Reconciling VAR-based Forecasts with Survey Forecasts*

Taeyoung Doh †
Federal Reserve Bank of Kansas City

A. Lee Smith‡
Federal Reserve Bank of Kansas City

Preliminary and Incomplete
August 6, 2018

Abstract

We propose a novel Bayesian approach for vector autoregressions (VARs) when survey forecasts are included in observed variables. In particular, our method poses a prior distribution on the consistency between the VAR model-based forecasts and survey forecasts. We demonstrate that the restrictions from the consistency of forecasts imply nonlinear cross-equation restrictions on VAR coefficients. By combining importance sampling with a maximum entropy prior for cross-equation restrictions, we obtain posterior draws of VAR parameters under cross-equation restrictions without much computational burden. We use two empirical applications to illusrate the implication of imposing cross-equation restrictions: 1) the evaluation of time-varying tail risks for inflation in a time-varying parameter VAR model, and 2) impulse-responses of macroeconomic variables to a monetary policy shock in a conventional monetary VAR.
Key Words: Vector Autoregression, Survey Forecasts, Bayesian Approach. JEL Classification: C11, C32, E31.

1 Introduction

Changing expectations about the future economic outlook are important for understanding macroeconomic dynamics based on people's intertemporal decisions. Consumption and investment as well as price-setting all depend on agents' beliefs of the future. Many empirical studies use forecasts based on a vector autoregression (VAR) as proxies for

*The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City, the Federal Reserve System.

†Economic Research Department, Federal Reserve Bank of Kansas City, 1 memorial Drive, Kansas City, MO 64198. Phone: (816) 881-2780, Email: Taeyoung.Doh@kc.frb.org.

‡Economic Research Department, Federal Reserve Bank of Kansas City, 1 memorial Drive, Kansas City, MO 64198. Phone: (816) 881-2294, Email: andrew.smith@kc.frb.org.

expectations (e.g. Campbell and Shiller (1988)). However, using multi-step ahead VAR forecasts as proxies for long-term expectations crucially depends on the assumption that the current and past values included in the VAR correctly represent all information available to agents, which may not hold when it includes only a small number of aggregate macro variables. Leduc and Sill (2013) argue that this problem can be alleviated by adding survey forecasts to the conventional VAR framework because they can provide an independent source of information on expectations beyond macroeconomic aggregates. In particular, survey forecasts for inflation turn out to be more accurate than other model-based forecasts without using survey data (Ang, Bekaert, and Wei (2007); Faust and Wright (2013)) and have been used to study the time variation in long-term inflation expectations (Clark and Davig (2011), Kozicki and Tinsley (2012), Mertens (2016)).

However, simply adding survey data into the VAR framework without imposing the consistency between the VAR forecast and the survey forecast can generate an internal inconsistency. As noted by Cogley (2005), if one variable in the VAR (e.g. long-term interest rate) can contain information on future expected values of another variable (e.g. short-term interest rate) in the VAR, we obtain two different forecasts for the same variable without imposing cross-equation restrictions on the VAR coefficients. Imposing such theoretical restrictions would reduce the dimension of VAR coefficients because not all VAR coefficients would be free parameters. For example, Kozicki and Tinsley (2012) impose the internal consistency between expected inflation based on an autoregressive (AR) model of inflation with time-varying intercepts and survey-based measures of expected inflation, expressing parameters determining the evolution of survey forecasts for inflation as explicit functions of parameters governing the evolution of the realized inflation. While imposing the internal consistency strictly is theoretically appealing, the practice may overly restrict the process representing the evolution of survey forecasts because there are no independent parameters to govern it.

In this paper, we propose a Bayesian approach to address the internal consistency issue. By incorporating cross-equation restrictions implied by the internal consistency into priors for VAR coefficients, we can nest the strict imposition of cross-equation restrictions as a special case. Our method shrinks VAR coefficients in the direction of satisfying cross-equation restrictions by constructing non-degenerate prior distributions for VAR coefficients. Since we add only prior information to the existing VAR framework, the posterior simulation does not change in any drastic way. We resample posterior draws obtained without using prior information from cross-equation restrictions by importance sampling weights based on prior information from cross-equation restrictions.[1] We call our method "model-consistent priors for survey expectations ".

Our paper is related to the literature on using informative priors to improve out-

[1]Our method of using importance sampling seems to be similar to the one used by Antolín-Díaz and Rubio-Ramírez (2018), who resample posterior draws of structural VARs to impose narrative sign restrictions. One main difference is that their importance sampling weights are determined by the probability distribution of exogenous shocks because narrative sign restrictions are imposed on the likelihood rather than prior distributions of parameters.

of-sample forecasting and the accuracy in the estimation of impulse response functions in VARs (e.g., Del Negro and Schorfheide (2004, DS hereafter), Giannone, Lenza, and Primiceri (2015, 2017, GLS hereafter)). These papers shrink VAR coefficients toward a more parsimonious specification. In DS (2004), the parsimonious specification is a dynamic stochastic general equilibrium (DSGE) model while it is the long-run dynamics of macroeocnomic variables from a structural model in GLS (2017) or a combination of naive benchmark models in GLS (2015). Our model-consistent priors will also shrink VAR coefficients in the direction of a more parsimonious specification in which coefficients in equations for survey expectations are determiend by coefficients in other equations via nonlinear restrictions. Our approach is more closely related to DS (2004) and GLS (2017) than GLS (2015) in that priors are derived from nonlinear restrictions based on economic theory. However, unlike DS (2004), we do not take a stance on an underyling equilibrium model that might have generated data used in the estimation of the VAR model. We only seek the internal consistency of forecasts inside the VAR model implied by the theory of rational expectations. Likewise, we do not take a stance on the long-run equilibrium relationship between different macroeconomic variables such as real output and real consumption in GLS (2017). Our nonlinear restrictions are derived from imposing the consistency of forecasts for the *same* variable. In this sense, our model-consistent priors for survey expectations are more flexible than the method used in the existing literature on imposing theory-based priors for VARs. The flexibility in our approach is comparable to Frey and Mokinski (2015) who also impose the consistency of forecasts based on a time-series model with survey forecasts in a probabilistic way. However, their time-series model for inflation does not allow that survey forecasts can predict inflation above and beyond lagged inflation data. In their specification, survey nowcasts help inflation forecasting by reducing the estimation uncertainty of parameters in the time-series model only because survey nowcasts and actual inflation depend on lagged inflation data in a similar way. In contrast, our specification allows the possibility that survey forecasts contain forward-looking information on future inflation.

We illustrate implications of imposing forecast consistency restrictions using a time-varying parameter (TVP) VAR model. The model includes survey forecasts for inflation with realized inflation, a real activity measures and a monetary policy instrument like in Clark and Davig (2011). We compute time-varying tail risks that have received much attention since the recession of 2007-09 (Anene and D'Amico (2017); Fleckenstein et al. (2017); Hattori, Schrimpf, and Sushko (2016)). Tail risks in in our model are quite comparable to estimates based on financial market data such as inflation swaps and options from Fleckenstein et al. (2017) although we do not use financial market at all. Imposing forecast consistency restrictions reduces tail risks of inflation for most of the sample period, reflecting the stability of inflation expectations in survey data during the recent decades. However, estimated deflation probability quickly spikes up and stays at a higher level for the period of 2008:Q2-2009:Q4 when forecast consistency restrictions are imposed. During this period, inflation expectations in survey data that have been stable for a quite long time fell somewhat. With consistency restrictions imposed, even small movements in survey forecasts seem to have bigger impacts on the model-based

projection of inflation. Since the volatility of survey forecasts since 1997 has been quite low, they should be able to povide more precise signals of future inflation conditional on changes. Imposing forecast consistency restrictions amplifies the influence of survey forecasts on the model-based forecast of future inflation and makes tail risk to be more sensitive to changes in survey forecasts.

We then show the implications of imposing forecast consistency to impulse responses. Using traditional monetary VAR models augmented with measures of short-term and long-term inflation expectations, we find that imposing the weak-consistency prior leads to more muted effects of inflation to a monetary policy shock and an amplification of the effects on real GDP. These results are consistent with a larger degree of nominal rigidities and therefore more monetary non-neutrality in a New-Keynesian DSGE model. For example, in the typical Calvo setting, if the average duration of prices increases then the price effects to a monetary policy shock decline and the effects on the real economy increase.

The rest of our paper is organized as follows. Section 2 illustrate the basic idea behind our method using a simple example of a bi-variate VAR(1) with one-period ahead survey forecasts. Section 3 provides details of our methodology in the context of a more general case with multi-horizon and multiple survey forecasts. Section 4 illustrates how to apply the methodology to conditional forecasts, such as impulse response functions from structural VAR models. Section 5 focuses on the empirical application of our method using the TVP-VAR model. Section 6 analyzes the effects of monetary policy in a VAR framework that imposes some form of consistency between VAR and survey forecasts. Section 7 concludes.

2 A simple example: VAR(1) with one-period ahead survey forecasts

Let's consider the following bivariate VAR(1) model for $y_t = [x_t\,,\, E_t^s(x_{t+1})]$. x_t is a variable of interest such as inflation or output growth and $E_t^s(x_{t+1})$ is the one-period ahead forecast from the available survey on x_t.

$$y_t = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} y_{t-1} + u_t = Ay_{t-1} + u_t\,,\; u_t \sim i.i.d.\mathcal{N}(0, \Sigma_u). \tag{1}$$

We assume the following Normal and Inverse-Wishart conjugate priors for ($\alpha = \text{vec}(A)\,,\, \Sigma_u$).

$$\begin{aligned} \Sigma_u &\sim IW(\Psi, d), \\ \alpha|\Sigma_u &\sim \mathcal{N}(\overline{\alpha}, \Sigma_u \otimes \Omega), \end{aligned} \tag{2}$$

where $\Psi, d, \overline{\alpha}$ and Ω are hyperparameters. Let's stack $[y_1, \cdots, y_T]'$ by Y and $[y_0, \cdots, y_{T-1}]$ by X. The ordinary least squares (OLS) estimator for A is $(X'X)^{-1}(X'Y)$ and the posterior mean of A is

$$\widehat{A} = (X'X + \Omega^{-1})^{-1}(X'Y + \Omega^{-1}\overline{A}), \tag{3}$$

where $\text{vec}(\overline{A}) = \overline{\alpha}$ and $\widehat{\alpha} = \text{vec}(\widehat{A})$. The posterior draws of Σ_u and α follow the Normal and Inverse-Wishart distributions.

$$\begin{aligned} \Sigma_u | y_0^T &\sim IW(\Psi + \widehat{u}'\widehat{u} + (\widehat{A} - \overline{A})'(\Omega)^{-1}(\widehat{A} - \overline{A}), T - 1 + d), \\ \alpha | \Sigma_u, y_0^T &\sim \mathcal{N}(\widehat{\alpha}, \Sigma_u \otimes (X'X + \Omega^{-1})^{-1}), \end{aligned} \tag{4}$$

where $\widehat{u} = Y - X\widehat{A}$.

Now we turn to imposing cross-equation restrictions implied by the internal consistency of forecasts. From the first equation of the VAR, we can generate the VAR-based forecast for x_{t+1} at time t.

$$E_t^{\text{VAR}}(x_{t+1}) = a_{11}x_t + a_{12}E_t^s(x_{t+1}). \tag{5}$$

Imposing the consistency restriction that $E_t^{\text{VAR}}(x_{t+1}) = E_t^s(x_{t+1})$ for all the values of x_t implies the following restrictions on VAR coefficients.

$$g(A) = e_1'A - e_2' = [0, 0] \Leftrightarrow [a_{11}, a_{12} - 1] = [0, 0]. \tag{6}$$

In this case, a_{21} and a_{22} remain free parameters. Under this restriction, $u_{1,t}$ is a martingale difference series of unanticipated shocks to x_t while $u_{2,t}$ represents a news shock for x_{t+1} at time t. By adopting the Bayesian approach, we can vary the tightness of these cross-equation restrictions to allow a possibility that survey forecasts may be an imperfect proxy for the VAR-based conditional forecast. Depending on the degree of tightness, we can consider the three cases as follows:

1. Strong Consistency: $g(A) = [0, 0]$. Therefore, $a_{11} = 0$ and $a_{12} = 1$.
2. Weak Consistency: $g(A) \sim \mathcal{N}(0, \lambda W^{-1})$. In other words, the average one-period ahead survey forecast is consistent with the average VAR-based one-period ahead forecast. But they may deviate from each other from time to time. This notion of weak consistency can be modelled by allowing non-degenerate prior distributions for $g(A)$. The normal distribution is the maximum entropy prior for $g(A)$ under constraints on first and second moments. So the prior under weak consistency is the distribution maximizing entropy under the constraint that the first moment

of $g(A)$ is zero and the second moment of $g(A)$ is λW^{-1}. Our choice of the maximum entropy prior is motivated by the fact that it minimizes the amount of prior information on the moments of $g(A)$ translated to the prior distribution, in the information-theoretic sense.[2]

3. No Consistency: $g(A)$ is not restricted and we can assume a diffuse prior for it.

Strong consistency and no consistency can be regarded as limiting cases when λ approaches 0 and ∞, respectively. In terms of A, the prior density function of $g(A)$ can be treated like the likelihood function for observations satisfying these restrictions. We use this property to calculate the following posterior density of α that satisfy cross-equation restrictions.

$$\begin{aligned} p(\alpha|\Sigma_u, y_0^T, g(A)) &= \frac{p(g(A))p(\alpha|y_0^T, \Sigma_u)}{\int p(g(A))p(\alpha|y_0^T, \Sigma_u,)dg(A)} \\ &= w(A)p(\alpha|y_0^T, \Sigma_u). \end{aligned} \tag{7}$$

We use importance weights ($w(A)$) to simulate posterior draws for α that respect prior information on cross-equation restrictions. Suppose we obtain M equally weighted posterior draws for α without using prior information on $g(A)$, $[\alpha^1, \cdots, \alpha^M]$. Now, we resample these draws according to new weights $[w(A^1), \cdots, w(A^M)]$.

3 A general case involving multi-horizon and multiple survey forecasts

Our methodology can be easily extended to a more general case involving multi-horizon and multiple survey forecasts. In this general case, consistency restrictions become nonlinear functions of VAR coefficients.

3.1 Multi-horizon survey forecasts

Survey forecasts typically involve multi-horizon forecasts as well as one-period ahead forecasts. We can derive implications of the consistency in forecasts for VAR coefficients similarly. Since the multi-step forecast in the VAR involves the repeated iteration of one-step ahead forecasts, the resulting cross-equation restriction shows up as a completely nonlinear function of all the VAR coefficients. For example, let's consider the following VAR(1) model with two-step ahead forecasts where $y_t = [x_t, E_t^s(x_{t+2})]'$. The VAR implied two-step ahead forecast is

$$E_t^{\text{VAR}}(x_{t+2}) = (a_{11}^2 + a_{12}a_{21})x_t + a_{12}(a_{11} + a_{22})E_t^s(x_{t+2}). \tag{8}$$

[2]See Robert (2001), Cover and Thomas (1991).

The strong consistency implies that the following two nonlinear restrictions should hold for VAR coefficients.

$$\begin{aligned} g_1(A) &= a_{11}^2 + a_{12}a_{21} = 0, \\ g_2(A) &= a_{12}(a_{11} + a_{22}) - 1 = 0. \end{aligned} \tag{9}$$

We can put together the two restrictions into a vector of nonlinear equations $g(A)$ as in the previous sectcion. Unlike the one-step ahead forecast case, solving for a_{11} and a_{12} as functions of a_{21} and a_{22} to impose the strong consistency is complicated because there might be cases of multiple solutions or no solution. This raises a particular challenge when we want to impose strong consistency restrictions as in Cogley (2005). However, this is not a serious issue for weak consistency because we set a prior for $g(A)$ rather than A.

3.2 Multiple survey forecasts

The VAR can include multiple survey forecasts of different horizons. For instance, y_t can contain a long-horizon survey forecast as well as a short-horizon one as follows:

$$y_t = Ay_{t-1} + u_t \,,\; y_t = [x_t, E_t^s(x_{t+1}), E_t^s(x_{t+2})]'. \tag{10}$$

In general, the number of cross-equation restrictions implied by consistency requirements is equal to the number of parameters describing the evolution of survey forecasts. So in the above three-variable VAR(1) with two survey forecasts, consistency requirements generate $2 \times 3 = 6$ cross-equations on 9 parameters in A. To generalize this, we look at the dimension of cross-equation restrictions from the following a VAR(P) with non-zero constants and n_s survey forecasts.

$$y_t = A_0 + A_1[y'_{t-1}, \cdots, y'_{t-p}] + u_t \,,\; y_t = [x_{1,t}, \cdots, x_{n_x,t}, E_t^s(x_{1,t+1}), \cdots, E_t^s(x_{n_x,t+h})]'. \tag{11}$$

The number of equation (n_y) is equal to the sum of the number of x_t variables (n_x) and the number of survey forecasts ($n_s <= n_x \times h$). Consistency requirements will impose $n_s \times (p+1)$ cross-equation restrictions on the $(n_x + n_s) \times (p+1)$ matrix $A = [A_0, A_1]$.

4 Conditional forecast consistency in impulse responses

While the methodology to this point has emphasized consistency between the unconditional VAR forecasts, the approach can also be applied to functions of conditional forecasts. Impulse responses to identified structural shocks represent one such conditional forecast that is of interest to macroeconomists. For illustration, consider the structural VAR using the notation used in Section 3.2 above:

$$By_t = BA_0 + BA_1[y'_{t-1}, \cdots, y'_{t-p}] + \varepsilon_t, \tag{12}$$

where ε_t are othrognal structural shocks, each with a standard deviation of one. These structural shocks are related to the reduced form residuals by way of the mapping $\varepsilon_t = B^{-1}u_t$. It will be convenient to define the cross-equation restrictions associated with expectational consistency in terms of the companion form of the VAR:

$$Y_t = \mathcal{A}_0 + \mathcal{A}_1 Y_{t-1} + \mathcal{B}^{-1}U_t, \tag{13}$$

where $Y_t = [y_t', \cdots, y_{t-p-1}']$, $U_t = [u_t; 0_{n_y(p-1)\times 1}]$, $\mathcal{A}_0 = [A_0; 0_{n_y(p-1)\times 1}]$, $\mathcal{A}_1 = [A_1; I_{n_y} \otimes I_{(p-1)}, 0_{n_y(p-1)\times 1}]$, and $\mathcal{B}^{-1} = [B^{-1}; 0_{n_y(p-1)\times n_y}]$.

Suppose that we are interested in the response to the structural innovation ε_t^1 occurring at time t. Then, the k-step ahead forecast of all the variables is given by :

$$E^{VAR}_{t|\varepsilon_t^1}(Y_{t+k}) = [I_{n_y p} + \mathcal{A}_1 + \cdots + \mathcal{A}_1^k]\mathcal{A}_0 + \mathcal{A}_1^{k+1}Y_{t-1} + \mathcal{A}_1^k\mathcal{B}^{-1}[1, 0, \cdots, 0]. \tag{14}$$

Define the k-step ahead impulse response of all the variables as the difference between this conditional forecast and the unconditional forecast made at time t:

$$\begin{aligned} IRF(k) &= E^{VAR}_{t|\varepsilon_t^1}(Y_{t+k}) - E^{VAR}_t(Y_{t+k}) \\ &= [I_{n_y p} + \mathcal{A}_1 + \cdots + \mathcal{A}_1^k]\mathcal{A}_0 + \mathcal{A}_1^{k+1}Y_{t-1} + \mathcal{A}_1^k\mathcal{B}^{-1}[1, 0, \cdots, 0] \\ &- [I_{n_y p} + \mathcal{A}_1 + \cdots + \mathcal{A}_1^k]\mathcal{A}_0 + \mathcal{A}_1^{k+1}Y_{t-1} \\ &= \mathcal{A}_1^k\mathcal{B}^{-1}[1, 0, \cdots, 0]. \end{aligned} \tag{15}$$

Let e_i be a $n_y p \times 1$ vector whose $i'th$ element is one and all other elements are zero. Then, if the variable ordered in position N_x is the same variable being foretasted one-period ahead in survey data in position N_s, the internal consistency of forecasts would require that:

$$e_{N_x}' IRF(k+1) = e_{N_s}' IRF(k). \tag{16}$$

Therefore, forecast consistency would require that:

$$e_{N_x}'\mathcal{A}_1^{k+1}\mathcal{B}^{-1}[1, 0, \cdots, 0] = e_{N_s}'\mathcal{A}_1^k\mathcal{B}^{-1}[1, 0, \cdots, 0]. \tag{17}$$

Notice that two features of this condition contrasts with the unconditional forecast consistency constraints presented in Section 3. First, even if consistency is imposed in response to every structural shock, forecast consistency in the impulse responses imposes no constraints on the VAR intercept term. This is because the intercept has no direct bearing on the impulse responses form the VAR. Of course, the unconditional forecasts depend crucially on A_0. Second, the forecast consistency criteria depends on the matrix B^{-1} which provides the identification of the structural shocks, or the mapping between reduced form VAR residuals and structural shocks. This feature suggests that the identification of structural shocks and the responses of data and survey forecasts of

that data are not independent.

Just as above, we will consider a weak-consistency prior denoted now by $g(A, B, h) \sim \mathcal{N}(0, \lambda W^{-1})$ that relaxes the strict requirement for Equation (17) to hold at every impulse response horizon h. The horizon h is a hyperparameter that calibrates the horizon over which the prior applies to the impulse responses. In our applications below, h aligns with the horizon over which impulse responses are plotted and displayed. In the application of this weak consistency prior, we study the impulse response functions to a monetary policy shock in a traditional monetary VAR augmented with survey forecasts of inflation.

5 Application I: TVP-VAR with survey forecasts of inflation

Kozicki and Tinsley (2012) estimate a time-varying parameter (TVP) AR(p) model for inflation together with survey forecasts. To efficiently use survey information on the estimation of time-varying coefficients, they impose consistency requirements that survey forecasts a noisy observation of forecasts implied by the TVP-AR(p) model. After introducing a noise term to survey forecasts, they essentially impose weak consistency requirements in our classification. The prior co-variance matrix for cross-equation restrictions in our framework is comparable to the variance matrix of the noise term in their formulation. In Kozicki and Tinsley (2012), only constant terms ("inflation endpoint") are allowed to drift over time while other coefficients in the AR(p) model remain time-invariant. Keeping AR coefficients time-invariant has some limitations in modelling possible time-variation in the relationship between inflation expectations and other variables. Changes in monetary policy regimes such as the adoption of explicit inflation target, if credible, will make inflation expectations less responsive to temporary shocks.[3] For this reason, we consider a more general TVP-VAR model in which not only constant terms but also other VAR coefficients drift over time. Our data include survey forecasts for short-term and long-term inflation and we impose weak consistency requirements for the model-based forecasts.

[3]Bundick and Smith (2018) show that the announcement of a numerical inflation target by the Federal Open Market Committee (FOMC) in 2012 reduced the sensitivity of the long-horizon forward inflation to temporary surprises in the realized inflation. Although their finding is based on the high-frequency financial market data, Doh and Oksol (2018) finds that the predictability of inflation including long-horizon inflation expectations from survey data also declined after the announcement of the explicit inflation target.

5.1 Data

As in Clark and Davig (2011), we use both 1-year and 10-year ahead forecasts for Consumer Price Index (CPI) inflation from the Survey of Professional Forecasters (SPF) as well as realized CPI inflation. Our measure of long-term inflation is the average expected inflation between the five quarter from now and the fourth quarter of ten years from the current year.[4] We include a measure of real economic activity (Chicago Fed National Activity Index, x_t) and a monetary policy indicator (the effective federal funds rate spliced together with Wu-Xia (2016) shadow rate (r_t) measure during the period when the short-term interest rate is constrained at the effective lower bound) on top of the three variables related to inflation. Figure 1 shows all the variables used in estimation. While the real economic activity measure does not show any trending behavior, all the other nominal variables exhibit downward trends since the early 1980s. The TVP-VAR model can accommodate time-varying trends in variables with the random-walk drifts of VAR coefficients.

5.2 TVP-VAR Model

We consider the following TVP-VAR(2) model with stochastic volatility for the five variables $y_t = [\pi_t^{S,L}, \pi_t^{S,S}, \pi_t, x_t, r_t]$:

[4]The 10-year forecast for CPI inflation from the SPF moves the forecast window only at the first quarter of each year. For example, the 10-year forecast for CPI inflation from the SPF at each quarter of 1990 is the expected inflation averaged over 1990:Q1-1999:Q4. In contrast, the 1-year forecast from the SPF moves the forecast window each quarter. Therefore, simply taking the difference between the 10-year forecast and the 1-year forecast will leave some overlapping with realized inflation data depending on the quarter of the year. We make additional adjustments using realized CPI inflation data and nowcasts observations to ensure the long-term forward inflation component does not overlap with other inflation-related variables in the VAR. For details, see the appendix.

$$
\begin{aligned}
y_t &= A_{0,t} + \sum_{j=1}^{2} A_{j,t} y_{t-j} + B^{-1} u_t \,,\; u_t \sim \mathcal{N}(0, \Sigma_{u,t}), \\
\Sigma_{u,t} &= \begin{pmatrix} \sigma_{1,t}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{2,t}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{5,t}^2 \end{pmatrix} \\
\tilde{y}_t &= \tilde{A}_{0,t} + \tilde{A}_{1,t} \tilde{y}_{t-1} + \tilde{u}_t, \\
\tilde{y}_t &= \begin{bmatrix} y_t \\ 0_{(5\times 5)} \end{bmatrix}, \tilde{A}_{0,t} = \begin{bmatrix} A_{0,t} \\ 0_{(5\times 5)} \end{bmatrix}, \tilde{A}_{1,t} = \begin{pmatrix} A_{1,t} & A_{2,t} \\ I_5 & 0_{(5\times 5)} \end{pmatrix}, \tilde{u}_t = \begin{bmatrix} B^{-1} u_t \\ 0_{(5\times 5)} \end{bmatrix}, \\
A_t &= [A_{0,t}, A_{1,t}, A_{2,t}]', \\
\mathrm{vec}(A_t) &= \mathrm{vec}(A_{t-1}) + \epsilon_t \,,\; \epsilon_t \sim \mathcal{N}(0, Q), \\
\ln(\sigma_{i,t}^2) &= \ln(\sigma_{i,t-1}^2) + e_{i,t} \,,\; e_{i,t} \sim \mathcal{N}(0, \sigma_{e,i}^2).
\end{aligned}
\tag{18}
$$

$\pi_t^{S,L}$ and $\pi_t^{S,S}$ are weighted averages of quarterly forward inflation at different horizons as defined below while π_t is realized CPI quarterly inflation.[5]

$$
\begin{aligned}
\pi_t^{S,L} &= \frac{\sum_{j=5}^{l(t)} E_t^s(\pi_{t+j})}{l(t) - 4}, \\
\pi_t^{S,S} &= \frac{\sum_{j=1}^{4} E_t^s(\pi_{t+j})}{4}.
\end{aligned}
\tag{19}
$$

Consistency requirements for $\pi_t^{S,L}$ and $\pi_t^{S,S}$ lead to 22 cross-equation restrictions $g(A)$:

[5]Since the forecast horizon of the 10-year forecast from the SPF changes only at the first quarter of each year, the number of quarterly forward inflation contained in the 10-year forecast varies depending on the quarter of year. In principle, we could treat four quarterly observations of the 10-year forecast of any given year as one quarterly observation for four different measures of long-term forecasts to keep the same forecast horizon for each measure of long-term forecasts. However, doing so will require us to run a mixed-frequency TVP-VAR that will substantially increase the dimension of parameters and latent variables. We leave this for the future research agenda.

$$
\begin{aligned}
g(A)_L &= \pi_t^{S,L} - \frac{\sum_{j=5}^{l(t)} E_t^{\text{VAR}}(\pi_{t+j})}{l(t)-4}, \\
&= e_1'\tilde{y}_t - \frac{[\sum_{h=5}^{l(t)} \sum_{j=0}^{h-1} \tilde{A}_{1,t}^j \tilde{A}_{0,t} + \sum_{h=5}^{l(t)} \tilde{A}_{1,t}^h \tilde{y}_t]}{l(t)-4}, \\
g(A)_S &= \pi_t^{S,S} - \frac{\sum_{j=1}^{4} E_t^{\text{VAR}}(\pi_{t+j})}{4}, \\
&= e_2'\tilde{y}_t - \frac{[\sum_{h=1}^{4} \sum_{j=0}^{h-1} \tilde{A}_{1,t}^j \tilde{A}_{0,t} + \sum_{h=1}^{4} \tilde{A}_{1,t}^h \tilde{y}_t]}{4}. \\
g(A) &= [g(A)_L, g(A)_S]',
\end{aligned}
\tag{20}
$$

where e_i is a 10×1 vector whose ith element is 1 while all the other elements are zeros. To calculate VAR-implied forecasts in the above equation, we assume that agents do not reflect future drifts in parameters when they form expectations. In other words,

$$
E_t^{\text{VAR}}(\prod_{k=1}^{h} \tilde{A}_{1,t+k}) = A_{1,t}^h \,,\; E_t^{\text{VAR}}(\prod_{k=1}^{h} \tilde{A}_{2,t+k}) = A_{2,t}^h \,,\; E_t^{\text{VAR}}(A_{0,t+h}) = A_{0,t}. \tag{21}
$$

This "anticipated utility "approximation works well for mean forecasts. Since we impose consistency requirements only on point forecasts from surveys, using the approximation might be less problematic.

5.3 Priors

To determine priors for initial states $(A_{0,0}, A_{1,0}, A_{2,0}, \Sigma_{u,0})$ and parameters (Q, B, σ_e^2), we split the training sample data from 1970:Q2 to 1981:Q2. Initial values of time-varying coefficients (A_0) and the covariance matrix of innovations to time-varying coefficients (Q) are calibrated from the Ordinary Least Squares (OLS) estimates of a time-invariant VAR(2) using the training sample data. The prior mean of A_0 are the OLS estimates of coefficients from the VAR(2) for the training sample data. The prior mean of Q is set to be proportional to the covariance matrix of the OLS estimates. More specifically, we rescale the covariance matrix of the OLS estimates to make the implied variation of the long-run mean of y_t plausible given the magnitude of the uncertainty about the the mean of y_t. More specifically, we generate A_{t+1} by adding one standard deviation shocks to all the components in A_t using the Cholesky decomposition of Q and calculate the implied variation of the long-run mean as follows:

$$
\Delta\bar{y}_{t+1} = [1_{1\times 5}, 0_{1\times 5}]((I_{10} - \tilde{A}_{1,t+1})^{-1}\tilde{A}_{0,t+1} - (I_{10} - \tilde{A}_{1,t})^{-1}\tilde{A}_{0,t}). \tag{22}
$$

We calibrate the prior mean of Q to match the implied $\Delta\bar{y}_{t+1}$ with the standard deviation of the constant mean of y_t under the assumption innovations to y_t are independently and identically distributed for the training-sample data. Similarly, priors for the initial values of stochastic volatilities ($\Sigma_{u,0}$) are calibrated from the estimated covariance matrix of the residuals in the time-invariant VAR(2). The prior for B is elicited by scaling up an identity matrix and the prior mean for σ_e^2 is set to 0.01, which implies small variations in the innovation to log volatility. As common in the TVP-VAR literature, we assume that the initial values of latent states and other parameters are a priori independent.

$$p(A_0, \Sigma_{u,0}, B, Q, \sigma_e^2) = p(A_0)p(\Sigma_{u,0})p(B)p(Q)p(\sigma_e^2). \tag{23}$$

To impose forecast consistency requirements, we need to calibrate hyperparameters such as λ and W that determine the maximum entropy prior for $g(A)$. Since the prior for $g(A)$ follows a normal distribution, W needs to match the covariance matrix of $g(A)$. We simulate multiple draws of A from the prior distribution and compute $g(A)$ for each draw. From these multiple realizations of $g(A)$, we compute the covariance matrix and set W equal to it. For λ, there is no a priori guideline to calibrate the value. As common in a hierarchical model, we estimate the TVP-VAR(2) model with different values of λ and pick up the value that gives the best fit for the sample data using a model selection criterion such as marginal likelihood.

5.4 Posterior simulation

We obtain the posterior output of latent states and parameters using Gibbs sampling. While the joint posterior distribution of $(A^T, \Sigma_u^T, Q, B, \sigma_e^2)$ is difficult to characterize analytically, the distribution of one component conditional on all the other components is either analytically tractable or easy to simulate.[6] Given hyperparameters (λ, W), the joint posterior distribution of latent states and parameters can be obtained as the product of conditional posterior distributions as follows:[7]

$$\begin{aligned} p_{\lambda,W}(A^T, \Sigma_u^T, B, Q, \sigma_e^2 | y^T) \propto\ & p(Q|A^T, \Sigma_u^T, B, \sigma_e^2, y^T)p(Q) \\ & \times p(\sigma_e^2|A^T, \Sigma_u^T, B, Q, y^T)p(\sigma_e^2) \\ & \times p(\Sigma_u^T|A^T, B, Q, \sigma_e^2, y^T)p(\Sigma_{u,0}) \\ & \times p(B|A^T, \Sigma_u^T, Q, \sigma_e^2, y^T)p(B) \\ & \times p_{\lambda,W}(A^T|\Sigma_u^T, B, Q, \sigma_e^2, y^T)p(A_0). \end{aligned} \tag{24}$$

In the last step, we first simulate M draws of A^T from $p(A^T|\Sigma_u^T, B, Q, \sigma_e^2, y^T)p(A_0)$ and then resample them using importance-sampling weights given by

[6] We use superscript to denote an array of observations up to the point at the superscript. For example, A^T represents an array of $[A_0, \cdots, A_T]$.

[7] The technical details of the implementation is explained in the appendix.

$$w(A^T(j)) = \frac{M \exp^{-0.5 g(A^T(j))'(\lambda W)^{-1} g(A^T(j))}}{\sum_{k=1}^{M} \exp^{-0.5 g(A^T(k))'(\lambda W)^{-1} g(A^T(k))}}. \quad (25)$$

5.5 Inflation Tail Risks

Since the Great recession of 2007-2009, deflation risk has received a lot of attention from financial markets and policymakers. Investors have used inflation-related derivatives to hedge against deflation risks and policymakers often justified the use of unconventional monetary policies such as large-scale asset purchases and forward guidance on the future path of the short-term interest rate to address impending deflation risks. Financial derivatives on inflation such as inflation swaps and options as well as survey forecasts provide information on future deflation risk. Fleckenstein et al. (2017) compute tail risks for inflation implied by estimating a no-arbitrage model for inflation swaps and options. Alternatively, we can use survey data on inflation to back out deflation probabilities. Survey data information can be combined with forecasts based on a time-series model of inflation as we do here in the TVP-VAR model or can be used to construct deflation probabilities by itself.

Table 1 compares their estimated deflation probabilities with those from the TVP-VAR model and the University of Michigan consumer survey. For the TVP-VAR model, we simulate future realization of inflation conditional on each draw of parameters assuming that they do not change in the futurre, which is consistent with our anticipated utility approximation in forecast consistency restrictions. For the University of Michgian survey, we calculate the percentage of respondents who anticipated prices would go down among all the respondents who provided answers on expected prices. To match the sampler period in Fleckenstein et al. (2017), we calcualte summary statistics for deflation probabilities for 2009:Q4-2015:Q4. Although the underlying source data are not overlapping, mean and median numbers are quite comparable, ranging from 10 to 20 percent. In terms of volatility and dispersion, the TVP-VAR model with resampling ranks first followed by estimates based on financial market data in Fleckenstein et al. (2017). The Univerisity of Michigan survey generates the smallest dispersion but this is partly attributable to the fact that we take into account only the cross-sectional dispersion in mean forecasts of each indiviual respondent by construction, ignoring future shock uncertainty. The TVP-VAR model with resampling generates median estimates closer to Fleckenstein et al. (2017) than the TVP-VAR model without resampling. However, resampling also increases skewness in the distribution of deflation probability with the distance between the maximum and the median substantially bigger than in estimates without resampling. The maximum deflation probability is much greater in estimates with resampling than otherwise. In contrast, the estimated probability of high inflation (inflation greater than 4 percent averaged over the forecast horizon) in Table 2 suggests that resampling attenuates the probability of high inflation.

Does this finding imply that resampling always overweights paramter draws increasing the probability of deflation while downweights parameter draws increasing the probability of high inflation? How do we reconcile this finding with the overall stability of survey-based inflation expectations shown in Figure 1? The time-varying tail risks from the TVP-VAR model in Figure 2 provides answers to these questions. For the early part of the sample (up to the mid 1980s), resampling actually amplifies the probability of high inflation. For the recent three decades, resampling generally implies lower tail risks either on the upside or on the downside except for the period from 2008:Q2 to 2009:Q4 related to the financial crisis and recession. It is only during this period that deflation probabilities implied by estimates with resampling overcedes deflation probabilities implied by estimates without resampling. Although survey-based inflation expectations have been quite stable for the recent two decades, they notably moved downwards during this short period relative to the remarkable stability for the recent two decades. Resampling tends to amplify the influence of movements in survey-based inflation expectations on realized inflation as it imposes forecast consistency restrictions. In other words, resampling makes the model-based inflation forecasts more sensitive to movements in survey-based inflation expectations. Therefore, resampling does not always overweight parameter draws increasing the probability of deflation. Indeed, estimates with resampling suggests lower probabilities of deflation since 2010 than those without resampling, consistent with the stability in survey-based inflation expectations.

6 Application II: Impulse Response to a Monetary Policy Shock

A large literature has sought to measure the effects of monetary policy shocks using recursive VAR models. This class of VAR models imposes timing restrictions on when the Federal Reserve reacts to information and when those actions transmit to the economy. For example, variables ordered ahead of the policy instrument are the key indicators which prompt a response form the Fed within the period. These variables often include measures of economic activity and price dynamics which are thought to respond to monetary policy with some lag. Other variables included in the VAR are assumed to respond within the period to the Fed's actions to capture the transmission of Fed policy to financial and money markets. While these assumptions have been criticized on various grounds and there have been some attempts to advance beyond the so-called "recursive" assumption to identify monetary policy shocks, it remains a benchmark model to capture the effects of monetary policy. For our purposes, since the identification is familiar to many, these models provide a useful framework to apply our methodology. Therefore, we augment these VAR models with survey forecasts of both near term and long-term inflation expectations to illustrate the weak-consistency prior.

The suite of monetary VAR models considered in Christiano, Eichenbaum, and Evans (1999) does not directly include survey measures of inflation expectations. However,

macroeconomists have long understood the importance of forward looking price behavior for successfully identifying monetary policy shocks. For example,Sims (1992) speculates that failure to account for the forward looking nature of central banks when contemplating the outlook for inflation could lead to misidentified monetary policy shocks. For example, if inflation is stable today, but the Fed anticipates inflation will rise in the future they begin to tighten policy. While a VAR with no forward looking elements may classify the policy tightening as a shock, in all actuality the Fed was reacting to the future state of the economy. Therefore, Sims speculated that the price puzzle, or the tendency of prices to initially rise following a monetary contraction, may stem from the possibility that, "monetary policy authorities have information about inflationary pressures better than what can be obtained from the variables in these [VAR] models." For this reason, it has become common to include an index of commodity prices to account for information about future inflation in the Fed's information set.

However, more recently, recursive monetary VAR models have sought to directly include forecasts of inflation in the model. For example, Leduc, Sill, and Stark (2007) include near-term inflation expectations in a small quarterly VAR model of inflation, inflation expectations, the unemployment rate, and short-term interest rates. They use the model to study the differential behavior of the inflation and monetary policy to a variety of shocks, including monetary policy, during the Great Inflation and after. Davig and Clark (2011) use a similar model that adds long-term inflation expectations and allows for time-variation in the parameters. They focus on the source of the decline in the volatility of inflation expectations in recent decades. More recently, Boivin, Kiley, and Mishkin (2010) study how the monetary transmission mechanism has evolved in recent decades. They conclude that an increase focus of policy makers on price stability has led to better anchored inflation expectations. Empirically, they find a particularly important role for inflation expectations in eliminating the price-puzzle in the more recent sample period even when commodity prices are included.

The widespread interest in the role of inflation expectations in identifying and accounting for the effects of monetary policy shocks, we augment two monetary VAR models with measures of short-term and long-term inflation expectations. In both models, we find that imposing the weak-consistency prior leads to more muted effects of inflation to a monetary policy shock and an amplification of the effects on real GDP. These results are consistent with a larger degree of nominal rigidities and therefore more monetary non-neutrality in a New-Keynesian DSGE model. For example, in the typical Calvo setting, if the average duration of prices increases then the price effects to a monetary policy shock decline and the effects on the real economy increase.

First we consider the federal funds benchmark model of Christiano, Eichenbaum, and Evans (1999). Their model includes real GDP, the GDP deflater, an index of commodity prices, the federal funds rate, non-borrowed reserves, total reserves, and the M2 measure of money. All variables are in log-levels except for the federal funds rate. We augment this benchmark model with actual CPI inflation, short-term inflation expectations, and long-term inflation expectations. We insert these variables in place of the GDP deflater.

Given our sample limitations, we estimate the VAR over 1967-Q1 to 1995:Q2, which is the end date used in Christiano, Eichenbaum, and Evans (1999) and Christiano, Eichenbaum, and Evans (2005). We use a non-informative Normal Inverse-Wishart prior so that the posterior estimates are centered at the OLS counterparts. We perform 10,000 posterior simulations, re-weighting the draws according to the restrictions that:

$$g(A,B)_L(k) = \frac{d}{d\varepsilon_t^{ff}}\pi_{t+k}^L - \frac{1}{32}\sum_{j=1}^{4}\left[\frac{d}{d\varepsilon_t^{ff}}\pi_{t+k+8+j}\right] \tag{26}$$

$$g(A,B)_S(k) = \frac{d}{d\varepsilon_t^{ff}}\pi_{t+k}^S - \frac{1}{4}\sum_{j=1}^{4}\left[\frac{d}{d\varepsilon_t^{ff}}\pi_{t+k+j}\right] \tag{27}$$

for $k = 1$ to 16.

The resulting impulse response functions of the unweighted draws and re-sampled posterior draws are shown in Figures 3 and 4 respectively. The red line shows the VAR-implied response of long-term and short-term inflation expectations respectively. Therefore, the distance between the red and blue lines is effectively a measure of the violation of strong consistency. As expected, imposing the weak consistency prior tilts the posterior in the direction of the parameter space where forecast consistency is closer to being satisfied. Compared to the unrestricted case, the resulting estimates of inflation and inflation expectations are more muted. Moreover, the estimated effects on the real economy are amplified when we impose the weak-consistency prior. All told, this suggests more evidence of monetary non-neutrality according to a measure like that used in Nakamura and Steinsson (2018).

The model of Christiano, Eichenbaum, and Evans (1999), though a commonly used benchmark, exhibits some of the puzzles common in the monetary VAR literature. For example, measures of non-borrowed and total reserves initially rise in response to a contraction monetary policy shock. Moreover, we end the estimation in 1995 as they do and the model can't be extended to the present period due to the zero lower bound and the negative values taken on by non-borrowed reserves. However, the model of Keating, Kelly, Smith, and Valcarcel (2018) was shown to delivery the salient features of monetary policy shock without generating common puzzles and can be extended to 2017-Q4. Therefore, in Figures 5 and 6 we apply our weak-consistency prior to a variant of their model that again replaces the GDP deflater with CPI inflation and both short-term and long-term inflation expectations. Again, we find that imposing the weak-consistency prior delivers impulse responses that show a more muted inflation reactions and stronger real effects.

7 Conclusion

Survey-based inflation forecasts for macroeconomic variables have been incorporated in VARs to accommodate forward-looking information that may not be included in the

current and past values of inflation. However, adding survey-based forecasts into VARs without considering cross-equation restrictions on coefficients generates the internal inconsistency of forecasts. In this paper, we propose a novel Bayesian approach to put a prior distribution on cross-equation restrictions. By combining importance sampling with a maximum entropy prior for cross-equation restrictions, we implement our approach without much computational burden. When applied to a TVP-VAR model for macroeconomic variables to calculate tail risks for inflation, imposing forecast consistency restrictions makes model-based forecasts of inflation more sensitive to movements in survey-based inflation expectations, with deflation probability jumping more during the period of the financial crisis and recession in 2008-09 but subsiding more quickly since then. We also show that imposing weak-consistency across VAR and survey-based measures of inflation expectations leads to more muted inflationary effects and greater real effects from monetary policy shocks.

References

Anene, D. and S. D'Amico., 2017 "A Tale of Four Tails: Inflation, the Policy Rate, Longer-Term Rates, and Stock Prices", Working Paper 2017-26, Federal Reserve Bank of Chicago.

Ang, A., G. Bekaert, and M. Wei., 2007 "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?", *Journal of Monetary Economics*, 54, 1163-1212.

Antolín-Díaz, J., and J. Rubio-Ramírez., 2018 "Narrative Sign Restrictions for SVARs", *American Economic Review*, forthcoming.

Boivin, J., M. T. Kiley, and F. S. Mishkin., 2010 "How has the monetary transmission mechanism evolved over time?." Handbook of monetary economics. Vol. 3. Elsevier, 369-422.

Bundick B., and A.L. Smith 2018 "Does Communicating a Numerical Inflation Target Anchor Inflation Expectations? Evidence & Bond Market Implications", Working Paper 18-01, Federal Reserve Bank of Kansas City.

Campbell, J. Y., and R. Shiller., 1988 "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies*, 1, 195-228.

Christiano, L. J., M. Eichenbaum, and C. L. Evans., 1999, "Monetary policy shocks: What have we learned and to what end?." *Handbook of Macroeconomics* vol. 1, 65-148.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans., 2005 "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of political Economy* 113, 1-45.

Clark, T.E., and T. Davig., 2011, "Decomposing the Declining Volatility of Long-term Inflation Expectations", *Journal of Economic Dynamics and Control*, 35, 981-999.

Cogley T., 2005, "Changing Beliefs and the Term Structure of interest Rates: Cross-equation Restrictions with Drifting Parameters", *Review of Economic Dynamics*, 8, 420-451.

Cover, T.M., and J.A. Thomas., 1991, *Elements of Information Theory*, John Wiley & Sons.

Del Negro M., and F. Schorfheide., 2004, "Priors from General Equilibrium Models for VARs", *International Economic Review*, 45, 647-673.

Doh, T., and A. Oksol., 2018, "Has the Anchoring of Inflation Expectations Changed in the United States during the Past Decade?", *Economic Review*, 103, 31-58, Federal Reserve Bank of Kansas City.

Faust, J., and J. Wright., 2013, "Forecasting Inflation", *Handbook of Economic Forecasting*, 2-56, edited by G. Elliott, C. Granger, and A. Timmermann, Elsevier.

Fleckenstein, M., F.A. Longstaff, and H. Lustig., 2017, "Deflation Risk", *Review of Financial Studies*, 30, 2719-2760.

Frey, C., and F. Mokinski., 2016, "Forecasting with Bayesian Vector Autoregressions Estimated Using Professional Forecasts", *Journal of Applied Econometrics*, 31, 1083-1099.

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

Giannone, D., M. Lenza, and G.E. Primiceri., 2017, "Priors for the Long Run", Staff Reports 832, Federal Reserve Bank of New York.

Hattori, M., A. Schrimpf, and V. Sushko., 2016, "The Response of Tail Risk Perceptions to Unconventional Monetary Policy", *American Economic Journal: Macroeconomics*, 8, 111-136.

Keating, J., L. Kelly, A. L. Smith, and V. Valcarcel, 2018. "A model of monetary policy shocks for financial crises and normal conditions", *Journal of Money, Credit, and Banking*, forthcoming.

Kozicki, S., and P.A. Tinsley., 2012, "Effective Use of Survey Information in Estimating the Evolution of Expected Inflation", *Journal of Money, Credit, and Banking*, 44, 145-169.

Leduc, S., and K. Sill., 2013, "Expectations and Economic Fluctuations: An Analysis Using Survey Data", *Review of Economics and Statistics*, 95, 1352-1367.

Leduc, S., K. Sill, and T. Stark., 2007, "Self-fulfilling expectations and the inflation of the 1970s: Evidence from the Livingston Survey." *Journal of Monetary Economics* 54.2 (2007): 433-459.

Mertens, E., 2016, "Measuing the Level and Uncertainty of Trend Inflation", *Review of Economics and Statistics*, 98, 950-967.

Nakamura, E., and J. Steinsson. 2018, "High frequency identification of monetary non-neutrality: The information effect." *Quarterly Journal of Economics.*

Robert, C., 2001, *The Bayesian Choice*, Springer.

Sims, C. A., 1992, "Interpreting the macroeconomic time series facts: The effects of monetary policy." *European Economic Review* 36, 975-1000.

Wu, J.C., and F.D. Xia., 2016, "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound", *Journal of Money, Credit, and Banking*, 48, 253-291.

Appendix

The derivation of the maximum entropy prior

Let π_0 be the Lebesgue measure on $\mathbb{R}$. Under the weak consistency, the prior $\pi(g)$ must satisfy the following two moment restrictions.

$$\int g\pi(g)dg = 0 \ , \ \int g^2\pi(g)dg = \lambda W^{-1}. \tag{28}$$

The entropy of $\pi(g)$ can be defined by the Kullback-Leibler distance between π and π_0. To maximize the entropy of $\pi(g)$ under moment restrictions, we need to solve the following maximization problem.

$$\pi^\star(g) = \text{argmax}_{\pi(g)} \int \pi(g) \ln(\frac{\pi(g)}{\pi_0(g)})dg - \mu_1(\int g\pi(g)dg) - \mu_2(\int g^2\pi(g) - \lambda W^{-1}dg). \tag{29}$$

The first-order condition for this is $\pi^\star(g) = \pi_0(g)e^{1+\mu_1 g+\mu_2 g^2}$. Hence, $\pi^\star \propto e^{\mu_1 g+\mu_2 g^2}$ and $\pi^\star$ is the normal distribution whose mean and variance are equal to $-\frac{\mu_1}{2\mu_2}$ and $-\frac{1}{2\mu_2}$. Since $\pi^\star$ must satisfy the above two moment restrictions, $\mu_1 = 0$ and $\mu_2 = -\frac{W}{2\lambda}$.

$$\pi^\star(g) \propto e^{-\frac{Wg^2}{2\lambda}} \rightarrow g \sim \mathcal{N}(0, \lambda W^{-1}). \tag{30}$$

Table 1: Summary Stastistics for Deflation Probabilities: 2009:Q4 - 2015:Q4

Horizon	Source	Mean	Standard Deviation	Minimum	Median	Maximum
1	Fleckenstein et al. (2017)	18.759	10.155	2.003	15.739	47.757
1	University of Michigan Survey	15.906	6.1249	6.06	14.1414	32.6531
1	TVP-VAR	22.19	9.12	9.16	19.38	49.02
1	TVP-VAR with resampling	15.35	10.94	3.44	13.30	56.28
2	Fleckenstein et al. (2017)	13.799	7.261	1.801	11.621	35.421
2	TVP-VAR	20.82	6.2	12.12	20.1	36.68
2	TVp-VAR with resampling	13.84	7.69	4.5	13.16	43.72

Notes: Deflation probabilities from Fleckenstein et al. (2017) are based on daily observations on inflation swaps and options from October 5, 2009 to October 28, 2015 while those from the University of Michigan survey are quarterly average values of monthly observations from October, 2009 to October, 2015.

Table 2: Summary Stastistics for High Inflation (>4 percent) Probabilities: 2009:Q4 - 2015:Q4

Horizon	Source	Mean	Standard Deviation	Minimum	Median	Maximum
1	Fleckenstein et al. (2017)	2.615	2.46	0.139	2.086	16.268
1	TVP-VAR	3.46	3.49	0.64	2.48	16.68
1	TVP-VAR with resampling	1.06	1.96	0.04	0.38	9.28
2	Fleckenstein et al. (2017)	3.089	2.394	0.295	2.702	15.315
2	TVP-VAR	3.74	2.98	0.7	2.94	14.16
2	TVp-VAR with resampling	1.07	1.66	0.1	0.46	7.86

Notes: Deflation probabilities from Fleckenstein et al. (2017) are based on daily observations on inflation swaps and options from October 5, 2009 to October 28, 2015.

Figure 1: Sample Data 1970:Q2-2017:Q4

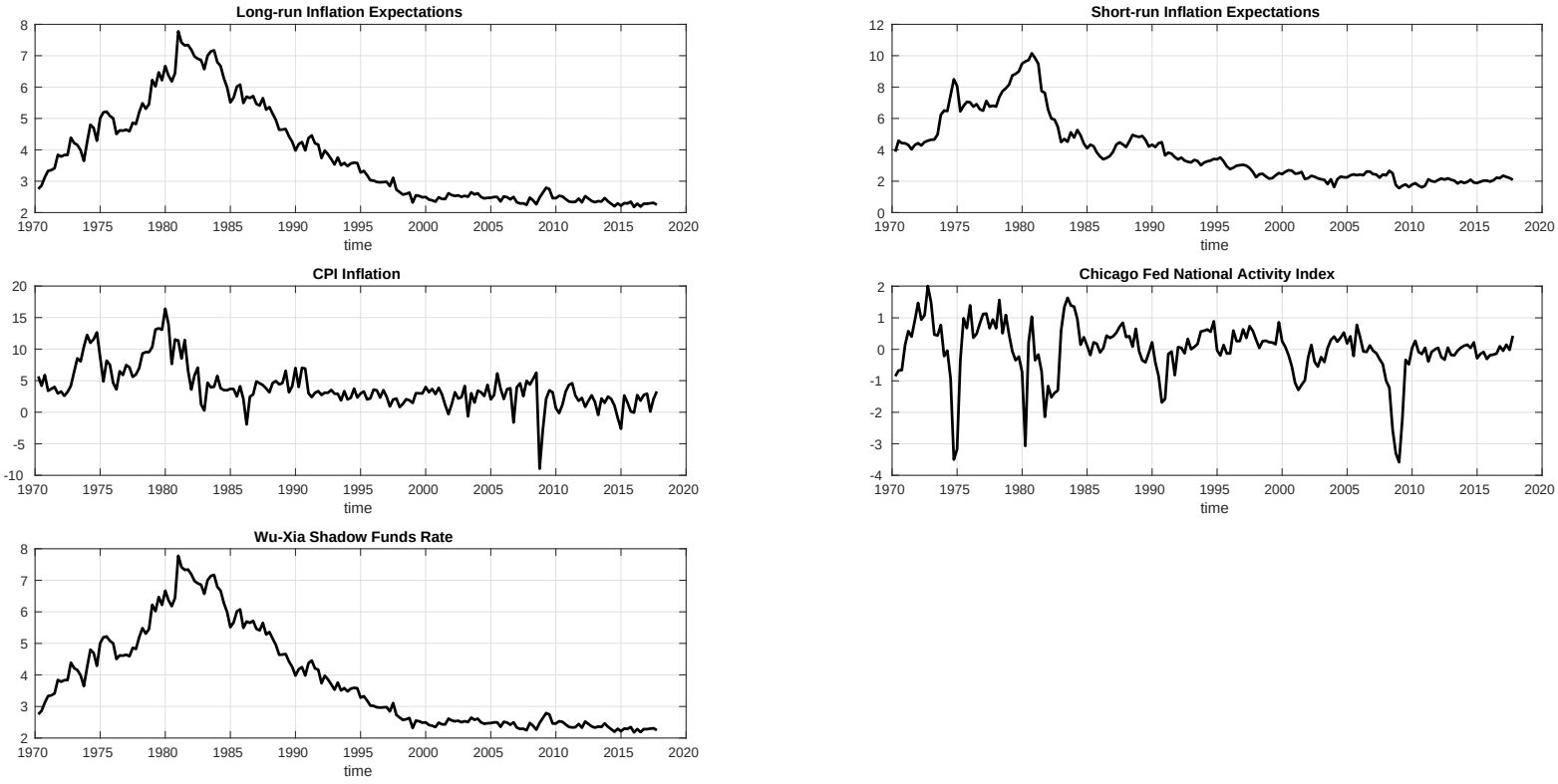


Figure 2: Inflation Tail Risks

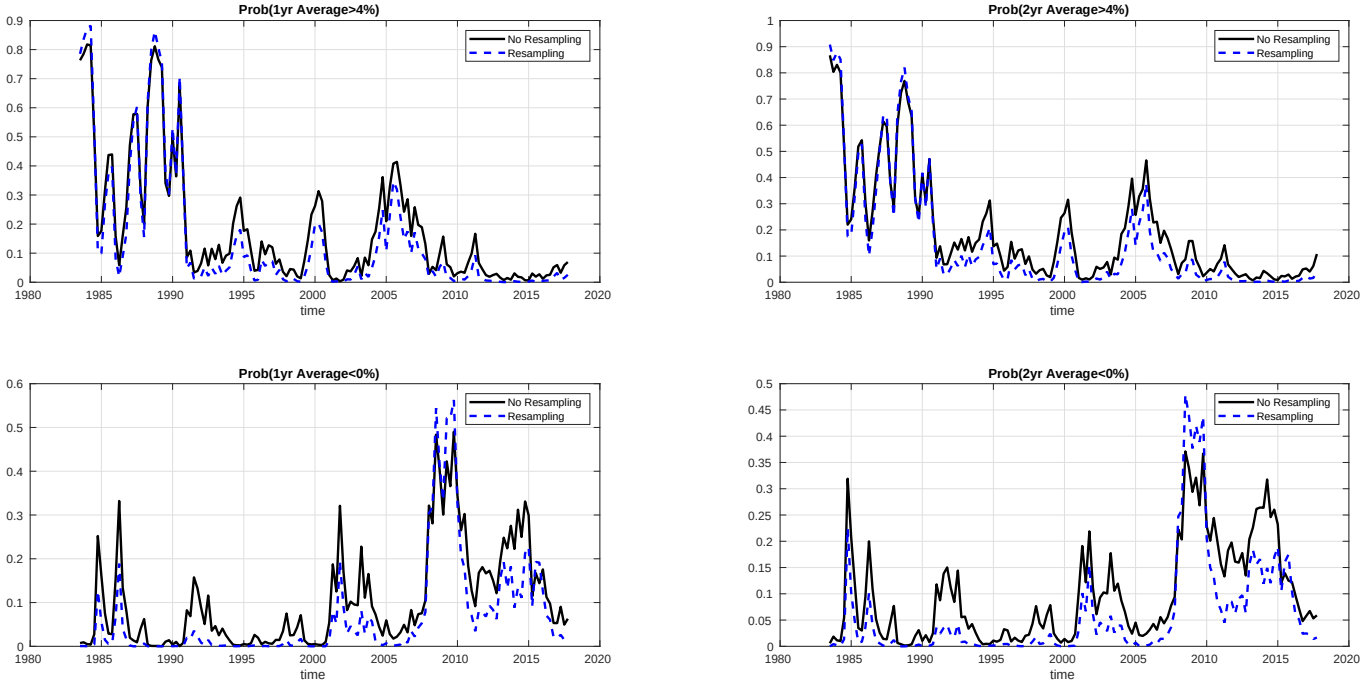
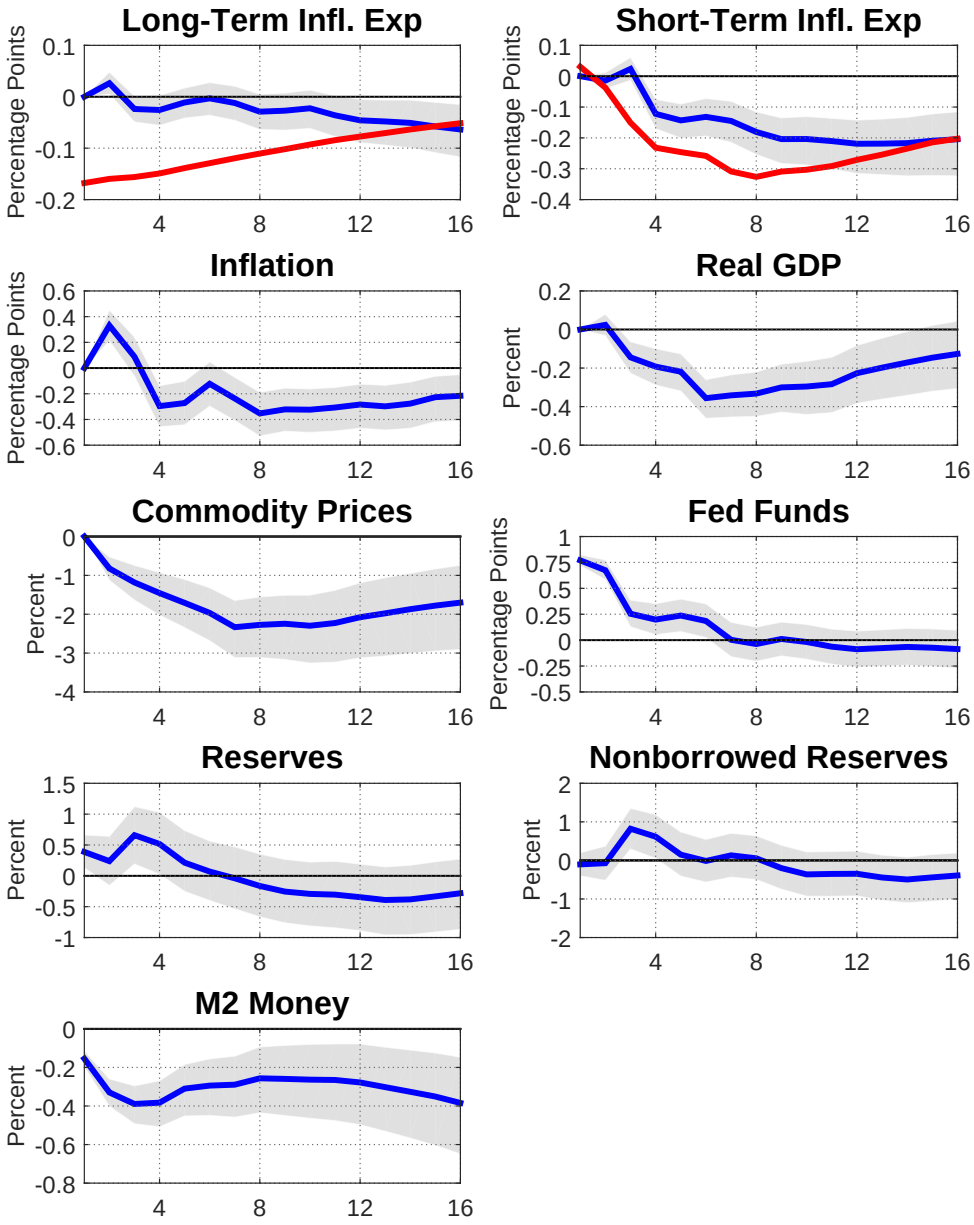
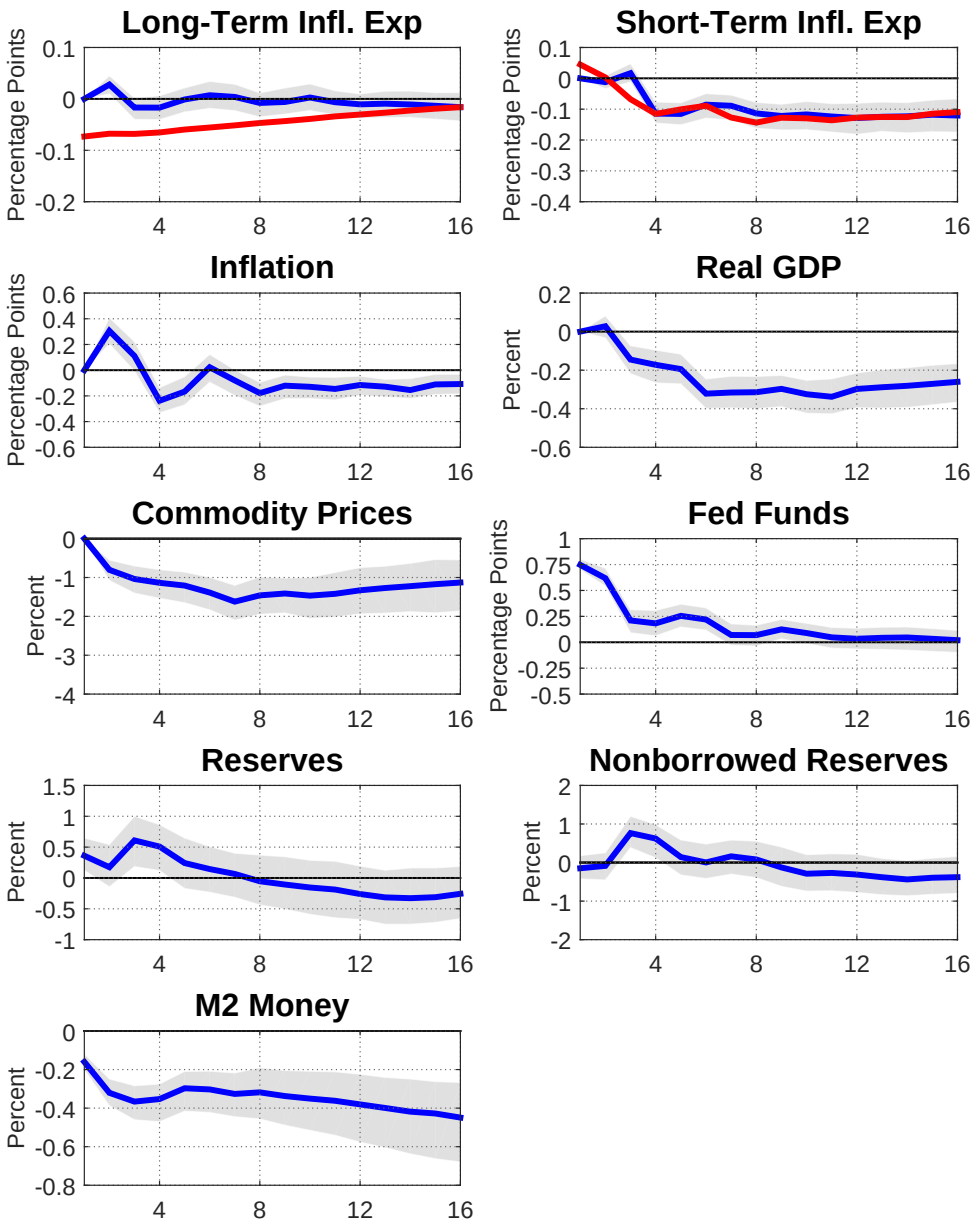


Figure 3: Christiano,Eichenbaum, and Evans (1999) Model Without Resampling



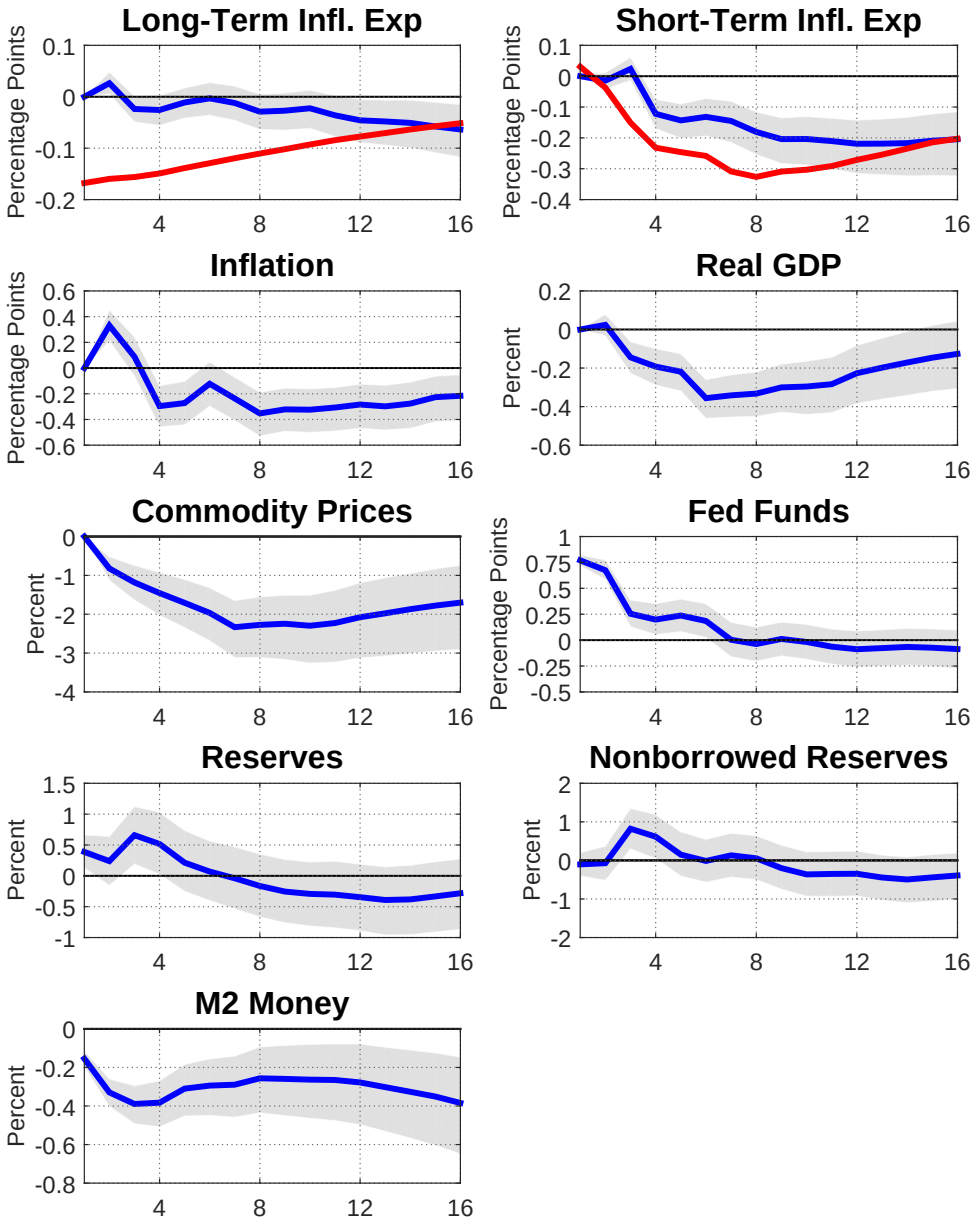
Notes: This figure shows the impulse responses to an identified monetary policy shocks without imposing the weak-consistency prior. The solid blue line is the median response and the shaded region is the 68% posterior interval. The red line shows the VAR-implied response of long-term and short-term inflation expectations respectively.

Figure 4: Christiano,Eichenbaum, and Evans (1999) Model With Resampling



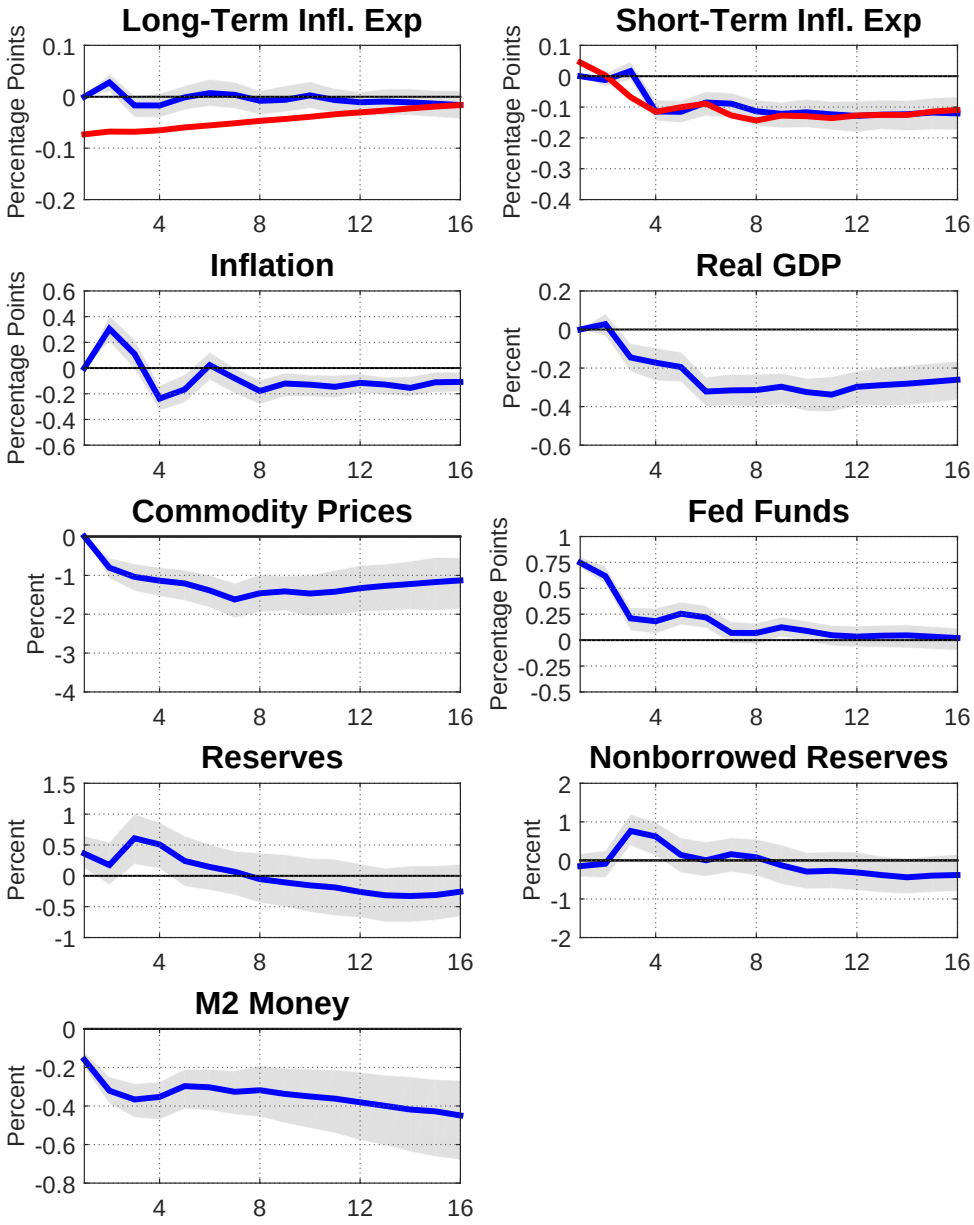
Notes: This figure shows the impulse responses to an identified monetary policy shocks imposing the weak-consistency prior with a value of $\lambda = 10$. The solid blue line is the median response and the shaded region is the 68% posterior interval. The red line shows the VAR-implied response of long-term and short-term inflation expectations respectively.

Figure 5: Keating, Kelly, Smith, Valcarcel (2018) Model Without Resampling



Notes: This figure shows the impulse responses to an identified monetary policy shocks without imposing the weak-consistency prior. The solid blue line is the median response and the shaded region is the 68% posterior interval. The red line shows the VAR-implied response of long-term and short-term inflation expectations respectively.

Figure 6: Keating, Kelly, Smith, Valcarcel (2018) Model Without Resampling



Notes: This figure shows the impulse responses to an identified monetary policy shocks imposing the weak-consistency prior with a value of $\lambda = 10$. The solid blue line is the median response and the shaded region is the 68% posterior interval. The red line shows the VAR-implied response of long-term and short-term inflation expectations respectively.