ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Binary Conditional Forecasts

Authors	Michael W. McCracken, Joseph McGillicuddy, and Michael T. Owyang
Working Paper Number	2019-029B
Revision Date	April 2021
Citable Link	https://doi.org/10.20955/wp.2019.029
Suggested Citation	McCracken, M.W., McGillicuddy, J., Owyang, M.T., 2021; Binary Conditional Forecasts, Federal Reserve Bank of St. Louis Working Paper 2019-029. URL https://doi.org/10.20955/wp.2019.029

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Binary Conditional Forecasts*

Michael W. McCracken†‡ Joseph T. McGillicuddy† Michael T. Owyang†

April 2021

Abstract

While conditional forecasting has become prevalent both in the academic literature and in practice (e.g., bank stress testing, scenario forecasting), its applications typically focus on continuous variables. In this paper, we merge elements from the literature on the construction and implementation of conditional forecasts with the literature on forecasting binary variables. We use the Qual-VAR [Dueker (2005)], whose joint VAR-probit structure allows us to form conditional forecasts of the latent variable which can then be used to form probabilistic forecasts of the binary variable. We apply the model to forecasting recessions in real-time and investigate the role of monetary and oil shocks on the likelihood of two U.S. recessions.

Keywords: Qual-VAR, recession, monetary policy, oil shocks
JEL Codes: C22, C52, C53

*The authors gratefully acknowledge helpful comments from participants at the 2018 Central Bank Forecasting conference at the Bank of England, the SNDE hosted by the Dallas Fed, the 2020 IWH-CIREQ-GW Macroeconometrics Workshop, and the 2nd Vienna Workshop on Economic Forecasting. Aaron J. Amburgey, Julie K. Bennett and Hannah G. Shell provided research assistance. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis, Federal Reserve System, or any of its staff.

†Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, USA.
‡corresponding author: Michael.W.McCracken@stls.frb.org

1 Introduction

Conditional forecasts are a standard tool that policymakers use to evaluate hypothetical scenarios.[1] Because most macroeconomic variables (e.g., GDP growth, inflation, etc.) are continuous, the focus of the literature has developed around conditional forecasts of continuous variables. However, some macroeconomic outcomes of interest are binary [e.g., sovereign debt defaults (Manasse and Roubini, 2005; IMF), bank failures (Curry, Elmer, and Fissel, 2004; FDIC), financial crises (Bussiere and Fratzscher, 2006; ECB), and phases of the business cycle (Chauvet and Potter, 2005; NY Fed)].

We develop a method for forming conditional forecasts of binary indicators using the Qual-VAR developed in Dueker (2005) and extensively investigated in El-Shagi and von Schweinitz (2016). The Qual-VAR augments a vector autoregression's (VAR) observables with a continuous latent variable that is deterministically related to the binary indicator. Similar to a probit, identification of the latent variable is achieved through the correlation between the binary events and the lagged observables. Because the model retains a VAR structure, much of the existing literature on conditional forecasts remains applicable. In particular, we construct the conditional forecasts of the latent in a Bayesian framework similar to that described in Antolin-Diaz, Petrella, and Rubio-Ramirez (2020) for VARs with strictly observable predictors, but with an added step in which the latent variable is drawn from an appropriately truncated normal distribution. Forecasts of the latent then map directly to probabilistic forecasts of the binary event.

While there exist other methods—e.g., the probit or logit—that can be used to form predictions of binary events, most do not capture the dynamic interactions of all the predictors—including the latent variable—across the prediction horizon.[2] Some of these models, however, are at least partially dynamic.[3] As we do in the Qual-VAR, Eichengreen, Watson, and Gross-

[1]Doan, Litterman, and Sims (1984; Minneapolis Fed) initially inspired the construction of conditional forecasts in vector autoregressions (VARs). Waggoner and Zha (1999; Atlanta Fed), Andersson, Palmqvist, and Waggoner (2010; Riksbank), Baumeister and Kilian (2014; IMF), Banbura, Giannone, and Lenza (2015; ECB) provide some examples of recent developments.

[2]See Lahiri and Yang (2013) for an extensive review.

[3]Dueker (1997) and Moneta (2005) add lags of the binary indicator to the set of predictors. Chauvet, Juhn, and Potter (2002) and Chauvet and Piger (2008) form binary predictions in a Markov-switching envi-

man (1985) permit lagged values of the latent variable to be part of the set of predictors but do not permit dynamic interaction among all the predictors over the prediction horizon. Finally, as noted by Fornari and Lemke (2010), their ProbVAR model is the Qual-VAR but only after shutting down the role of lagged values of the latent variable in the set of predictors.

Although we offered a number of potential applications above, our focus in this paper is the prediction of business cycle phases. These phases can be of particular interest to policymakers and firm decision makers, as economic dynamics may vary across the business cycle [e.g., Morley and Piger (2012) and others].[4] Moreover, previous studies have found differences in the efficacy of both monetary and fiscal policy across business cycle regimes [Tenreyro and Thwaites (2016) for monetary policy, and Auerbach and Gorodnichenko (2013) for fiscal policy]. If policymakers could predict when a turning point is likely, they could be more expedient about implementing appropriate countercyclical policy.

We provide two types of experiments, each of which is designed to address a different question related to the conditional forecasting of recessions. The first uses conditional forecasting to manage the ragged-edge problem when forecasting recessions based solely on historical correlations among the series in the VAR—no particular cause of a recession is hypothesized. The second is a pair of counterfactual scenario forecasts in which we hypothesize a (specific) recession was caused by a certain structural shock (e.g., monetary policy, oil prices). Following the language delineated in Antolin-Diaz, Petrella, and Rubio-Ramirez (2020) we will refer to these two types of conditional forecasts as conditional-on-observables and structural scenarios respectively.

Our first application relates to real-time recession forecasting. Most macroeconomic data are released with lags that may not coincide across predictors. In our framework, we form probability-based backcasts and forecasts of a recession using the ragged edge of data available in the current vintage. In contrast, most methods used for binary prediction work absent issues associated with the timing of data releases.[5] In this application, the ragged edge issue

ronment. Kauppi and Saikkonen (2008) use a dynamic probit in which lagged binary indicators and lagged event probabilities are part of the set of predictors.

[4]We do not explicitly model state changes in economic dynamics as these make the model nonlinear and complicate both estimation and prediction. We leave these for future research.

[5]Chauvet and Piger (2008) discuss this issue. Their published results avoid the problem by dropping all

is particularly acute because turning points (i.e., peak or trough) are typically announced with a significant lag. Thus, timely backcasts of recession dates are also desirable and these are easily produced in a conditional-on-observables forecasting framework.

We then consider two structural scenarios. Unlike the preceding exercise, these applications are based on the effects of structural shocks. First, we revisit a conditional forecasting exercise conducted by Leeper and Zha (2003), who evaluate the effect of monetary policy shocks on the 1990-1991 U.S. recession. Second, we revisit a conditional forecasting experiment conducted by Hamilton (2009a,b), who investigates whether a sharp rise in oil prices was a leading cause of the Great Recession. The original analyses are motivated in terms of the onset of recessions but the applications are based on the growth rate or the path of GDP. Our model, on the other hand, computes the effect of these scenarios directly on the probability of recessions.

Using our model for the Leeper and Zha experiment, we find that tighter monetary policy would have made the 1990-1991 recession even more certain. Moreover, by the October 1990 FOMC meeting, an aggressively looser policy would have been unlikely to prevent the 1990-1991 recession. In our version of the Hamilton experiment, we find that oil shocks may have played a substantial role in the onset of the Great Recession; however, other factors—like the financial crisis—had a more important effect on its duration.

The remainder of the paper proceeds as follows. Section 2 provides a description of the Qual-VAR model, its estimation, and how we construct our forecasts. Sections 3 and 4 discuss our applications. Section 5 concludes.

2 Empirical Approach

At the forecast origin T, we want to predict the h-period-ahead value of an observed binary indicator $S_{T+h} \in \{0, 1\}$, given a history of observable predictors $\mathbf{X}_T = \{X_t\}_{t=1}^{T}$ and binary indicators $\mathbf{S}_T = \{S_t\}_{t=1}^{T}$. In addition to $\mathbf{X}_T$ and $\mathbf{S}_T$, we may have information about a subset of the future realizations of X_{T+i}, $i = 1, ..., h$. For example, we may know the future path of

releases until each series has an observation in a given month. In a footnote, they state that they experimented with ragged edge issues and obtained similar, but unreported, results.

a policy instrument. Alternatively, we may seek to predict the binary indicator conditional on a scenario comprised of future realizations of the observables believed to affect S_{T+h}. In what follows, we delineate a method for producing these conditional forecasts.

2.1 Model

To generate the h-period-ahead forecast, we propose a model for a one-period-ahead forecast formed at time t that can be iterated h periods ahead. First, define a continuous latent variable, y_t^*, that is linearly related to lags of an $(n-1)$ vector of observables X_t, as well as lags of itself:

$$y_t^* = B_{y0} + B_{yy}(L)\, y_{t-1}^* + B_{yx}(L)\, X_{t-1} + u_t^y, \tag{1}$$

where B_{y0} denotes an intercept term, the $B_{ij}(L)$'s represent finite-ordered lag polynomials, and $u_t^y \sim i.i.d.N(0,1)$. The binary classification variable, S_t, takes the values 0 or 1 based on the sign of the continuous latent variable such that

$$\begin{array}{lll} S_t = 1 & \text{if} & y_t^* < 0 \\ S_t = 0 & & otherwise \end{array}. \tag{2}$$

Equations (1) and (2) define a dynamic probit for the binary indicators, S_t [see Eichengreen, Watson, and Grossman, 1985]. Omitting the second term on the right-hand side of (1) yields the static probit model. The presence of the lags of the latent variable in the probit equation makes the system dynamic. Assuming the lagged coefficients are positive, the model exhibits more persistence than a standard probit. The probability that $S_{t+1} = 1$, conditional on information at time t, takes the form

$$\Pr\left[S_{t+1} = 1 | \mathbf{X}_t, \mathbf{y}_t^*\right] = 1 - \Phi\left(B_{y0} + B_{yy}(L)\, y_t^* + B_{yx}(L)\, X_t\right), \tag{3}$$

where $\Phi(.)$ is the standard normal CDF and $\mathbf{y}_t^* = \{y_\tau^*\}_{\tau=1}^t$. The binary dynamic probit includes an intercept that allows the unconditional persistence probabilities to vary across the two regimes, $S_t = 0$ and $S_t = 1$.

Iterative multistep forecasting using the dynamic probit requires a method for forming predictions for the future values of the observables, $X_{t+1}, ..., X_{t+h}$. Accordingly, we assume that X_t follows an autoregressive process that also depends on lags of the latent variable, y^*. We can form a stacked vector Y_t containing both the latent variable and the observables into a VAR(P):

$$Y_t = \begin{bmatrix} y_t^* \\ X_t \end{bmatrix} = \begin{bmatrix} B_{y0} \\ B_{x0} \end{bmatrix} + \begin{bmatrix} B_{yy}(L) & B_{yx}(L) \\ B_{xy}(L) & B_{xx}(L) \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} u_t^y \\ u_t^x \end{bmatrix}, \tag{4}$$

where P is the lag order of the VAR polynomials, $u_t = [u_t^y, u_t^{x\prime}]'$ is $i.i.d.$ $\sim N(0, \Sigma)$ with $\Sigma_{yy} = 1$. Let $W_t = [Y_{t-1}', ..., Y_{t-P}']'$. We can then write (4) as

$$Y_t = B_0 + BW_t + u_t, \tag{5}$$

where $B = [B_1, ..., B_P]$.

2.2 Estimation and Conditional Forecasting

The model specified above can be estimated by combining standard Bayesian techniques used to estimate polychotomous models [e.g., Albert and Chib, 1993; Tanner and Wong, 1987] with techniques used to estimate VARs [e.g., Giannone, Lenza, and Primiceri, 2015]. We employ the Gibbs sampler [Gelfand and Smith, 1990; Carter and Kohn, 1994] to alternately generate draws from the joint distribution of the model parameters, the latent variable, and the conditional forecasts.[6] Our reported results are based on 10,000 draws after burning the first 10,000 draws.

In what follows, let $\widehat{Y}_{T+1,T+h}$ represent the vector of the conditional forecasts (including the latent variable) and conditioning information (where applicable) for periods $T+1$ to $T+h$.

[6]The conditioning set contains information that may inform the parameter draws of the VAR. Consistent with Waggoner and Zha (1999), we use conditioning information along with the draws of the conditional forecasts to draw the model parameters.

2.2.1 Drawing $B, B_0, \Sigma | \mathbf{y}_T^*, \mathbf{X}_T, \widehat{Y}_{T+1,T+h}$

For the VAR parameters, we adopt a standard Minnesota-type, normal-Inverse-Wishart prior of the form:

$$\Sigma \sim IW\left(\Psi, d\right),$$

$$\text{vec}(B)|\Sigma \sim N\left(0, \Sigma \otimes \Theta\left(\lambda\right)\right),$$

$$B_0 \sim N\left(0, \Theta_0\right),$$

where d and Ψ are, respectively, the degrees of freedom and scale matrix of the Inverse-Wishart distribution and λ scales the covariances across the VAR parameters. To remain consistent with the probit, the first element of the covariance matrix—corresponding to the latent variable—is parameterized to have unit mean.

Some of the VAR prior parameters are predetermined. For example, we set $d = n + 2$ and we assume Ψ is a diagonal matrix with elements set to the residual variance of an AR(1) process for the corresponding variable in the VAR. Because the VAR includes a stationary latent variable, it will be convenient to transform all of the other variables in the system to be stationary. In this case, we assume at the outset that all of the VAR coefficients are prior mean zero (rather than prior mean one as in a random walk prior). The prior on the constant term B_0 is diffuse.

We further assume that $\Theta\left(\lambda\right)$ is a $k \times k$ matrix, $k = nP$, parameterized such that the prior covariance of the VAR coefficients takes the following form:

$$cov\left(\left(B_s\right)_{ij}, \left(B_r\right)_{hm} | \Sigma\right) = \begin{cases} \lambda^2 \frac{1}{s^2} \frac{\Sigma_{ih}}{\psi_j/(d-n-1)} & \text{if} \quad m = j \text{ and } r = s \\ 0 & otherwise \end{cases},$$

where λ governs the overall tightness of the prior by controlling the scale of the variances and covariances of the VAR coefficients. Thus, while the coefficients $B_1, ..., B_P$ are assumed to be independent of each other, coefficients associated with the same variable are allowed to be contemporaneously correlated across different equations. In general, the prior imposes a tighter variance on the distant lags.

While one could set λ ex ante, Giannone, Lenza, and Primiceri (2015; GLP) show that it is possible to impose an optimal amount of shrinkage by choosing λ to maximize the marginal data density.[7] In their algorithm, GLP draw λ with $\boldsymbol{B} = [B_0, B]$ and Σ by augmenting the sampler with a Metropolis step. Their prior for λ is a Gamma distribution with a mode of 0.2 and standard deviation of 0.4.

Given the posterior of the hyperparameter λ, the in-sample augmented data $\mathbf{y}_T^*$, and the conditioning information (including the conditional forecasts from the last iteration, $\widehat{Y}_{T+1,T+h}$), we obtain the posterior distribution of the VAR parameters ($\mathbf{B}$ and Σ) by drawing from normal-Inverse-Wishart posterior distributions implied by the conjugate prior. Because the latent variable has a fixed unit variance, we then rescale the first row and column of each draw of Σ so that $\Sigma_{yy} = 1$.

2.2.2 Drawing $\{y_t^*\}_{t=1}^T | B, B_0, \Sigma, \mathbf{X}_T, \mathbf{S}_T$

Because of the autoregressive properties of the model, the draw of the latent variable through the forecast origin is broken into three steps: (i) $\{y_t^*\}_{t=1}^P$; (ii) $\{y_t^*\}_{t=P+1}^{T-P}$; and (iii) $\{y_t^*\}_{t=T-P+1}^T$. The first step draws from conditional distributions that depend on the assumed initial conditions. While these conditional distributions can be computed, Dueker (2005) argues that they are slow and inefficient to sample.

We sample $\{y_t^*\}_{t=1}^P$ from independent Metropolis-Hastings steps, where the candidates are drawn from random-walk truncated normal proposal distributions:

$$y_t^{*(i+1)} \sim TN(y_t^{*(i)}, \sigma^2_{y^*_{P+1}} | S_t),$$

where $y_t^{*(i)}$ is the previous draw of y_t^*, $\sigma^2_{y^*_{P+1}}$ is the conditional variance of y^*_{P+1}, and the distribution is truncated consistent with the definition of S_t in (2).

[7]In their paper, GLP use a nonstationary prior that sets prior mean of the VAR coefficients to 1. In addition to optimizing over λ, GLP maximize the marginal data density with respect to two other prior hyperparameters—one that governs the sum of the VAR coefficients and one that accounts for possible cointegration. Because we transform the VAR to stationarity for the latent variable, we omit these two hyperparameters.

For $\{y_t^*\}_{t=P+1}^{T-P}$, we can draw each y_t^* from its full conditional distribution. Note that the posterior for y_t^* at any time $t = P+1, ..., T-P$, conditional on S_t, is truncated normal and depends on the draws of y_τ^* at times $\tau = t-P+1, ..., t-1, t+1, ..., t+P$. Obviously, the period-$t$ draw depends on the P lags prior to t through the lag polynomials. The value of the latent at time t also affects the likelihood at periods $t+1$ through $t+P$. We draw y_t^* sequentially from

$$y_t^{*(i+1)} \sim p\left(y_t^* | y_{t-P+1}^{*(i+1)}, ..., y_{t-1}^{*(i+1)}, y_{t+1}^{*(i)}, ..., y_{t+P}^{*(i)}, S_t, X_{t-P+1}, ..., X_{t+P}, \boldsymbol{B}^{(i+1)}, \Sigma^{(i+1)}\right)$$

for $t = P+1, ..., T-P$. Dueker (2005) derives the exact conditional densities and we refer the reader there.

For $\left\{y_t^{*(i+1)}\right\}_{t=T-P+1}^{T}$, El-Shagi and von Schweinitz (2016) exploit the VAR structure to compute the conditional mean of the latent variable. However, their draw does not take into account all of the available information in the time T dataset—specifically, the values of the observables $X_t, ..., X_T$ and the previous Gibbs draws of the latent $y_{t+1}^{*(i)}, ..., y_T^{*(i)}$ when drawing $y_t^{*(i+1)}$. Rather than implementing a Metropolis-Hastings step here, we utilize the conditional forecasting densities in the next section, treating future values of X_t and $y_t^{*(i)}$ as conditioning data. Thus, we draw

$$y_t^{*(i+1)} \sim TN\left(\overline{y}_t^*, \sigma_{T-P}^2 | y_{t-P+1}^{*(i+1)}, ..., y_{t-1}^{*(i+1)}, y_{t+1}^{*(i)}, ..., y_T^{*(i)}, S_t, X_{t-P+1}, ..., X_T, \boldsymbol{B}^{(i+1)}, \Sigma^{(i+1)}\right),$$

for $t = T-P+1, ..., T$, where $\overline{y}_t^*$ is the mean of the conditional forecast density developed below and σ_{T-P}^2 is the conditional variance carried over from the $T-P$ period in the $\{y_t^*\}_{t=P+1}^{T-P}$ step.

2.2.3 Conditional/Scenario Forecasting

Forming an h-period-ahead forecast of the binary indicator, S_{T+h}, follows directly from forming an h-period-ahead forecast of the latent variable y_{T+h}^*. Based on the VAR specification of our model, producing a forecast of y_{T+h}^* is subsumed by the construction of the entire path of forecasts $\hat{Y}_{T+1,T+h}$ of $Y_{T+1,T+h} = (Y'_{T+1}, ..., Y'_{T+h})'$. Let $\bar{C}$ denote a $k_0 \times nh$ selection matrix for which $\bar{C}Y_{T+1,T+h}$ is the future path of conditioning observables. Given the ith Gibbs draw of the VAR parameters and latent variables, we can obtain conditional forecasts as draws from the posterior distribution

$$\hat{Y}^{(i)}_{T+1,T+h} \sim p(Y_{T+1,T+h}|\mathbf{X_T}, \mathbf{y}^{*(i)}_T, \bar{C}Y_{T+1,T+h}, \boldsymbol{B}^{(i)}, \Sigma^{(i)}).$$

For brevity, in what follows, we omit the superscript (i) denoting the MCMC iteration.

Because our model for Y is a VAR with Gaussian errors, we can apply existing results developed in Waggoner and Zha (1999) and Antolin-Diaz, Petrella, and Rubio-Ramirez (2020) that characterize the appropriate posterior distribution $p(Y_{T+1,T+h}|\mathbf{X_T}, \mathbf{y}^*_T, \bar{C}Y_{T+1,T+h}, \boldsymbol{B}, \Sigma)$. Let $\hat{Y}^c_{T+1,T+h}$ and $\hat{Y}^u_{T+1,T+h}$ denote the conditional and unconditional point forecasts of $Y_{T+1,T+h}$, respectively. The unconditional forecast $\hat{Y}^u_{T+1,T+h}$ is simply the standard iterated multistep forecast from a VAR based on the recursive nature of the model.

Given $\hat{Y}^u_{T+1,T+h}$, and our assumed Gaussian VAR structure, the conditional forecast $\hat{Y}^c_{T+1,T+h}$ is drawn from a normal distribution with a mean and variance that depend on $\hat{Y}^u_{T+1,T+h}$, the conditioning paths, the type of conditioning (conditional-on-observables vs. structural scenario), the moments of the variables in the model (which are functions of $\boldsymbol{B}$ and Σ), and the forecast horizon. When structural scenarios are used, it is also assumed that we know the $n \times n$ orthogonal matrix Q that identifies the structural shocks, $\varepsilon_{t+1} = Q\Sigma^{-1/2}u_{t+1}$.[8] Define $\Phi_j\tilde{\Sigma}^{1/2} = \Phi_j\Sigma^{1/2}Q^{-1}$ as the matrix of orthogonalized structural impulse responses after j periods and let

$$M = \begin{pmatrix} \tilde{\Sigma}^{1/2} & 0 & & 0 & 0 \\ \Phi_1\tilde{\Sigma}^{1/2} & \tilde{\Sigma}^{1/2} & & 0 & 0 \\ & & ... & \tilde{\Sigma}^{1/2} & 0 \\ \Phi_{h-1}\tilde{\Sigma}^{1/2} & \Phi_{h-2}\tilde{\Sigma}^{1/2} & & \Phi_1\tilde{\Sigma}^{1/2} & \tilde{\Sigma}^{1/2} \end{pmatrix}.$$

[8]In our first application we use the Cholesky decompostion to identify the structural shocks and hence $Q = I$. More generally, Q may need to be estimated and we do so in our second application - details are provided there. Since estimation of Q depends on the specific application, for brevity we take it as given. See Antolin-Diaz, Petrella, and Rubio-Ramirez (2018) for an example where Q is estimated using sign restrictions.

The conditional forecast, $\hat{Y}^c_{T+1,T+h}$, is distributed $N(\mu_Y, \Sigma_Y)$ with

$$\mu_Y = \hat{Y}^u_{T+1,T+h} + MD^*(f_{T+1,T+h} - C\hat{Y}^u_{T+1,T+h}),$$
$$\Sigma_Y = MM' + MD^*(\Omega_f - DD')D^{*\prime}M',$$

and $D^* = D'(DD')^{-1}$, where the definitions of D, $f_{T+1,T+h}$, C, and Ω_f vary depending on whether the forecasts are conditional-on-observables or form a structural scenario. When the forecasts are conditional-on-observables, we set $f_{T+1,T+h} = \bar{C}Y_{T+1,T+h}$, $C = \bar{C}$, $\Omega_f = 0$, and $D = \bar{C}M$. Note that, in this case, the choice of Q is irrelevant.

On the other hand, for a structural scenario, we first need to identify which future structural shocks are determining the scenario. To do so, define $\varepsilon_{T+1,T+h} = (\varepsilon'_{T+1}, ..., \varepsilon'_{T+h})'$ as the vector of all future structural shocks across the forecast horizon and let Ξ denote the $k_s \times nh$ selection matrix that defines $\Xi\varepsilon_{T+1,T+h}$ as the future structural shocks that are *not* determining the scenario. These shocks will be drawn from their unconditional, standard normal distribution, while those not selected will be restricted so that the scenario paths are satisfied.[9] Then, we set $f_{T+1,T+h} = (Y'_{T+1,T+h}\bar{C}', \hat{Y}^{u\prime}_{T+1,T+h}M^{-1\prime}\Xi')'$, $C = (\bar{C}', M^{-1\prime}\Xi')'$, $\Omega_f = diag(0_{k_0}, I_{k_s})$, and $D = CM$. Note that when no conditioning occurs, $\bar{C} = 0$ and we obtain the posterior associated with the unconditional forecast $\hat{Y}^u_{T+1,T+h}$.

Regardless of which form of conditional forecast is used, this procedure allows us to form a conditional point forecast of the latent variable y^*_{T+h} by selecting the appropriate element $\hat{y}^{*c}_{T+h|T}$ from μ_Y. If we let Σ_{y^*} denote the corresponding element of Σ_Y, we can construct a probability forecast that y^*_{T+h} is negative using $1 - \Phi(\hat{y}^{*c}_{T+h|T}\Sigma^{-1/2}_{y^*})$. Forming a conditional forecast of the binary indicator S_{T+h} itself can then be made based on the magnitude of $1 - \Phi(\hat{y}^{*c}_{T+h|T}\Sigma^{-1/2}_{y^*})$ relative to a prespecified threshold.

Because we construct a conditional forecast for each saved Gibbs iteration, we can collect all of the conditional forecasts for each horizon. This collection allows us to form a density

[9] See Antolin-Diaz, Petrella, and Rubio-Ramirez (2018) for more discussion on the generality of this approach under various rank conditions.

forecast for the observed variables and recession probability for each horizon.

3 Real-time Recession Prediction

In this section, we apply our model to the problem of forming probabilistic forecasts of a recession. Because the goal is ex ante prediction, we focus exclusively on conditional-on-observables forecasting, imposing no structure on the sources of future shocks.

3.1 Data

We use real-time, vintage data associated with six monthly-frequency U.S. series: (i) nonfarm payroll employment (EMP; in log first differences), (ii) industrial production (IP; in log first differences), (iii) real manufacturing and trade sales (MTS; in log first differences), (iv) real personal income less transfers (PIX; in log first differences), (v) the effective federal funds rate (FFR; in levels), and (vi) the term spread (TS; level of 10y - level of 1y). The National Bureau of Economic Research Business Cycle Dating Committee (NBER) lists the first four as the most important monthly series for contemporaneous evaluation of the business cycle phase. We also include FFR and TS in our experiments as they are considered useful predictors of future recessions [e.g., Estrella and Mishkin, 1998; Wright, 2006].

For each of these series, we have monthly vintages of data starting in 1976:11 and extending through 2018:12. Each vintage has observations dating back to 1967:01. A large portion of the vintages for the four macroeconomic series were obtained from Jeremy Piger's website.[10] More recent vintages were added using the ALFRED database hosted at the Federal Reserve Bank of St. Louis. The interest rate series are monthly averages of daily series and are obtained from the FRED database also hosted at the Federal Reserve Bank of St. Louis. In all of our experiments, we treat EMP, IP, and PIX as being released with a one month lag. In contrast, MTS is released with a two month lag. The two financial series are for the current month.

[10]https://pages.uoregon.edu/jpiger/research/published-papers/
See Chauvet and Piger (2008) for more details on the construction of these series.

In addition to these observable series, we use historical states of the business cycle, as defined by the NBER, as our measure of S_t. Unlike conventional series that are observed at a regular frequency, the NBER only announces the business cycle *turning points* (i.e., peaks or troughs). Moreover, these announcements are made after-the-fact and with a varying delay. For example, in October 2018, the last NBER announcement was made in September 2010 declaring June 2009 as the business cycle trough. Thus, we only know with certainty the values of $S_t \in \{0, 1\}$ through June 2009. Even so, it seems unnecessarily restrictive to consider forecasts of the U.S. business cycle, made from an October 2018 forecast origin, without treating (at least some of) the values of S_t since the last announcement as anything but an ongoing expansion. There will be, however, periods of uncertainty about the current (or even recent past) state of the economy at some forecast origins—especially when a turning point has occurred or is imminent but has not yet been announced.

To avoid introducing ad hoc measurement errors into the most recent values of S_t, we take a conservative approach, similar to that suggested by Giusto and Piger (2017), when incorporating these values. After any turning point, we treat the next period's business cycle phase as known. After a trough, we assume that the announcement delay for a new peak is no longer than 12 months.[11] That is, for vintages after the NBER announcement, we treat the periods from the trough to 12 months before the forecast origin as known expansion periods and the more recent periods as unknown.[12] After a peak, we assume that the last known value of S_t is the first month of the recently identified recession.[13]

3.2 Experiments

For a fixed forecast origin and corresponding vintage of data T, let $T^* \leq T$ denote the most recent date for which the state of the business cycle is known as discussed in the previous

[11]Giusto and Piger (2017) note that 12 months is the longest historical lag of the NBER's announcement of a new business cycle peak. Thus, this seems like a reasonable threshold to consider the state of the business cycle unknown when the last known turning point is a trough.

[12]If the forecast origin is within the first 12 months after the turning point, we assume that every period after the onset of the new regime (i.e., the period after the turning point) is unknown.

[13]This heuristic is more conservative than Giusto and Piger (2017), who assume that the minimum length of a recession is six months.

section. Then, $\Delta T^* = T - T^*$ reflects the number of periods prior to the forecast origin that the state of the business cycle is unknown. Using vintage T data, we estimate the Qual-VAR using data through T^*. We then use the conditional-on-observables framework to construct forecasts of the latent variable for periods $h = -\Delta T^* + 1, ..., 0, ..., 24$, given all data available through period T. Note that our first forecast horizon will (presumably) be negative, corresponding to a backcast. As described in the previous section, these predictions of the latent variable, $\hat{y}^{*c}_{T+h|T}$, along with the relevant variance, Σ_{y^*}, allow us to form a forecast of the probability of a recession:

$$\begin{aligned} \Pr\left[S_{T+h} = 1\right] &= \Pr\left[y^*_{T+h} < 0\right] \\ &= 1 - \Phi(\hat{y}^{*c}_{T+h|T} \Sigma^{-1/2}_{y^*}). \end{aligned}$$

For comparison, we benchmark our results to a direct multistep probit model using the same set of macroeconomic predictors as before. For the forecasts from the probit model, we manage the ragged edge of the data releases in two ways. First, the probit parameters are estimated separately for each horizon using the sample for which the NBER recession dates are known—i.e., using data only through T^*. Second, we form the forecasts using a data sample through the most recent date for which all of the predictors have been released, denoted $\widetilde{T}$, where $T^* \leq \widetilde{T} \leq T$. Let $\Delta\widetilde{T}$ represent the number of periods from the forecast origin since all of the data were observed. Thus, the probit forecasts will take the form

$$\Pr\left[S_{T+h} = 1\right] = \begin{cases} 1 - \Phi(\gamma_1 (L) X_{T+h-1}) & \text{for} \quad h = -\Delta T^* + 1, ..., -\Delta\widetilde{T} + 1 \\ 1 - \Phi(\gamma_j (L) X_{\widetilde{T}}) & \text{for} \quad j = h + \Delta\widetilde{T},\ h = -\Delta\widetilde{T} + 2, ..., 24 \end{cases}$$

where $\gamma_j(L)$ is a lag polynomial obtained from estimating the j-period-ahead direct multistep probit. Notice that the first row represents backcasts for the period after the last (assumed) known business cycle phase to the period after all of the predictors are observed. The second row represents both backcasts and forecasts for periods thereafter.

For the probit, the number of lags is recursively chosen using BIC, capped by the order of the Qual-VAR.[14] We select the number of lags separately for each j and estimate by maximum likelihood.

3.3 Forecast Evaluation

Here, we describe how we evaluate the accuracy of the Qual-VAR and probit recession probabilities at each horizon. For a fixed horizon, we compute the receiver operating characteristic (ROC) curve, which reflects the trade-off between the false positive rate (FPR) and true positive rate (TPR) as a function of a probability threshold, κ. For each Gibbs draw, at each forecast origin, and across the range of $\kappa \in [0, 1]$, we predict $\hat{S}_{t+h} = 1$ if $1 - \Phi(\hat{y}^{*c}_{T+h|T}\Sigma^{-1/2}_{y^*}) > \kappa$ and $\hat{S}_{t+h} = 0$ otherwise. The ROC curve is the graph of $(FPR(\kappa), TPR(\kappa))$ on the unit square, where $FPR(\kappa)$ and $TPR(\kappa)$ are determined by the percentage of the predictions over all forecast origins that coincide with the realized value. For a given value of the $FPR(\kappa)$, a higher value of $TPR(\kappa)$ indicates a better predictive model; thus, a common measure of predictive quality is the area under the ROC curve (AUROC). Higher values of AUROC reflect greater predictive ability, while an AUROC equal to one-half suggests the model predicts no better than a coin toss.[15]

Because we compute the AUROC for each Gibbs draw, we can also compute the 70-percent coverage interval as a measure of uncertainty for the AUROC.

3.4 Results

Starting with the first vintage in 1976:11 and proceeding across all vintages up to 2016:12, we construct forecasts of recession probabilities at horizons ranging as far back as $h = -30$ through $h = 24$, where $h = 0$ corresponds to the forecast origin.[16,17] We estimate two different

[14]BIC selects 4 lags for the baseline Qual-VAR and 3 for the one that adds financial series.
[15]See Berge and Jorda (2011) for further discussion on ROC curves.
[16]The final forecast origin of 2016:12 is determined by the fact that our longest horizon is 24 months.
[17]In principle, the longest backcast in the sample could be $h = -30$, resulting from the gap between the peak preceding the Great Recession and subsequent NBER announcement of a June 2009 trough made in September 2010. Thus, when forecasting from August 2010—just before the announcement—the last known value of S_t

forecasting specifications for the Qual-VAR and direct multistep probit: one using only the four baseline macroeconomic variables and one using the same four macroeconomic variables plus the two financial series.

Before assessing the relative accuracy of the Qual-VAR and probit recession forecasts, we examine the recession probabilities obtained from the Qual-VAR at various horizons. Figure 1 plots the time series of real-time median recession probabilities from the Qual-VAR, along with the 70 percent coverage intervals at horizons $h = -6, 0$, and 6 months from the forecast origin. Shaded bars indicate NBER recession and are aligned with the respective forecast horizon rather than forecast origin. The left column of panels corresponds to the baseline model; the right column corresponds to the model that also includes financial series.

Perhaps as one would expect, the probabilities are sharper when backcasting than when forecasting. Particularly for the baseline model, at $h = 6$ the probabilities are often near the historical average for the sample (roughly 15 percent) with a few limited spikes that only arise after the onset of a recession. When $h = -6$ the probabilities are much closer to zero or one and tend to rise prior to (rather than after) the onset of the recession. Holding the horizon constant, the probabilities are often sharper when we add the financial series. This is strikingly so when $h = 6$ but is also arguably so for the now- and back-casting horizons $h = 0$ and -6. This result may be expected as the macroeconomic data we use are contemporaneous indicators, while financial series tend to be leading indicators.

Figure 2 plots the AUROC paths for both the Qual-VAR and the probit using the baseline model (top panel) and the model augmented with the two financial series (bottom panel). Consider the first panel. For backcasts (horizons less than zero), both the Qual-VAR and the probit perform well with AUROC levels above 0.9, but the Qual-VAR performs uniformly better than the probit.[18] The (pointwise) 70-percent confidence bands for the Qual-VAR AUROC all exclude the AUROC from the probit model. For the positive horizons, the Qual-VAR clearly outperforms the probit for the first three months but beyond that, the differences are small. More than six months out, neither model provides AUROC values that are above

is January 2008, the first month of the recession.

[18] Although we construct backcasts for horizons as far back as $h = -30$, we start each plot at $h = -11$, because beyond that the number of backcasts per horizon is too small to perform meaningful analysis.

0.5 and hence have no predictive content.

In the second panel, the models include the federal funds rate and the term spread as predictors. For these models, the AUROC paths are largely comparable to those in the first panel, especially for the backcasts, suggesting that the financial variables contain little additional predictive content, over real variables, for negative forecast horizons. This is not the case for positive horizons. Here, we find that both models have AUROC values well above 0.5 through $h = 24$, suggesting that financial variables do add predictive content at positive forecast horizons as far out as 24 months. Given the extensive literature—starting with Estrella and Mishkin (1998)—on the predictive content of the term spread for recessions, this result is not at all surprising. It is worth noting that, at most positive horizons, the Qual-VAR performs substantially better than the probit, suggesting some benefits to using the Qual-VAR rather than the static probit when forecasting recessions.

3.5 Additional Results

The NBER suggests that real GDP growth is also a useful indicator of the business cycle phase. Here, we compare the Qual-VAR to the common rule-of-thumb for identifying recession phases of at least two consecutive quarters of negative real GDP growth. To convert to the quarterly frequency, we estimate the Qual-VAR from the previous section using quarterly averages of the monthly series and add real GDP growth to both the baseline model and the model with financials.

The quarterly Qual-VAR is analogous to the monthly version with quarterly forecast origins/horizons and the maximum announcement lag for NBER dating of peaks converted to 4 quarters. Quarterly frequency NBER recession dates are obtained from the official website. Accuracy of the recession probability forecasts are evaluated using AUROC over the same sample as the monthly exercise. The forecast origins are chosen so that the financial series are defined by the last business day of the current calendar quarter. Note that because we conduct forecasts as if we are observing these series on the last day of each quarter, the base set of variables are all observed with a one-quarter lag while the financials are observed through the current quarter. This implies that there is no ragged edge of data releases for the baseline model (other than with respect to the latent variable) while one does arise for the model including financials and this is managed using the same conditional-on-observables method that was used for the monthly models.

As mentioned above, for this exercise, we compare accuracy of the Qual-VAR to the rule-of-thumb of two consecutive quarters of negative real GDP growth identifying a recession. Rule-of-thumb forecasts are constructed using a BVAR consisting of the same series as the Qual-VAR but without the latent variable. The BVAR is estimated using the same priors and code as the Qual-VAR with the obvious difference that draws and priors associated with the latent are absent. Recession probabilities are constructed as the percent of (saved) Gibbs draws for which real GDP growth is predicted to be negative at horizons h and $h+1$. Forecasts from the benchmark manage the data release lags using conditional-on-observables in the same way as the Qual-VAR.

Figure 3 plots the AUROC paths for both the Qual-VAR and the benchmark BVAR using the baseline model (top panel) and the model augmented with the two financial series (bottom panel). In both panels, the quarterly frequency version of the Qual-VAR performs very similarly to the monthly frequency version at all horizons. Adding real GDP as a predictor does not improve accuracy over the monthly Qual-VAR. When comparing Figures 2 and 3, the clearest difference is performance of the benchmark. Compared to the static probit, the BVAR performs better at most non-negative horizons and even more so when the benchmark contains the two financial series. At these horizons, the benchmark typically performs as well as the Qual-VAR, as the BVAR's recession probabilities often lie within the Qual-VAR's 70-percent coverage intervals.

The largest differences arise for the backcast. Here, the Qual-VAR performs substantially better. To be fair, the BVAR is not designed for backcasting. The negative horizon recession forecasts are based solely on whether the relevant vintage reports negative real GDP growth for two consecutive quarters.

4 Scenario Analysis of Recessions

In this section, we revisit two applications in which conditional forecasting was used as a tool for identifying a causal link to the onset of a recession: Leeper and Zha (2003) for monetary policy shocks and Hamilton (2009a,b) for oil shocks. The models in the original papers are not explicitly designed to predict recessions. Instead, recession probabilities are based on an ad hoc linkage between (negative) GDP growth and an NBER-dated recession. On the other hand, the Qual-VAR explicitly links the recession to the sign of the latent variable for all monthly forecast horizons.

4.1 Monetary Scenario

Leeper and Zha (2003) estimate a structural VAR and use it to assess the effect of monetary policy interventions on the likelihood of the U.S. recession of July 1990 through March 1991.[19] They consider two conditioning scenarios in which the agent is producing forecasts at the October 2, 1990 FOMC meeting: the ex post "Actual" fed funds rate path (8.11 percent, 7.81 percent, 7.31 percent, and 6.91 percent) and a "Tighter" counterfactual path (8.70 percent, 8.95 percent, 8.95 percent, and 8.95 percent) across October 1990 through January 1991. In addition to their two experiments, we consider an aggressively "Looser" scenario to identify whether monetary policy could have lowered the likelihood of—or reduced the duration of—this recession. For this scenario, the counterfactual path is 1 percentage point lower than the "Actual" scenario.

Leeper and Zha then construct recession probabilities for calendar years 1991, 1992, and 1993, where a recession occurs when total real GDP growth is negative for the year.[20] For our version of their model, we report these recession probabilities and probabilities based on the GDP growth rule-of-thumb mentioned above.[21]

[19] The NBER announced the peak and trough in April 1991 and December 1992, respectively.

[20] Our results will differ from what they obtain (their Table 2, p. 1695) for several reasons. First, our Qual-VAR augments their VAR with the latent variable. Second, we use a more recent data vintage: Leeper and Zha's data are taken from a vintage from the latter months of 1998. Finally, based on BIC, our model uses 4 lags of the predictors rather than the 13 chosen by Leeper and Zha. The goal is not to replicate their results so much as to show that our methodology may lead to very different probabilistic forecasts of a recession.

[21] Because the model is monthly, we chose to treat "two consecutive quarters" as equivalent to "5 of 6

Leeper and Zha's dataset consists of six series: three production series [a monthly-frequency measure of real GDP (log-level); the consumer price index (CPI; log-level); and the unemployment rate (levels)], two monetary policy series [the money supply (M2; log-level) and the federal funds rate (FFR; levels)], and one expectations series [a commodity price index (log-level)].[22] The latent variable in our VAR requires stationary data; thus, we use first or second differences of some series rather than levels or log-levels.[23] All series are taken from an April 2019 vintage and are obtained from the FRED database, except for the commodity price index and the PMI, which are obtained from Haver.

The structural shocks are identified using a Cholesky decomposition, where the latent recession indicator is ordered first and the rest of the variables are ordered as listed above. As in Leeper and Zha, we treat the monetary policy shocks as those associated with the federal funds rate. Our identification is common in the literature and implies: (i) monetary shocks respond contemporaneously to the production variables (as suggested by a Taylor-type rule) but not vice versa and (ii) financials respond contemporaneously to policy but not vice versa. Thus, the identification accounts for the stylized fact that macroeconomic variables respond to policy only at (long and variable) lags and also suggests that the Fed does not include financials in its policy rule.

Following Leeper and Zha (2003), we estimate the model parameters (including the latent variable for the Qual-VAR) using the full sample of the April 2019 vintage: January 1959 through March 2019. We construct the scenario forecasts as if the forecasting agent were doing so on the October 2, 1990 FOMC meeting.[24] As in the recession-dating exercises, at this forecast origin, we assume that the state variable is known with a 12-month lag: Thus, the latent variable is only known to the forecaster through October of 1989. We then construct

consecutive months."

[22]Monthly real GDP is constructed from a Chow-Lin (1971) interpolation with AR(1) errors using monthly data on total industrial production; civilian employment for age sixteen years or older; retail sales deflated by consumer prices; real personal consumption expenditures; and the National Association of Purchasing Managers' Composite Index [see Leeper, Sims, and Zha (1996)].

[23]Specifically, we model monthly GDP in log first differences, CPI in log first differences, the unemployment rate in first differences, M2 in log second differences, FFR in levels, and the commodity price index in log first differences.

[24]That is, we assume to have information through the end of September 1990, ignoring data-release lags for all series except the NBER business cycle indicator.

our scenario forecast in a two-step fashion. First, we use the observables through September 1990 to form backcasts of the latent variable. Then, we produce forecasts of the variables in the system—including the latent variable—conditional on the monetary shocks associated with the "Actual", "Tighter", or "Looser" paths, while allowing the other structural shocks to retain their unconditional distribution.

Figure 4 provides results for each scenario. In the top two panels, we plot median paths and 70-percent coverage intervals for the federal funds rate (top left) and the latent variable (top right).[25] For each variable, there are four paths: one for each of the three scenarios and one for the baseline, unconditional forecast. In the bottom three panels, we report recession probabilities. The first two of these probabilities (left bottom) are calculated as the percentage of Gibbs draws for which the indicated ad hoc connection between GDP and recessions is satisfied. The third set of probabilities (right bottom) are based on the posterior distribution of probabilities $1 - \Phi(\hat{y}^{*c}_{T+h|T} \Sigma_{y^*}^{-1/2})$ constructed using draws of the latent variable.

Consistent with the results in Leeper and Zha, the "Tighter" scenario leads to a higher recession probability than the "Actual" scenario. However, the recession probabilities measured by our latent variable are an order of magnitude higher than those based on GDP growth. Notably, the latent variable approach finds a much higher probability of a recession starting in 1990 and continuing into 1991, consistent with the date that the NBER eventually announced. As expected, "Looser" monetary policy lowers the likelihood of the recession. However, we find that, by the October 1990 FOMC meeting, monetary policy was unlikely to have prevented the recession from either occurring or extending into the following year.

4.2 Oil Scenario

Hamilton (2009a,b) uses an ARX model to determine whether the U.S. would have been in a recession over the period 2007:Q4 though 2008:Q3 had there been no increase in oil prices between 2007:Q4 and 2008:Q3. He augments a quarterly-frequency autoregressive model of

[25] The median path of the latent need not resemble the unobserved true latent variable. El-Shagi and von Schweinitz (2016) highlight this point based on extensive simulations. For this reason, we report the 70% coverage intervals rather than just the median paths.

real GDP growth with lags of net oil price increases—the latter variable being defined as the percentage change in real crude oil PPI during the quarter if oil prices made a new three-year high and zero otherwise. Two sets of GDP growth forecasts are constructed: (i) a purely autoregressive forecast that omits net oil and (ii) a "dynamic conditional forecast" that conditions on the ex post realized values of the net oil price measure.

Similar to Leeper and Zha, Hamilton frames the question in terms of the likelihood of a recession; however, the actual application is a real GDP forecasting exercise. From a motivational standpoint, this meshes well with our goal of distinguishing recession forecasts from real GDP forecasts. Even so, the exercise turns out to be a good example of a limitation of our model. Recall that we assume that the variables in the Qual-VAR are Gaussian. While we could argue real GDP growth is Gaussian, no such argument could be made for net oil price increases. This series is always non-negative and frequently takes the value of zero. Thus, we cannot simply estimate a trivariate version of our Qual-VAR consisting of the latent variable, real GDP growth, and the net oil price measure.

Instead, we substitute real crude oil PPI for net oil prices and conduct a structural scenario using the Qual-VAR.[26] We identify structural oil supply shocks applying the method of external instruments (Stock and Watson, 2018), using an instrument constructed by Känzig (2019a) based on high-frequency changes to oil futures around OPEC announcements. For each Gibbs draw, we follow the procedure described by Känzig (2019b; pp. 1-4) to estimate the third column of $\tilde{S} = \Sigma^{1/2} Q^{-1}$, associated with real crude oil PPI. The remaining columns are estimated by imposing the restriction that $\tilde{S}\tilde{S}' = \Sigma$.[27]

As in the previous exercise, we estimate the model parameters and latent variable using the August 2019 vintage of the observables spanning 1959:Q1 to 2019:Q2.[28] We then construct scenario forecasts originating at the end of 2007:Q3, ignoring data-release lags for all series except the NBER business cycle indicator, which are assumed known only with a four quarter

[26]CPI was used to deflate nominal crude oil PPI. All data was obtained from the FRED database.

[27]The first two columns of $\tilde{S}$ are not uniquely identified by this restriction. Because the structural shocks for the latent and real GDP growth are drawn from their unconditional distribution, the fact that $\tilde{S}\tilde{S}' = \Sigma$ suffices for our analysis.

[28]Based on BIC, our model uses 4 lags of the predictors.

lag. The scenario forecast is constructed in two steps. First, we form a backcast of the latent variable, using a conditional-on-observables framework up to the forecast origin. Second, we forecasts all three variables, including the latent, conditional on the oil supply shocks associated with the ex post realized values of the real crude oil PPI series between 2007:Q4 and 2008:Q3. In this step, we also restrict the structural shocks associated with the latent variable and real GDP growth to be from their unconditional distributions.

Figure 5 provides the results from the oil scenario forecasting exercise. The top two panels plot median paths and 70-percent coverage intervals for the real crude oil PPI series (top left) and the latent variable (top right). For each variable there are two paths: one for the scenario and one for the baseline, unconditional forecast. The bottom left panel plots the median paths and 70-percent coverage intervals for the log-level of real GDP. The bottom right panel shows the recession probabilities based on the posterior distribution of probabilities $1 - \Phi(\hat{y}^{*c}_{T+h|T}\Sigma_{y^*}^{-1/2})$ constructed using draws of the latent variable.

In large part, our results reinforce the main message in Hamilton. Relative to the unconditional forecast, the scenario associated with the rise in real oil prices leads to a substantially higher likelihood of a recession occurring. This occurs quickly with the median recession probability rising to 60-percent as early as 2007:Q4 and peaking at 80-percent in 2008:Q2. In addition, our median forecast suggests a slowdown in real GDP that closely aligns with the ex post realized values for the first four quarters. After that, the median forecast of real GDP diverges from the realized values suggesting that other factors, like the financial crisis, were more important as the recession progressed.

5 Conclusion

In this paper, we develop methods for forming conditional forecasts of binary outcomes. We apply recent innovations on conditional forecasting in Bayesian VARs by Waggoner and Zha (1999) and, in particular, Antolin-Diaz, Petrella, and Rubio-Ramirez (2020) to the Qual-VAR developed in Dueker (2005). The Qual-VAR is a standard VAR in which one series is a latent Gaussian variable that is deterministically related to the binary outcome in a manner similar to a probit.

A main focus of this paper is showing how the method can be used to form conditional forecasts of binary events given information about the path of future observables or also given information about the structural shocks determining that path. For the former, we apply the methodology to real-time forecasting of U.S. recessions using vintage data. This is a natural application for our method not only because the state of the business cycle is binary, but also because data release lags and the NBER announcement lag are both easily managed in a conditional forecasting framework. Using the AUROC as a measure of model accuracy, we find that the model performs well backcasting the state of the business cycle and, if the federal funds rate and a term spread are included, the model does a reasonable job forecasting the state of the business cycle at horizons as long as two years. Comparable results are obtained regardless of whether a monthly or quarterly frequency version of the model is used.

A distinguishing feature of our model is it can be used to form structural, scenario-based forecasts of the binary event. We are therefore able to address counterfactual questions related to the causes of, for example, recessions. This question has certainly been asked by Leeper and Zha (2003), exploring the role of monetary policy interventions, and Hamilton (2009a,b), investigating the role of oil price shocks, among others. What makes our results distinct from theirs is that we are able to form predictions of the binary event itself, rather than an approximation based on whether or not real GDP growth happens to be negative. In our version of the Leeper and Zha exercise, we find a much higher likelihood of monetary policy inducing a recession when that event is defined using the latent recession indicator than when the event is defined using negative real GDP growth. In our version of the Hamilton exercise, we find that the sharp rise in oil prices likely played a role in causing the Great Recession though unmodeled factors, such as the financial crisis, were substantially more important for its duration.

References

Albert, J.H. & Chib, S. (1993), Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association*, 88, 669-679.

Andersson, M.K., Palmqvist, S., & Waggoner, D.F. (2010), Density-Conditional Forecasts in Dynamic Multivariate Models, Sveriges Riksbank Working Paper Series 247.

Antolin-Diaz, J., Petrella, I., & Rubio-Ramirez, J.F. (2020), Structural Scenario Analysis with SVARs, *Journal of Monetary Economics*, forthcoming.

Auerbach, J. & Gorodnichenko, Y. (2013), Output Spillovers from Fiscal Policy, *American Economic Review*, 103, 141-46.

Banbura, M., Giannone, D., & Lenza, M. (2015), Conditional Forecasts and Scenario Analysis with Vector Autoregressions for Large Cross-Sections, *International Journal of Forecasting*, 31, 739-756.

Baumeister, C. & Kilian, L. (2014), Real-Time Analysis of Oil Price Risks Using Forecast Scenarios, *IMF Economic Review*, 62, 119-145.

Berge, T.J. & Jorda, O. (2011), Evaluating the Classification of Economic Activity into Recessions and Expansions, *American Economic Journal: Macroeconomics*, 3, 246-277.

Bussiere, M. & Fratzscher, M. (2011), Towards a New Early Warning System of Financial Crises, *Journal of International Money and Finance*, 25, 953-973.

Carter, C.K. & Kohn, R. (1994), On Gibbs Sampling for State Space Models, *Biometrika*, 81, 541-553.

Chauvet, M. & Piger, J. (2008), A Comparison of the Real-Time Performance of Business Cycle Dating Methods, *Journal of Business and Economic Statistics*, 26, 42-49.

Chauvet, M. & Potter, S. (2005), Forecasting Recessions Using the Yield Curve, *Journal of Forecasting*, 24, 77-103.

Chauvet, M., Juhn, C., & Potter, S. (2002), Markov Switching in Disaggregate Unemployment Rates, *Empirical Economics*, 27, 205-232.

Chow, G.C. & Lin, A. (1971), Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series, *Review of Economics and Statistics*, 53, 372-375.

Curry, T.J., Fissel G.S., & Elmer, P.J. (2004), Can the Equity Markets Help Predict Bank Failures?, FDIC Working Paper No. 04-3.

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

Dueker, M. (1997), Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions, Federal Reserve Bank of St. Louis *Review*, March, 41-51.

Dueker, M. (2005), Dynamic Forecasts of Qualitative Variables: A Qual VAR Model of U.S.

Recessions, *Journal of Business and Economic Statistics*, 23, 96-104.

Eichengreen, B., Watson, M.W., & Grossman, R.S. (1985), Bank Rate Policy Under the Interwar Gold Standard: A Dynamic Probit Model, *Economic Journal*, 95, 725-745.

El-Shagi, M. & von Schweinitz, G. (2016), Qual VAR Revisited: Good Forecast, Bad Story, *Journal of Applied Economics*, 19, 293-322.

Estrella, A. & Mishkin, F.S. (1998), Predicting U.S. Recessions: Financial Variables as Leading Indicators, *Review of Economics and Statistics*, 80, 45-61.

Fornari, F. & Lemke, W. (2010), Predicting Recession Probabilities with Financial Variables over Multiple Horizons, ECB Working Paper Series 2010-1255.

Gelfand, A.E. & Smith, A.F.M. (1990), Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, 85, 398-409.

Giannone, D., Lenza, M., & Primiceri, G. E. (2015), Prior Selection for Vector Autoregressions, *Review of Economics and Statistics*, 97, 436-451.

Giusto, A. & Piger, J. (2017), Identifying Business Cycle Turning Points in Real Time with Vector Quantization, *International Journal of Forecasting*, 33, 174-184.

Hamilton, J.D. (2009a), Causes and Consequences of the Oil Shock of 2007-08, *Brookings Papers on Economic Activity*, Spring, 215-259.

Hamilton, J.D. (2009b), Oil Prices and the Economic Recession of 2007-08, *VoxEU.org*, June 16.

Känzig, D.R. (2019a), The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements, manuscript, London Business School.

Känzig, D.R. (2019b), The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements Online Appendix, London Business School.

Kauppi, H. & Saikkonen, P. (2008), Predicting U.S. Recessions with Dynamic Binary Response Models, *The Review of Economics and Statistics*, 90, 777-791.

Lahiri, K. & Yang, L., (2013), Forecasting Binary Outcomes, *Handbook of Economic Forecasting*, 2, 1025-1106.

Leeper, E.M., Sims, C.A., & Zha, T. (1996), What Does Monetary Policy Do?, *Brookings Papers on Economic Activity*, 2, 1-63.

Leeper, E.M. & Zha, T. (2003), Modest policy interventions, *Journal of Monetary Economics*, 50, 1673-1700.

Manasse, P. & Roubini, N. (2005), "Rules of Thumb" for Sovereign Debt Crises, IMF Working Paper 05/42.

Moneta, F. (2005), Does the Yield Spread Predict Recessions in the Euro Area?, *International Finance*, 8, 263-301.

Morley, J. & Piger, J. (2012), The Asymmetric Business Cycle, *The Review of Economics and Statistics*, 94, 208-221.

Stock, J.H. & Watson, M.W. (2018), Identification and Estimation of Dynamic Causal Effects in Macroeconomics using External Instruments, *The Economic Journal*, 128, 917-948.

Tanner, M.A. & Wong, W.H. (1987), The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association*, 82, 528-540.

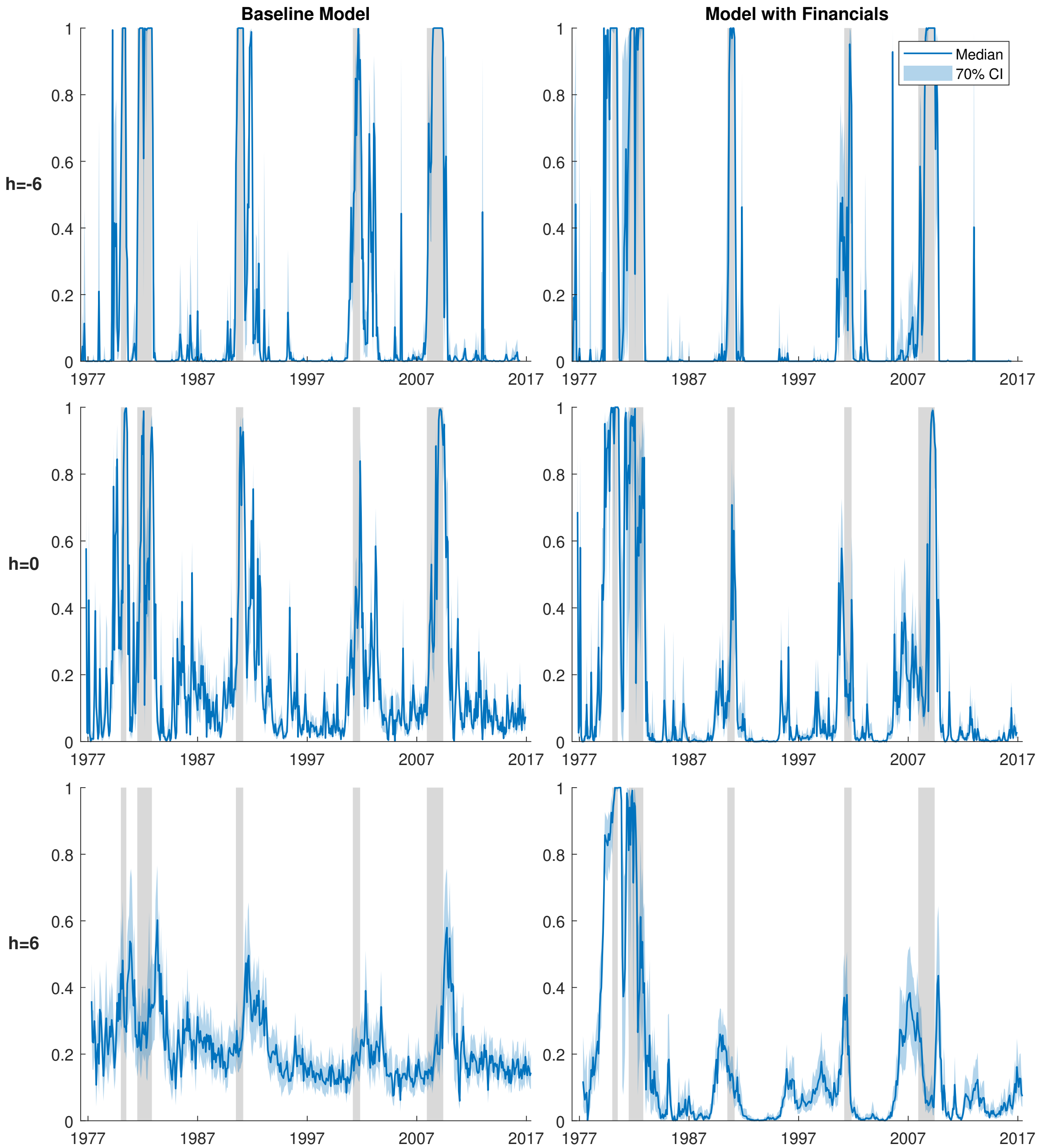
Tenreyro, S. & Thwaites, G. (2016), Pushing on a String: US Monetary Policy Is Less Powerful in Recessions, *American Economic Journal: Macroeconomics*, 8, 43-74.

Waggoner, D.F. & Zha, T. (1999), Conditional Forecasts in Dynamic Multivariate Models, *The Review of Economics and Statistics*, 81, 639-651.

Wright, J.H. (2006), The Yield Curve and Predicting Recessions, FEDS Working Paper 2006-7.

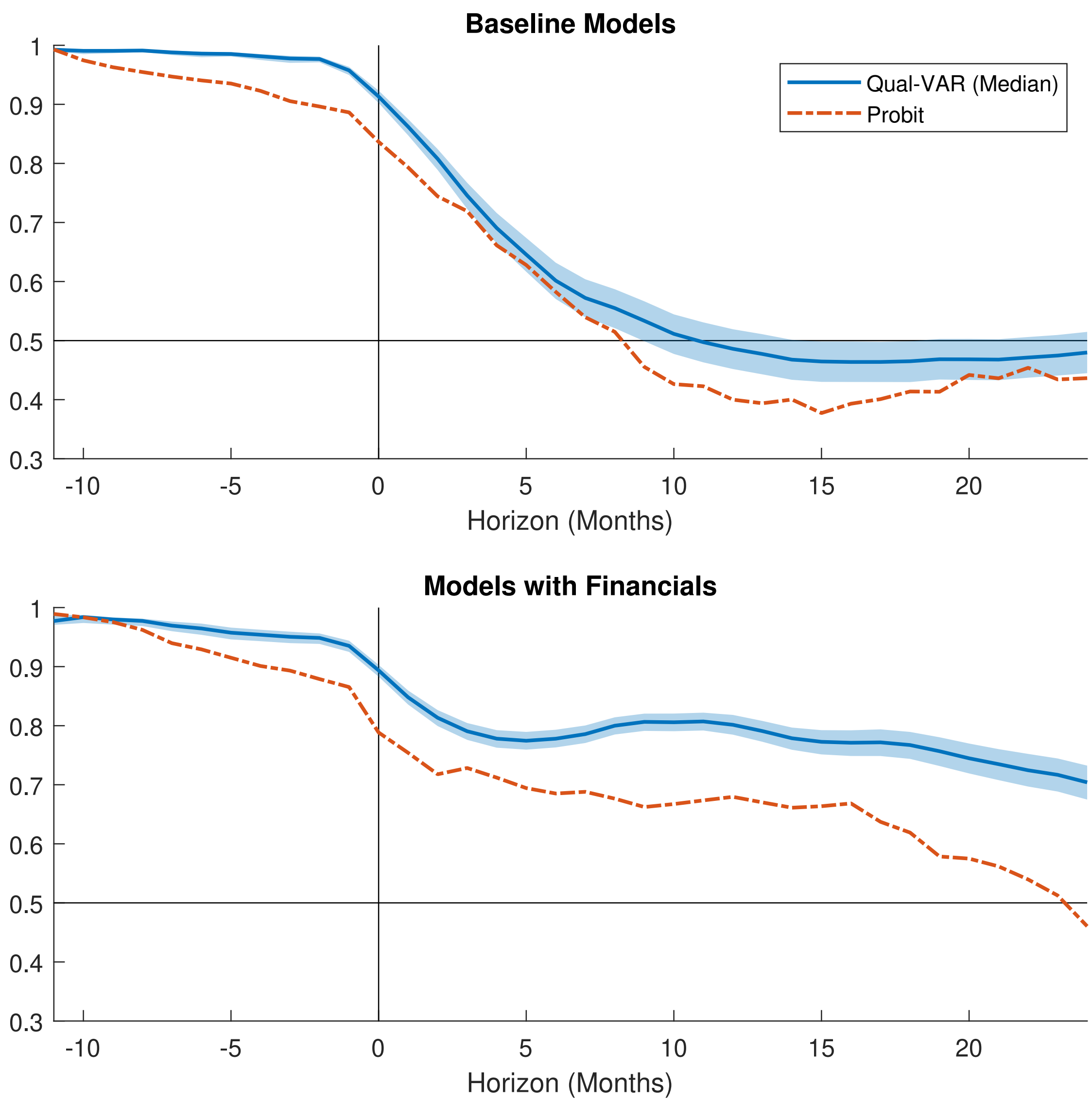
Figures

Figure 1: Qual-VAR Real-Time Recession Probability Forecasts



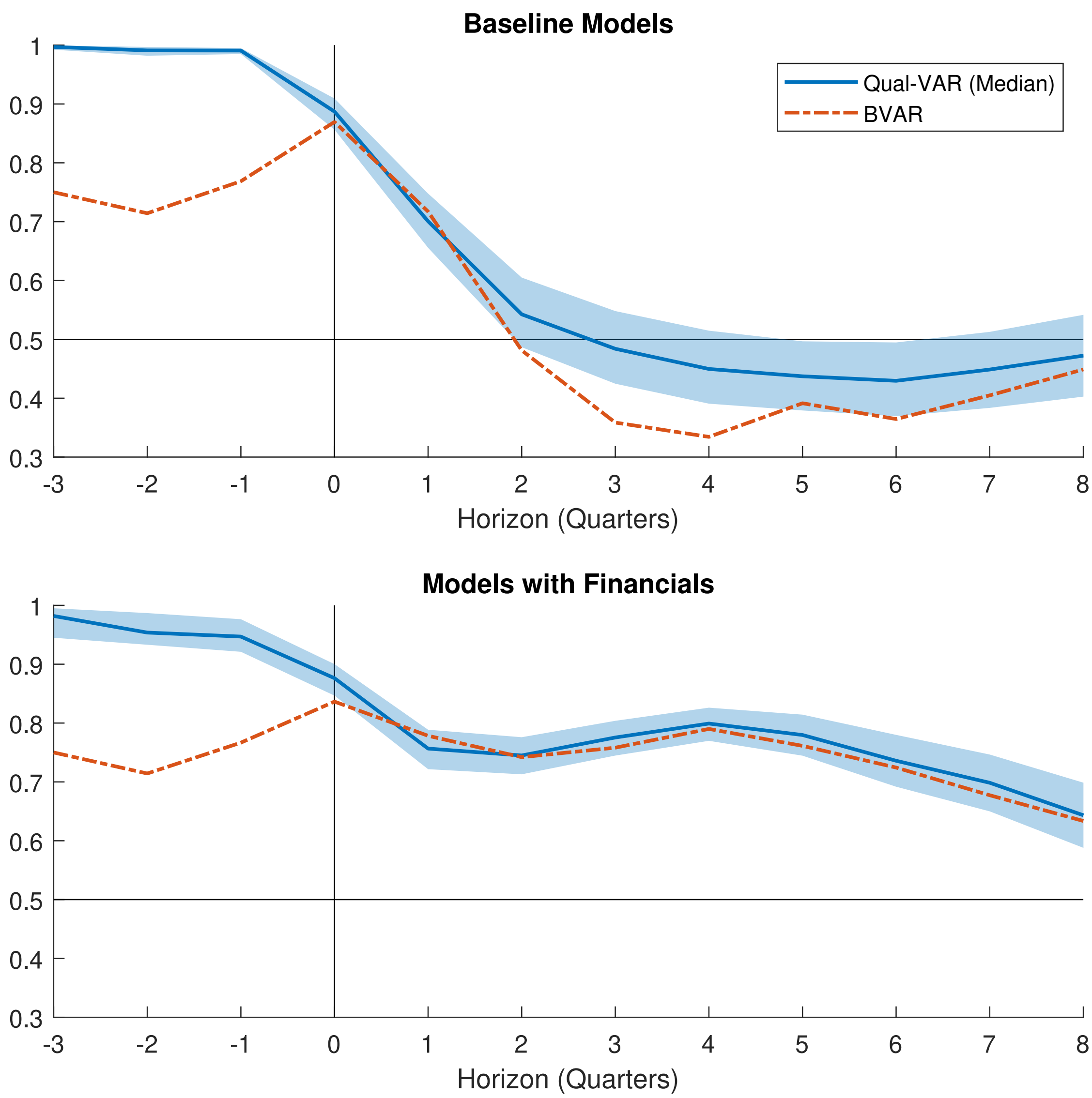
Notes: The top, middle, and bottom panels display forecasts at horizons of -6, 0, and 6 months from the forecast origin, respectively. The baseline model includes as predictors employment, industrial production, real manufacturing and trade sales, and real personal income less transfers. The model with financials augments this set of predictors with the federal funds rate and the term spread. Gray shaded bars indicate NBER recessions. The x-axis displays the period being forecast (as opposed to the forecast origin). The shaded areas represent the 70% coverage intervals. See Section 3 for further details regarding the methodology.

Figure 2: AUROC by Horizon, Monthly



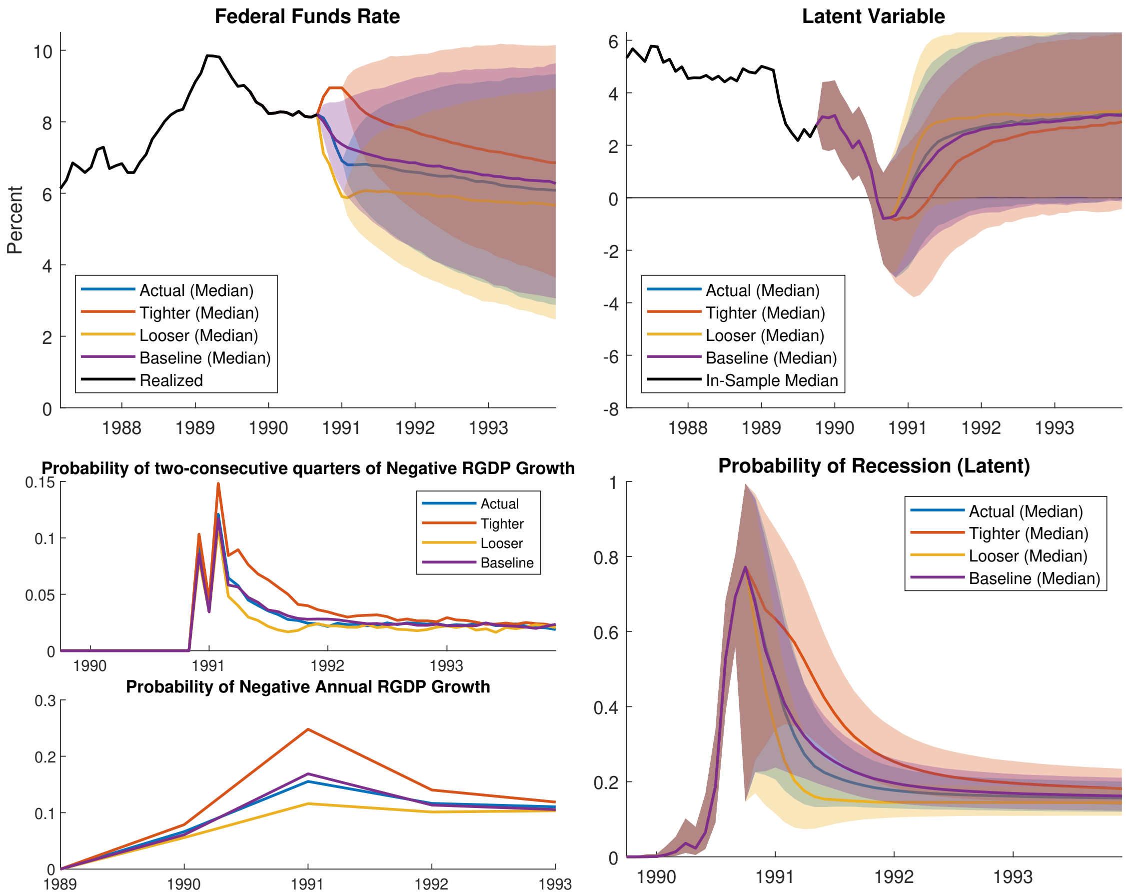
Notes: "AUROC" stands for "area under the receiver operating characteristic curve," and the shaded areas represent the 70% coverage intervals. Horizon is number of months from the forecast origin. The baseline models include as predictors employment, industrial production, real manufacturing and trade sales, and real personal income less transfers. The models with financials augment this set of predictors with the federal funds rate and the term spread. See Section 3 for further details regarding the methodology.

Figure 3: AUROC by Horizon, Quarterly



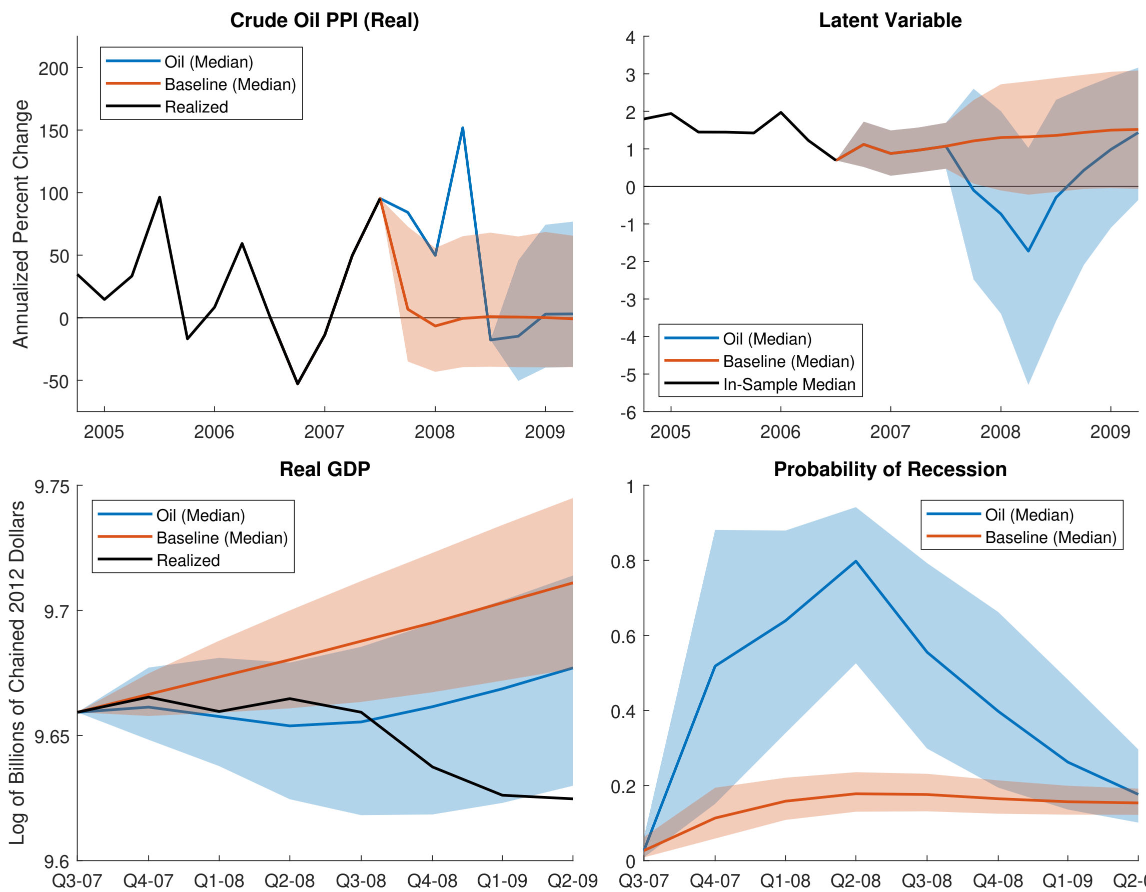
Notes: "AUROC" stands for "area under the receiver operating characteristic curve," and the shaded areas represent the 70% coverage intervals. Horizon is number of months from the forecast origin. The baseline models include as predictors real GDP, employment, industrial production, real manufacturing and trade sales, and real personal income less transfers. The models with financials augment this set of predictors with the federal funds rate and the term spread. See Section 3 for further details regarding the methodology.

Figure 4: Monetary Scenario



Notes: The shaded areas represent the 70% coverage intervals. The "Actual" scenario conditions on the ex-post realized federal funds rate path of 8.11, 7.81, 7.31, and 6.91 percent for months October 1990 to January 1991; the "Tighter" scenario conditions on a hypothetical federal funds rate path of 8.70 percent in October 1990 and 8.95 percent in November 1990 to January 1991; the "Looser" scenario conditions on a hypothetical federal funds rate path that is 100 basis points below the "Actual" path; and the "Baseline" scenario is simply the unconditional forecast. The top panels display monthly forecasts for the federal funds rate and latent variable through December 1993. The bottom panels display predicted recession probabilities. Those in the lower left-hand panel are the percentage of draws for which conditional forecasts of real GDP growth are either negative for two-consecutive quarters or negative (in sum) for a given year. Those in the lower right-hand corner are based on the posterior distribution of recession probabilities constructed using conditional forecasts of the latent variable at each monthly horizon. Backcasts of the latent variable are constructed conditional on observables from November 1989 to September 1990. See Section 4.1 for specifics regarding the model and other details.

Figure 5: Oil Scenario



Notes: The shaded areas represent the 70% coverage intervals. The "Oil" scenario consists of conditioning on the ex-post realized value of real Crude Oil PPI from 2007:Q4 to 2008:Q3, while the "Baseline" scenario is simply the unconditional forecast. The first three panels display quarterly forecasts for real Crude Oil PPI, the latent variable, and real GDP (in log-levels) through 2009:Q2. The fourth panel plots recession probabilities constructed using conditional forecasts of the latent variable at each quarterly horizon. Backcasts of the latent variable are constructed conditional on observables from 2006:Q4 to 2007:Q3. See Section 4.2 for specifics regarding the model and other details.