

NO. 976
AUGUST 2021

A Large Bayesian VAR of the United States Economy

Richard K. Crump | Stefano Eusepi | Domenico Giannone |
Eric Qian | Argia Sbordone

A Large Bayesian VAR of the United States Economy

Richard K. Crump, Stefano Eusepi, Domenico Giannone, Eric Qian, and Argia Sbordone

Federal Reserve Bank of New York Staff Reports, no. 976

August 2021

JEL classification: C32, C11, C53, C54, E32, E37

Abstract

We model the United States macroeconomic and financial sectors using a formal and unified econometric model. Through shrinkage, our Bayesian VAR provides a flexible framework for modeling the dynamics of thirty-one variables, many of which are tracked by the Federal Reserve. We show how the model can be used for understanding key features of the data, constructing counterfactual scenarios, and evaluating the macroeconomic environment both retrospectively and prospectively. Considering its breadth and versatility for policy applications, our modeling approach gives a reliable, reduced form alternative to structural models.

Key words: bayesian vector autoregressions, conditional forecasts, scenario analyses, financial conditions index

Crump, Sbordone: Federal Reserve Bank of New York (emails: richard.crump@ny.frb.org, argia.sbordone@ny.frb.org). Eusepi: University of Texas at Austin (email: stefano.eusepi@austin.utexas.edu). Giannone: Amazon (email: dgiannon2@gmail.com). Qian: Princeton University (email: ericqian@princeton.edu) Giannone's contribution to the paper was completed prior to the author joining Amazon. This publication and its contents are not related to Amazon and do not reflect the position of the company and its subsidiaries.

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the author(s) and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the author(s).

To view the authors' disclosure statements, visit
https://www.newyorkfed.org/research/staff_reports/sr976.html.

1 Introduction

Vector autoregressive (VAR) techniques—popularized by a series of papers by Chris Sims—are a powerful tool in macroeconomics (Sims, 1972, 1980b,a). These atheoretic dynamic systems were developed in response to a growing skepticism toward the “incredible identifying restrictions” that plagued large-scale macroeconomics models (see Christiano, 2012). VARs enable researchers to forecast time series, evaluate economic models, and produce counterfactual policy experiments (Sims, 1980b). However, because of the high level of generality and the ability to capture numerous complex dynamic relationships among variables, standard VAR models suffer from the “curse of dimensionality” where by the number of parameters can quickly outstrip the number of observations as variables are added to the model. The introduction of Bayesian methods was crucial in addressing the resulting parameter instability. In the 1980s, Robert Litterman began using a small 6-variable Bayesian vector autoregressive model (BVAR) for macroeconomic forecasting at the Federal Reserve Bank of Minneapolis by using a prior that shrinks the variables¹ toward a random walk (Litterman, 1980). This simple 6-variable model was a milestone for the advancement of big data in macroeconomics, setting the stage for the future development of richer and more accurate models. Sims and Zha (1999) built on Litterman’s advances by introducing credible error bands for forecasts and impulse response functions, essential for fully characterizing uncertainty. Furthermore, Waggoner and Zha (1999) developed Bayesian methods for characterizing the distribution of conditional forecasts within a VAR modeling framework.

Similar to dynamic stochastic general equilibrium (DSGE) models, VARs were developed to overcome the critique of Lucas (1976) to large-scale macroeconomic models. Indeed, VARs are used to provide guidance and validation for micro-founded macroeconomic models (Rotemberg and Woodford, 1997; Christiano, Eichenbaum, and Evans, 2005; Del Negro, Schorfheide, Smets, and Wouters, 2007). Despite the theoretical appeal of these modelling

¹The model variables were real GNP, GNP deflator, real business fixed investment, 3-month Treasury bill rate, unemployment rate, and money supply.

frameworks, policymakers often rely on large-scale structural models of the economy in the tradition of the Cowles foundation. The most prominent example is FRB/US, the workhorse macro model used by Federal Reserve. As discussed in Tetlow (2015), “the relevance of the model is unquestionable: since its introduction in July 1996, the FRB/US model has been used continuously for communicating ideas to the Board of Governors and the Federal Open Market Committee (FOMC). All of the staff’s alternative scenarios focusing on domestic economic issues during this period were conducted using the model; forecast confidence intervals are computed using FRB/US, as are optimal policy exercises that are presented to the FOMC” (Tetlow, 2015, p. 115). Similar models are also adopted in other policy institutions such as the European Central Bank’s Multi Country (MC) model, which “at its core, [it] is designed along the lines of the Federal Reserve’s FRB/US model” (Constancio, 2017).

Why are policy institutions still relying on these large-scale models? It is likely because of their ability to jointly model a large number of economic and financial variables, encompassing more facets of the economy. However, recent advances in Bayesian methods have made estimation of large-scale, reduced-form models feasible (e.g. De Mol, Giannone, and Reichlin, 2008; Bańbura, Giannone, and Reichlin, 2010; Koop, 2013; Bańbura, Giannone, and Lenza, 2015; Carriero, Clark, and Marcellino, 2015; Giannone, Lenza, and Primiceri, 2015). In this paper, we develop a comprehensive modeling framework based on a large BVAR of the U.S. economy paired with a conditional forecasting approach, which provides an intuitive, flexible, and powerful framework for policy analyses. The model includes 31 economic and financial variables spanning the universe of key indicators as used in the Federal Reserve Board’s Tealbook A² (judgmental forecast) and the FRB/US model. Importantly, this breadth of indicators allows for a single model to be used for a wide slate of applications, which enables increased transparency of communication and comparability of results across empirical exercises. To demonstrate that our large-scale BVAR model is well-suited

²Discussion of the Tealbook and Greenbook may be found at https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm.

to economic forecasting applications, we show that the real-time forecasting performance of the model is comparable to professional forecasters and Greenbook forecasts.

We center our BVAR policy framework around the concept of conditional forecasting. The appeal of relying on conditional forecasts is that they connect the long tradition of scenario analyses conducted at central banks to structural analyses based on estimated impulse response functions to identified shocks common in the academic literature ([Angelini, Lalik, Lenza, and Paredes \(2019\)](#)). The difference between conditional and unconditional forecasts in the VAR can be interpreted as the impulse response function to the combination of shocks that is the most likely to generate the specified scenario. We show that one can use carefully designed scenarios to narrow down the set of possible shocks, which makes conditional forecasts a useful alternative to existing approaches in VAR modeling. To emphasize the link between our approach and the traditional VAR toolkit, we begin by showing generalized impulse responses in our model and draw parallels to Cholesky and [Blanchard and Quah \(1989\)](#) identification schemes. Recursive identification schemes can be replicated exactly through the appropriate conditional forecasting exercise. Similarly, identification via longer-run restrictions can be replicated in a directly interpretable way by conditioning on long-horizon forecasts ([Uhlig, 2004](#)). The conditioning approach can be expanded to accommodate uncertainty in the conditional paths; we use entropic tilting to shift the predictive density in order to match aspects calibrated from external information. This is a natural approach, for example, to anchor long-run inflation forecasts in the presence of a central bank target for inflation. Here, one may not want to impose that the target is exactly achieved in the medium or longer term.

To demonstrate the flexibility and power of our modeling framework, we cover a wide variety of policy-relevant applications. First, we leverage the ability to accommodate a large number of variables – including a wide variety of financial indicators – to decompose the 2015:Q2 episode of financial turmoil into, loosely speaking, a supply shock (from lower oil and commodity prices) and a demand shock (from tighter financial conditions). These

transmission mechanisms are clear only after explicitly modeling the dynamic interactions of a large number of financial and economic indicators.

We apply the large BVAR to the study of fiscal policy and the real economy, estimating the effect of the 2017 Tax Cuts and Jobs Act (TCJA) on economic activity using conditional forecasts based on the path of cyclically-adjusted tax revenue. Adding tax variables, we demonstrate the ease by which the BVAR can be expanded for specific applications. Here we show how to use revenue projections from external sources, to inform the scenario analyses. We also show how additional scenario analyses can allow to estimate the required elasticities via generalized impulse response functions. We find that with the TCJA, the level of output is more than 1% higher after three years relative to the unconditional forecast.

We can also use the model to assess possible changes in the structure of the economy (see also [Giannone, Lenza, and Reichlin \(2019\)](#)). We fit the model to the pre-2008 data and examine the distribution of inflation and asset prices conditional on the realized paths of real economic activity in the post-crisis period. Finally, we use the philosophy of conditional analyses to construct an internally consistent financial conditions index (FCI) which measures (in GDP units) the effect of financial shocks on macroeconomic activity. Specifically, we define the index as the counterfactual history of real GDP growth in the presence of only financial shocks and the absence of macroeconomic shocks. The large BVAR is well-suited for the task because it summarizes the dynamics of a large number of financial and economic indicators, allowing for heterogeneous transmission mechanisms of financial shocks. Although our FCI co-moves somewhat with alternative measures, it shows distinct behavior in sub-periods, which suggests that our approach is capturing different aspects of the relation between the real economy and financial conditions.

We view our BVAR framework as a natural complement to DSGE models for policymakers. Analyses from the BVAR require the maintained assumption of unchanged expectations formation and policy relative to the conditions that prevailed historically and therefore are best suited to analyze “modest” policy interventions (see [Leeper and Zha, 2003](#)). In con-

trast, structural models such as DSGE models are more suitable for studying changes in the conduct of systematic policies that are large and persistent. However, these models also have meaningful limitations. For example, they rely on a number of behavioral assumptions, such as rational expectations and Ricardian equivalence, and specific parameterizations of preferences and technology. They also are typically limited in scope and size as compared to the BVAR and their extensions are often very costly: adding new variables and sectors requires, in most cases, a full re-specification of the structure of the economy. Our methodology based on VARs, in contrast, allows for straightforward modifications enabling real-time policy analysis and inference.

The remainder of the paper is structured as follows. In Section 2 we introduce the modeling and estimation framework along with the economic and financial time series we use in our empirical applications. Section 3 provides a summary of the different approaches to conditional forecasting and scenario analyses that the BVAR accommodates. We then introduce a number of empirical applications of our approach in Section 4. Section 5 concludes. Additional empirical results along with full details of data sources are relegated to the Appendix.

2 Model and Data

2.1 Data

Bayesian shrinkage allows us to handle a large number of series. Our baseline dataset contains 31 quarterly variables, with the goal of capturing U.S. macroeconomic and financial conditions. To do this, we begin with the variables selected in the Federal Reserve Board of Governor’s Tealbook A Greensheets as a guide.³ The Greensheets contain projections for a set of economic indicators and are produced before each regularly scheduled FOMC meeting

³Additional details are available at the Federal Reserve Board of Governors’ site https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm

by the Federal Reserve Board of Governors research staff. For our purposes, the Greensheets are a useful starting point for variable selection as they are relevant in informing discussions among policymakers. They include real indicators like real GDP, real personal consumption expenditures, real business fixed investment, and real federal government consumption. They also include price indicators like headline CPI, core CPI, headline PCE Price Index, core PCE Index, and the GDP deflator. Other indicators like the unemployment rate, housing starts, and the industrial production are also included.

We also add widely-tracked indicators that summarize financial conditions—like the 2-year Treasury yield, 10-year Treasury yield, Moody’s Aaa Corporate bond yield, Moody’s Baa Corporate bond yield, trade-weighted dollar, and stock prices. To avoid complications with the zero lower bound, we omit the federal funds rate. Our primary dataset is fully described in Table 1, which also displays data sources, units, statistical transformations, and the chosen prior. Variables natively available at higher frequencies are averaged in order to obtain a quarterly time series, and values are only used when the whole quarter’s data are available. Note that while our analysis of the Tax Cuts and Jobs Act described in the Scenario Analysis section includes four additional tax variables, for parsimony we omit these from the baseline (the dynamics of the tax base and tax revenue are not material for the other scenarios).

For our analysis, we use a non-stationary VAR. Rates enter the model in levels, while most other variables enter in log levels. For interpretation and analysis, log-level forecasts are transformed to growth rates ex-post—either quarter over quarter or year over year depending on the specific application.

Series	Source	Units	Transformation	SPF	Greenbook	Prior
Real Gross Domestic Product	BEA	SAAR, Bil.Chn.2012\$	100 × log	X	X	RW
Real Personal Consumption Expenditures	BEA	SAAR, Bil.Chn.2012\$	100 × log	X	X	RW
Real Private Residential Fixed Investment	BEA	SAAR, Bil.Chn.2012\$	100 × log	X	X	RW
Real Private Nonresidential Fixed Investment	BEA	SAAR, Bil.Chn.2012\$	100 × log	X	X	RW
Real Exports of Goods & Services	BEA	SAAR, Bil.Chn.2012\$	100 × log			RW
Real Imports of Goods & Services	BEA	SAAR, Bil.Chn.2012\$	100 × log			RW
Real Govt. Consumption Expenditures & Gross Investment	BEA	SAAR, Bil.Chn.2012\$	100 × log	X	X	RW
Real Fed. Consumption Expenditures & Gross Investment	BEA	SAAR, Bil.Chn.2012\$	100 × log			RW
Gross Domestic Product: Chain Price Index	BEA	SA, 2012=100	100 × log	X	X	RW
Personal Consumption Expenditures: Chain Price Index	BEA	SA, 2012=100	100 × log	X	X	RW
PCE less Food & Energy: Chain Price Index	BEA	SA, 2012=100	100 × log	X	X	RW
CPI-U: All Items	BLS	SA, 1982-84=100	100 × log	X	X	RW
CPI-U: All Items Less Food & Energy	BLS	SA, 1982-84=100	100 × log	X	X	RW
Business Sector: Compensation Per Hour	BLS	SA, 2012=100	100 × log			RW
All Employees: Total Nonfarm	BLS	SA, Thous	100 × log			RW
Civilian Unemployment Rate: 16 yr +	BLS	SA, %	Raw	X	X	RW
Industrial Production Index	Federal Reserve Board	SA, 2012=100	100 × log	X	X	RW
Capacity Utilization: Manufacturing [SIC]	Federal Reserve Board	SA, Percent of Capacity	Raw			RW
Housing Starts	Census Bureau	SAAR, Thous.Units	100 × log	X	X	RW
Real Disposable Personal Income	BEA	SAAR, Bil.Chn.2012\$	100 × log			RW
University of Michigan: Consumer Sentiment	University of Michigan	NSA, Q1-66=100	Raw			RW
2-Year Treasury Note Yield at Constant Maturity	Federal Reserve Board	% p.a.	Raw			RW
5-Year Treasury Note Yield at Constant Maturity	Federal Reserve Board	% p.a.	Raw			RW
10-Year Treasury Note Yield at Constant Maturity	Federal Reserve Board	% p.a.	Raw	X		RW
Moody's Seasoned Aaa Corporate Bond Yield	Federal Reserve Board	% p.a.	Raw	X		RW
Moody's Seasoned Baa Corporate Bond Yield	Federal Reserve Board	% p.a.	Raw	X		RW
Nom. Trade-Weighted US\$ Index vs Major Currencies	Federal Reserve Board	Mar-73=100	100 × log			RW
Stock Price Index: Standard & Poor's 500 Composite	Standard & Poor's	1941-43=10	100 × log			RW
Spot Oil Price: West Texas Intermediate	EIA/WSJ	\$/Barrel	100 × log			RW
S&P GSCI Non-Energy Commodities Nearby Index	Standard & Poor's	AVG, Jan-02-70=100	100 × log			RW
Standard & Poor's 500 Composite Realized Volatility	Standard & Poor's	1941-43=10	Raw			WN

Table 1: Description of data: The above table gives the series, source, units, and transformation. Rates enter in levels. Other series enter in scaled log levels.

2.2 Model and inference

To describe the dynamic interactions among a large set of variables, we specify a very general VAR model:

$$y_t = c + B_1 y_{t-1} + \dots + B_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma), \quad (1)$$

where B_1, \dots, B_p are matrices of VAR coefficients, y_t is an n -dimensional vector of variables, and ε_t is an n -dimensional vector of innovations. For a large number of series, estimating the model using traditional VAR methods runs the risk of over-parametrization, as the number of unrestricted parameters is high. The resulting model would likely be unreliable due to overfitting. To account for this concern, we use a Bayesian approach. This procedure addresses the “curse of dimensionality” by automatically selecting the degree of shrinkage, using tighter priors when the number of unknown coefficients relative to available data is high, and looser priors otherwise. As discussed by [De Mol, Giannone, and Reichlin \(2008\)](#); [Bańbura, Giannone, and Reichlin \(2010\)](#) informative priors make inference viable and reliable in high-dimensional macroeconomic systems that are characterized by substantial comovement.

Inference is conducted following [Giannone, Lenza, and Primiceri \(2015\)](#). The priors belong to the normal-inverse-Wishart family. letting $\beta \equiv \text{vec}([C, B_1, \dots, B_p]')$, we can write the prior as

$$\beta \mid \Sigma \sim N(b, \Sigma \otimes \Omega) \quad (2)$$

$$\Sigma \sim IW(\Psi; d) \quad (3)$$

where elements Ψ, d, b and Ω are functions of a lower dimensional vector of hyperparameters γ . For a prior on the covariance matrix of innovations Σ , the degrees of freedom are set to $n + 2$.

We use the Minnesota prior, originally described in [Litterman \(1979\)](#), which is based

on the assumption that each variable follows an independent random walk process potentially with drift, a parsimonious yet “reasonable approximation for the behavior of economic variables.” This belief can be expressed in terms of the moments of the VAR coefficients:

$$\mathbb{E}[(B_s)_{ij}|\Sigma, \lambda, \Psi] = \begin{cases} 1, & \text{if } i = j \text{ and } s = 1. \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

$$\text{cov}[(B_s)_{ij}, (B_r)_{hm}|\Sigma, \lambda, \Psi] = \begin{cases} \lambda^2 \frac{1}{s^2} \frac{\Sigma_{ih}}{\Psi_{jj}}, & \text{if } m = j \text{ and } s = r. \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

In other words, Equation (4) shows that the expectation of the VAR coefficient corresponding to the own variable’s first lag is 1, while all others are zero. Equation (5) indicates that for greater lags s , prior uncertainty decreases, while the parameter λ controls the overall degree of shrinkage. The term Σ_{ih}/Ψ_{jj} accounts for the relative scale of each variable.

We also include the “sum-of-coefficients prior.” Consider a VAR in error correction form

$$\Delta y_t = c + \Pi y_{t-1} + A_1 \Delta y_{t-1} + \dots + B_{\tilde{p}} \Delta y_{t-\tilde{p}} + \varepsilon_t,$$

where $\tilde{p} = p - 1$, $A_s = -B_{s+1} - \dots - B_p$, $s = 1, \dots, \tilde{p}$, and $\Pi = B_1 + \dots + B_p - I_n$. Our prior shrinks Π to zero, which is simply a VAR in first differences, implying the presence of a unit root in each equation. We introduce the hyperparameter μ which controls the degree of shrinkage on the parameter Π .⁴ The hyperparameters λ, Ψ and μ are treated as unknown and we conduct formal inference according to the theoretically grounded approach of [Giannone, Lenza, and Primiceri \(2015\)](#).

The draws from the posterior are computed using the algorithm proposed by [Gian-](#)

⁴This prior is rudimentary because it is not centered on cointegrating relationships that are predicted by economic theory. [Giannone, Lenza, and Primiceri \(2019\)](#) show how to adapt it to incorporate theoretical predictions about the long run. The use of these priors tend to improve predictive accuracy especially in the long run. However, implementing them is challenging in our high-dimensional setup. Hence, we leave this extension for future work.

none, Lenza, and Primiceri (2015). This allows us to incorporate uncertainty on the hyper-parameters.

2.3 Forecasting and scenario analysis

The unconditional predictive density is defined as

$$p(y_{T+1}, \dots, y_{T+h} | y_{1:T}) = \int p(y_{T+1}, \dots, y_{T+h} | y_{1:T}, \theta) p(\theta | y_{1:T}) d\theta$$

where we have collected all the model coefficients in θ . In practice, to compute the predictive density, we iterate the following two steps. First, draw the coefficients from their posterior, call it $\theta^{(m)}$. Second, to draw from $p(y_{T+1}, \dots, y_{T+h} | y_{1:T}, \theta^{(m)})$, we generate draws of the forecast errors $\varepsilon_{T+1}, \dots, \varepsilon_{T+h}$ and simulate the model forward. As we will see below, this can be also be done using a simulation smoother.

We will also compute scenarios in the form of conditional predictive densities

$$p(y_{T+1}, \dots, y_{T+h} | y_{1:T}, \mathcal{C}_T) = \int p(y_{T+1}, \dots, y_{T+h} | y_{1:T}, \mathcal{C}_T, \theta) p(\theta | y_{1:T}) d\theta$$

where \mathcal{C}_T is a conditioning set; scenarios on specific future realizations of a subset of observables are:

$$\mathcal{C}_T = \{\tilde{y}_{j_s, T+s}, s = 1, 2, \dots; \quad j_s \in \mathcal{J}_s \subset \{1, \dots, n\}\},$$

where $\{\mathcal{J}_s : s = 1, 2, \dots\}$ collects the variables in the system which are involved in the scenario.

In practice, to compute the predictive density, iterate the following two steps. First, draw the coefficients from their posterior, again call it $\theta^{(m)}$. Second, to draw from $p(y_{T+1}, \dots, y_{T+h} | y_{1:T}, \mathcal{C}_T, \theta^{(m)})$, we use Kalman filtering techniques. Effectively, variables without information on their future path can be treated as missing data. As a consequence, this step is easily and efficiently implemented using the the simulation smoother.

To produce conditional and unconditional forecasts, we follow the approach of [Bańbura, Giannone, and Lenza \(2015\)](#) and use Kalman filtering techniques that makes the computations viable and efficient in our large dimensional system. The VAR in equation (1) can be cast as a linear state space model. Define the measurement equation as

$$y_t = C_t S_t \tag{6}$$

and the transition equation as

$$S_{t+1} = G_t S_t + w_t, \quad w_t \sim \mathcal{N}(0, H_t). \tag{7}$$

where y_t is an n -dimensional vector of observables and S_t is an m -dimensional vector of states. w_t is a vector of i.i.d. errors. Then, for p total lags, define

$$S_t \equiv \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \\ c \end{pmatrix}, \quad C_t \equiv \begin{pmatrix} I_n & 0_{n \times np} \end{pmatrix}, \quad G_t \equiv \begin{pmatrix} B_1 & B_2 & \dots & B_p & 0_n \\ I_n & 0_n & \dots & 0_n & 0_n \\ \vdots & \ddots & \dots & \vdots & \vdots \\ 0_n & \dots & I_n & 0_n & 0_n \\ 0_n & \dots & 0_n & 0_n & I_n \end{pmatrix}, \quad H_t \equiv \begin{pmatrix} \Sigma & \dots & 0_n \\ \vdots & \ddots & \vdots \\ 0_n & \dots & 0_n \end{pmatrix}. \tag{8}$$

The size of C_t adapts to the presence of “missing” observations used in the generation of scenarios. To condition on a linear combination of variables (rather than individual variables), the state-space model just needs to be modified by adding an additional entry in the observation equation reflecting the desired linear combination.

Then, with the VAR in its state space representation, draws from the state vector S_t can

be obtained using a simulation smoother⁵ where

$$S_{t|T} \sim p(S_t|y_t, C_t, G_t, H_t), \quad t = 1, \dots, T. \quad (9)$$

Next, draws from the unconditional predictive density can be computed using the simulation smoother applied to a dataset appended with missing values at the end of the sample.

A useful way to report scenarios is to compute the difference between the expectation under the conditional and the unconditional forecasts. The former is referred to as the “scenario forecast” whereas the latter is referred to as the “baseline forecast.” Define the counterfactual point forecast conditional on the model parameters as

$$\mathbb{E}[y_{T+s}|y_{1:T}, \mathcal{C}_T, \theta] = \int y_{T+s} p(y_{T+s}|y_{1:T}, \mathcal{C}_T, \theta) dy_{T+s},$$

and the baseline point forecast as

$$\mathbb{E}[y_{T+s}|y_{1:T}, \theta] = \int y_{T+s} p(y_{T+s}|y_{1:T}, \theta) dy_{T+s}.$$

In other words, the unconditional forecast corresponds to the situation in which $\mathcal{C}_T = \emptyset$.

To compute these objects, there is no need to draw from the predictive distribution, they can just be computed using the Kalman smoother. The difference gives:

$$\mathbb{D}[y_{T+s}|y_{1:T}, \mathcal{C}_T, \theta] = \mathbb{E}[y_{T+s}|y_{1:T}, \mathcal{C}_T, \theta] - \mathbb{E}[y_{T+s}|y_{1:T}, \theta].$$

In order to account for uncertainty on the parameters θ , we will compute the difference, $\mathbb{D}[y_{T+s}|y_{1:T}, \mathcal{C}_T, \theta]$, for different draws of θ from its posterior $p(\theta|y_{1:T})$.

This object is strictly related to the impulse response function (IRF), which relates predictions of the variable of interest conditional on forecast errors rather than on observables.

⁵We use the simulation smoother of [Durbin and Koopman \(2002\)](#) for which the computational complexity grows with the number of variables, which is smaller than the number of states in our context.

The IRF is defined as:

$$\frac{\partial y_{i,T+s}}{\partial u_{j,T}} = \underbrace{\mathbb{E}[y_{i,T+s}|y_{1:T-1}, u_{j,T} = 1, \theta]}_{\text{Counterfactual}} - \underbrace{\mathbb{E}[y_{i,T+s}|y_{1:T-1}, \theta]}_{\text{Baseline}}$$

where $u_{j,T}$ is a structural shock identified using economic considerations, usually linear combinations of forecast errors $u_{j,T} = \delta' \varepsilon_T$ for some δ .

To understand the connection between the two objects, we can write the scenarios in terms of forecast errors:

$$\mathbb{E}[y_{T+s}|y_{1:T}, \mathcal{C}_T, \theta] =: \mathbb{E}[y_{T+s}|y_{1:T}, \mathcal{S}_T, \theta]$$

where

$$\mathcal{S}_T = \{\delta'_s \tilde{\varepsilon}_{T+s}, s = 1, 2, \dots\}$$

and where \mathcal{S}_T is the combination/sequence of shocks that is the most likely to have generated the path for the observables defined in \mathcal{C}_T . A sufficient number of restrictions can narrow down the type of shocks that are most likely the source of the scenario (specific examples are given below).

The simplest case is the generalized IRF (GIRF) introduced by [Koop, Pesaran, and Potter \(1996\)](#). In this case the mapping between conditioning on observables and conditioning on shocks is immediate. In terms of observables the GIRF is defined as $\mathcal{C}_T = \{\tilde{y}_{j,T+1}\}$ where $\tilde{y}_{j,T+1} = \mathbb{E}[y_{j,T+1}|y_{1:T}, \theta] + 1$. In terms of shocks, it is defined as $\mathcal{S}_T = \{\tilde{\varepsilon}_{j,T+1}\}$ where $\tilde{\varepsilon}_{j,T+1} = \tilde{y}_{j,T+1} - \mathbb{E}[y_{j,T+1}|y_{1:T}, \theta] = 1$.

Another case when the mapping between counterfactuals based on shocks and counterfactuals based on observables is in simple recursive identification. Consider a monetary policy shock to the policy instrument (say the r th variable in the system) identified by assuming that macroeconomic variables (say the first $r - 1$ without loss of generality) are slow, that is, that they cannot react contemporaneously. The remaining variables are instead fast finan-

cial variables that can react immediately to the monetary policy shock. Monetary policy is assumed to react only to the slow variables but with a lag of at least a period relative to the fast ones. In terms of shocks, the scenario is constructed by perturbing a linear combination of the forecast errors obtained using a Cholesky decomposition. In terms of observables, it is implemented as follows:

$$\mathcal{C}_T = \{\tilde{y}_{j,T+1}; \quad j = 1, \dots, r\}$$

$$\tilde{y}_{j,T+1} = \mathbb{E}[y_{j,T+1}|y_{1:T}, \theta] \text{ for } j = 1, \dots, r - 1$$

$$\tilde{y}_{r,T+1} = E(y_{j,T}|y_{1:T}, \theta) + 1$$

A special case is when r is equal to 1. In that situation, we obtain the generalized impulse response function. In the paper we construct several scenarios and connect them with different identification schemes used in the structural VAR literature.

Before concluding this section, we should note that we have defined the scenarios as predictions conditional on alternative events in the future. This is the most common and natural way to think about counterfactuals. However, the same approach can be used to condition on an alternative past. Computations are trivial once the problem is cast as projections in conditioning sets that have missing data.

2.4 Forecast Evaluation

The BVAR coupled with the conditional forecasting methodology provides a powerful tool for scenario analysis. However, these results are trustworthy to the extent that unconditional forecasts are reasonable. We find that despite the large number of parameters to estimate, the BVAR produces out-of-sample predictions comparable to those of professional forecasters.

For such comparisons, we begin by constructing a quarterly vintage dataset from 2005:Q1 to 2018:Q3 to account for both real-time data availability and data revisions. The vintage for each quarter corresponds to the due date of the Survey of Professional Forecasters, which

is typically a week after the first GDP release of the previous quarter. The model itself, including its hyperparameters, is estimated recursively on a sample from 1973:Q1 to the previous quarter. While the number of lags is fixed at five, we find that empirical performance is not sensitive to lag selection; recall that under our Bayesian framework, longer lags shrink toward zero. Variable selection is also fixed, a reasonable assumption considering that our specification is primarily taken from the Greenbook/Tealbook, whose variables have been essentially stable over time.

Our first source is the Survey of Professional Forecasters (SPF), which began in 1968 and has been managed by the Federal Reserve Bank of Philadelphia since 1990. Surveys are released in the middle of each quarter, and respondents are asked to provide quarterly point forecasts from zero to four quarters ahead for a selection of macroeconomic indicators. These forecasts are public, timely, well-used, and well-studied. Our second source is the Federal Reserve Board of Governors Greenbook (“The Greenbook”), which is a collection of forecasts of macroeconomic indicators produced by the Board’s research staff before every regular FOMC meeting. These forecasts help inform monetary policy decisions and are subject to a five-year publication delay. We use forecasts contained in the Federal Reserve Bank of Philadelphia’s Greenbook Data Set, using the Greenbook date closest to the SPF due date.

For comparability, our forecast evaluation focuses on marginal point performance. Professional forecasts, like those found in the SPF or Greenbook, are typically made for a single variable and rarely provide even univariate density forecasts. For policy analyses however, it is imperative that a model can provide good joint forecasts, as they can be used for scenario analyses through conditional predictive densities. The ultimate aim is to construct internally consistent narratives for both current conditions and the macroeconomic outlook.

We illustrate this point in Figure 1, which plots the joint one-year ahead forecasts made in 2008:Q4 for the unemployment rate and quarterly PCE inflation generated by our model (using real-time data). The black contours in the center panel display the joint predictive

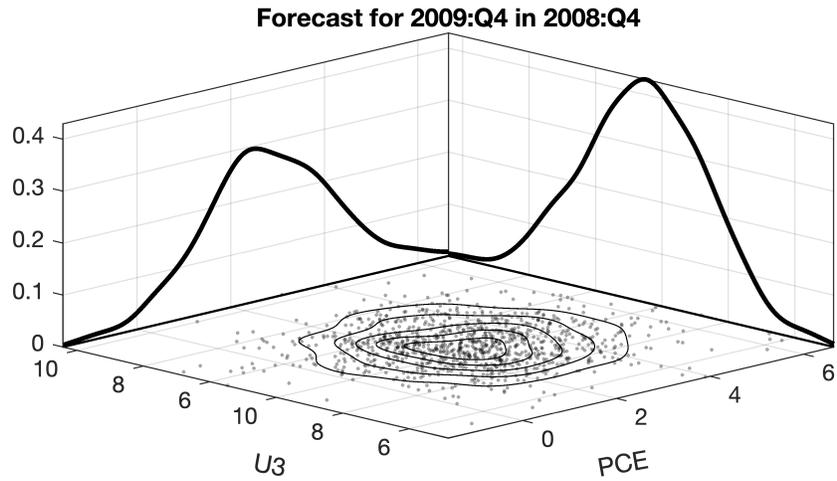


Figure 1. Joint forecasts in the Great Recession. Using the large BVAR, the above plot show the real-time distributional joint forecasts for quarterly PCE inflation and the unemployment rate made during 2008:Q4 for the 2009:Q4 reference period. The center dots and contours indicate the posterior predictive density. The side panels show the marginal distributions.

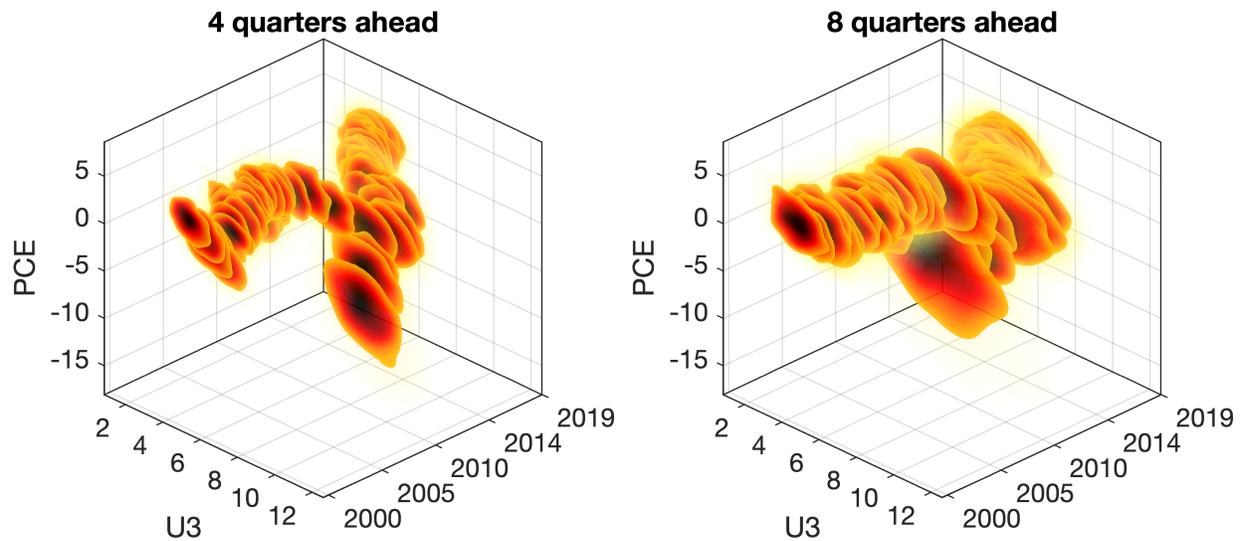


Figure 2. Joint forecasts. The above panels show joint density out of sample forecasts using the BVAR model for PCE inflation and unemployment rate. Darker areas indicate regions with higher probability, while lighter and transparent areas indicate regions with lower probability. Forecasts for one year ahead and two years ahead are shown in the two panels.

density, while the side panels show the marginal distributions. The model produces forecasts that are compatible with a Phillips curve-like relationship, where forecasts for a higher unemployment rate coincide with forecasts of lower inflation. These joint dynamics would be missed by only focusing on the marginal distributions displayed in the side panels.

Stitching together real-time forecasts, Figure 2 shows the same relationship over time. The subpanels show density forecasts for two horizons (one- and two-years ahead of the current quarter). Darker areas indicate regions of higher probability while lighter areas show regions of lower probability. As the horizon increases, the distributional forecasts become more disperse, reflecting greater uncertainty. As the forecast distributions evolve over time, we observe that inflation forecasts and unemployment forecasts co-move in a natural way reflecting the model’s internal consistency.

We next move to the evaluation of point forecasts. Figures 3 and 4 compare the empirical error distribution of the median BVAR, median SPF, and Greenbook forecasts for the current quarter (“nowcast”) and four quarters ahead. Errors are computed as the difference between the forecast and the latest release’s realization, scaled by the standard deviation of the BVAR’s empirical errors for visibility. Each dot gives a forecast error for vintages from 2005:Q1-2018:Q3. The bars indicate the inner 90% of the empirical error distribution, where the center hash gives the median. Due to the five-year publication lag of the Greenbook, the evaluation sample of the Greenbook is shorter than that of the SPF.

By exploiting monthly information, the SPF and the Greenbook have an advantage over the BVAR in producing current period forecasts. Given the timing of the SPF’s survey, professional forecasters have access to the previous month’s labor report as well as an additional month of financial and survey data. In contrast, only data corresponding to the previous quarter enter the BVAR. The advantage is reflected in the tighter error distributions of the SPF and Greenbook in Figure 3. Two approaches can be used in this context. One possibility is to incorporate external nowcasts by either treating them as data or more formally by tilting the predictive distributions produced by the VAR (see Krüger, Clark, and

Ravazzolo, 2017; Tallman and Zaman, 2020). Similar accuracy can, however, be achieved without judgment by using nowcasting techniques to incorporate high-frequency information in real time. These methods were originally developed using dynamic factor models (see Giannone, Reichlin, and Small, 2008; Bok, Caratelli, Giannone, Sbordone, and Tambalotti, 2018) and show that dynamic factor models can produce forecasts comparable to the Greenbook, SPF, and the first release of GDP produced by the BEA. Recently, they have also been adapted and successfully applied to a large BVAR (see Cimadomo, Giannone, Lenza, Sokol, and Monti, 2020). Although these extensions can be straightforwardly implemented, they venture beyond the scope of this paper.

For longer horizons the forecasts produced by our model are competitive. Consider the error distributions for one-year ahead forecasts in Figure 4. For some indicators (like investment and government spending), the BVAR’s error distribution is tighter, while for others, professional forecasters’ error distribution is tighter (like consumption and CPI). Despite the large number of parameters to be estimated in such a large system, the BVAR yields predictions comparable to those made by professional forecasters, lending credibility to the scenario analysis results. The BVAR’s accurate forecasts for our U.S. system are consistent with results for similar models of the UK and Euro area economies. [Domit, Monti, and Sokol \(2019\)](#) show that a Bayesian VAR outperforms the Bank of England’s COMPASS DSGE model in forecasting GDP and produces similar results in forecasting inflation. [Angelini, Lalik, Lenza, and Paredes \(2019\)](#) show that a multi-country Bayesian VAR of the four largest euro area countries produce accurate conditional forecasts and can mimic Eurosystem projections.

3 Scenario Analyses

In this section, we use conditional forecasting to analyze several scenarios. The breadth in subject material showcases the versatility of the large BVAR for policy applications. We

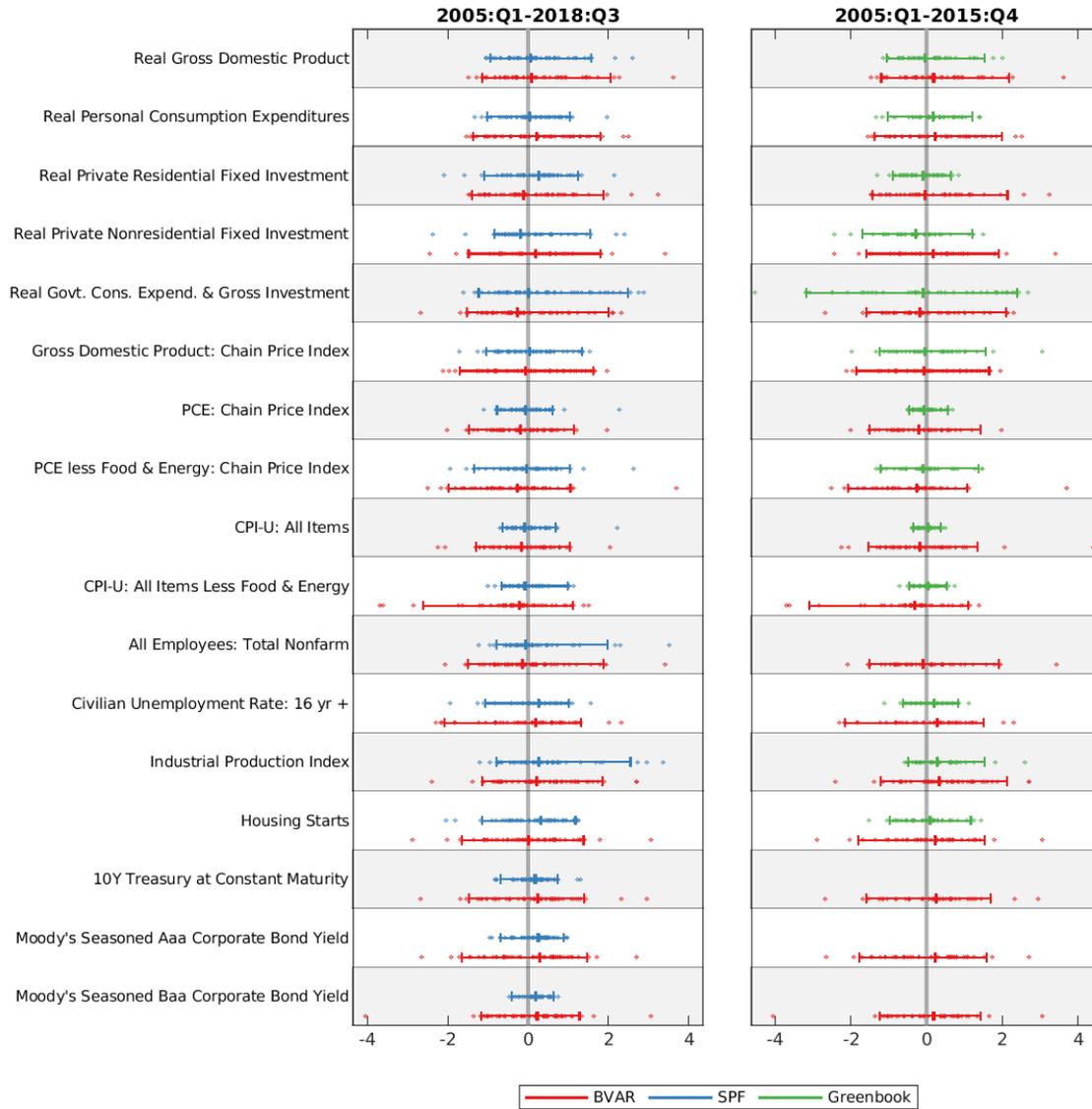


Figure 3. Empirical BVAR error distribution compared to the SPF and Greenbook (0 quarters ahead). Errors are computed as the difference between the forecast and the latest release’s realization, scaled by the standard deviation of the BVAR’s empirical errors for visibility. Each dot gives a forecast error for vintages from 2005:Q1-2018:Q3. The bars indicate the inner 90% of the empirical error distribution, where the center hash gives the median.

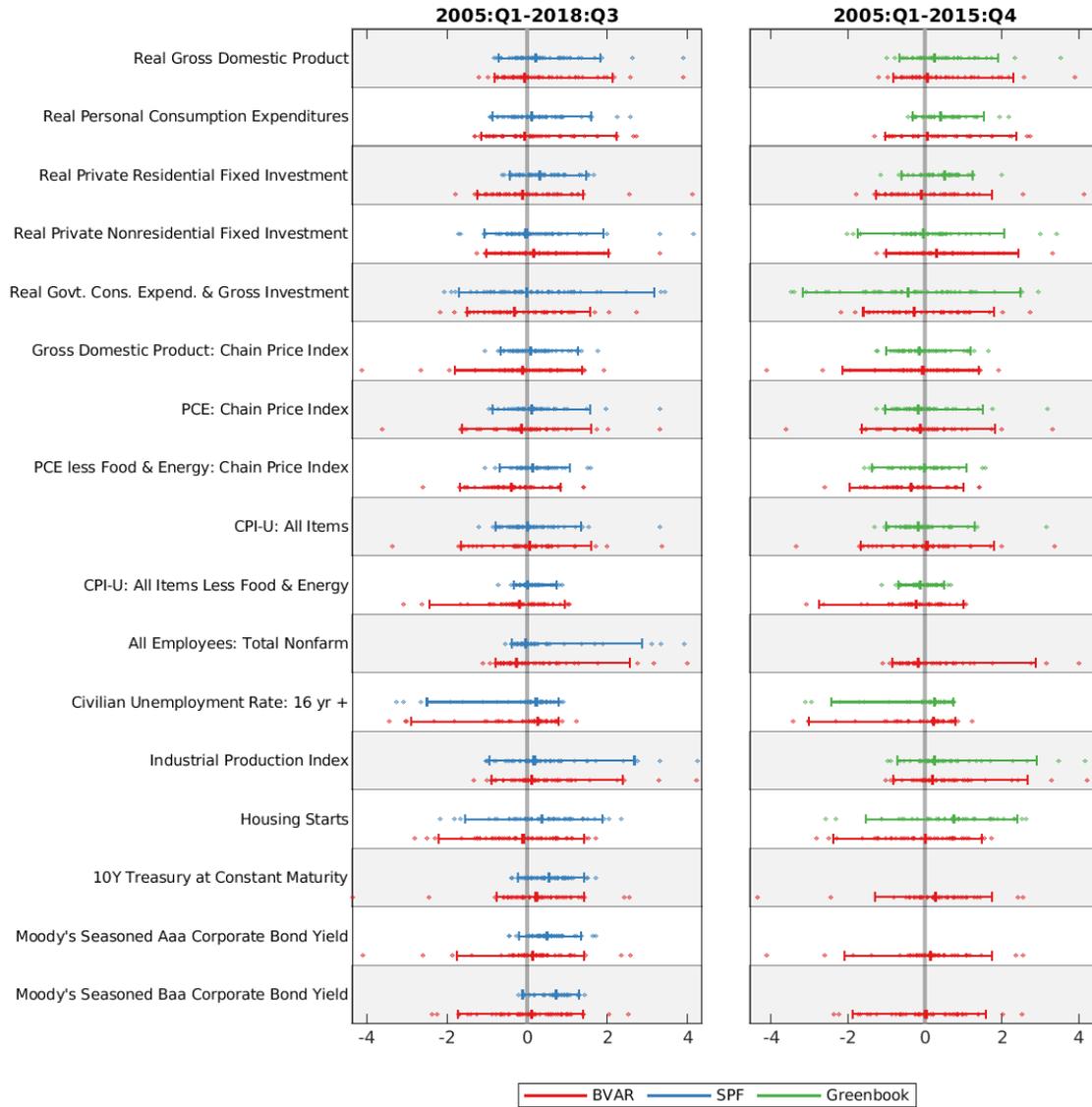


Figure 4. Empirical BVAR error distribution compared to the SPF and Greenbook (4 quarters ahead). Errors are computed as the difference between the forecast and the latest release’s realization, scaled by the standard deviation of the BVAR’s empirical errors for visibility. Each dot gives a forecast error for vintages from 2005:Q1-2018:Q3. The bars indicate the inner 90% of the empirical error distribution, where the center hash gives the median.

begin with two simple exercises on real GDP to illustrate our empirical approach. In the first, we increase the level GDP by one percent relative to the unconditional forecast. This scenario is analogous to a residual impulse response function in the traditional VAR context with unorthogonalized innovations. In the second, we increase the level of real GDP 8-12 quarters ahead by one percent—analogue to a [Blanchard and Quah \(1989\)](#) long-run identification strategy.

We then continue with two exercises on the 2-year Treasury yield. Just as before, we begin with an unorthogonalized IRF, increasing the Treasury yield by 25 basis points one period ahead, leaving all other variables unconstrained. Next, we perform the same exercise, but we constrain the economic variables to their unconditional forecast in the first period, analogous to a Cholesky identification strategy; economic activity responds to changes in financial conditions with delay.

Afterward, we explore several more articulated applications. First, we evaluate the effects of the 2015 turbulence in financial markets on macroeconomic outcomes. The large set of financial indicators allows us to decompose the event into a positive “supply” shock and negative “demand” shock. Second, we model the effects of the 2017 Tax Cuts and Jobs Act (TCJA). The primary challenge to identification is the endogeneity of output and tax revenue—this precludes using a simple shock to tax revenue for producing valid inference. Our approach is to condition on the linear combination of tax revenue and tax base, using the static budget impact of TCJA as an input.

3.1 Increase in Real GDP

To illustrate how conditional forecasting can be recast to an impulse response setting, we begin by mimicking an unorthogonalized impulse response function. We form conditional forecasts by increasing the one-period ahead forecasts for real GDP by one percent, while leaving all other variables unconstrained.

To illustrate, see [Figure 5](#), which shows the level of real GDP along with our proposed

scenarios. The top chart displays the unconditional forecasts (blue bands) and conditional forecasts (dotted red lines), where the bands correspond to 90% and 68% coverage intervals respectively. The solid lines correspond to the posterior median. To compute the resulting impulse response function, we simply take the difference between the conditional and unconditional scenarios (bottom chart).

Figure 6 displays the responses of 17 selected macroeconomic (top left panel), prices (top right panel), labor (bottom left panel), and asset price (bottom right panel) variables. Relative to the unconditional scenario, a one percent increase in the level of real GDP gives a higher level of GDP, higher PCE chain price index, higher core PCE chain price index, higher 2-year Treasury yield, and higher Aaa/Baa bond yields. The unemployment rate decreases on impact, but returns to the level of the unconditional scenario after three years.

As discussed in [Bańbura, Giannone, and Lenza \(2015\)](#), the difference between conditional and unconditional forecasts in the exercise above is equivalent to a generalized impulse response function to GDP, which corresponds to the a shock to GDP using a recursive identification with GDP ordered first. If there exists one and only one exogenous shock that moves GDP, the exercise above would have the structural meaning. However, there is no economic reason why this would be the case, hence the scenario is driven by the combination of shocks that explains the bulk of forecast errors in GDP.⁶ Next, we show that the careful design of the scenarios can narrow down the set of possible shocks to more meaningful economic concepts.

For our next exercise, we condition on a one percent increase in the level of real GDP 8-12 quarters ahead, while leaving all other variables unconstrained. The motivating example for this scenario is [Blanchard and Quah \(1989\)](#) who identify the long-run shock as the driver of a permanent increase in the level of GDP. Our scenario is a “medium-run” version that considers

⁶The generalized IRF reflects the effects of the linear combination of structural shocks that explains most of the forecast error of GDP growth. As shown by [Del Negro, Lenza, Primiceri, and Tambalotti \(2020\)](#), these are similar to the IRF of a typical business cycle shock, defined by [Giannone, Lenza, and Reichlin \(2019\)](#) as the linear combination of structural shocks that drives the bulk of GDP variation at business cycle frequencies. For a recent application of business cycle shocks, see [Angeletos, Collard, and Dellas \(2020\)](#), [Ascari and Fosso \(2021\)](#), and [Caldara, Scotti, and Zhong \(2021\)](#).

the effect of the combination of shocks that are the most likely to lift the level of GDP two-years ahead. Figure 7 shows the responses for selected variables. By construction, the level of real GDP is one percent higher relative to the unconditional, as marked by the blue dots in the top left panel. At impact and along the path, the headline PCE price index, core PCE price index, and civilian unemployment rate decrease. The two-year Treasury is nearly unchanged at impact, but increases in the long run. The BAA and AAA yields decrease immediately, but are unchanged relative to the unconditional scenario in the long-run. These results indicate that supply shocks are the dominant force driving these scenarios. Relatedly, these are the types of shocks that [Blanchard and Quah \(1989\)](#) endeavored to isolate.

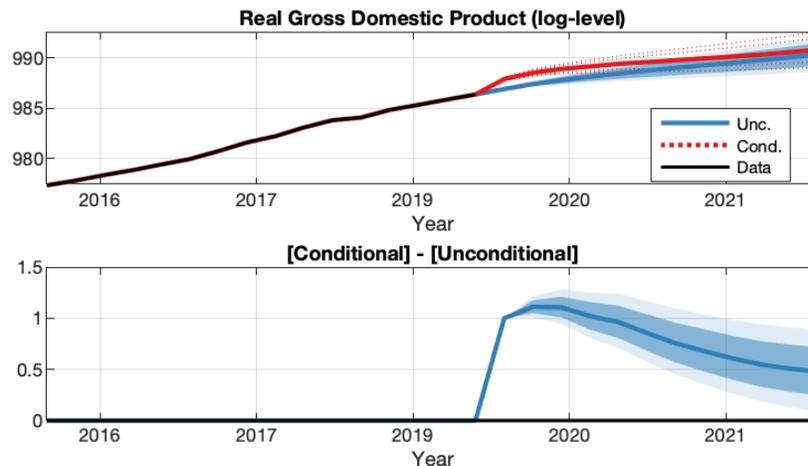


Figure 5. Conditional (dotted red lines) and unconditional forecasts (blue bands) for an increase in the level of real GDP one quarter ahead. The shaded bands correspond to 90% and 68% coverage intervals. The top row gives the unconditional and conditional forecasts in log levels. The bottom row gives the difference between the conditional and unconditional forecasts, where the blue dot indicates the conditioning assumption (a one percent increase in the level of GDP).

3.2 Increase in 2-year Treasury rate

Next, we present two exercises involving the 2-year Treasury rate. First, we condition on a 25 basis point increase in the 2 Year Treasury Rate one quarter ahead, while leaving all other variables unconstrained (analogous to an unorthogonalized IRF). As discussed earlier, this

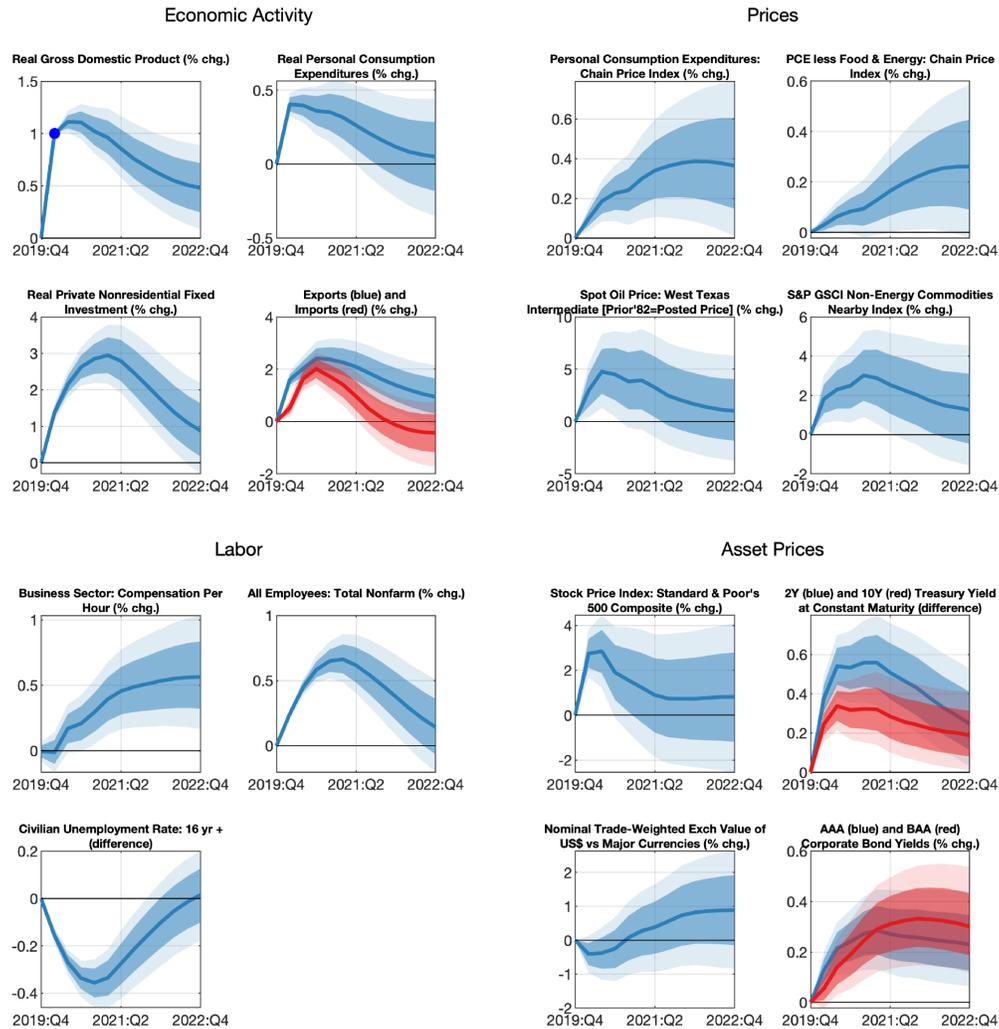


Figure 6. Responses to a 1% increase in the level of real GDP one quarter ahead. The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

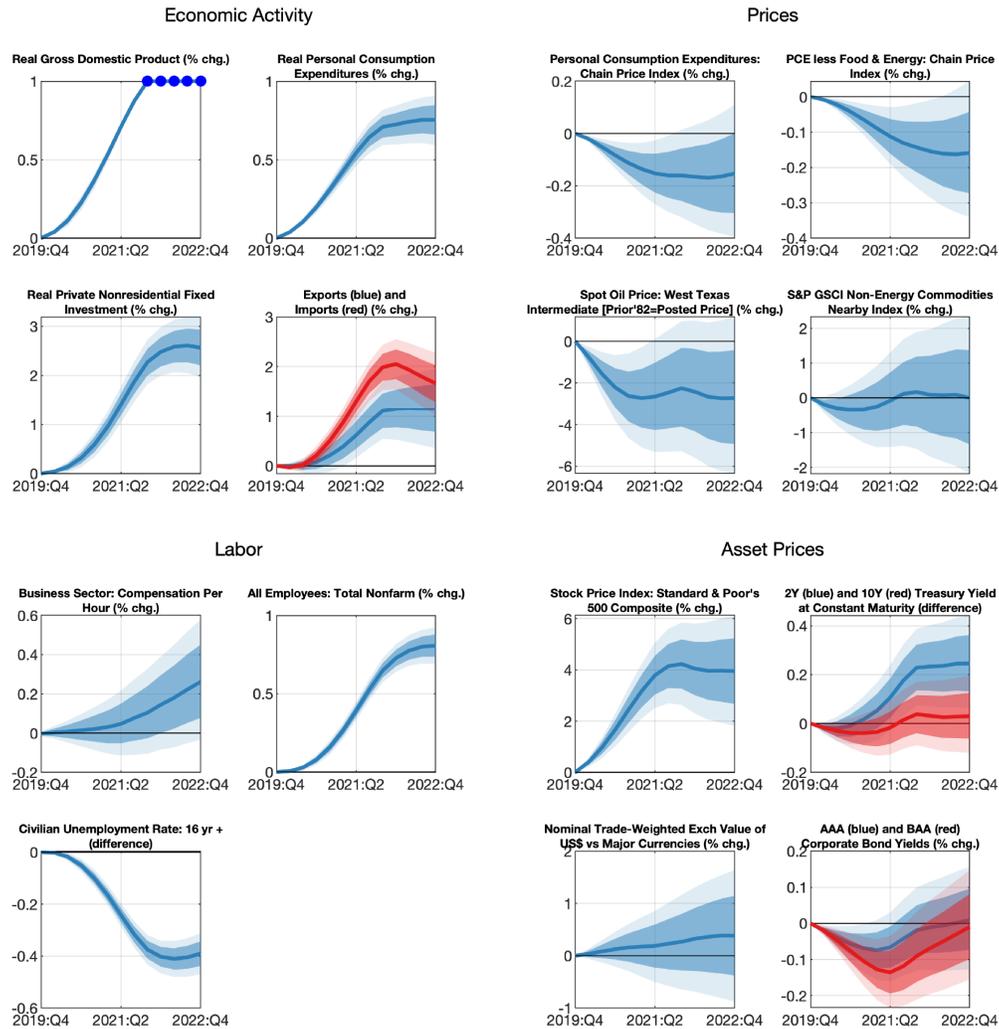


Figure 7. Responses to an increase in the level of real GDP 8-12 quarter ahead. The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

scenarios is driven by the combination of shocks that, based on historical correlations, is the most likely to increase the short term interest rate. In a second exercise we narrow down the set of possible shocks by including additional conditional assumptions. We again condition on a 25 basis point increase in the 2-year Treasury yield one quarter ahead, but we fix the value of one-quarter ahead forecasts for economic variables equal to their unconditional forecast. After taking the difference, the response of economic variables at the impact is zero. This exercise is grounded in the reasoning used for recursive identification: financial conditions respond to economic conditions contemporaneously, while economic conditions respond to financial conditions with a lag. Indeed, the difference between scenario and baseline is equivalent to the IRF of a monetary policy shock identified recursively, with slow moving economic variables and fast moving financial variables (see [Bańbura, Giannone, and Lenza, 2015](#)). If exogenous monetary policy shifts materialize in a lift of the short end of the yield curve without affecting the slow variables, the exercise has a structural interpretation and provides an assessment of the transmission mechanism of monetary policy shocks ([Sims, 1980b](#); [Rotemberg and Woodford, 1997](#); [Christiano, Eichenbaum, and Evans, 1999](#)). In the general case, the scenario will capture the transmission of the combination of shocks that are the most likely to shift the short-term interest rate without affecting immediately the macroeconomy.

Figure 8 shows responses in the unconstrained scenario, while Figure 9 shows responses under the restricted scenario. The constrained responses are in line with what one would expect from a monetary policy shocks: Economic activity deteriorates, stock prices go down, corporate financial conditions deteriorate as the corporate bond yields increase more than Treasuries, and the exchange rate appreciates. The unconstrained scenario is difficult to interpret since it is driven by a multiplicity of shocks of different types. Indeed, the forecast errors in the short-term rate reflect both exogenous monetary policy shifts and the systematic policy response to the evolution of the economy. For example, at impact and in the short-run, real GDP is higher relative to the unconditional scenario, and is lower in

the long-run. Without restrictions, the initial increase can be a consequence of the data correlation structure; higher than expected Treasury rates are related to higher output. In contrast, with the contemporaneous restriction of economic variables in Figure 9, the response dynamics differ. At impact, real GDP's response is identical to the unconditional scenario by construction (as noted by the blue dots at 0%). After impact, the level of real GDP is immediately lower relative to the unconditional scenario. The resulting dynamics of the unemployment rate are similar. Under the unrestricted scenario, the unemployment rate decreases immediately after impact before increasing in the medium run. Under the restricted economic scenario, the unemployment rate increases two periods ahead. These results are consistent with those obtained from Cholesky identification in smaller (see [Stock and Watson, 2001](#), for a survey) and larger systems (see [Bańbura, Giannone, and Reichlin, 2010](#); [Giannone, Lenza, and Primiceri, 2015](#); [Giannone, Lenza, and Reichlin, 2019](#)).

3.3 Assessing the effect of the 2015 financial market turmoil

In 2015:Q2, U.S. financial conditions sharply tightened: stock market volatility spiked, the stock market declined, and yields increased. Concurrently, commodity prices declined and the dollar appreciated relative to other currencies. By explicitly modeling the dynamic interactions among a large set of macroeconomic and financial variables, the Bayesian VAR presents a natural tool for assessing the macroeconomic effects of this episode. We find that the 2015:Q2 episode is, loosely, a composition of a supply shock (from lower oil and commodity prices) and a demand shock (from tighter financial conditions).

We begin by estimating the usual model on the full set of 31 macroeconomic and financial variables from 1973:Q1 to 2019:Q1. While our scenario is taken from events occurring in 2015, we take the underlying correlations as time-invariant. Then, just as before, we construct scenarios. Our baseline is the unconditional forecast based on economic and financial data through 2015:Q2, which represents a world without financial market turbulence. Then, in our first scenario, we also condition on the quarterly realizations of financial variables for

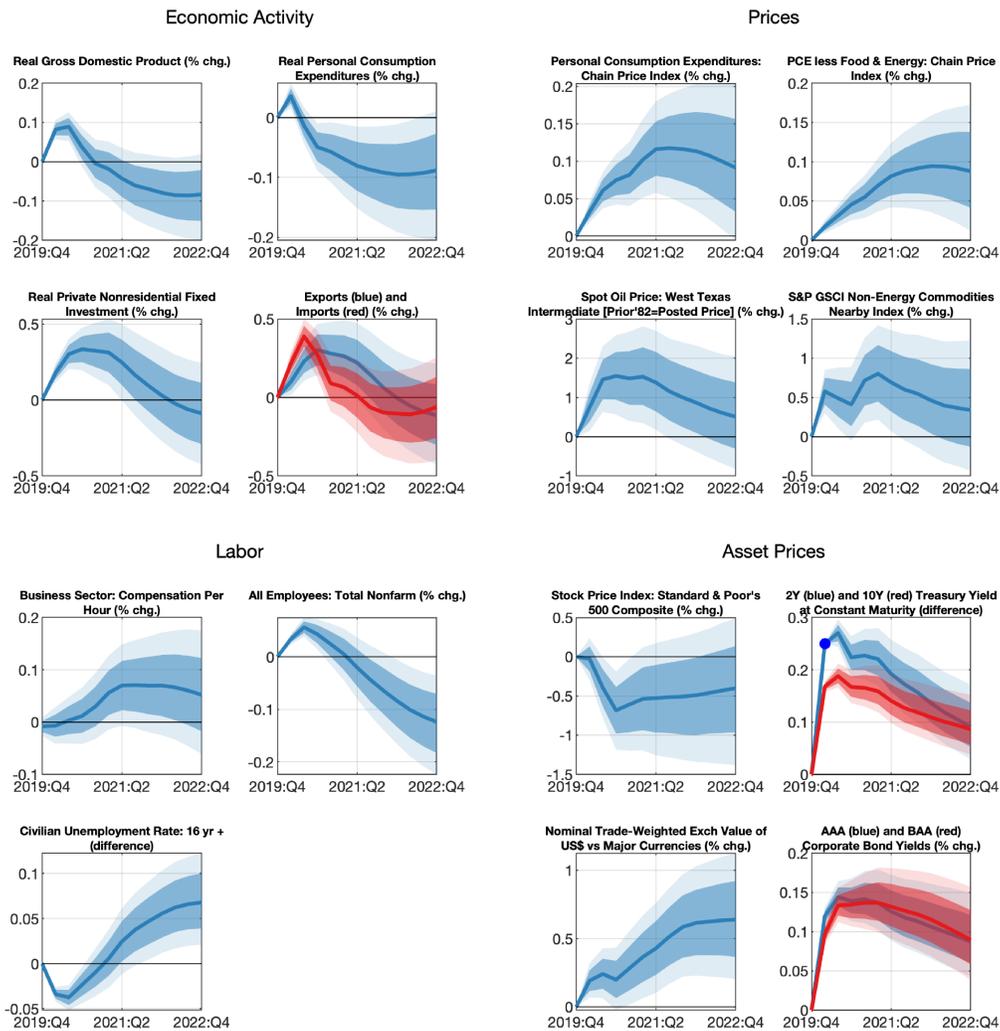


Figure 8. Responses to a 25 basis point increase in the 2-year Treasury yield one quarter ahead. The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

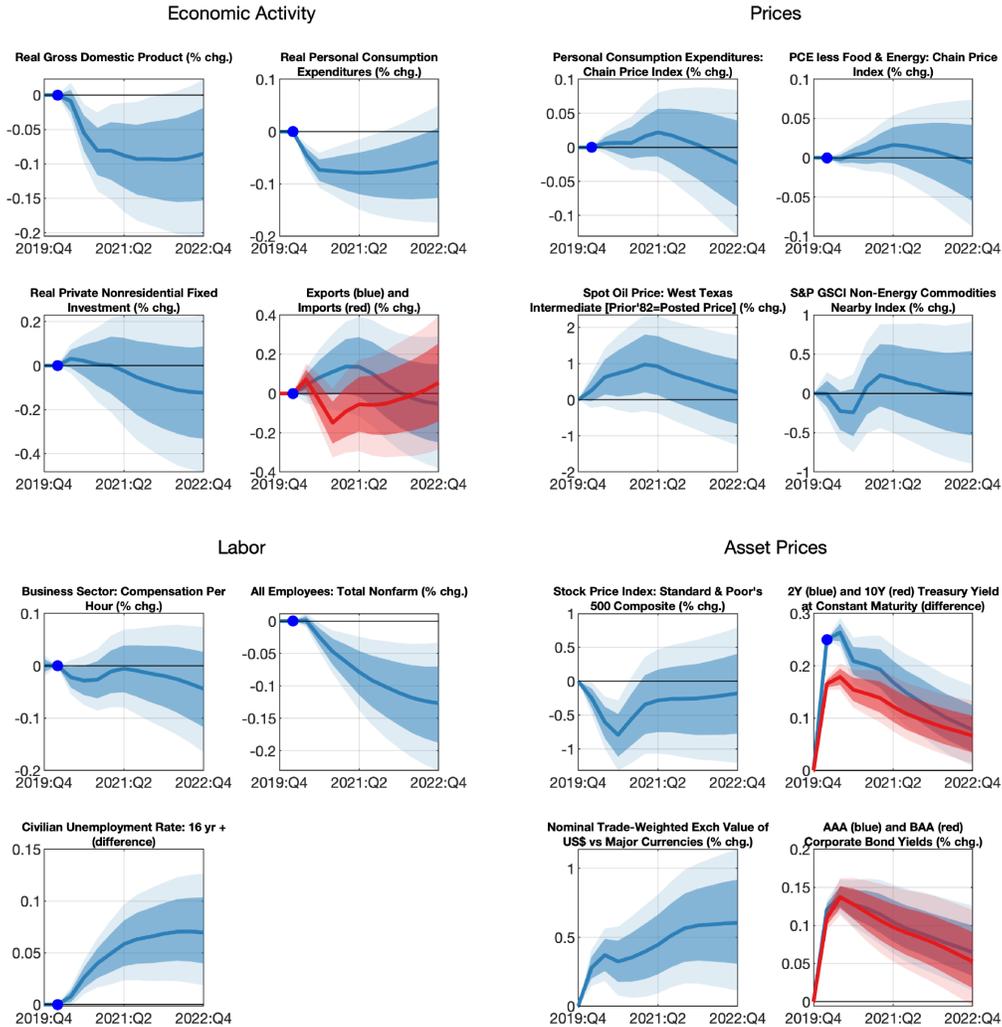


Figure 9. Responses to a 25 basis point increase in the 2-year Treasury yield one quarter ahead, holding constant the path of macroeconomic variables. The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

2015:Q3. The counterfactual gives a world with “turbulent” financial markets. We label this the “financial turmoil” scenario.

Figure 10 shows the resulting impulse response functions. The indicators corresponding to real economic activity behave as one would expect. GDP growth decreases for the first three quarters, but recovers two years ahead. The unemployment rate increases by 0.4% relative to the baseline three quarters ahead, but recovers fully after two years. Prices decline over the three-year horizon, with a consistently lower price levels relative to the baseline in the headline and core PCE price indices. More strikingly, oil and commodity prices sharply decline two quarters ahead and remain below the baseline. Financial conditions remain tight. The S&P 500 sharply declines relative to the baseline, and doesn’t fully recover after three years.

The concurrent declines in the stock market and oil/commodity prices hint at the existence of two transmission mechanisms. Tighter financial conditions make borrowing more difficult, suppressing output. Lower oil and commodity prices lower manufacturing costs, increasing output. To distinguish the two mechanisms, we construct an additional scenario. We condition only on bonds, the stock price, realized volatility, and the exchange rate for 2015:Q3; we do not condition on commodity or oil prices. This scenario would produce impulse responses in the absence of a positive supply shock stemming from reduced oil and commodity prices; we would expect differing responses if the financial market turmoil were indeed the composition of two distinct shocks. We will call this the “no supply shock” scenario.

Figure 11 shows the resulting impulse response functions. Compared to the “financial turmoil” scenario, the decline in real GDP three-quarters ahead is deeper and more persistent. While in the “financial turmoil” scenario it recovers to the baseline after two years, in the “no supply shock” scenario it takes three years to recover. The labor market tells a similar story. The unemployment rate under the “financial turmoil” scenario peaks at 0.4% above the baseline, while in the “no supply shock” scenario peaks at 0.5% above the baseline. While

prices are lower than the baseline in the “no supply shock” scenario, they are still higher than those found in the “financial turmoil” scenario. The headline PCE chain price index for one period ahead, for instance, is about 0.4% lower than baseline in the “financial turmoil” scenario but is about 0.2% lower under the “no supply shock” scenario.

In summary, without conditioning on commodity/oil prices, we see a deeper decline in output relative to the “financial turmoil” scenario and a smaller decline in prices—providing evidence that tighter financial conditions can be viewed as a negative “demand” shock. Including the decline in oil and commodity prices results in lower prices and higher output, offering evidence of a positive “supply” shock. By explicitly modeling the dynamic interactions of a large number financial and economic indicators can the two transmission mechanisms be brought to light.

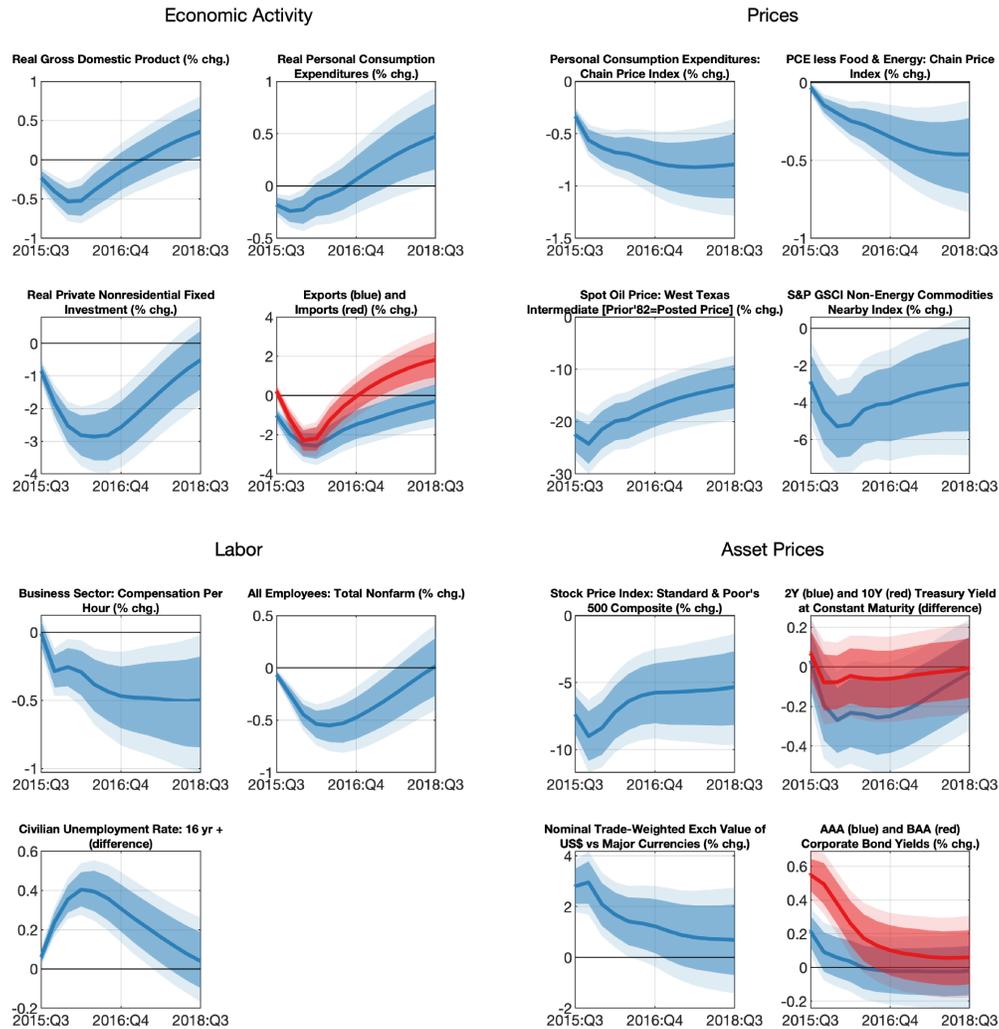


Figure 10. Responses to conditioning on 2015:Q3 financial variables (“financial turmoil” scenario). The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

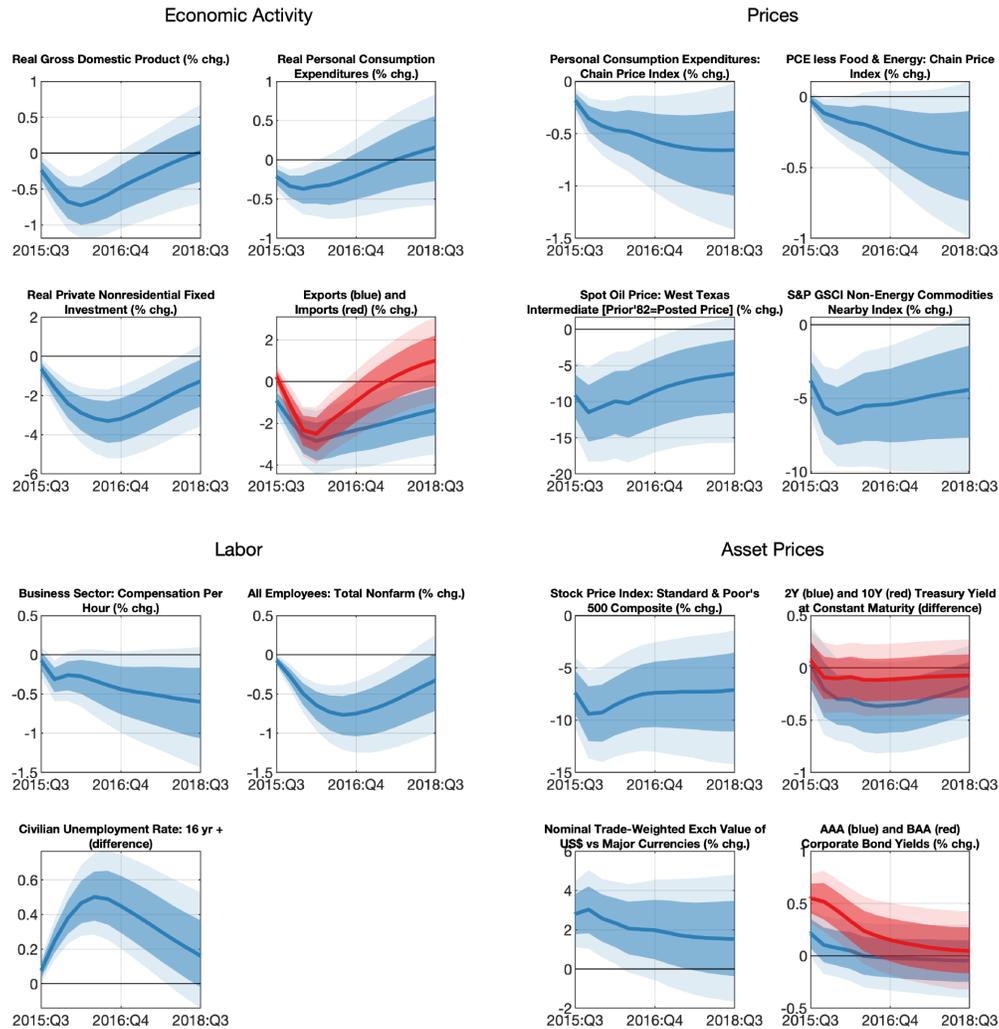


Figure 11. Responses to conditioning on 2015:Q3 non-oil or commodity price financial variables (“no supply shock” scenario). The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

3.4 Tax Cuts and Jobs Act

The Tax Cuts and Jobs Act (TCJA) was signed into law on December 22, 2017. The legislation cut the statutory tax rate for both individuals and businesses until 2025. Moreover, it lowered the corporate tax rate from 35% to 21%. Most of the changes at the individual level come from a lower statutory marginal tax rate coupled with a higher standard deduction and AMT exemptions. Table 2 gives the estimated budget impact of the legislation for the 2018, 2019, and 2020 fiscal years, *which deliberately do not incorporate macroeconomic feedback effects*. These estimates are in part constructed using confidential tax return-level information from the IRS, and are, by the Congressional Budget Act of 1974, the official revenue estimates for Congressional tax legislation. Statically, TCJA was projected to lower both individual and business tax revenue for 2018-2020.

	Fiscal Year	2018	2019	2020
Individual Tax Reform	Billions \$	-75	-189	-172
	As % of GDP in 2017	-0.8	-1.2	-1.1
Business Tax Reform	Billions \$	-129	-134	-113
	As % of GDP in 2017	-7.4	-5.8	-4.9

Table 2: Static Budget Impact of TCJA

A key challenge for identifying the changes in tax policy is the endogeneity of output and taxation over time. Higher output leads to a larger tax base, which increases tax revenue. More explicitly, the relationship can be summarized by the following system:

$$\Delta y_{t+h} = \zeta \Delta T_t + e_{t+h} \tag{10}$$

$$\Delta T_t = \theta \Delta y_t + \Delta \tau_t. \tag{11}$$

The first equation gives a relation between future output y_{t+h} and current growth in tax revenue ΔT_t . However, tax revenue itself, as highlighted in the second equation, is

directly related to economic activity contemporaneously. Through changes in its tax base, the federal government’s tax revenue varies depending on the business cycle. Any scheme identifying the effect of tax policy must take into account these dynamics; otherwise, the tax policy treatment would be endogenous to the business cycle. The relationship specified in equations (10) and (11) hints at an identification strategy used in the structural VAR literature and in policy circles (Blanchard and Perotti, 2002; Perotti, 2007; Caldara and Kamps, 2017). An external estimate of elasticity θ can be used to obtain τ_t , allowing for the estimation of ζ . Equivalently, equipped with θ , we can cleanse ΔT_t of its dependence on the business cycle—ultimately allowing to estimate the effect of a change in tax revenue on output.

Using a similar approach, we use our model to examine the effect of the TCJA on a large number of economic and financial indicators. We begin by augmenting the baseline 31-variable model with tax receipts from corporate income, tax receipts from personal income (the sum of personal current taxes and contributions for government social insurance), corporate tax base (corporate profits less corporate profits with inventory value adjustment: Federal Reserve Banks), and personal income tax base (the sum of personal income and contributions for government social insurance less government social benefit payments).⁷ We split tax revenue into its corporate and personal income components to model the differing corporate/personal income tax policies. Then, we cyclically adjust personal and corporate tax revenue below:

$$\varepsilon_t^p = \log T_t^p - \theta_p \log B_t^p \tag{12}$$

$$\varepsilon_t^c = \log T_t^c - \theta_c \log B_t^c \tag{13}$$

for tax base B_t and tax revenue T_t . Thus, the Tax Cuts and Jobs Act materializes as an

⁷An advantage of our framework is that accommodating these additional indicators is straightforward; if we were to perform a similar exercise with a structural approach, we would likely need to modify relationships equation-by-equation.

exogenous shift to ε_t^p and ε_t^c .

Just like the simpler scenarios, our procedure has two steps. First, we produce unconditional forecasts from the BVAR. These unconditional forecasts represent the evolution of the indicators in our sample in the absence of the TCJA treatment. Second, we produce conditional forecasts. Again, the budget table created by Joint Committee on Taxation produces projections *without macroeconomic feedback effects*. We can therefore directly use these projections for our cyclically-adjusted tax revenue. Our scenario is then produced by conditioning on the linear combination of tax revenue and the tax base, using the static estimates from the budget table. TCJA's effect materializes as the difference between the counterfactual and baseline forecasts.

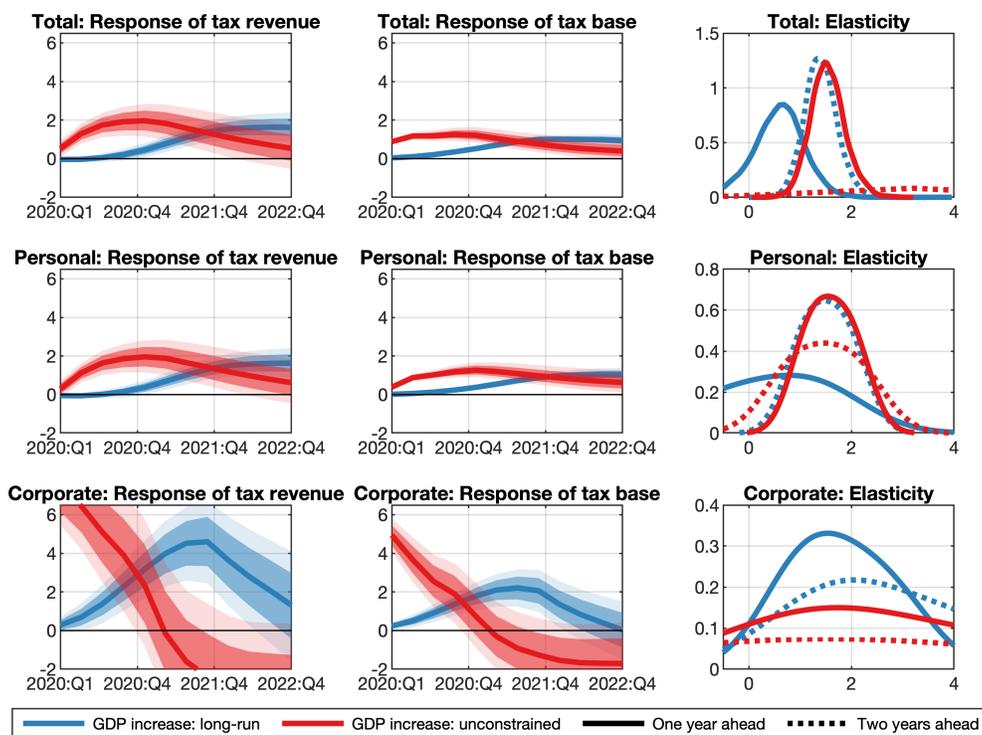


Figure 12. Responses and elasticities. The responses to tax revenue (first column), responses to tax base (second column), and their implied elasticities (third column) one year (solid line) and two years ahead (dotted line) are plotted above. The fan chart bands give 68% and 90% intervals. The results from the unconstrained IRF scenario are displayed in red and the long-run constrained IRF scenario are displayed in blue.

The key parameter of interest for quantifying the effect of TCJA is the elasticity $\theta_p = \theta_c = \theta$. As a baseline specification, we follow the convention of [Caldara and Kamps \(2017\)](#) by using an elasticity of $\theta \approx 2$. As a rough check, we also examine the elasticities implied by our two GDP scenarios, using the large BVAR augmented with personal and corporate tax variables. We begin by computing impulse response functions of personal and corporate tax revenue to an increase in real GDP, as constructed in the previous exercises (see the first two columns of [Figure 12](#)). The red lines indicate responses to the unconstrained GDP increase scenario and the blue lines indicate responses to the long-run GDP increase scenario, where bands correspond to 68% and 90% intervals. Then, we generate the resulting elasticities by dividing the responses for each horizon. The right-most column gives the elasticities one-year ahead and two-years ahead. Considering the relative spread of elasticities across specifications, this exercise should be interpreted as a rough check on θ drawn from external sources.

To assess the implications of differing elasticities, [Table 3](#) gives the mean and standard deviations of the cumulative effects on growth three years ahead for a grid of corporate tax and individual tax rate elasticities. After cumulating, the estimated effect of TCJA on real GDP is qualitatively similar, with a 1.25 percent increase at $\theta_c = \theta_p = 2$. Note that the response differs depending on the chosen value of θ , where lower elasticities result in a lower cumulative effect. For context, [Caldara and Kamps \(2017\)](#) estimate θ to be around 2.2 using a 6-variable VAR. [Figure 13](#) shows impulse response functions for TCJA for the baseline scenario. Output increases through three years ahead, stabilizing at around 1.25% above the unconditional forecast. Consumption follows a similar pattern, stabilizing at around 1.0% above the unconditional forecast. The responses of the labor-market variables give evidence of a tightening in labor markets, with nonfarm payroll employment increasing to around 0.8% above the baseline and the unemployment rate declining 0.4% relative to the baseline.

		Elasticity of Corporate Tax (θ_c)		
		1	2	3
θ_p	1	0.31 (0.71)	0.72 (0.72)	0.89 (0.65)
	2	0.80 (0.58)	1.25 (0.59)	1.26 (0.53)
	3	0.75 (0.49)	1.17 (0.49)	1.14 (0.43)

Table 3: Cumulative effect of TCJA on real GDP three years ahead for varying personal and corporate tax elasticities. Mean effects are above the standard deviation (in parentheses).

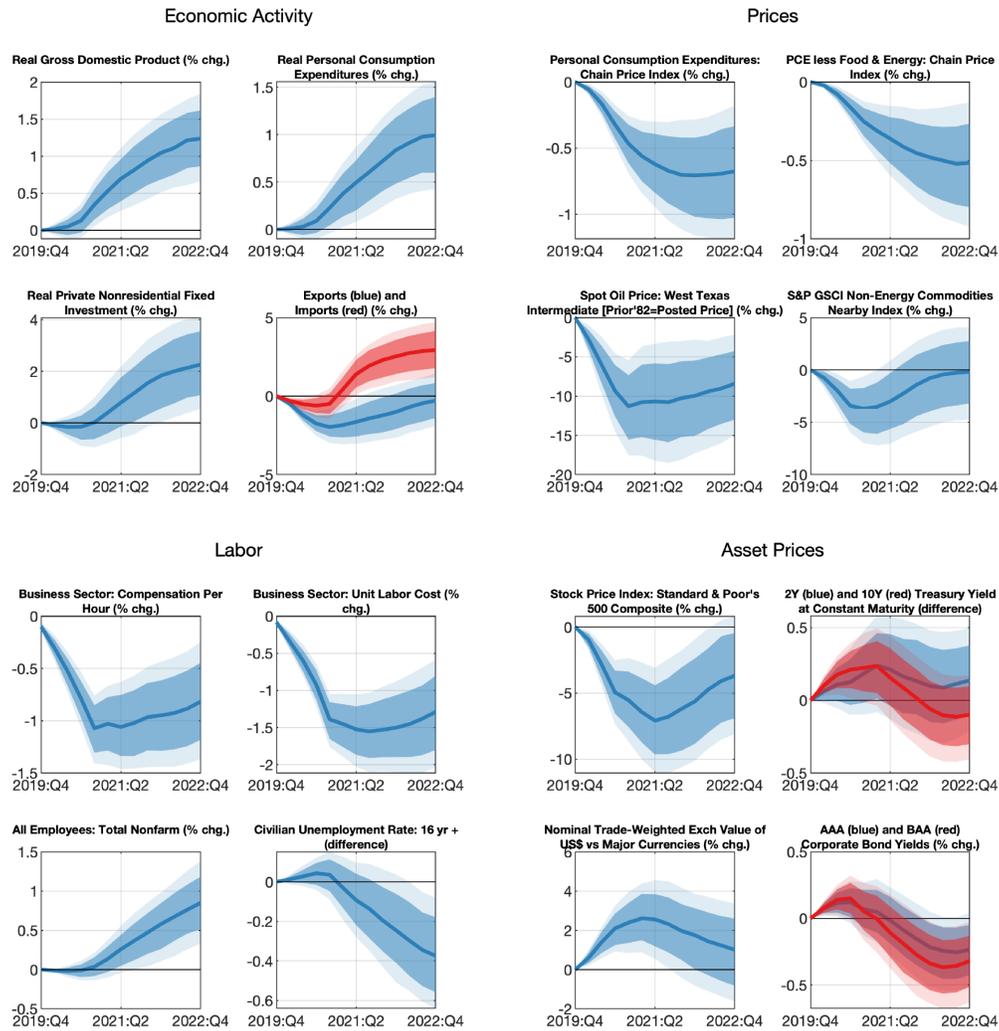


Figure 13. Impulse response functions for an elasticity $\theta = 2$. The solid line shows the median. The shaded areas denote 68% and 90% coverage intervals.

4 Beyond Scenario Analyses

In this section, we show the versatility of the large Bayesian VAR for policy applications beyond modeling scenarios. First, we examine the extent to which future dynamics of real activity inform the paths of financial variables after the Great Recession. Comparing conditional forecasts from models estimated on the data before the crisis and for the entire sample, we also evaluate the presence of structural breaks. Second, we demonstrate a method for incorporating information on long-run inflation expectations into forecasts. Using entropic tilting, we discipline long-run forecasts by imposing median four-quarter PCE inflation as 2 percent. Third, we close with a new index of financial conditions. The index gives a measure of the extent to which financial shocks affect macroeconomic activity.

4.1 Structural breaks and conditioning on real activity

In this exercise, we use conditional forecasting to examine the informativeness of future real economic activity on the future dynamics of headline PCE inflation, the 2-year Treasury rate, and stock prices. Moreover, we can gather evidence of structural changes in dynamics by tracking the evolution of forecasts after estimating the model on pre-financial crisis data and the whole sample. Our goal is to examine the extent to which pre-crisis correlations and observed real economic data can track the evolution of prices and financial data. Conditional on future economic activity, if forecasts of a model estimated on pre-crisis data track poorly with the realized values, then this would give evidence of a structural break.⁸

More specifically, we use in-sample (IS) and out-of-sample (OOS) specifications for our exercises. For our in-sample specification, we estimate the model on data from 1973:Q1 to 2019:Q2. These results are in-sample in the sense that forecasts are generated for dates contained in this period. To gather evidence of structural changes, the out-of-sample (OOS) specification is estimated on data from 1973:Q1-2007:Q2. With the resulting VAR param-

⁸A similar exercise was conducted by [Giannone, Lenza, and Reichlin \(2019\)](#) for the euro area.

eters, we generate forecasts over the period 2007:Q3-2019:Q2, where we treat the paths of selected real activity variables as known. By comparing the in-sample with out-of-sample results, we can assess whether structural economic changes from the pre- to the post-crisis period affect economic relationships. If the underlying relationships differ between the two periods, we would expect inaccuracies in the out-of-sample forecasts.

The top panel of Figure 14 shows the path of headline PCE inflation conditioning only on GDP and the unemployment rate. The bottom panel conditions on fourteen real economic activity variables. In the figure, the black line gives the data, the solid blue line gives the median and shades denote 90% and 68% bands. The dotted red line displays the posterior median for the out-of-sample forecasts. Conditioning on the future paths of fourteen real economic activity indicators (bottom panel), both the in-sample and out-of-sample paths track realized headline PCE inflation quite closely. In contrast, conditioning on just the paths of real GDP and the unemployment rate (top panel), much of the high frequency dynamics are lost. The exercise demonstrates the degree to which real economic activity can inform inflation dynamics.

Figure 15 shows the same exercise for the quarterly growth of the exchange rate (top panel) and stock prices (bottom panel). Just as with inflation, both the in-sample and out-of-sample results track closely with the realized data, showing the extent to which dynamics in financial indicators are closely related to real economic activity and suggesting the absence of a post-crisis structural break.

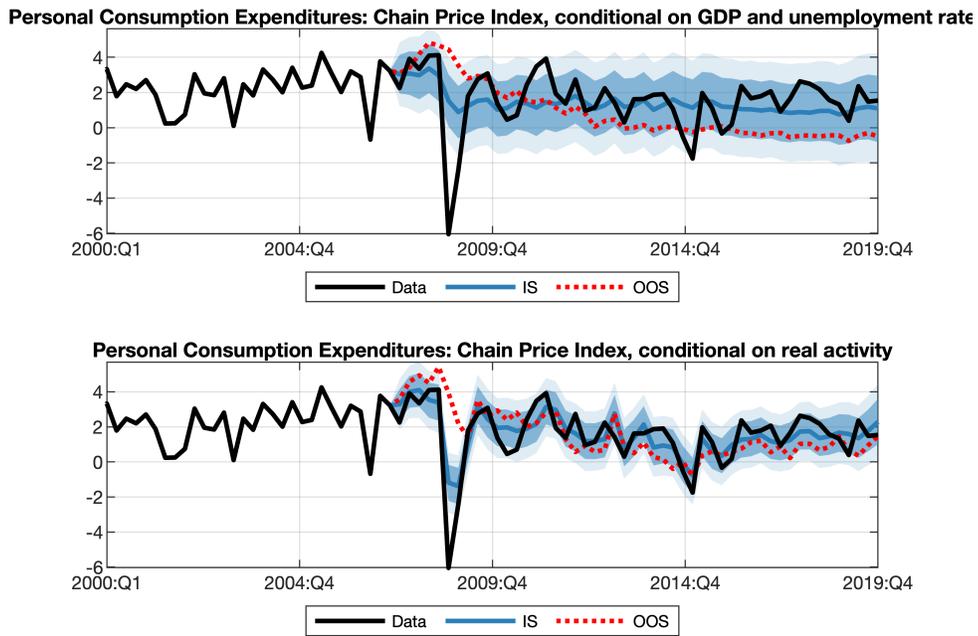


Figure 14. Conditioning on real activity. The panels give headline PCE inflation (black line), 90% and 68% bands for the in-sample forecasts (blue shades), and posterior median of the out-of-sample forecasts (dotted red line). The top panel gives the paths for PCE inflation after conditioning on the realized values of GDP and unemployment rate. The bottom panel gives the paths for PCE inflation after conditioning on the realized values of fourteen real activity indicators.

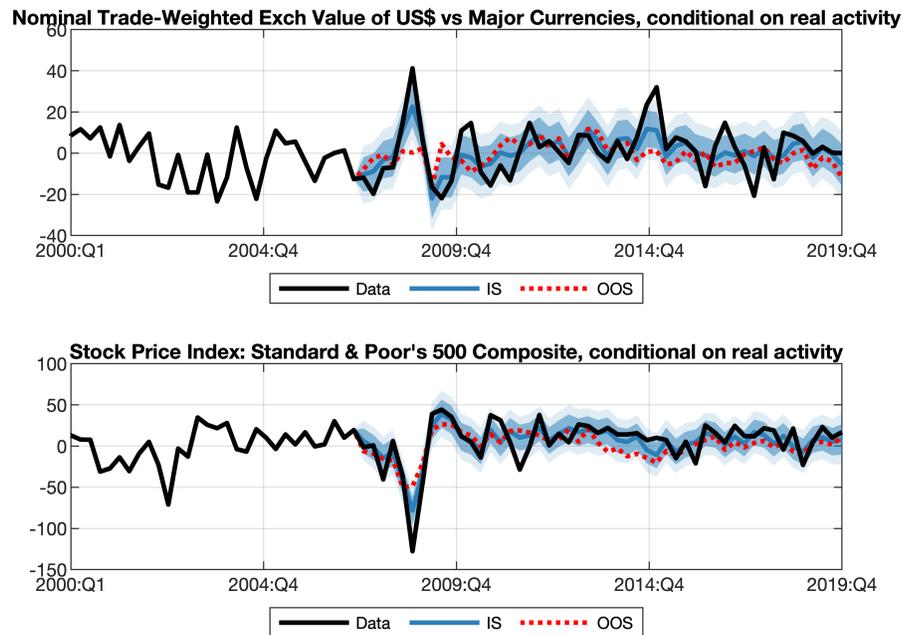
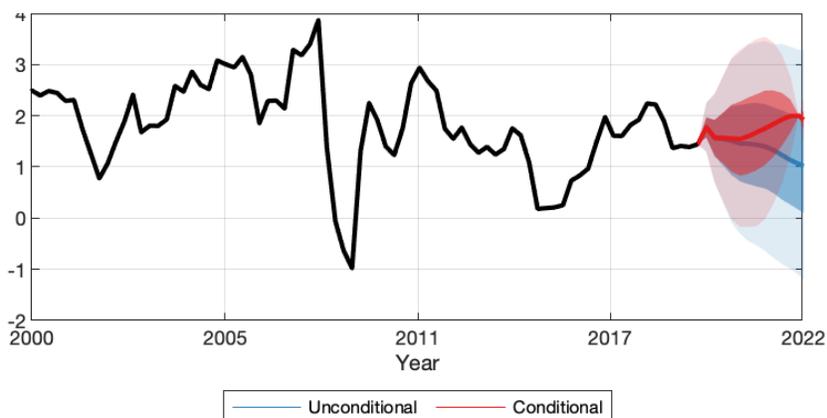
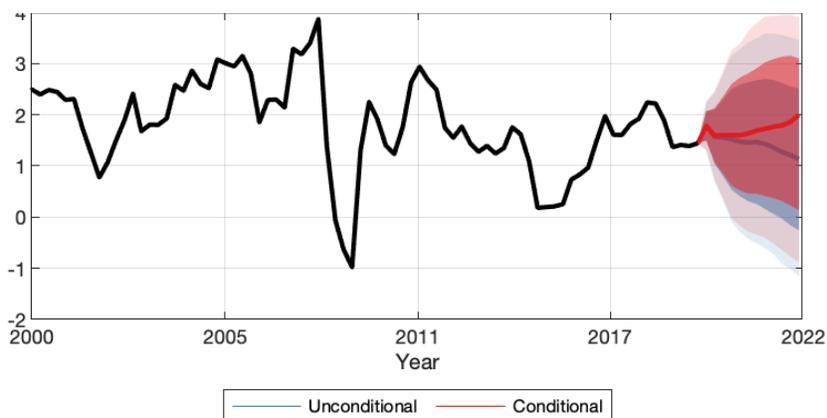


Figure 15. Conditioning on real activity. The panels give the realization (black line), 90% and 68% bands for the in-sample forecasts (blue shades), and posterior median of the out-of-sample forecasts (dotted red line) after conditioning on fourteen real activity indicators.

4.2 Incorporating Judgment



(a) Hard conditioning



(b) Conditioning with uncertainty

Figure 16. Comparison of anchoring schemes. The solid line shows the median. The shades denote 68% and 90% coverage intervals. The top panel compares unconditional distributional forecasts for headline PCE inflation (blue) with those produced by imposing a value of 2.0% 3-5 years ahead. The bottom panel compares unconditional distributional forecasts for headline PCE inflation (blue) with those produced by imposing a median value of 2.0% 3-years ahead through entropic tilting.

In 1998, the Reserve Bank of New Zealand became the first central bank to explicitly target inflation. In the years that followed, the central banks of other countries—Canada, U.K., the United States, and Sweden—adopted similar frameworks. While many banks do not adopt numerical targets, the target is a guide for the long-run steady state rate of inflation.

Incorporating information on inflation targeting in to macroeconomic forecasting can substantially improve forecasts. [Villani \(2009\)](#) uses an informative steady state prior in a Bayesian framework for stationary and cointegrated VAR models, and finds forecasting gains in Swedish macroeconomic data. In empirical settings, this methodology has been applied successfully to large systems ([Österholm and Zettelmeyer, 2008](#); [Adolfson, Andersson, Lindé, Villania, and Vredina, 2007](#); [Beechey and Österholm, 2014](#)). Consistent with the theoretical literature, informative priors in reduced form settings yield forecasts with more stability ([Orphanides and Williams, 2004, 2007](#); [Beechey and Österholm, 2010](#)).

To model the inflation anchoring problem we use entropic tilting—a technique for modifying a baseline distribution to match a particular set of moment conditions. [Robertson, Tallman, and Whiteman \(2005\)](#) introduced this technique and [Krüger, Clark, and Ravazzolo \(2017\)](#) and [Tallman and Zaman \(2020\)](#) used the approach to incorporate judgmental information about the nowcast and the long run. Our approach provides two main contributions. First, we apply the methodology to incorporate beliefs about inflation anchoring. We formalize the belief that inflation will be 2% in the long-run with some uncertainty. Second, while past work has focused on applications to smaller systems, we show that the approach is sensible for high-dimensional models. Third, we show that the framework can be applied to including judgment in forecasts more generally. Matching the median long-run forecasts of the January 2009 FOMC Summary of Economic Projections (SEP), we show how multivariate tilting can be used to study joint dynamics.

Our goal is to obtain forecasts where the long-run distribution (characterized as 12 quarters ahead) of the four-quarter growth rate of the headline PCE price index has a median of 2.0%. Following the setup of [Krüger, Clark, and Ravazzolo \(2017\)](#), we start by producing unconditional forecasts of our baseline large Bayesian VAR model:

$$f := \{y^{(j)}\}_{j=1}^J$$

where y is $n \times 1$ vector of forecasts for n variables, j is the index for an MCMC draw, and J gives the total number of MCMC draws kept after burn-in. In computing moments of the unconditional distribution, each draw j has a flat weight of $1/J$. We aim to reweight f to satisfy the following moment condition

$$E(g(\mathbf{y})) = \bar{g}.$$

where $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ for m total moment conditions and n variables. Then the “entropic tilting” problem can be characterized as

$$\min_{\tilde{f} \in \mathbb{F}} KLIC(\tilde{f}, f) \text{ subject to } E_{\tilde{f}}g(\mathbf{y}) = \bar{g}. \quad (14)$$

\mathbb{F} gives the class of all distributions that can be constructed by reweighting the draws of f . $KLIC$ gives the Kullback-Leibler divergence between \tilde{f} (constructed with reweighted draws) and f (constructed with flat weights) characterized below:

$$KLIC(\tilde{f}, f) = \sum_{j=1}^J \tilde{w}_j \log(J\tilde{w}_j).$$

Weight \tilde{w} gives the weights used to produce \tilde{f} . Then, the tilting solution is given by

$$w_j^* = \frac{\exp(\gamma^{*'}g(\mathbf{y}_j))}{\sum_{j=1}^J \exp(\gamma^{*'}g(\mathbf{y}_j))} \quad (15)$$

$$\gamma^{*'} = \underset{\gamma}{\operatorname{argmin}} \sum_{j=1}^J \exp(\gamma'(g(\mathbf{y}_j) - \bar{g})) \quad (16)$$

To target the median, we use the following function g and \bar{g}

$$g(y^{(j)}) = 1[s'y_j \leq 2.0], \quad \bar{g} = 0.5$$

where s is unit vector that selects the element of y_j corresponding to headline PCE inflation.

Intuitively, the procedure selects for the closest distribution to f (among distributions that can be constructed through reweighting) satisfying our moment condition.

To illustrate the value of modeling forecast uncertainty, consider an alternative conditioning scheme. Figure 16a shows the result of conditioning on *exactly* 2% headline PCE inflation 12-20 periods ahead. The resulting bands reflect the degree of uncertainty from this belief—incredulously, the interquartile range 11 periods ahead is actually smaller than the uncertainty 6 periods ahead. Failing to account for uncertainty in the long-run ultimately affects the uncertainty in the medium-run. Instead, for Figure 16b, the bands more accurately reflect future uncertainty—incorporating the belief that long run PCE inflation forecasts will have a median of two percent, but the exact value is uncertain. This framework can be easily extended to incorporate prior beliefs for long-run behavior of other series, like the natural rate of interest, natural rate of unemployment, and potential output.

With each regularly scheduled meeting of the Federal Open Market Committee (FOMC), each meeting participant submits projections for real GDP growth (Q4/Q4), the unemployment rate (Q4), and inflation (Q4/Q4) for 1-3 years ahead. These projections use information available through the end of the meeting period, and reflect predictions based on appropriate monetary policy (defined as “future policy that, based on current information, is deemed most likely to foster outcomes for economic activity and inflation that best satisfy the participant’s interpretation of the Federal Reserve’s dual objectives of maximum employment and price stability”). The central tendencies of these projections are publicly released through the Summary of Economic Projections (SEP) following each meeting. Table 4 shows the SEP’s central tendency (range excluding the three highest and three lowest projections) at the January 2009 FOMC meeting for real GDP growth, the unemployment rate, and headline and core PCE inflation. The median is also reported for more recent meetings. Participants provide projections for 2009, 2010, and 2011. The central tendency suggests negative output growth (-1.3% to -0.5%) and low inflation (0.3% to 1.0%) for 2009, with higher output (3.8% to 5.0%) and higher inflation (0.9-1.7%) expected in 2011.

The long-run projections of the SEP can be used to inform the long-run dynamics of the BVAR and to examine implications for forecasts in the medium to short-term. To illustrate, we estimate the BVAR and produce unconditional forecasts using information available as of January 30, 2009, the date of the first release of GDP for the 2008:Q4 reference period. Then, to include judgment, we match the median of the BVAR forecasts to the center of the January 2009 SEP central tendency for the 2011 reference period. These projections were released to the public on January 28, 2009—two days prior to the release of GDP for 2008:Q4. We match the unemployment rate, PCE inflation, core PCE inflation, and real GDP growth.

Variable	Central tendency			
	2009	2010	2011	Longer Run
Change in real GDP	-1.3 to -0.5	2.5 to 3.3	3.8 to 5.0	2.5 to 2.7
October projection	-0.2 to 1.1	2.3 to 3.2	2.8 to 3.6	n.a.
Unemployment rate	8.5 to 8.8	8.0 to 8.3	6.7 to 7.5	4.8 to 5.0
October projection	7.1 to 7.6	6.5 to 7.3	5.5 to 6.6	n.a.
PCE inflation	0.3 to 1.0	1.0 to 1.5	0.9 to 1.7	1.7 to 2.0
October projection	1.3 to 2.0	1.4 to 1.8	1.4 to 1.7	n.a.
Core PCE inflation	0.9 to 1.1	0.8 to 1.5	0.7 to 1.5	
October projection	1.5 to 2.0	1.3 to 1.8	1.3 to 1.7	

Table 4: Summary of Economic Projections (January 2009). Above shows the central tendencies (range excluding the three highest and three lowest projections for each variable for each year) for the change in real GDP (Q4/Q4), the unemployment rate, PCE inflation (Q4/Q4), and core PCE inflation (Q4/Q4).

Figure 17 shows draws of the joint predictive distribution of PCE inflation and the unemployment rate for 2011:Q4. The point size corresponds directly to the weights assigned under the entropic tilting procedure, matching the median of the BVAR forecast distribution to the SEP’s central tendency. We immediately see that to match the median the procedure assigns higher weights to high unemployment rate draws. Moreover, the procedure assigns higher weights to draws in a “low” inflation band (around 0%-1.25%) and “high” inflation band (around 3%-4.5%).

Figure 18 illustrates the results for the joint predictive density of the unemployment rate and PCE inflation for 2009:Q4 and 2011:Q4. The black lines show the BVAR’s uncondi-

tional predictive density (joint distribution in the center and the marginal distributions on the sides). The red lines show the “tilted” distributions, matching the SEP’s central tendency for the 2011:Q4 reference period. The blue dot in the center panel gives the latest data, also indicated by the vertical lines on the marginal distributions. In the left panel, the tilted predictive densities are comparable. Evident from the marginal densities, the unemployment rate and PCE inflation are slightly higher. Moreover, the tilted density places higher probability on the realization. The right panel shows the joint forecasts three years ahead, matching the horizon used for tilting. The unconditional forecasts are unimodal, while the tilted forecasts have multiple modes. The marginal densities show modes for a high/low unemployment states and high/low inflation states. The realization is near the high inflation and high unemployment rate mode. These are a reflection of the weighting procedure, where differential weights are given to draws in different regions of the support. Conditional on the model and the data, these results suggest that multiple modes are necessary to generate forecasts consistent with the SEP.

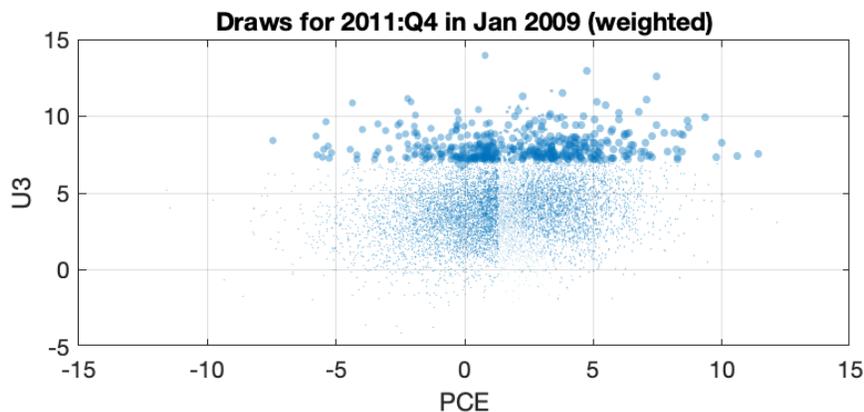


Figure 17. Joint draws of PCE inflation and the unemployment rate for 2011:Q4 in January 2009 by weights. Each point is proportional to the weight assigned by the entropic tilting procedure, matching the median of the BVAR forecast distribution to the SEP’s central tendency.

The tilting framework could easily be extended to match a forecaster’s perception of risk and uncertainty. While the SEP point estimates reflect the most likely outcome, they do not summarize the full range and likelihood of potential scenarios. If provided, this

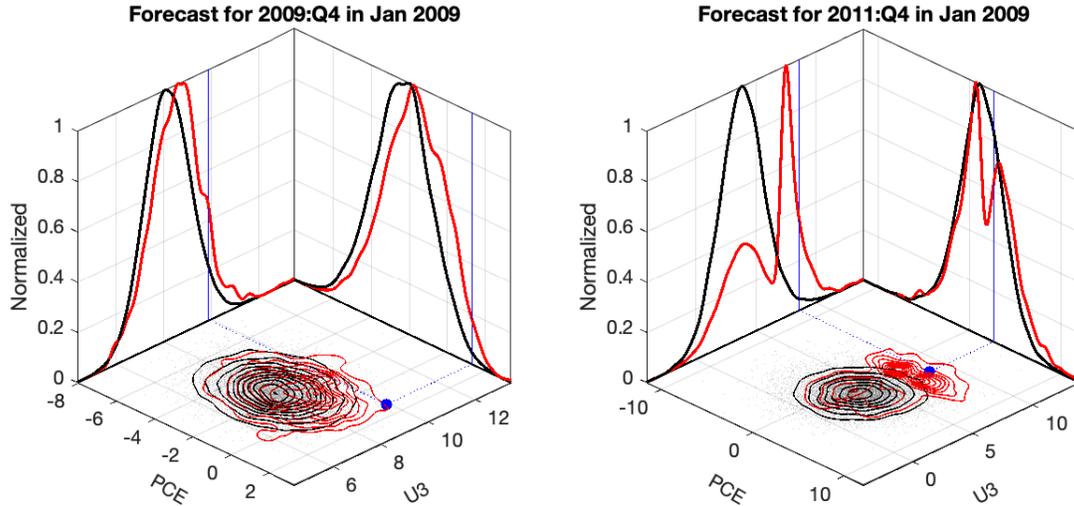


Figure 18. Joint forecasts of the unemployment rate and PCE inflation as of January 30, 2009. The left and right figures respectively show forecasts for PCE inflation and the unemployment rate for 2009:Q4 and 2011:Q4. The black lines show the BVAR’s unconditional forecast distributions (joint distribution in the center and the marginal distributions on side). The red lines show the same for the “tilted” distribution, matching the central tendency for the 2011:Q4 reference period of the January 2009 SEP. The blue dot in the center panel shows the realization, also indicated by the vertical blue lines on the marginal distributions.

information could be encoded into BVAR forecasts by tilting higher moments, like variance and skewness. The inclusion in the December 2020 SEP of figures with diffusion indices over time to summarize forecast uncertainty shows the interest of the policy community in providing not only point forecasts, but also forecast distributions.

4.3 A BVAR Financial Conditions Index

The Great Recession underscores the importance of understanding how financial conditions can inform and affect the future state of the economy. Severe financial disruptions in 2007 preceded sharp and lasting deteriorations in economic conditions, with real GDP recovering to pre-recession levels only in 2011:Q3, nearly four years after its peak. By measuring how current financial conditions can influence future economic outcomes, financial conditions indices (FCIs) are an important tool for policymakers in assessing macrofinancial risks. In this section, we briefly survey methodologies of several prominent FCIs and propose a new

FCI constructed using our model. Simulating the counterfactual path of GDP growth in the absence of economic shocks, our approach combines the ease of economic interpretation found in structural approaches with the breadth of indicators prevalent in statistical approaches.

The Bank of Canada’s MCI was one of the first broad measures of financial conditions. The index is constructed as a weighted sum of the short-term interest rate and the exchange rate, where weights are constructed from a macroeconomic model as measuring the relative impact of changes in the variable on output (Freedman, 1994). In determining the weights, the index’s “weighted sum” approach makes explicit the connection between financial conditions and real economic activity. Following the Bank of Canada’s MCI, the Goldman Sachs FCI (GSFCI) is another well-established example of the “weighted sum” approach. The index includes the federal funds rate, nominal 10-year Treasury yield, corporate spread, TED spread, equity prices, and trade-weighted dollar (Hatzius and Stehn, 2018). Using a small-scale macroeconomic model, the weights are derived using the estimated partial impact of each variable on real GDP growth over the following four quarters. Other examples of the “weighted sum” approach are the OECD financial conditions index and the Citi financial conditions index (Davis, Kirby, and Warren, 2016; D’Antonio, 2008).

In contrast, some indices use statistical techniques to extract a common factor from a set of financial indicators, typically called the “principal components approach” (Hatzius, Hooper, Mishkin, Schoenholtz, and Watson, 2010). One such example is the Federal Reserve Bank of Chicago’s National Financial Conditions Index (NFCI), which accounts for the dynamic and cross-sectional correlations of a large number of indicators through a dynamic factor model (Brave and Butters, 2011). A key feature of the index is its representation of financial markets; it includes 28 money market indicators, 27 indicators for debt and equity markets, and 45 indicators of the banking system. The NFCI ultimately gives a measure for gauging financial stability, where its history tracks major events in U.S. financial history. Other examples include the IMF’s Financial Soundness Indicators, Kansas City Fed Financial Stress Index (KCFSI), and the St. Louis Fed Financial Stress Index (Reinbold and Paulina,

2017; Cardarelli, Lall, and Elekdag, 2009; Keeton and Hakkio, 2009).

By tacitly assuming dynamic homogeneity (that indicators have identical lead-lag behavior in response to a set of shocks), both “principal components” and “weighted sum” approaches fail to model the heterogenous shock transmission mechanisms in financial markets. To illustrate, consider the functional form of a typical dynamic factor model, like the one underlying the NFCI:

$$x_{i,t} = \lambda_i f_t + e_{i,t} \quad (17)$$

$$f_t = \sum_{h=1}^l a_h f_{t-h} + \nu_t, \quad (18)$$

where f_t gives the latent financial factor, $x_{i,t}$ represents a financial variable, $e_{i,t}$ gives its associated measurement error, l is the number of lags, and ν_t gives the shocks. There are two notable consequences of this functional form: (1) there exists a single type of latent shock (ν_t) and (2) the shock, up to scale, propagates to the observables $x_{i,t}$ under identical lead-lag behavior. In other words, a financial shock today affects all indicators uniformly (up to scale λ_i) in the same period. Homogeneity in both the source of shocks and their transmission mechanism is untenable for the goal of accurately characterizing how financial markets affect economic activity. In a similar spirit as the factor model’s restricted measurement equation (equation (18)), the “weighted sum” approach doesn’t model the lead-lag relationship among its components.

The index we constructed using our large Bayesian VAR explicitly models the dynamics of economic/financial variables and can be viewed as a combination of both approaches. Similar to the “principal components” approach, our index uses statistical methods to summarize the dynamics of a large number of indicators, avoiding ad-hoc choices in variable selection. The BVAR provides a means for modeling the complex linkages among a large number of financial and economic indicators, allowing for heterogenous transmission channels of financial shocks. Similar to the “weighted sum” approach, our index makes explicit the connection between

macroeconomic and financial indicators through structural assumptions, where we impose that current financial conditions are (a) affected by current/past macroeconomic conditions and (b) affect future macroeconomic conditions. The resulting index is effectively a structural conditional forecast of real GDP—allowing for a full characterization of uncertainty in shocks, parameters, and hyperparameters through confidence bands. Moreover, natively existing in GDP space, the index is easy to interpret without any auxiliary equations, a particularly salient feature for policy applications.

Formally, we begin by estimating the model under the full sample (1973:Q1-2019:Q4) using our standard large dataset, where the economic variables are ordered ahead of the s financial variables. We then recast our problem as a structural conditional forecasting exercise, using the methodology described in [Bańbura, Giannone, and Lenza \(2015\)](#). Define a vector of orthonormal structural shocks:

$$\nu_t = (\nu_{1,t}, \dots, \nu_{n,t})', \quad E(\nu_t \nu_t') = I_n.$$

Then, for lower triangular rotation matrix H , the innovations ε_t are related to the structural shocks by $\varepsilon_t = H\nu_t$ and $HH' = \Sigma$. From variable ordering, we can eliminate shocks corresponding to economic variables from our conditional forecasts by replacing the first $n - s$ columns of matrix H with zeros. Call this new matrix \bar{H} and $\bar{\Sigma} = \bar{H}\bar{H}'$. By setting the correlation between economic and financial variables to zero, the resulting variance covariance matrix $\bar{\Sigma}$ takes the form

$$\bar{\Sigma} = \begin{pmatrix} 0_{n-s \times n-s} & 0_{n-s \times s} \\ 0_{s \times n-s} & K_{s \times s} \end{pmatrix}.$$

Under the revised $\bar{\Sigma}$, innovations to economic variables affect neither economic variables nor financial variables. Finally, to construct our index, we generate conditional forecasts by conditioning on the entire history of financial variables and the first 5 observations of economic variables, using the modified variance covariance matrix $\bar{\Sigma}$ to isolate the financial

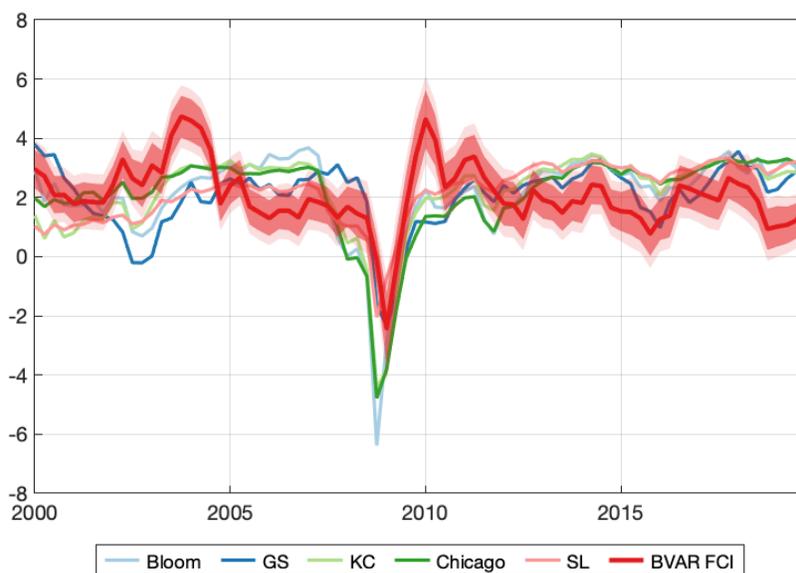


Figure 19. Comparison of financial conditions indices. The above figure compares the BVAR FCI (dark red) to the Bloomberg FCI (light blue), Goldman Sachs FCI (dark blue), Kansas City Fed FSI (light green), Chicago Fed NFCI (dark green), and the St. Louis Fed FSI (light red). For interpretation, each index is projected onto quarterly GDP growth for a common evaluation period (2000:Q1-2019:Q4). The BVAR FCI, already in GDP units, is left untransformed.

shocks. The resulting conditional forecast of real GDP growth can be interpreted as its history subject only to financial shocks.

Figure 19 shows several prominent FCIs plotted against the BVAR FCI. For comparability, each index is projected onto real GDP growth for a common evaluation period (2000:Q1-2019:Q4). All of the indices decline during the Great Recession, and sharply increase immediately afterwards. Consistent with the other indicators, the BVAR FCI is relatively stable in the post-crisis period. In contrast to other indices, the magnitude of its increase in 2010:Q1 is much greater. Next, for its conceptual similarity with the BVAR FCI, we focus on the Goldman Sachs FCI in Figure 20. For comparability, the GS FCI is multiplied by -1. The double axes are necessary because the GS FCI is not natively available in GDP units. The dynamics of both indicators, particularly during the Great Recession and the post-recovery period are more correlated, whereas the two indices differ in the 1990s.

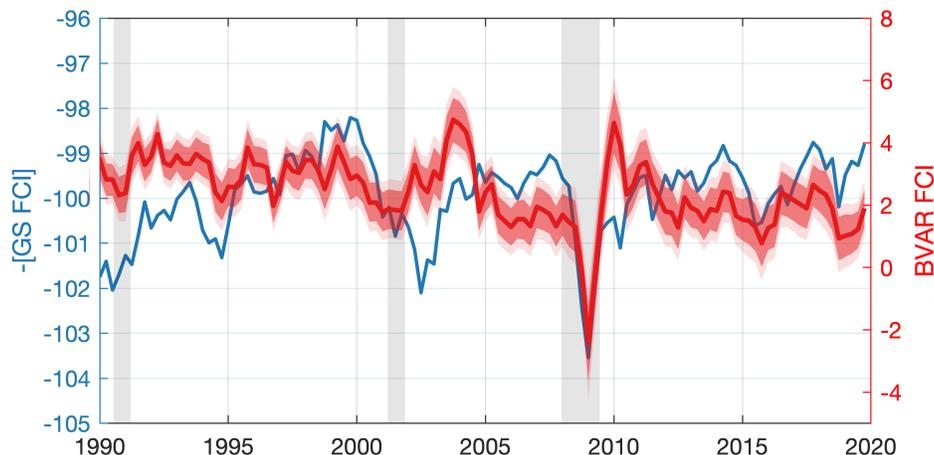


Figure 20. Comparison of BVAR FCI with GS FCI.

Figure 21. BVAR FCI and GS FCI. The figure above compares the Goldman Sachs FCI (blue) with the BVAR FCI (red). To compare dynamics, the GS FCI is multiplied by -1.

While the GS FCI displays a small upward trend over the sample, the BVAR FCI appears stationary.

Figure 22 illustrates the difference that structural conditioning makes when constructing a financial conditions index. The blue and red shades indicate conditional forecasts under traditional and structural conditioning, respectively. Under the traditional conditioning scheme, a much larger proportion of the variation in real GDP growth is ascribed to variation in financial conditions, evidenced by how closely the blue bands track with real GDP growth. Without restrictions to the variance-covariance matrix, some of the variation in the financial indicators is polluted by economic shocks.

Stability in real-time is imperative for policymakers, where judgment is often made with preliminary macroeconomic data. To evaluate the index's statistical stability and robustness to data revisions, we recreate the index in real-time. More specifically, for each quarter from 2000:Q1-2019:Q4, we estimate the BVAR and compute the index with data available as of the Survey of Professional Forecasters due date. The idea is to simulate the information a policymaker would see in real time. To illustrate, Figure 23 shows the latest index (in red). The blue bands show the latest FCI bands for each vintage (ex-ante). Despite differing

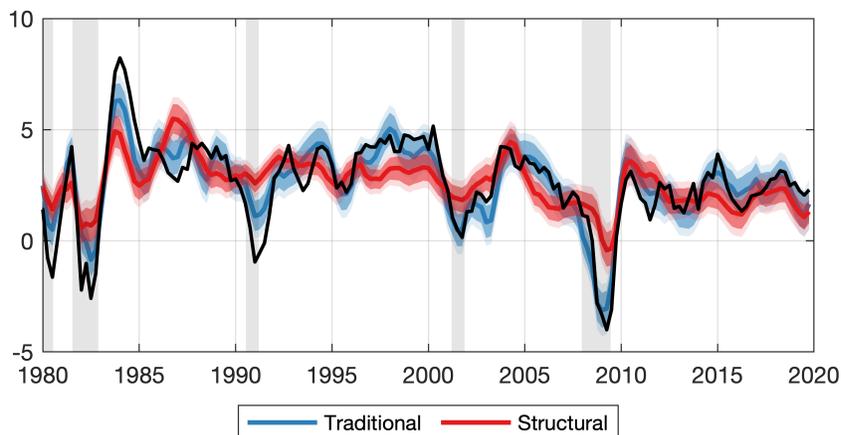


Figure 22. Comparison of structural and reduced form conditioning. The blue and red shades show 68% and 90% bands under reduced form (blue) and structural (red) conditioning. The black line gives realized quarterly real GDP growth.

information sets and data revisions, the ex-ante and ex-post bands roughly coincide.

5 Conclusion

We have shown that our large-scale Bayesian VAR produces reliable predictions of the joint distribution of the large set of macroeconomic and financial indicators monitored by the Federal Reserve Staff and professional forecasters. These predictions are competitive with those produced by professional forecasters and also those by Federal Reserve staff that are judgmentally-adjusted predictions from large-scale macroeconomic models (FRB/US).

We illustrate how the BVAR modeling framework can be used for policy analysis utilizing economic scenarios constructed by conditioning the predictive distribution on well-crafted assumptions about future values of appropriately-chosen series in the model. We show that these scenarios are closely related to structural analyses conducted in the academic literature using structural VAR (SVAR) models. We also show how conditional forecasts can be modified to incorporate expert judgment – e.g., about long-run inflation. We also construct an index of financial conditions by generating forecasts conditioning on the entire

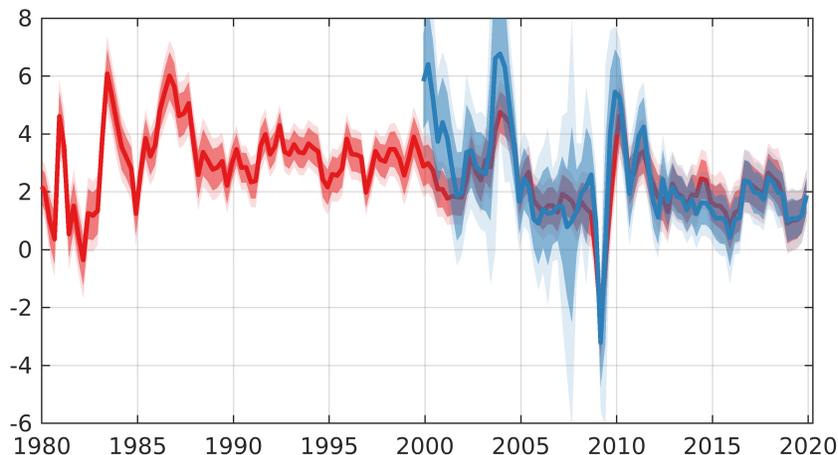


Figure 23. BVAR FCI in real-time. The red bands show the latest FCI (50% and 68% bands). The blue bands show the ex-ante FCI using real-time data, where the latest available FCI bands are plotted for each vintage.

history of the financial variables only.

Small scale structural VARs have been extensively used to validate and cross-check the predictions of micro-founded macroeconomic models. The results of this paper suggest promise for a more extensive use of VAR models for policy evaluation through scenario analyses, that can be an alternative to large-scale and comprehensive models of the macroeconomy.

Because VAR and micro-founded models are intimately related, they can be combined for prediction, structural analysis, and policy evaluation by exploiting their relative strengths and complementary traits: the flexibility and generality of the VAR on the one hand, and the ability of structural models to tell stories and guide the evaluation of alternative policies on the other hand.

Scenarios are simple conditional forecasts. We have seen that by appropriately designing the conditioning assumption it is possible to narrow down the set of underlying primitive factors, which can help providing a “more structural” interpretation of the dynamics.

Still, this approach has some limits, and in order to interpret the scenarios in terms of structural shocks and derive policy implications, it may be useful to read the results through

the lens of a micro-founded model. [Del Negro \(2019\)](#) shows how to combine DSGE and VAR models for scenario-based policy analyses. This allows one to derive implications for model concepts (such as r^* and potential output) and study scenarios under alternative policies (e.g., price-level targeting, optimal policy). Such scenarios can be obtained by filtering the structural shocks from the DSGE model that match selected conditional forecasts obtained from the BVAR. One can then proceed to evaluate the predicted path of key “unobserved” variables such as the natural rate of interest using these identified shocks. This is a promising avenue for future research.

References

- ADOLFSON, M., M. K. ANDERSSON, J. LINDÉ, M. VILLANIA, AND A. VREDINA (2007): “Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks,” *International Journal of Central Banking*.
- ANGELETOS, G.-M., F. COLLARD, AND H. DELLAS (2020): “Business-Cycle Anatomy,” *American Economic Review*, 110(10), 3030–3070.
- ANGELINI, E., M. LALIK, M. LENZA, AND J. PAREDES (2019): “Mind the gap: A multi-country BVAR benchmark for the Eurosystem projections,” *International Journal of Forecasting*, 35(4), 1658–1668.
- ASCARI, G., AND L. FOSSO (2021): “The Inflation Rate Disconnect Puzzle: On the International Component of Trend Inflation and the Flattening of the Phillips Curve,” Discussion Paper 2021-13, Centre for Macroeconomics.
- BAÑBURA, M., D. GIANNONE, AND M. LENZA (2015): “Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections,” *International Journal of Forecasting*, 31(3), 739–756.
- BAÑBURA, M., D. GIANNONE, AND L. REICHLIN (2010): “Large Bayesian vector auto regressions,” *Journal of Applied Econometrics*, 25(1), 71–92.
- BEECHEY, M., AND P. ÖSTERHOLM (2010): “Forecasting inflation in an inflation-targeting regime: A role for informative steady-state priors,” *International Journal of Forecasting*, 26(2), 248–264, Special Issue: Bayesian Forecasting in Economics.
- (2014): “Central Bank Forecasts of Policy Interest Rates: An Evaluation of the First Years,” *Economic Notes*, 43(1), 63–78.
- BLANCHARD, O., AND R. PEROTTI (2002): “An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output,” *The Quarterly Journal of Economics*, 117(4), 1329–1368.
- BLANCHARD, O. J., AND D. QUAH (1989): “The Dynamic Effects of Aggregate Demand and Supply Disturbances,” *American Economic Review*, 79(4), 655–673.
- BOK, B., D. CARATELLI, D. GIANNONE, A. M. SBORDONE, AND A. TAMBALOTTI (2018): “Macroeconomic Nowcasting and Forecasting with Big Data,” *Annual Review of Economics*, 10(1), 615–643.

- BRAVE, S., AND R. A. BUTTERS (2011): “Monitoring financial stability: A financial conditions index approach,” *Federal Reserve Bank of Chicago Economic Perspectives*, 35(1).
- CALDARA, D., AND C. KAMPS (2017): “The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers,” *The Review of Economic Studies*, 84(3), 1015–1040.
- CALDARA, D., C. SCOTTI, AND M. ZHONG (2021): “Macroeconomic and Financial Risks: A Tale of Mean and Volatility,” Working paper.
- CARDARELLI, M. R., M. S. LALL, AND S. ELEKDAG (2009): “Financial Stress, Downturns, and Recoveries,” IMF Working Papers 2009/100, International Monetary Fund.
- CARRIERO, A., T. E. CLARK, AND M. MARCELLINO (2015): “Bayesian VARs: Specification Choices and Forecast Accuracy,” *Journal of Applied Econometrics*, 30(1), 46–73.
- CHRISTIANO, L., M. EICHENBAUM, AND C. EVANS (2005): “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, 113(1), 1–45.
- CHRISTIANO, L. J. (2012): “Christopher A. Sims and Vector Autoregressions,” *The Scandinavian Journal of Economics*, 114(4), 1082–1104.
- CHRISTIANO, L. J., M. EICHENBAUM, AND C. EVANS (1999): “Monetary policy shocks: what have we learned and to what end?,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 2, pp. 65–148. Elsevier.
- CIMADOMO, J., D. GIANNONE, M. LENZA, A. SOKOL, AND F. MONTI (2020): “Nowcasting with large Bayesian vector autoregressions,” Working Paper Series 2453, European Central Bank.
- CONSTANCIO, V. (2017): “Developing models for policy analysis in central banks,” Opening speech by Vítor Constancio, Vice-President of the ECB, at the Annual Research Conference, Frankfurt am Main.
- D’ANTONIO, P. (2008): “Appendix,” in *A View of the U.S. Subprime Crisis*, ed. by R. DiClemente, and K. Schoenholtz, EMA Special Report, Citigroup Global Markets Inc. World Scientific.
- DAVIS, E. P., S. KIRBY, AND J. WARREN (2016): “The Estimation of Financial Conditions Indices for the Major OECD Countries,” Economics Department Working Papers 1335, OECD.
- DE MOL, C., D. GIANNONE, AND L. REICHLIN (2008): “Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?,” *Journal of Econometrics*, 146(2), 318–328.

- DEL NEGRO, M. (2019): “Building the NY Fed Research Staff Forecast Distribution,” Joint research day on quantitative tools for monitoring macroeconomic and financial conditions, Federal Reserve Banks of Atlanta and New York.
- DEL NEGRO, M., M. LENZA, G. E. PRIMICERI, AND A. TAMBALOTTI (2020): “What’s Up with the Phillips Curve?,” *Brookings Papers on Economic Activity*, 51(1), 301–373.
- DEL NEGRO, M., F. SCHORFHEIDE, F. SMETS, AND R. WOUTERS (2007): “On the Fit of New Keynesian Models,” *Journal of Business & Economic Statistics*, 25, 123–143.
- DOMIT, S., F. MONTI, AND A. SOKOL (2019): “Forecasting the UK economy with a medium-scale Bayesian VAR,” *International Journal of Forecasting*, 35(4), 1669–1678.
- DURBIN, J., AND S. J. KOOPMAN (2002): “A simple and efficient simulation smoother for state space time series analysis,” *Biometrika*, 89(3), 603–616.
- FREEDMAN, C. (1994): “The use of indicators and of the monetary conditions index in Canada,” in *Frameworks for Monetary Stability: Policy Issues and Country Experiences*, ed. by T. J. Balino, and C. Cottarelli. IMF.
- GIANNONE, D., M. LENZA, AND G. E. PRIMICERI (2015): “Prior Selection for Vector Autoregressions,” *The Review of Economics and Statistics*, 97(2), 436–451.
- GIANNONE, D., M. LENZA, AND G. E. PRIMICERI (2019): “Priors for the Long Run,” *Journal of the American Statistical Association*, 114(526), 565–580.
- GIANNONE, D., M. LENZA, AND L. REICHLIN (2019): “Money, Credit, Monetary Policy, and the Business Cycle in the Euro Area: What Has Changed Since the Crisis?,” *International Journal of Central Banking*, 15(5), 137–173.
- GIANNONE, D., L. REICHLIN, AND D. SMALL (2008): “Nowcasting: The real-time informational content of macroeconomic data,” *Journal of Monetary Economics*, 55(4), 665–676.
- HATZIUS, J., P. HOOPER, F. S. MISHKIN, K. L. SCHOENHOLTZ, AND M. W. WATSON (2010): “Financial Conditions Indexes: A Fresh Look after the Financial Crisis,” Working Paper 16150, National Bureau of Economic Research.
- HATZIUS, J., AND S. J. STEHN (2018): “The Case for a Financial Conditions Index,” Global economics paper, Goldman Sachs Economic Research.

- KEETON, W. R., AND C. S. HAKKIO (2009): “Financial stress: what is it, how can it be measured, and why does it matter?,” *Economic Review*, 94(Q II), 5–50.
- KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): “Impulse response analysis in nonlinear multivariate models,” *Journal of Econometrics*, 74(1), 119–147.
- KOOP, G. M. (2013): “Forecasting with Medium and Large Bayesian VARS,” *Journal of Applied Econometrics*, 28(2), 177–203.
- KRÜGER, F., T. E. CLARK, AND F. RAVAZZOLO (2017): “Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts,” *Journal of Business & Economic Statistics*, 35(3), 470–485.
- LEEPER, E. M., AND T. ZHA (2003): “Modest policy interventions,” *Journal of Monetary Economics*, 50(8), 1673–1700.
- LITTERMAN, R. (1979): “Techniques of forecasting using vector autoregressions,” Working Paper 115, Federal Reserve Bank of Minneapolis.
- LITTERMAN, R. B. (1980): “A Bayesian procedure for forecasting with vector autoregression,” Working paper, Massachusetts Institute of Technology, Department of Economics.
- LUCAS, R. J. (1976): “Econometric policy evaluation: A critique,” *Carnegie-Rochester Conference Series on Public Policy*, 1(1), 19–46.
- ORPHANIDES, A., AND J. C. WILLIAMS (2004): “Imperfect Knowledge, Inflation Expectations, and Monetary Policy,” in *The Inflation-Targeting Debate*. National Bureau of Economic Research, Inc.
- (2007): “Robust monetary policy with imperfect knowledge,” *Journal of Monetary Economics*, 54(5), 1406–1435, *Carnegie-Rochester Conference Series on Public Policy: Issues in Current Monetary Policy Analysis* November 10-11, 2006.
- ÖSTERHOLM, P., AND J. ZETTELMEYER (2008): “The Effect of External Conditions on Growth in Latin America,” *IMF Staff Papers*, 55(4), 595–623.
- PEROTTI, R. (2007): “In Search of the Transmission Mechanism of Fiscal Policy,” Working Paper 13143, National Bureau of Economic Research.
- REINBOLD, B., AND R.-E. PAULINA (2017): “Financial Conditions Indexes,” *Economic Synopses*, 17.

- ROBERTSON, J., E. TALLMAN, AND C. WHITEMAN (2005): “Forecasting using relative entropy,” *Journal of Money, Credit and Banking*, 37(3), 383–401.
- ROTEMBERG, J., AND M. WOODFORD (1997): “An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy,” in *NBER Macroeconomics Annual 1997, Volume 12*, pp. 297–361. National Bureau of Economic Research, Inc.
- SIMS, C. A. (1972): “Money, Income, and Causality,” *The American Economic Review*, 62(4), 540–552.
- (1980a): “Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered,” *The American Economic Review*, 70(2), 250–257.
- SIMS, C. A. (1980b): “Macroeconomics and Reality,” *Econometrica*, 48(1), 1–48.
- SIMS, C. A., AND T. ZHA (1999): “Error Bands for Impulse Responses,” *Econometrica*, 67(5), 1113–1155.
- STOCK, J. H., AND M. W. WATSON (2001): “Vector autoregressions,” *Journal of Economic Perspectives*, 15(4), 101–115.
- TALLMAN, E. W., AND S. ZAMAN (2020): “Combining survey long-run forecasts and nowcasts with BVAR forecasts using relative entropy,” *International Journal of Forecasting*, 36(2), 373–398.
- TETLOW, R. J. (2015): “Real-Time Model Uncertainty in the United States: ‘Robust’ Policies Put to the Test,” *International Journal of Central Banking*, 11(2), 113–155.
- UHLIG, H. (2004): “Do Technology Shocks Lead to a Fall in Total Hours Worked?,” *Journal of the European Economic Association*, 2(2-3), 361–371.
- VILLANI, M. (2009): “Steady-state priors for vector autoregressions,” *Journal of Applied Econometrics*, 24(4), 630–650.
- WAGGONER, D. F., AND T. ZHA (1999): “Conditional Forecasts in Dynamic Multivariate Models,” *The Review of Economics and Statistics*, 81(4), 639–651.

A Appendix

A.1 Data

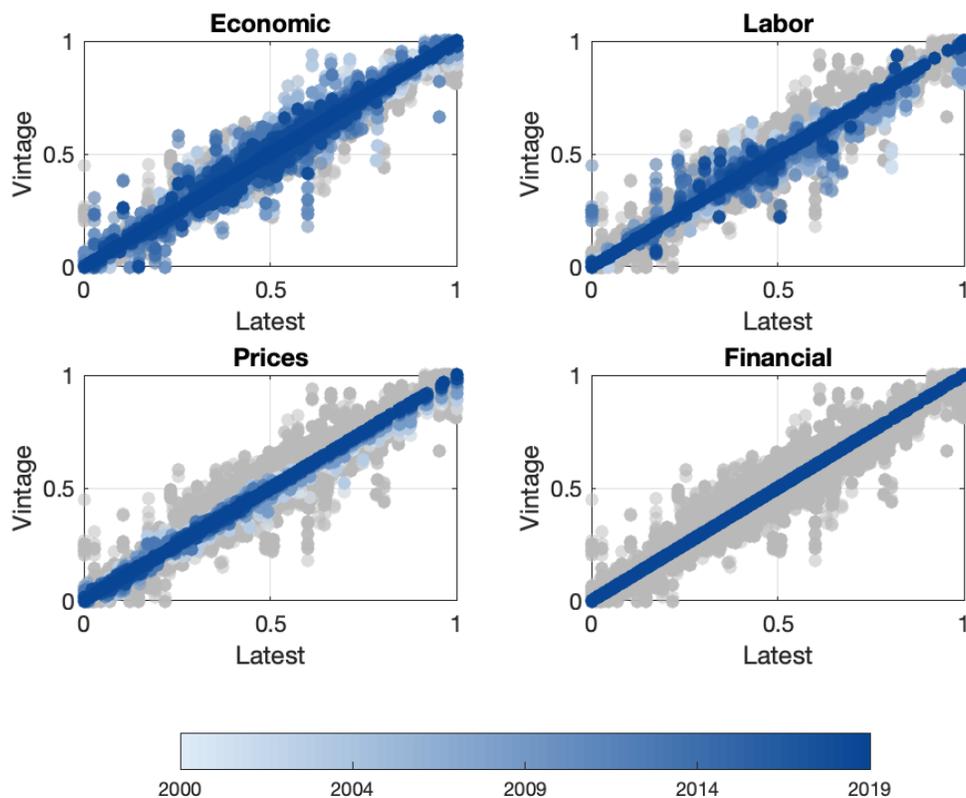


Figure 24. Comparison of data revisions. The above figure highlights the magnitude of data revisions by category. Rates are taken as levels, and other series enter as quarterly growth rates. After normalizing to the interval $[0, 1]$, the latest value is plotted the vintage value. The shades of blue indicate the vintage. In cases of no data revisions, the points fall on the 45 degree line.

The effect of data revisions in the macroeconomic context is nontrivial, and is particularly important to account for comparing forecasts to those produced in real time (like the SPF). To illustrate, figure 24 displays the magnitude of data revisions over time. Vintage data are plotted against the latest data, where rates are untransformed and the other data enter as quarterly growth rates. For visibility, the scale is normalized to the range $[0, 1]$. The shades of blue indicate the vintage date, where darker shades are for later vintages. For

series without data revisions, such as those in the Financial category, the observations lie on the 45 degree line; the latest data are identical to the data first released. For the three other panels corresponding to non-financial data however, the pattern differs. For earlier vintages, observations are farther from the 45 degree line. These subplots illustrate that accounting for data revisions for evaluating model-based forecasts is crucial—recursive evaluation using the latest data would effectively result in substantially different information sets.