.57 |

**Notes:** This table holds the forecast evaluation metrics defined in Section 2.3 constructed for in-sample predictions of the business cycle phase (expansion or recession) produced separately for NBER defined expansion and recession months over the period August 1960 to June 2011. The in-sample predictions are constructed from the posterior predictive density of two versions of the model in (2), which differ on the inclusion of state-level predictors.

**Table 3**
**Full Sample Predictor Inclusion Probabilities**

| | $h=0$ | $h=1$ | $h=2$ | $h=3$ |
|---|---|---|---|---|
| *National-Level Predictors* | | | | |
| Federal Funds Rate | 1.00 | 1.00 | 1.00 | 1.00 |
| S&P 500 Return | 1.00 | 1.00 | 1.00 | 1.00 |
| Term Spread | 0.11 | 0.43 | 0.33 | 1.00 |
| Employment Growth | 0.59 | 0.14 | 0.05 | 0.08 |
| Industrial Production Growth | 1.00 | 1.00 | 0.55 | 0.04 |
| *State-Level Employment Growth* | | | | |
| California | 0.46 | 1.00 | 0.76 | 1.00 |
| Connecticut | 0.26 | 1.00 | 0.12 | 0.29 |
| Florida | 1.00 | 1.00 | 1.00 | 1.00 |
| Illinois | 0.46 | 1.00 | 0.52 | 0.87 |
| Iowa | 1.00 | 0.19 | 0.15 | 0.04 |
| Kentucky | 0.25 | 0.76 | 0.13 | 0.34 |
| Minnesota | 0.87 | 0.2 | 0.01 | 0.08 |
| Nebraska | 0.02 | 1.0 | 1.0 | 1.0 |
| New Jersey | 0.93 | 0.12 | 0.06 | 0.02 |
| Ohio | 1.00 | 0.13 | 0.07 | 0.11 |
| Oregon | 1.00 | 0.04 | 0.03 | 0.02 |
| Pennsylvania | 0.98 | 1.00 | 0.62 | 0.86 |
| Virginia | 1.00 | 0.47 | 0.21 | 0.01 |
| Washington | 0.21 | 0.90 | 0.60 | 0.89 |
| West Virginia | 0.58 | 0.92 | 0.00 | 0.01 |
| **Number of Variables** | **17.7** | **17.5** | **9.4** | **11.1** |

**Notes:** This table holds the posterior probability that specific variables are included in the model given by (2), based on estimation over the period from August 1960 to June 2011. For state-level variables, only those states with inclusion probabilities above 50% for at least one forecast horizon are reported.

**Table 4**
**Out-of-Sample Forecast Evaluation Metrics**

| | **National-Level Predictors** | | | **National and State-Level Predictors** | | |
|---|---|---|---|---|---|---|
| Forecast Horizon | $CSP$ | $QPS$ | $LPS$ | $CSP$ | $QPS$ | $LPS$ |
| $h = 0$ | 0.92 | 0.11 | 0.18 | 0.94 | 0.08 | 0.15 |
| $h = 1$ | 0.91 | 0.13 | 0.20 | 0.95 | 0.09 | 0.26 |
| $h = 2$ | 0.90 | 0.14 | 0.22 | 0.90 | 0.15 | 0.30 |
| $h = 3$ | 0.90 | 0.15 | 0.24 | 0.89 | 0.18 | 0.39 |

**Notes:** This table holds the forecast evaluation metrics defined in Section 2.3 constructed for out-of-sample forecasts of the business cycle phase (expansion or recession) produced over the period January 1979 to June 2011. The out-of-sample predictions are constructed from the posterior predictive density of two versions of the model in (2), which differ on the inclusion of state-level predictors.

**Table 5**
**Out-of-Sample Forecast Evaluation Metrics - Expansion vs. Recession Months**

| | National-Level Predictors | | | National and State-Level Predictors | | |
|---|---|---|---|---|---|---|
| Forecast Horizon | $CSP$ | $QPS$ | $LPS$ | $CSP$ | $QPS$ | $LPS$ |
| *Expansion Months* | | | | | | |
| $h = 0$ | 0.96 | 0.06 | 0.10 | 0.95 | 0.07 | 0.13 |
| $h = 1$ | 0.95 | 0.07 | 0.12 | 0.96 | 0.07 | 0.20 |
| $h = 2$ | 0.95 | 0.07 | 0.12 | 0.95 | 0.09 | 0.19 |
| $h = 3$ | 0.97 | 0.07 | 0.12 | 0.95 | 0.09 | 0.25 |
| *Recession Months* | | | | | | |
| $h = 0$ | 0.68 | 0.41 | 0.60 | 0.91 | 0.15 | 0.24 |
| $h = 1$ | 0.64 | 0.47 | 0.69 | 0.88 | 0.23 | 0.61 |
| $h = 2$ | 0.61 | 0.55 | 0.81 | 0.61 | 0.51 | 0.94 |
| $h = 3$ | 0.54 | 0.64 | 0.98 | 0.52 | 0.71 | 1.19 |

**Notes:** This table holds the forecast evaluation metrics defined in Section 2.3 constructed for out-of-sample forecasts of the business cycle phase (expansion or recession) produced separately for NBER defined expansion and recession months over the period January 1979 to June 2011. The out-of-sample predictions are constructed from the posterior predictive density of two versions of the model in (2), which differ on the inclusion of state-level predictors.

**Table 6**
**One-Month Ahead Forecasts: 2008-2009 Recession**

| Date | NBER Recession Indicator | National-Level Predictors | National and State-Level Predictors |
|---|---|---|---|
| November 2007 | 0 | 0.08 | 0.05 |
| December 2007 | 0 | 0.08 | 0.02 |
| January 2008 | 1 | 0.06 | 0.00 |
| February 2008 | 1 | 0.36 | 0.38 |
| March 2008 | 1 | 0.19 | **0.61** |
| April 2008 | 1 | 0.42 | **0.90** |
| May 2008 | 1 | 0.16 | **0.66** |
| June 2008 | 1 | 0.15 | 0.10 |
| July 2008 | 1 | 0.30 | **0.92** |
| August 2008 | 1 | **0.80** | **1.00** |
| September 2008 | 1 | **0.76** | **1.00** |
| October 2008 | 1 | **0.88** | **1.00** |
| November 2008 | 1 | **1.00** | **1.00** |
| December 2008 | 1 | **1.00** | **1.00** |
| January 2009 | 1 | **1.00** | **1.00** |
| February 2009 | 1 | **1.00** | **1.00** |
| March 2009 | 1 | **1.00** | **1.00** |
| April 2009 | 1 | **1.00** | **1.00** |
| May 2009 | 1 | **1.00** | **1.00** |
| June 2009 | 1 | **0.66** | **0.90** |
| July 2009 | 0 | 0.17 | 0.20 |
| August 2009 | 0 | 0.35 | 0.26 |

**Notes:** This table holds the one-month ahead, out-of-sample forecasts of the business cycle phase (expansion and recession) around the 2008-2009 NBER-defined recession. The forecasts are constructed from two versions of the model in (2), which differ on the inclusion of state-level predictors.

**Table 7**
**Average Predictor Inclusion Probabilities for Recursive Estimations**

| | $h = 0$ | $h = 1$ | $h = 2$ | $h = 3$ |
|---|---|---|---|---|
| *National-Level Predictors* | | | | |
| Federal Funds Rate | 0.69 | 0.62 | 0.61 | 0.19 |
| S&P 500 Return | 0.97 | 1.00 | 0.99 | 0.99 |
| Term Spread | 0.39 | 0.64 | 0.74 | 0.99 |
| Employment Growth | 0.25 | 0.39 | 0.15 | 0.08 |
| Industrial Production Growth | 0.64 | 0.39 | 0.08 | 0.03 |
| *State-Level Employment Growth* | | | | |
| Alaska | 0.05 | 0.09 | 0.15 | 0.51 |
| California | 0.63 | 0.89 | 0.82 | 0.99 |
| Connecticut | 0.15 | 0.79 | 0.17 | 0.38 |
| Florida | 0.63 | 0.36 | 0.18 | 0.10 |
| Illinois | 0.71 | 0.78 | 0.43 | 0.65 |
| Iowa | 0.77 | 0.24 | 0.33 | 0.05 |
| Nebraska | 0.26 | 0.80 | 0.74 | 0.46 |
| Pennsylvania | 0.85 | 0.53 | 0.68 | 0.97 |
| **Number of Variables** | **17.02** | **15.94** | **10.98** | **9.55** |

**Notes:** This table holds the average posterior probability that specific variables are included in the model given by (2), where averaging is across the multiple recursive estimations beginning over the period August 1960-December 1978, and ending with the period August 1960-mid 2011, with the exact ending month dependent on the forecasting horizon. For state-level variables, only those states with average inclusion probabilities above 50% for at least one forecast horizon are reported.

**Table 8**
**Out-of-Sample Forecast Evaluation Metrics - Recession Months**
**All National- and State-Level Predictors Included**

| Forecast Horizon | $CSP$ | $QPS$ | $LPS$ |
|---|---|---|---|
| $h = 0$ | 0.82 | 0.23 | 0.43 |
| $h = 1$ | 0.86 | 0.20 | 3.25 |
| $h = 2$ | 0.70 | 0.47 | 2.26 |
| $h = 3$ | 0.48 | 0.83 | 3.94 |

**Notes:** This table holds the forecast evaluation metrics defined in Section 2.3 constructed for out-of-sample forecasts of the business cycle phase (expansion or recession) computed separately for NBER defined recession months over the period January 1979 to June 2011. The out-of-sample predictions are constructed from the posterior predictive density of two versions of the model in (2), where all possible covariates, both national- and state-level, are included.

**Table 9**
**Out-of-Sample Forecast Evaluation Metrics with "Real-Time" Data**

| | National-Level Predictors | | | National and State-Level Predictors | | |
|---|---|---|---|---|---|---|
| Forecast Horizon | *CSP* | *QPS* | *LPS* | *CSP* | *QPS* | *LPS* |
| *All Months* | | | | | | |
| $h = 1$ | 0.81 | 0.24 | 0.35 | 0.85 | 0.24 | 0.47 |
| *Expansion Months* | | | | | | |
| $h = 1$ | 1.00 | 0.01 | 0.03 | 0.97 | 0.04 | 0.06 |
| *Recession Months* | | | | | | |
| $h = 1$ | 0.50 | 0.64 | 0.89 | 0.67 | 0.57 | 1.15 |

**Notes:** This table holds the forecast evaluation metrics defined in Section 2.3 constructed for out-of-sample forecasts of the business cycle phase (expansion or recession) produced over the period January 1979 to June 2011. Forecasts are constructed using "real-time" data as it appeared at the time the forecast would have been produced. The out-of-sample predictions are constructed from the posterior predictive density of two versions of the model in (2), which differ on the inclusion of state-level predictors.