FORECAST ERROR VARIANCE DECOMPOSITIONS WITH LOCAL PROJECTIONS

Yuriy Gorodnichenko

University of California – Berkeley

and NBER

Byoungchan Lee

University of California – Berkeley

March 15, 2019

Abstract: We propose and study properties of an estimator of the forecast error variance decomposition in the local projections framework. We find for empirically relevant sample sizes that, after being bias-corrected with bootstrap, our estimator performs well in simulations. We also illustrate the workings of our estimator empirically for monetary policy and productivity shocks.

1. Introduction

Macroeconomists have been long interested in estimating dynamic responses of output, inflation, and other aggregates to structural shocks. While many analyses use vector autoregressions (VARs) or dynamic stochastic general equilibrium (DSGE) models to construct estimated responses, an increasing number of researchers focus on a single structural shock and employ single-equation methods to study the dynamic responses. This approach allows concentrating on well-identified shocks and leaving other sources of variation unspecified. In addition, these approaches often impose no restrictions on the shape of the impulse response function. As a result, the local projections (LP) method (Jordà 2005, Stock and Watson 2007) has gained prominence in applied macroeconomic research.

The properties of impulse responses estimated with these methods are well studied (see, e.g., Coibion 2012, Kilian and Kim 2011), but little is known about how one can reliably estimate the quantitative significance of shocks in the single-equation framework. While some methods for constructing the forecast error variance decompositions (FEVDs) have been suggested, it usually has been done without investigation of their econometric properties, especially for empirically relevant sample sizes.[1] As a result, the vast majority of studies using single-equation approaches do not report the FEVD for the variable of interest, and hence one does not know if a given shock accounts for a large share of variation for the variable.[2] This practice contrasts sharply with the nearly universal convention to report FEVDs in VARs and DSGE models. In this paper, we propose and study finite-sample and asymptotic properties of a method to construct forecast error variance decompositions in the local projections framework.

We show that local projections lead to a simple and intuitive way to assess the contribution of identified shocks to the variation of forecast errors at different horizons. While there are several options to implement this insight, we mostly focus on an estimator based on the coefficient of determination, or R^2. To illustrate the properties of this method, we use several data generating processes (DPGs), including the Smets and Wouters (2007) model. These DGPs cover main profiles

---

[1] For example, Jordà (2005) suggests an estimator close in spirit to LP-A and LP-B estimators that we cover in Section 3.1 and Appendix B. Our baseline estimator of FEVDs performs better than these estimators for empirically relevant sample sizes. Another method is to compute FEVDs by using VARs that directly include a structural shock (Plagborg-Møller and Wolf 2018). While this method identifies the same population FEVDs, it requires a large number of lags (Baek and Lee 2019), a feature that may be too costly in practice given the curse of dimensionality in VARs and the noise generated by many estimated parameters.

[2] Coibion et al. (2017) is among the very few papers reporting FEVDs based on the local projection method.

of FEVDs documented in previous works. We show that estimated contributions to the variation of forecast errors may be biased in small samples and one should use bootstrap to correct for possible biases in the FEVDs estimated by local projections. We also show that, in simulations, our estimator performs better than alternative approaches based on sums of squared estimates of impulse responses. We further illustrate the performance of our method with actual data and commonly used identified shocks. In short, our contribution is to develop a new estimator of FEVDs and to assess finite-sample properties of our estimator and alternative estimators.

We assume in this paper that the researcher has a series of identified shocks. However, these shocks may be measured with error in practice because, e.g., they are estimated rather than directly observed. We show that our estimator of FEVDs is *downward* biased when the shocks are imperfectly observed. Thus, our point estimates are conservative and likely provide a *lower* bound. In a concurrent and complementary work, Plagborg-Møller and Wolf (2017) provide set-identified FEVDs given measurement errors in the local projections framework. Their partially identified untestable bounds could be useful tools for the researcher who is interested in *upper* bounds of the FEVDs.

The rest of the paper is structured as follows. Section 2 lays out a basic setting to derive the estimator. Section 3 introduces our estimator and illustrates its econometric properties. Section 4 presents simulation results for bivariate and multivariate settings. Section 5 applies our method to measuring the contribution of monetary policy and productivity shocks to the forecast error variance of output and inflation in the local projections framework. Section 6 concludes.

2. Basics of the forecast error variance decomposition

Consider a generic setup encountered in studies using local projections. Let y_t be an endogenous variable of interest. An identified white-noise shocks series $\{z_t\}$ has mean zero and variance σ_z^2. We assume that variation in y due to z is represented by $\psi_z(L)z_t = \sum_{i=0}^{\infty} \psi_{z,i} z_{t-i}$, where coefficients $\{\psi_{z,i}\}$ provide us with the impulse response function of y to z.

The forecast error for the *h*-period ahead value of the endogenous variable is given by

$$f_{t+h|t-1} \equiv (y_{t+h} - y_{t-1}) - P[y_{t+h} - y_{t-1}|\Omega_{t-1}], \qquad (1)$$

where $P[y_{t+h} - y_{t-1}|\Omega_{t-1}]$ is the projection of $y_{t+h} - y_{t-1}$ on the information set $\Omega_{t-1} \equiv \{\Delta y_{t-1}, z_{t-1}, \Delta y_{t-2}, z_{t-2}, \dots\}$. To keep the exposition as simple as possible, we focus only on a

single shock and a single endogenous variable for now, but in Section 3.6 we consider the case where the information set includes other ("control") variables. We can decompose the forecast errors due to innovations in z and other sources of variation as follows:

$$f_{t+h|t-1} = \psi_{z,0} z_{t+h} + \cdots + \psi_{z,h} z_t + v_{t+h|t-1}, \tag{2}$$

where $v_{t+h|t-1}$ is the error term due to innovations orthogonal to $\{z_t, z_{t+1}, \ldots, z_{t+h}\}$ and Ω_{t-1}.

Following Sims (1980), we can define the population share of the variances explained by the contemporaneous and future innovations in z_t to the total variations in $f_{t+h|t-1}$:

$$s_h = \frac{Var(\psi_{z,0} z_{t+h} + \cdots + \psi_{z,h} z_t)}{Var(f_{t+h|t-1})}. \tag{3}$$

In what follows, we propose and evaluate a method to estimate s_h based on equation (3).

Note that, if we use definitions of Plagborg-Møller and Wolf (2017), the object of our analysis is the forecast variance ratio. Although this definition of s_h seems natural, one should bear in mind several caveats. First, s_h depends on Ω_t: adding more control variables changes the population parameter s_h (see Section 3.6). Second, the forecast error variance decomposition for a structural VAR model or a DSGE model is usually defined given an information set which includes *all* structural shocks, while s_h above is purely based on the observables. These two definitions might not coincide if two information sets differ. For example, if a data generating process is not invertible for structural shocks (the shocks are not recoverable from the history of observable variables), forecast variance ratio is different from variance decomposition (see Plagborg-Møller and Wolf (2017) for details on this point).

3. Estimator

In this section, we introduce our estimator of FEVDs using the coefficient of determination, or R^2 of local projections. We discuss asymptotic properties of our estimator and address issues that may be encountered in practice. Those issues include measurement errors in z_t, small-sample refinements with a focus on biases, and other control variables in the information set.

3.1. R^2 method

Let $Z_t^h = (z_{t+h}, \ldots, z_t)'$. It can be shown with some algebra that equation (3) can be written as

$$s_h = \frac{Cov(f_{t+h|t-1}, Z_t^h)[Var(Z_t^h)]^{-1}Cov(Z_t^h, f_{t+h|t-1})}{Var(f_{t+h|t-1})}. \tag{4}$$

In the numerator, the first Cov term is a row vector, the Var in the middle is a matrix, and the last Cov is a column vector. This quantity can be understood as an R^2 of the population projection of $f_{t+h|t-1}$ on Z_t^h, or the probability limit of sample R^2's. This observation suggests a natural estimator of s_h. First, the forecast errors for each horizon h are estimated using local projections. Second, the estimated forecast errors for the horizon h at time t are regressed on shocks that happen between t and $t + h$. The R^2 in this regression is an estimate of s_h.

More precisely, the estimated forecast error $\hat{f}_{t+h|t-1}$ is the residual of the following regression:

$$y_{t+h} - y_{t-1} = c_h + \sum_{i=1}^{L_y} \gamma_i^h \Delta y_{t-i} + \sum_{i=1}^{L_z} \beta_i^h z_{t-i} + f_{t+h|t-1}, \tag{5}$$

which is an approximation to $y_{t+h} - y_{t-1} = c_h + \sum_{i=1}^{\infty} \gamma_i^h \Delta y_{t-i} + \sum_{i=1}^{\infty} \beta_i^h z_{t-i} + f_{t+h|t-1}$ in population. Then we run the following regression and calculate its R^2:

$$\hat{f}_{t+h|t-1} = \alpha_{z,0} z_{t+h} + \cdots + \alpha_{z,h} z_t + \tilde{v}_{t+h|t-1}. \tag{6}$$

Thus, our estimator $\hat{s}_h^{R2}$ is R^2 of equation (6) which, by construction, is between 0 and 1. Note that $\alpha_{z,i}$ in equation (6) corresponds to the impulse response coefficient $\psi_{z,i}$. Because $\hat{f}_{t+h|t-1}$ in equation (6) is a residual of an OLS regression with an intercept in equation (5) and the mean of z_t is zero, an intercept term in equation (6) is not required. Moreover, the population mean of both $f_{t+h|t-1}$ and Z_t^h are zeros, and so both centered and non-centered R^2's are the same in the population. We report results for the non-centered R^2 below, but properties are similar when we use the centered R^2.

Note that one may implement this estimator by augmenting equation (5) with shocks $z_t, \dots, z_{t+h}$ and calculating the partial R^2. This modification ensures that any predictable variation in $z_t, \dots, z_{t+h}$ is removed. In practice, this step likely makes little difference since z_t is typically constructed in a way such that z_t is not predictable by lags of macroeconomic variables.

LP-A and LP-B estimators of s_h. While we concentrate on the R^2 estimator, there are other options for estimating s_h. For example, note that s_h admits the following representations:

$$s_h = \frac{(\sum_{i=0}^{h} \psi_{z,i}^2)\sigma_z^2}{Var(f_{t+h|t-1})} \tag{7}$$

4

$$= \frac{\left(\sum_{i=0}^{h} \psi_{z,i}^2\right)\sigma_z^2}{\left(\sum_{i=0}^{h} \psi_{z,i}^2\right)\sigma_z^2 + Var\left(v_{t+h|t-1}\right)}. \tag{7'}$$

Thus, one may estimate s_h by plugging estimates of $\psi_{z,i}$'s, σ_z^2, $Var\left(f_{t+h|t-1}\right)$, or $Var\left(v_{t+h|t-1}\right)$ into either (7) or (7').

We estimate $\psi_{z,h}$ with local projections by running the following regression:

$$y_{t+h} - y_{t-1} = c_h^{\mathrm{LP}} + \sum_{i=1}^{L_y} \gamma_i^{h,LP} \Delta y_{t-i} + \sum_{i=0}^{L_z} \beta_i^{h,LP} z_{t-i} + r_{t+h|t-1}, \tag{8}$$

where $\hat{\beta}_0^{h,LP}$ is an estimator of $\psi_{z,h}$ (Jordà 2005). Note that, in contrast to equation (5), equation (8) includes the current value of z_t. Since we can estimate σ_z^2 directly from the time series of z, we can estimate $\left(\sum_{i=0}^{h} \psi_{z,i}^2\right)\sigma_z^2$ in equation (7) or (7'). For the denominator in equation (7), we note that the residual in equation (8) can be related to the forecast error $f_{t+h|t-1}$ in equation (5). By comparing equations (5) and (8), it becomes clear that $f_{t+h|t-1} = \beta_0^{h,LP} z_t + r_{t+h|t-1}$ for each h. Therefore, we can construct estimates of the forecast errors, denoted by $\hat{f}_{t+h|t-1}^{LP}$, by adding $\hat{\beta}_0^{h,LP} z_t$ to $\hat{r}_{t+h|t-1}$. Then we can compute $\widehat{Var}(\hat{f}_{t+h|t-1}^{LP})$ that is an estimate of the denominator in equation (7), where $\widehat{Var}(\cdot)$ denotes a sample variance. We now define a local projection estimator of FEVDs, which we call "LP-A" estimators, as

$$\hat{s}_h^{LPA} = \frac{\left(\sum_{i=0}^{h}\left\{\hat{\beta}_0^{i,LP}\right\}^2\right)\hat{\sigma}_z^2}{\widehat{Var}\left(\hat{\beta}_0^{h,LP} z_t + \hat{r}_{t+h|t-1}\right)}, \tag{9}$$

where $\hat{\sigma}_z^2 \equiv \widehat{Var}(z_t)$.

Although simple, the LP-A estimator does not guarantee that the estimated s_h is between 0 and 1. A simple solution to this issue is to split the denominator into variation due to z and due to v so that $\left(\sum_{i=0}^{h} \psi_{z,i}^2\right)\sigma_z^2$ appears in both the numerator and the denominator as in equation (7'). Note that

$$\hat{v}_{t+h|t-1} = \hat{f}_{t+h|t-1}^{LP} - \hat{\beta}_0^{h,LP} z_t - \hat{\beta}_0^{h-1,LP} z_{t+1} - \cdots - \hat{\beta}_0^{0,LP} z_{t+h}$$

$$= \hat{r}_{t+h|t-1} - \hat{\beta}_0^{h-1,LP} z_{t+1} - \cdots - \hat{\beta}_0^{0,LP} z_{t+h},$$

and that $\widehat{Var}\left(\hat{r}_{t+h|t-1} - \hat{\beta}_0^{h-1,LP} z_{t+1} - \cdots - \hat{\beta}_0^{0,LP} z_{t+h}\right)$ is an estimate of $Var\left(v_{t+h|t-1}\right)$. We use this quantity to define another local projection estimator of FEVDs, or "LP-B":

$$\hat{s}_h^{LPB} = \frac{\left(\sum_{i=0}^{h}\left\{\hat{\beta}_0^{i,LP}\right\}^2\right)\hat{\sigma}_z^2}{\left(\sum_{i=0}^{h}\left\{\hat{\beta}_0^{i,LP}\right\}^2\right)\hat{\sigma}_z^2 + \widehat{Var}\left(\hat{r}_{t+h|t-1} - \hat{\beta}_0^{h-1,LP} z_{t+1} - \cdots - \hat{\beta}_0^{0,LP} z_{t+h}\right)}. \tag{10}$$

The LP-A and LP-B estimators are based on a single regression (8) for each horizon, while the R2 estimator requires two regressions (5) and (6). While the LP-A and the LP-B are in some sense simpler (they estimate only one equation and they correspond more closely to the conventional way to compute FEVD, that is, use squares of estimated impulse responses to compute variance contributions), we find that the R2 estimator has weakly better finite-sample performances. To preserve space, we focus on the R2 estimator in the rest of the paper and relegate the details for the LP-A and LP-B estimators to Appendix B.

3.2. Asymptotics

To derive the asymptotic properties of our R2 estimator, we begin with the case where the forecast errors are observable, not generated. Then we show that using the estimated forecast errors does not alter the asymptotic distribution. Readers more interested in the implementation of the estimator may want to skip to the next subsection.

For now, we suppose that $f_h = \left(f_{T|T-h-1}, f_{T-1|T-h-2}, \dots, f_{L_{max}+h+1\,|L_{max}}\right)'$ is observable for any $h \geq 0$, where $L_{max} = \max\{L_z, L_y\}$. We write $Z_t^h = (z_{t+h}, \dots, z_t)'$ for all t and h and define a matrix $\mathbf{Z}_h = \left(Z_{T-h}^h, Z_{T-1}^h, \dots, Z_{L_{max}+1}^h\right)'$. The (non-centered) R^2 of the regression of $f_{t+h|t-1}$ on Z_t^h is given by $\left(f_h' P_{\mathbf{Z}_h} f_h\right)/(f_h' f_h)$, where $P_{\mathbf{Z}_h} = \mathbf{Z}_h(\mathbf{Z}_h' \mathbf{Z}_h)^{-1}\mathbf{Z}_h'$.

Let $\theta_0 = \left(\theta_{1,0}', \theta_{2,0}', \theta_{3,0}'\right)'$, where $\theta_{1,0} = \left(E[Z_t^h Z_t^{h'}]\right)^{-1}\left(E[Z_t^h f_{t+h|t-1}]\right) = \left(\psi_{z,0}, \psi_{z,1}, \dots, \psi_{z,h}\right)'$, $\theta_{2,0} = E[Z_t^h f_{t+h|t-1}] = \theta_{1,0}\sigma_z^2$, and $\theta_{3,0} = E[f_{t+h|t-1}^2] \equiv \sigma_{f,h}^2$. A method of moments estimator $\hat\theta = \left(\hat\theta_1', \hat\theta_2', \hat\theta_3'\right)'$ is as follows: $\hat\theta_1 = (\mathbf{Z}_h' \mathbf{Z}_h)^{-1}(\mathbf{Z}_h' f_h)$, $\hat\theta_2 = \frac{Z_h' f_h}{T_h}$, $\hat\theta_3 = \frac{f_h' f_h}{T_h}$, where $T_h = T - (L_{max} + h)$. For $\xi(\theta) = \xi(\theta_1, \theta_2, \theta_3) = \frac{\theta_2' \theta_1}{\theta_3}$, we have $s_h = \xi(\theta_0)$ and $\frac{f_h' P_{\mathbf{Z}_h} f_h}{f_h' f_h} = \xi(\hat\theta)$. Therefore, we first derive the asymptotic distribution of $\sqrt{T}(\hat\theta - \theta_0)$ and then apply the delta method to obtain the asymptotic distribution of $\sqrt{T}\left(\xi(\hat\theta) - \xi(\theta_0)\right) = \sqrt{T}\left(\frac{f_h' P_{X_h} f_h}{f_h' f_h} - s_h\right)$.

The moment conditions above can be summarized as $E[g_{t+h}(\theta)] = 0$, where

$$g_{t+h}(\theta) \equiv g(f_{t+h|t-1}, Z_t^h, \theta) = \begin{pmatrix} Z_t^h \left(f_{t+h|t-1} - \left(Z_t^h \right)' \theta_1 \right) \\ Z_t^h f_{t+h|t-1} - \theta_2 \\ f_{t+h|t-1}^2 - \theta_3 \end{pmatrix}. \qquad (11)$$

It is clear that the conditions are satisfied only when $\theta = \theta_0$ and the system is just-identified. Therefore, $\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}\left(0, \ (G_{h,R^2})^{-1} \Omega_{h,R^2} (G'_{h,R^2})^{-1} \right)$, where '$\xrightarrow{d}$' denotes convergence in distribution, $G_{h,R^2} = E[\nabla_\theta g_{t+h}(\theta_0)]$, $\Omega_{h,R^2} = \sum_{l=-\infty}^{\infty} \Gamma(l)$, and $\Gamma(l)$ is the autocovariance of $g_{t+h}(\theta_0)$ at lag l (Hansen 1982). With some algebra, we can further show that $G_{h,R^2} = -diag(\sigma_z^2 I_{h+1}, I_{h+2})$ where $diag(A, B)$ is the block diagonal matrix whose diagonal components are A and B in order, and I_h is the h-dimensional identity matrix .

Now we define Δ_{h,R^2} as $\frac{\partial \xi(\theta_0)}{\partial \theta'} = \frac{1}{\theta_{3,0}} \left(\theta'_{2,0}, \theta'_{1,0}, -s_h \right)$. By combining the above derivations, we derive the asymptotic distribution of $\frac{f'_h P_{Z_h} f_h}{f'_h f_h}$.

Proposition 1. Let $f_h = \left(f_{T|T-h-1}, f_{T-1|T-h-2}, \dots, f_{L_{max}+h+1 \,|L_{max}} \right)'$ and $Z_h = \left(Z_{T-h}^h, Z_{T-1}^h, \dots, Z_{L_{max}+1}^h \right)'$ for all $h \geq 0$, where $L_{max} = \max\{L_z, L_y\}$. The R^2 of the regression of $f_{t+h|t-1}$ on Z_t^h is given by $(f'_h P_{Z_h} f_h)/(f'_h f_h)$, where $P_{Z_h} = Z_h (Z'_h Z_h)^{-1} Z'_h$. Furthermore, the following holds:

$$\sqrt{T} \left(\frac{f'_h P_{Z_h} f_h}{f'_h f_h} - s_h \right) \xrightarrow{d} \mathcal{N}\left(0, \ \Delta_{h,R^2} (G_{h,R^2})^{-1} \Omega_{h,R^2} (G'_{h,R^2})^{-1} \Delta'_{h,R^2} \right), \qquad (12)$$

where $\Delta_{h,R^2} = \frac{1}{\sigma_{f,h}^2} \left(\psi_{z,0}\sigma_z^2, \dots, \psi_{z,h}\sigma_z^2, \psi_{z,0}, \dots, \psi_{z,h}, -s_h \right)$, $G_{h,R^2} = -diag(\sigma_z^2 I_{h+1}, I_{h+2})$, and Ω_{h,R^2} is the long-run variance of $g_{t+h}(\theta_0)$ in equation (11). We denote the variance in equation (12) by $V_{h,R^2} = \Delta_{h,R^2} (G_{h,R^2})^{-1} \Omega_{h,R^2} (G'_{h,R^2})^{-1} \Delta'_{h,R^2}$.

However, f_h is not directly observable in practice. We use its estimate $\hat{f}_h$ instead, which is based on equation (5). Next, we show that the feasible estimator $\frac{\hat{f}'_h P_{Z_h} \hat{f}_h}{\hat{f}'_h \hat{f}_h}$ has the same asymptotic variance V_{h,R^2} in Proposition 1.

To separate issues from truncation and estimation of the forecast errors, we now assume that L_y and L_z are large enough, and the population residual of equation (5) is the true forecast

error. In other words, we assume that $(z_t, \Delta y_t)'$ follows a finite-order Markov process and focus on the variability in $\hat{f}_h$ due to the estimation of the forecast errors.

For a simple notation, we rewrite equation (5) as $y_{t+h} - y_{t-1} = W'_{t-1}\phi + f_{t+h|t-1}$, where $W_{t-1} \equiv \left(1, \Delta y_{t-1}, \dots, \Delta y_{t-L_y}, z_{t-1}, \dots, z_{t-L_z}\right)'$ and $\phi \equiv \left(c_h, \gamma_1^h, \dots, \gamma_{L_y}^h, \beta_1^h, \dots, \beta_{L_z}^h\right)'$. For $\hat{\phi}$ being the OLS estimator of ϕ, we have $\hat{f}_{t+h|t-1} = f_{t+h|t-1} - W'_{t-1}(\hat{\phi} - \phi)$. By stacking up and defining W matrix accordingly, we obtain $\hat{f}_h = f_h - W(\hat{\phi} - \phi)$.

A feasible estimator $\tilde{\theta}$ of θ_0 based on $\hat{f}_h$ is given by $\tilde{\theta}_1 = (Z'_h Z_h)^{-1}(Z'_h \hat{f}_h)$, $\tilde{\theta}_2 = \frac{Z'_h \hat{f}_h}{T_h}$, and $\tilde{\theta}_3 = \frac{\hat{f}'_h \hat{f}_h}{T_h}$. We will show that $\tilde{\theta} = \hat{\theta} + O_p\left(\frac{1}{T_h}\right)$, and therefore the feasible estimator $\tilde{\theta}$ converges to the infeasible estimator $\hat{\theta}$ fast enough not to change the asymptotic distribution of $\sqrt{T}(\tilde{\theta} - \theta)$, and more specifically, the asymptotic variance. Note that $\tilde{\theta}_1 = \left(\frac{Z'_h Z_h}{T}\right)^{-1}\left(\frac{Z'_h \hat{f}_h}{T}\right) = \hat{\theta}_1 - \left(\frac{Z'_h Z_h}{T}\right)^{-1}\left(\frac{Z'_h W}{T}\right)(\hat{\phi} - \phi) = \hat{\theta}_1 - O_p(1)\left(E[Z_t^h W'_{t-1}] + O_p\left(\frac{1}{\sqrt{T}}\right)\right)O_p\left(\frac{1}{\sqrt{T}}\right)$, which follows from the law of large numbers, the central limit theorem, and standard asymptotics of OLS estimators. Because $W_{t-1} \in \Omega_{t-1}$, $Z_t^h = (z_{t+h}, \dots, z_t)'$ is orthogonal to W_{t-1}. In other words, $E[Z_t^h W'_{t-1}] = 0$. Thus, $\tilde{\theta}_1 = \hat{\theta}_1 - O_p(1)O_p\left(\frac{1}{\sqrt{T}}\right)O_p\left(\frac{1}{\sqrt{T}}\right) = \hat{\theta}_1 + O_p\left(\frac{1}{T}\right)$. One can similarly show that $\tilde{\theta}_2 = \hat{\theta}_2 + O_p\left(\frac{1}{T}\right)$ and $\tilde{\theta}_3 = \hat{\theta}_3 + O_p\left(\frac{1}{T}\right)$ using $E[Z_t^h W'_{t-1}] = 0$ and $E\left[f_{t+h|t-1} W'_{t-1}\right] = 0$. We summarize these results in the following proposition.

Proposition 2. Suppose that $(z_t, \Delta y_t)'$ follows a finite-order Markov process, and therefore the true residual in equation (5) coincides with the population forecast error for large enough L_y and L_z. In this case, the feasible R^2 estimator has the same asymptotic distribution as the infeasible estimator in Proposition 1. That is,

$$\sqrt{T}\left(\frac{\hat{f}'_h P_{Z_h}\hat{f}_h}{\hat{f}'_h \hat{f}_h} - s_h\right) = \sqrt{T}\left(\frac{f'_h P_{Z_h} f_h}{f'_h f_h} - s_h\right) + o_p(1) \xrightarrow{d} \mathcal{N}\left(0, \ V_{h,R^2}\right), \qquad (13)$$

where $V_{h,R^2} = \Delta_{h,R^2}\left(G_{h,R^2}\right)^{-1}\Omega_{h,R^2}\left(G'_{h,R^2}\right)^{-1}\Delta'_{h,R^2}$ is the asymptotic variance in Proposition 1.

3.3. Measurement errors

Empirically identified shocks z_t could be measured with errors since e.g., these shocks are often estimates rather than direct observations. One may handle this issue by considering noisy measures of underlying structural shocks as external instruments as is the case in Plagborg-Møller and Wolf (2017) who derive partial-identification results and set-identified s_h.

Our approach is different. Given measurement errors, we show in Appendix D that asymptotic biases of our estimators are negative. Therefore, our methods *underestimate* the true s_h without further refinements to tackle measurement errors. Furthermore, although shocks are often estimated and thus are generated regressors, the researcher is often interested in testing the null of no responses (i.e., $s_h = 0$), and there is no need to adjust inference for this exercise (Pagan 1984).

Specifically, we can decompose the true shock into two parts as $z_t = z_t^o + z_t^u$, where superscripts o and u denote observable and unobservable components, respectively. We assume that $\{(z_t^o, z_t^u)'\}$ is a white noise process with $\sigma_o^2 = Var(z_t^o)$, $\sigma_u^2 = Var(z_t^u)$, and $\rho_{o,u} = corr(z_t^o, z_t^u)$. For example, a measurement error m_t can be modelled as $z_t^o = z_t + m_t$ and $z_t^u = -m_t$, and so $\rho_{o,u} < 0$. Denote the full information set with $\Omega_{t-1} = \{z_{t-1}^o, z_{t-1}^u, \Delta y_{t-1}, \dots\}$ for now and the econometrician's information set with $\Omega_{t-1}^e \equiv \{z_{t-1}^o, \Delta y_{t-1}, \dots\}$. The econometrician's forecast error $f_{t+h|t-1}^e$ is given by $f_{t+h|t-1}^e = y_{t+h} - y_{t-1} - P[y_{t+h} - y_{t-1}|\Omega_{t-1}^e]$. Note that we project $y_{t+h} - y_{t-1}$ on Ω_{t-1}^e, while the full-information forecast error $f_{t+h|t-1}$ is based on Ω_{t-1}. Finally, the econometrician's regressor is denoted by $Z_t^{h,e} = (z_{t+h}^o, \dots, z_t^o)'$.

Proposition 3. Given the assumptions above, the followings hold for any $|\rho_{o,u}| \leq 1$.

(a) $Var(f_{t+h|t-1}^e) \geq Var(f_{t+h|t-1})$.

(b) $Var(\psi_{z,0} z_{t+h} + \dots + \psi_{z,h} z_t)$

$$= Cov(f_{t+h|t-1}^e, Z_t^{h,e})[Var(Z_t^{h,e})]^{-1} Cov(Z_t^{h,e}, f_{t+h|t-1}^e) + \sum_{i=0}^h \psi_{z,i}^2 (1 - \rho_{o,u}^2)\sigma_u^2.$$

(c) $s_h = \dfrac{Var(\psi_{z,0} z_{t+h} + \dots + \psi_{z,h} z_t)}{Var(f_{t+h|t-1})} \geq \dfrac{Cov(f_{t+h|t-1}^e, Z_t^{h,e})[Var(Z_t^{h,e})]^{-1} Cov(Z_t^{h,e}, f_{t+h|t-1}^e)}{Var(f_{t+h|t-1}^e)}.$

For a formal proof, please see Appendix D. Proposition 3(a) covers the forecast error variance, which is the denominator of s_h in equation (4). The result implies that the econometrician's forecast error variance is greater than that based on the full information set. Furthermore, one can show that the equality holds only for (uninteresting) special cases such as $\psi_z(L) = 0$, $\rho_{o,u} = \pm 1$,

and $\sigma_u^2 = 0$. We discuss the numerator of s_h in Proposition 3(b). When estimated without taking z_t^u into consideration, the econometrician's numerator in equation (4) is less than that under the full information set by $\sum_{i=0}^{h} \psi_{z,i}^2 \left(1 - \rho_{o,u}^2\right)\sigma_u^2$. Similarly, the difference reduces when $\psi_{z,i}$'s are close to 0, when the observable component and the unobservable component are highly correlated, and when the variance of the unobservable component σ_u^2 is small. Because the econometrician's denominator is greater and the numerator is less than those based on the full information set, the econometrician's FEVDs are *downward* biased to zero as illustrated in Proposition 3(c). In other words, our point estimate is conservative in favor of the hypothesis $s_h = 0$.

3.4.   Small-sample refinements

While $\hat{s}_h^{R2}$ is asymptotically unbiased as illustrated in Proposition 1 and 2, there may exist substantial finite-sample biases. Note that the OLS estimator in equation (6) is obtained by maximizing the sum of explained variation, or R^2, which may lead to an upward bias in $\hat{s}_h^{R2}$ (Cramer 1987).

To correct for potential small-sample biases in the estimates of s_h and to enhance coverage rates for confidence bands, we employ a VAR-based bootstrap, where the VAR includes two variables $(z_t, \Delta y_t)'$. We use a VAR-based bootstrap to address challenges associated with bootstrapping highly persistent data but researches may utilize alternative approaches.[3,4]

We now discuss the details of the bootstrap procedure. First, we need to choose the order of the VAR model L_{VAR}. In simulations below, we rely on the Hannan-Quinn information criterion (HQIC) for the purpose. We simulate the estimated $VAR(L_{VAR})$ model $Y_t = \hat{\mu} + \hat{\Phi}_1 Y_{t-1} + \cdots + \hat{\Phi}_{L_{VAR}} Y_{t-L_{VAR}} + \epsilon_t$ to generate artificial time series B times, where $Y_t \equiv (z_t, \Delta y_t)'$. And we use this model to compute s_h^*, the true contribution of z to the forecast error variance of y at the horizon h for this data generating process. For each $b \leq B$, we randomly choose t between $1 + L_{VAR}$ and T to initiate the simulation. Then $(z_t, \Delta y_t)', \dots, \left(z_{t-L_{VAR}}, \Delta y_{t-L_{VAR}}\right)'$ are used as $Y_0^{(b)}, \dots, Y_{-L_{VAR}}^{(b)}$. Given the initial condition, we randomly draw $\left\{\epsilon_t^{(b)}\right\}$ from the estimated reduced form residuals

---

[3] One may use alternative implementations of bootstrap to refine asymptotic inference. We tried the block bootstrap for local projections following Kilian and Kim (2011). However, this block bootstrap method performs worse than the VAR-based bootstrap in simulations. Results are in Appendix E1.

[4] Our bootstrap procedure implicitly assumes homoscedasticity of shocks. If a researcher suspects important heteroskedasticity in shocks, one should use alternative bootstrap methods (e.g. Gonçalves and Kilian 2004). An extensive discussion of practical considerations for various bootstrap methods is in Kilian and Lütkepohl (2017, Ch. 12).

$\{\hat{\epsilon}_t\}$ with replacement. Using the estimated model with the above initial conditions and the shuffled residuals, we obtain the simulated series $\left\{\left(z_t^{(b)}, \Delta y_t^{(b)}\right)'\right\}$, where the first T_{BurnIn} number of observations are discarded as burn-in. We apply our estimator to $\left\{\left(z_t^{(b)}, \Delta y_t^{(b)}\right)'\right\}$ and obtain the bootstrap estimate $\hat{s}_h^{R2,(b)}$ for each b. Then we estimate the bias in $\hat{s}_h^{R2}$ with $bias_h \equiv \frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - s_h^*$ and compute bias-corrected estimates $\hat{s}_h^{R2,BC} \equiv \hat{s}_h^{R2} - bias_h$. The procedure is similar for VARs.

3.5. Standard errors and confidence intervals

We have several options to construct standard errors and confidence intervals. For example, one may directly estimate V_{h,R^2} in equations (12) and (13) and derive a symmetric confidence interval based on the estimated V_{h,R^2} (see Appendix A for details including implementation of pre-whitening following Andrews and Monahan 1992). While this works asymptotically, its finite-sample performance is not better than bootstrap confidence intervals as discussed in Appendix E3. Furthermore, the estimated standard errors are often spiky across h's, which induce non-smooth and erratic confidence bands.

Therefore, we employ a different approach for the simulations and the application in this paper. To study finite-sample properties of our estimator, we rely on the distribution of the bootstrap estimates $\hat{s}_h^{R2,(b)}$. The standard error can be easily obtained from a standard deviation of $\hat{s}_h^{R2,(b)}$ across B replications. Constructing a symmetric confidence interval is also straightforward. On the other hand, one may want to take the shape of the bootstrap distribution into consideration when constructing confidence intervals. Let $\hat{q}_{h,\alpha/2}^{R2}$ and $\hat{q}_{h,1-\alpha/2}^{R2}$ refer to the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of the distribution of $\hat{s}_h^{R2,(b)} - \frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)}$. Then the $100(1 - \alpha)\%$ confidence interval is given by $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC}, \ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right]$. Note that we consider the distribution of $\hat{s}_h^{R2,(b)} - \frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)}$ to make the confidence interval centered around the estimated FEVD with bias-correction.

3.6. Extension

While our analysis has focused on the bivariate case, this framework can be readily generalized to include more controls in equation (5):

$$y_{t+h} - y_{t-1} = \sum_{i=1}^{L_z} \beta_i^h z_{t-i} + \sum_{i=1}^{L_C} C_{t-i}' \Gamma_i^h + f_{t+h|t-1}, \tag{14}$$

where C_t is the vector of control variables which may include lags of additional variables and structural shocks other than z_t. In the base case, C_t consists only of Δy_t. Note that for VAR-based bootstraps, one has to include z_t and *all* variables in C_t to simulate data.[5]

One should bear in mind that, although including or excluding C_t or changing the composition of variables in C_t should make little difference of impulse responses estimated with local projections (provided z_t is uncorrelated with other shocks), what goes in C_t is potentially important for FEVDs. Intuitively, by including more controls in C_t (that is, information set Ω_t expands), we (weakly) reduce the size of the forecast error, and hence the amount of variation to be explained shrinks. In other words, the regressand in equation (6) and therefore s_h change with the list of variables in C_t. Thus, one should not be surprised to observe that the share of variation explained by shocks $\{z_t, \dots, z_{t+h}\}$ may be sensitive to C_t.

Similar to the simple case considered in Section 3.1, for equation (6), one may want to use residuals from projecting z_t on lags of z_t and C_t rather than the "raw" shock z_t. For example, when the Cholesky orderings are an identifying assumption, such a procedure is essential to guarantee that forecastable movements in $z_t, \dots, z_{t+h}$ are not used to account for variation in $\hat{f}_{t+h|t-1}$. In practice, however, shocks z_t are constructed in ways to ensure that z_t is not predictable by current values and lags of macroeconomic variables. As a result, we find in our simulations and applications that purifying structural shocks in this manner makes little difference.

---

[5] As the number of variables in C_t increases, the number of parameters in the VAR increases rapidly. When C_t is a large vector, or when a VAR is not a good representation of the DGP for control variables, VAR-based bootstrap might not be an appealing option. In this case, one may consider other forms of bootstrap (e.g., block bootstrap). Alternatively, one may correct for biases by simulating asymptotic distributions of primitive quantities in equations (3), (7), and (7') such as $\hat{\psi}_{z,i}, \hat{\sigma}_z^2$, and $\widehat{Var}(\hat{v}_{t+h|t-1})$. By considering s_h as a non-linear function of those parameters, such simulations would detect biases due to the non-linearity. See Appendices A and B for implementation and F and G for the results.

3.7.　　Taking stock: A cookbook for FEVDs

To summarize our discussion so far, we suggest that the researcher should take the following steps to estimate FEVDs:

> **Step 1**:　Estimate the forecast errors for the horizon h from local projections (5) or (14) depending on the information set.

> **Step 2**:　Regress the estimated forecast errors on the shocks from t to $t + h$ as in equation (6). The R^2 of this regression measures the share of the forecast error variance explained by the shock at the horizon h.

> **Step 3**: To improve the small-sample performance of the estimator, a bias-correction step is recommended for empirically relevant sample sizes. One may rely on a VAR-based bootstrap to do so, where the lag order can be selected via an information criterion.

> **Step 4**: For inference, we can calculate the standard error from either the analytical expression for the asymptotic variance in equation (12) or the distribution of the bootstrap estimates in Step 3. Similarly, we may construct the confidence interval by using either the standard error or the quantiles of the bootstrap distribution. We recommend using bootstrap to construct confidence bands, but one may choose a different approach depending on the data generating process and the sample size.

4. Simulations

This section presents two sets of simulations. The first set shows results for the baseline bivariate case and studies the performances of R2 methods and VARs for various profiles of the contribution of z to the forecast error variance of y at different horizons. The second set uses the estimated Smets and Wouters (2007) model to investigate the performance in a setting with many control variables.

For each data generating process (DGP), we simulate data 2,000 times. When we employ bootstrap to correct for biases, the number of bootstrap replications is set to B=2,000 and $T_{BurnIn} = 100$. As a benchmark, we also report results based on a corresponding VAR. This benchmark corresponds to the practice of including shocks into VARs directly (e.g., Basu et al. 2006, Ramey 2011, Barakchian and Crowe 2013, Romer and Romer 2004, 2010). For the

simulations below, we order z_t as the first variable in VARs as is the case in Section 3.4. We choose the Hannan-Quinn information criterion (HQIC) as our benchmark criterion to determine the number of lags in VAR. To make VAR and LP models comparable, we use HQIC number of lags in the VAR for L_z and L_y (Plagborg-Møller and Wolf 2018). Results are similar when we use higher-order VARs, where the lag order is selected by Akaike information criterion instead of HQIC (Appendix E2).

The sample size for simulated data is $T = 160$, which is common in applied macroeconomic analyses. Results for other sample sizes are reported in Appendices F and G. The coverage rates are calculated as $\Pr\left(\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC} \leq s_h \leq \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right)$ where $\alpha = 0.1$, and therefore the nominal coverage rate is 90%.[6]

4.1.　　Bivariate Data Generating Processes

We study three data generating processes (DGPs) to cover different shapes of s_h. The basic structure is as follows:

$$y_t = \psi_z(L)z_t + u_t,$$
$$u_t = p_t + a_t,$$
$$\left(\Delta p_t - g_y\right) = \rho_p\left(\Delta p_{t-1} - g_y\right) + e_t^p, \qquad e_t^p \sim iid\ N\left(0, \sigma_p^2\right),$$
$$a_t = \rho_a a_{t-1} + e_t^a, \qquad e_t^a \sim iid\ N(0, \sigma_a^2),$$
$$z_t \sim iid\ N(0, \sigma_z^2),$$

where z, e^p and e^a are mutually independent. p and a are permanent and transitory components of u. To find the value of s_h based on $\Omega_{t-1} = \{\Delta y_{t-1}, z_{t-1}, \Delta y_{t-2}, z_{t-2}, \dots\}$, we need to find the population $MA(\infty)$ representation of $\Delta u_t = g_y + \sum_{i=0}^{\infty} \psi_{e,i} e_{t-i}$, where $\{e_t\}$ is a zero mean white-noise series with variance σ_e^2, $\sum_{i=0}^{\infty} \psi_{e,i}^2 < \infty$, and $e_t \in \Omega_t$. We assume that $\psi_{e,0} = 1$ without loss of generality, and the Wold Decomposition implies that such representation exists uniquely. Because z_t and e_t are uncorrelated at all leads and lags, we can write s_h in equation (3) in terms of $\{\psi_{z,i}\}, \{\psi_{e,i}\}, \sigma_z^2$, and σ_e^2. Appendix C discusses how one can use a Kalman filter to derive $\{\psi_{e,i}\}$ and σ_e^2 from $\rho_p, \rho_a, \sigma_p^2$, and σ_a^2.

---

[6] We also considered percentile-t bootstrap and found similar results.

DGP1 is characterized by hump-shaped ψ_z and s_h. We assume that $\{\psi_z(L)z_t\}$ follows an $MA(100)$ process with the maximum response of 3 after 8 periods.[7] The resulting profile of s_h is consistent with e.g., predictions about how monetary shocks contribute to variation in output: there is little to no response of output in the short-run due to various rigidities, then the response is strong in the medium-run, and the long-run response is zero due to nominal neutrality (e.g., Christiano et al. 2005). DGP2 has a strong response of y to z only in the short-run, and thus the shape of s_h is downward-sloping. This profile is consistent with e.g., how temporary fiscal shocks influence output: the effect of a government spending increase or a tax cut is large on impact but then the effect gradually wears out (e.g., Smets and Wouters 2007). Finally, DGP3 assumes $\lim\limits_{h\to\infty} \psi_{z,h} > 0$, so that z has persistent effects on y and the shape of s_h is upward-sloping. This profile is consistent with e.g., models emphasizing that technology shocks are a key (or even exclusive) source of variation in output at long horizons (e.g., Blanchard and Quah 1989). Table 1 reports parameter values for each DGP. Figure 1 plots true impulse responses of y to z (Panel A) and the contribution of z to forecast error variances of y at different horizons (Panel B).

For DGP1, we find (Table 2) that local projections capture the hump-shaped impulse response correctly but $\hat{s}_h^{R2}$ without bias-correction fails to match the hump-share dynamics of s_h: $\hat{s}_h^{R2}$ tends to monotonically increase with the horizon. When we use a VAR to estimate impulse responses and FEVDs, the VAR misses the hump both in the impulse response and FEVDs as HQIC selects too few lags (on average the number of lags is 1.24). Confidence bands yield poor coverage rates. This performance reflects the fact that, by construction, z contributes little to the forecast error variation in y for this DGP at short horizons with $h \leq 4$. Since s_h is between zero and one, we effectively have estimates close to the boundary, and therefore standard methods are likely to fail. While bootstrap appears to provide some improvement (e.g., the bias at long horizons as $h \geq 12$ when z accounts for a larger share of the forecast error variance in y is corrected),[8] it does not perform consistently better in terms of the coverage rates because the parameter is at the boundary. When we allow z to explain 5 percent or more of the forecast error variance in y at short horizons, bootstrap brings coverage rates close to nominal (results are available upon request).

---

[7] This value and pattern are motivated by a 3 percent response of real GDP to a 100bp monetary policy shock estimated in Coibion (2012).

[8] The bias can be further reduced by using higher values of L_z and L_y by reducing errors in $\hat{f}_{t+h|t-1}$ due to the truncation.

Note that, although the VAR estimators ($\hat{s}_h^{VAR}$) are strongly biased, they tend to have smaller variances so that the root mean squared error (RMSE) is similar in magnitude to that of the $\hat{s}^{R2}$ estimator. The large RMSEs underscore difficulties in estimating R^2 (Cramer 1987) and hence s_h.

Because DGP2 permits an exact, finite-order VAR representation, [9] $\hat{s}_h^{VAR}$ has good properties in terms of bias, RMSE, and coverage rates (Table 3). The local projections recover the impulse responses properly, but the estimates of FEVDs again overstate the contribution of z to the unforecasted variation in y at long horizons as $h \geq 12$. Note that bootstrap can correct for this bias. Given that the VAR nests the DGP and that the VAR is more parsimonious than local projections, the VAR has a better performance than the $\hat{s}^{R2}$ estimators.

In the case of DGP3, z has long-lasting effects on y and the VAR underestimates the responses at long horizons as $h \geq 16$ in small samples. Impulse responses estimated with local projections perform better but also exhibit a downward bias at long horizons. In a similar spirit, $\hat{s}_h^{VAR}$ shows a strong downward bias and $\hat{s}_h^{R2}$ is downward biased by a smaller, but still considerable amount (this is the case even after we use bootstrap to correct for possible biases). This performance reflects the fact that HQIC chooses a low number of lags (1.29 lags on average across simulations). As a result, VARs used to simulate bootstrap samples fail to capture the degree of persistence in the data. To demonstrate the importance of the lag order, we report results (Table 4) when we use VAR(5) and VAR(10) for bootstrap. As the number of lags increases, we observe some improvement (e.g., the remaining bias in the bias-corrected $\hat{s}_h^{R2}$ is smaller for VAR(10) than VAR(5)), but these enhancements are achieved at the price of higher variances in the estimates (e.g., the RMSEs of the bias-corrected $\hat{s}_h^{R2}$ are similar for both VARs used for bootstrap). These results suggest that one may want to overfit VAR for persistent processes at the bootstrap stage.

In summary, we find for small samples that estimating s_h precisely is not easy. Nonetheless, we also note that the $\hat{s}_h^{R2}$ estimator performs reasonably well across the DGPs and that bootstrap helps to improve the estimator's properties. In contrast, VARs that include structural shocks z_t tend to perform poorly when a DGP is not nested in a small-order VAR.

---

[9] Given the parameter values in Table 1, $\Delta y_t = g_y + (1 - L)(1 - 0.9L)^{-1}z_t + (1 - 0.9L)^{-1}e_t^p$. By pre-multiplying $(1 - 0.9L)$, we have $\Delta y_t = 0.1g_y + 0.9\Delta y_{t-1} - z_{t-1} + z_t + e_t^p$.

4.2. Smets-Wouters model

While the bivariate DGPs provide important insights on how the R^2 estimator performs, researchers face potentially more complex DGPs and often have more information in practice. In this section, we use the Smets and Wouters (2007) model to study the performance of our estimator in an environment with multiple shocks and many control variables.

As discussed above, different information sets determine different population s_h. In the simulations, we assume that the researcher is interested in explaining variation in output and that the researcher observes output growth rate, inflation, federal funds rate, and monetary policy shocks.[10] This choice of variables is motivated by the popularity of small VARs which include output, inflation, and a policy rate to study the effects of monetary policies on the economy. In this exercise, the shock is ordered first because the Smets-Wouters model allows contemporaneous responses of macroeconomic variables to policy shocks. When estimating impulse responses using local projections, we augment equation (14) with $\psi_{z,h} z_t$ on the right-hand side.

We find (Figure 2) that local projections correctly recover the responses of output to monetary policy shocks, while a low order VAR (lag length is chosen with HQIC) fails to capture the transitory effect of monetary shocks on output. Consistent with our bivariate analysis, $\hat{s}_h^{R2}$ increase with the horizon while the true s_h exhibits hump-shaped dynamics. s_h estimated with a VAR also fails to capture the true dynamics as $\hat{s}_h^{R2}$ flattens out after about $h = 5$. Similar to our results in the previous section, we find that bias correction helps $\hat{s}_h^{R2}$ to recover the true hump-shaped profile of s_h. Coverage rates are close to nominal at all horizons after bias-correction. Again, although the VAR estimator of s_h is strongly biased, the variance of the estimator is low so that RMSEs are broadly similar across methods. We conclude that our proposed methods to estimate FEVDs work reasonably well in more complex settings.

---

[10] For this information set, we construct the true FEVD using a stationary Kalman filter similar to the method in Appendix C. We also tried various combinations of shocks and endogenous variables in the information set and found similar results. Figures for inflation and results with large samples are in Appendix G. Note that monetary policy shocks are *nearly* invertible in the Smets-Wouters model (see Wolf 2017 for more details). While this may be a problem if we use shocks identified and recovered from a DSGE model, the spirit of our exercise is to assume that we have access to other information (as in e.g. Romer and Romer 2004) so that we can observe monetary policy shocks directly.

5. Application

To illustrate the properties of our estimators, we use two structural shocks identified in the literature. The first shock is the monetary policy (MP) innovation identified as in Romer and Romer (2004) and extended in Coibion et al. (2017). The second shock is the total factor productivity (TFP) change identified as in Fernald (2014).[11] The sample autocorrelations and the sample partial autocorrelations at non-zero lags are close to zero for both shocks, that is, the shocks are white noises. The correlation between the shocks is -0.059. Our objective is to quantify the contribution of these shocks to the variation of output and inflation. The sample covers 1969Q1-2008Q4 which excludes the period of binding zero lower bound. The set of variables for local projections includes inflation (annualized growth rate of GDP deflator, i.e., $400\Delta\ln(P_t)$), annual GDP growth rate ($400\Delta\ln(Y_t)$), federal funds rate, and both identified shocks. We set $L_C = L_z = 4$ in equation (14) and add control variables similarly when estimating impulse responses. In the benchmark VAR, we have all five variables and allow four lags.[12]

Consistent with previous studies, we find (Figures 3 and 4) that a contractionary monetary policy shock lowers output and prices, and that a positive TFP shock raises output and lowers prices. Impulse responses estimated with a VAR and local projections are similar at horizons $h \leq L_{VAR} = 4$. However, the estimated impulse responses differ at longer horizons, and therefore the peak effects and the overall shapes are different. The VAR estimates of the FEVDs suggest that TFP (MP) shocks account for approximately 10 (3.5) percent of the forecast error variances of output at horizons longer than 2 years. For inflation, MP shocks contribute up to 19 percent of the variation in the forecast error of inflation at the 5-year horizon and little variation at shorter horizons while the contribution of TFP shocks is generally small. Bias-correction makes no material difference for the forecast error variance decomposition estimates for all cases but one: the bias-corrected estimate of the contribution of MP shocks to the variation in the forecast error of inflation at the 5-year horizon increases to 32 percent.

The local projections estimates of the contribution of the two shocks to the forecast error variances of output are much larger than the VAR estimates. Moreover, bias-correction tends to generate lower contributions, consistent with simulations. For example, monetary policy shocks account for 18 percent of the forecast error variance of output according to the R2 estimate (28

---

[11] Appendix H presents results for military spending shocks constructed in Ramey and Zubairy (2018).

[12] The ordering of variables in the VAR is TFP measure (from Fernald 2014), output growth rate, inflation, monetary policy innovations (from Coibion et al. 2017), and fed funds rate. For the VAR-based analysis, we follow the practice and compute FEVDs using shocks in these variables where shocks are identified recursively from reduced-form residuals.

percent without bias-correction) and only 3.5 percent according to the VAR estimate at the 5-year horizon. Similarly, the VAR estimate of the contribution of MP shocks to inflation at the 5-year horizon is less than 20 percent, which is a surprising result given Milton Friedman's "inflation is always and everywhere a monetary phenomenon." In contrast, the R2 estimate of the same FEVD with bias-correction amounts to 44 percent. Also, while the profile of $\hat{s}_h^{VAR}$ for output is generally flat after $h = 5$, $\hat{s}_h^{R2}$ has richer dynamics. This is consistent with what we find in our simulations for DGP1: when the true s_h is close to zero for small h's, $\hat{s}_h^{VAR}$ fails to match the shape, while $\hat{s}_h^{R2}$ is much more successful. The profiles of $\hat{s}_h^{R2}$ and $\hat{s}_h^{VAR}$ for output also differ remarkably for TFP shocks. While $\hat{s}_h^{R2}$ increases in h, $\hat{s}_h^{VAR}$ flattens around 10 percent after $h = 10$. At the 5-year horizon, TFP shocks contribute to 28 percent of the forecast error variance of output based on the R2 estimate after bias-correction, where the VAR estimate without bias-correction is only 11 percent.

6. Concluding remarks

Single-equation methods can offer flexibility and parsimony that many economists seek. The increasing popularity of these methods, specifically the local projections, calls for further development of these tools. An important limitation for practitioners using this framework has been a lack of simple tools with well-known econometric properties especially in small samples to assess quantitative significance of a given set of shocks, that is, the contribution of the shocks to the forecast error variance of the variable of interest. We propose a method to provide such a metric. In a series of simulation exercises, we document that our method has good small-sample properties. We also show that conventional approaches to assess the quantitative significance of two popular structural shocks (monetary policy shocks and total factor productivity shocks) could have understated the importance of these two shocks.

References

Andrews, Donald W.K., and J. Christopher Monahan, 1992. "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator," *Econometrica*, 60(4): 953-966.

Baek, ChaeWon, and Byoungchan Lee, 2019. "ABCs of Understanding Single Equation Regressions," manuscript.

Barakchian, S. Mahdi, and Christopher Crowe, 2013. "Monetary policy matters: Evidence from new shocks data," *Journal of Monetary Economics*, 60(8): 950-966.

Basu, Susanto, John G. Fernald, and Miles S. Kimball, 2006. "Are Technology Improvements Contractionary?" *American Economic Review,* 96(5): 1418–1448.

Blanchard, Olivier, and Danny Quah, 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79(4): 655-673.

Cramer, Jan Solomon, 1987. "Mean and variance of R2 in small and moderate samples," *Journal of Econometrics,* 35(2-3): 253-266.

Coibion, Olivier, 2012. "Are the Effects of Monetary Policy Shocks Big or Small?" *American Economic Journal: Macroeconomics,* 4(2): 1–32.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia, 2017. "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 88(C): 70–89.

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): 1-45.

Fernald, John, 2014. "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," Working Paper 2012-19.

Gonçalves, Sílvia, and Lutz Kilian, 2004. "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form," *Journal of Econometrics*, 123(1): 89-120.

Hansen, Lars P., 1982. "Large sample properties of generalized method of moments estimators," *Econometrica,* 50(4): 1029-1054.

Jordà, Oscar, 2005. "Estimation and Inference of Impulse Responses by Local Projections," A*merican Economic Review*, 95(1): 161–182.

Kilian, Lutz, and Yun Jung Kim, 2011. "How reliable are local projection estimators of impulse responses?" *Review of Economics and Statistics,* 93(4): 1460-1466.

Kilian, Lutz, and Helmut Lütkepohl, 2017. *Structural Vector Autoregressive Analysis*, Cambridge University Press.

Pagan, Adrian, 1984. "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, 25(1): 221-247.

Plagborg-Møller, Mikkel, and Christian K. Wolf, 2017. "Instrumental Variable Identification of Dynamic Variance Decompositions," manuscript.

Plagborg-Møller, Mikkel, and Christian K. Wolf, 2018. "Local Projections and VARs Estimate the Same Impulse Responses," manuscript.

Ramey, Valerie A., 2011. "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126(1): 1–50.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy,* 126(2): 850-901.

Romer, Christina, D., and David H. Romer, 2004. "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94(4): 1055-1084.

Romer, Christina D., and David H. Romer, 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100(3): 763-801.

Sims, Christopher A., 1980. "Macroeconomics and reality." *Econometrica,* 48(1): 1-48.

Smets, Frank, and Rafael Wouters, 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American Economic Review,* 97(3): 586-606.

Stock, James, and Mark Watson, 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Banking and Credit*, 39(1): 3–33.

Wolf, Christian K., 2017. "Masquerading Shocks in Sign-Restricted VARs," manuscript.

Table 1. Parameter values for data generating processes (DGPs) used in simulations.

	$\psi_z(L)$	σ_z	g_y	ρ_p	σ_p	ρ_a	σ_a
DGP1	Hump-shaped	1	0.5	0.9	0.5	0.9	3
DGP2	$(1-0.9L)^{-1}$	3	0.5	0.9	1.5	-	-
DGP3	$(1-L)^{-1}(1-0.9L)^{-1}$	1	0.5	0.5	2	0.9	3

Table 2. Simulation results for DGP 1.

	Horizon h					
	0	4	8	12	16	20
Impulse Response						
True	0.00	1.39	3.00	2.06	0.88	0.29
Local projections	0.00	1.39	3.00	2.05	0.87	0.29
VAR(HQIC)	0.00	0.18	0.24	0.25	0.25	0.25
Forecast Error Variance Decomposition						
True	0.00	0.04	0.19	0.21	0.18	0.14
Average estimate						
R2	0.01	0.06	0.20	0.25	0.26	0.27
VAR(HQIC)	0.01	0.02	0.02	0.02	0.03	0.03
Root mean squared error						
R2	0.01	0.05	0.11	0.15	0.19	0.22
VAR(HQIC)	0.01	0.03	0.17	0.20	0.16	0.13
Coverage (90 % level, asymptotic)						
R2	0.99	0.81	0.69	0.65	0.63	0.61
VAR(HQIC)	0.99	0.75	0.06	0.06	0.07	0.10
Forecast Error Variance Decomposition (bias-corrected, VAR(HQIC))						
True	0.00	0.04	0.19	0.21	0.18	0.14
Average estimate						
R2	0.00	0.02	0.13	0.16	0.13	0.11
VAR(HQIC)	0.00	0.00	0.01	0.01	0.01	0.01
Root mean squared error						
R2	0.01	0.05	0.12	0.16	0.17	0.18
VAR(HQIC)	0.01	0.04	0.19	0.21	0.17	0.14
Coverage (90 % level, asymptotic)						
R2	0.99	0.95	0.64	0.64	0.72	0.81
VAR(HQIC)	1.00	0.53	0.06	0.05	0.07	0.09

Notes: The table reports the performance of estimators introduced in Section 3 for DGP1. The sample size is $T = 160$, and the number of simulations is 2,000. R2 and VAR stand for $\hat{s}_h^{R2}$ and $\hat{s}_h^{VAR}$ estimators of forecast error variance decompositions. The lag order is selected by the Hannan-Quinn information criterion (HQIC). Confidence intervals for the bias-corrected R2 estimator are given by $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC}, \ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right]$ as discussed in Section 3.5, where $\alpha = 0.1$. Confidence intervals for the other estimators are constructed similarly.

Table 3. Simulation results for DGP 2.

	Horizon h					
	0	4	8	12	16	20
Impulse Response						
True	3.00	1.97	1.29	0.85	0.56	0.36
Local projections	2.99	1.83	1.07	0.57	0.22	0.06
VAR(HQIC)	2.96	1.93	1.33	0.95	0.71	0.56
Forecast Error Variance Decomposition						
True	0.80	0.25	0.10	0.05	0.03	0.02
Average estimate						
R2	0.79	0.26	0.15	0.14	0.15	0.19
VAR(HQIC)	0.80	0.27	0.12	0.08	0.06	0.05
Root mean squared error						
R2	0.03	0.11	0.12	0.14	0.17	0.21
VAR(HQIC)	0.03	0.08	0.06	0.06	0.05	0.05
Coverage (90 % level, asymptotic)						
R2	0.90	0.89	0.89	0.82	0.73	0.67
VAR(HQIC)	0.88	0.90	0.92	0.96	0.97	0.98
Forecast Error Variance Decomposition (bias-corrected, VAR(HQIC))						
True	0.80	0.25	0.10	0.05	0.03	0.02
Average estimate						
R2	0.81	0.24	0.09	0.03	0.01	0.00
VAR(HQIC)	0.80	0.25	0.10	0.05	0.03	0.02
Root mean squared error						
R2	0.03	0.10	0.09	0.09	0.10	0.12
VAR(HQIC)	0.03	0.07	0.06	0.05	0.04	0.04
Coverage (90 % level, asymptotic)						
R2	0.92	0.90	0.97	0.97	0.95	0.94
VAR(HQIC)	0.88	0.89	0.91	0.96	0.99	0.99

Notes: The table reports the performance of estimators introduced in Section 3 for DGP2. The sample size is $T = 160$, and the number of simulations is 2,000. R2 and VAR stand for $\hat{s}_h^{R2}$ and $\hat{s}_h^{VAR}$ estimators of forecast error variance decompositions. The lag order is selected by the Hannan-Quinn information criterion (HQIC). Confidence intervals for the bias-corrected R2 estimator are given by $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC}, \ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right]$ as discussed in Section 3.5, where $\alpha = 0.1$. Confidence intervals for the other estimators are constructed similarly.

	Horizon h					
	0	4	8	12	16	20
Impulse Response						
True	1.00	4.10	6.13	7.46	8.33	8.91
Local projections	0.98	3.93	5.75	6.86	7.46	7.70
VAR(5)	0.93	3.71	4.70	4.94	5.04	5.08
VAR(10)	0.91	3.65	5.33	6.05	6.17	6.27
Forecast Error Variance Decomposition (bias-corrected, VAR(5))						
True	0.06	0.29	0.47	0.58	0.65	0.70
Average estimate						
R2	0.06	0.26	0.41	0.49	0.55	0.57
VAR(5)	0.06	0.24	0.32	0.36	0.37	0.38
Root mean squared error						
R2	0.04	0.12	0.16	0.19	0.21	0.23
VAR(5)	0.04	0.12	0.20	0.27	0.32	0.35
Coverage (90 % level, asymptotic)						
R2	0.81	0.82	0.82	0.84	0.83	0.83
VAR(5)	0.87	0.81	0.66	0.50	0.38	0.32
Forecast Error Variance Decomposition (bias-corrected, VAR(10))						
True	0.06	0.29	0.47	0.58	0.65	0.70
Average estimate						
R2	0.07	0.29	0.46	0.56	0.62	0.65
VAR(10)	0.06	0.27	0.41	0.49	0.53	0.55
Root mean squared error						
R2	0.05	0.12	0.16	0.19	0.21	0.23
VAR(10)	0.04	0.11	0.16	0.19	0.22	0.24
Coverage (90 % level, asymptotic)						
R2	0.74	0.78	0.79	0.80	0.82	0.82
VAR(10)	0.87	0.85	0.83	0.81	0.78	0.75

Notes: The table reports the performance of estimators introduced in Section 3 for DGP3. The sample size is $T = 160$, and the number of simulations is 2,000. R2 and VAR stand for $\hat{s}_h^{R2}$ and $\hat{s}_h^{VAR}$ estimators of forecast error variance decompositions. L_z and L_y are selected by the Hannan-Quinn information criterion (HQIC) and L_{VAR} is either 5 or 10. Confidence intervals for the bias-corrected R2 estimator are given by $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC}, \ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right]$ as discussed in Section 3.5, where $\alpha = 0.1$. Confidence intervals for the other estimators are constructed similarly.

Figure 1. Population impulse responses and forecast error variance decompositions for each DGP.
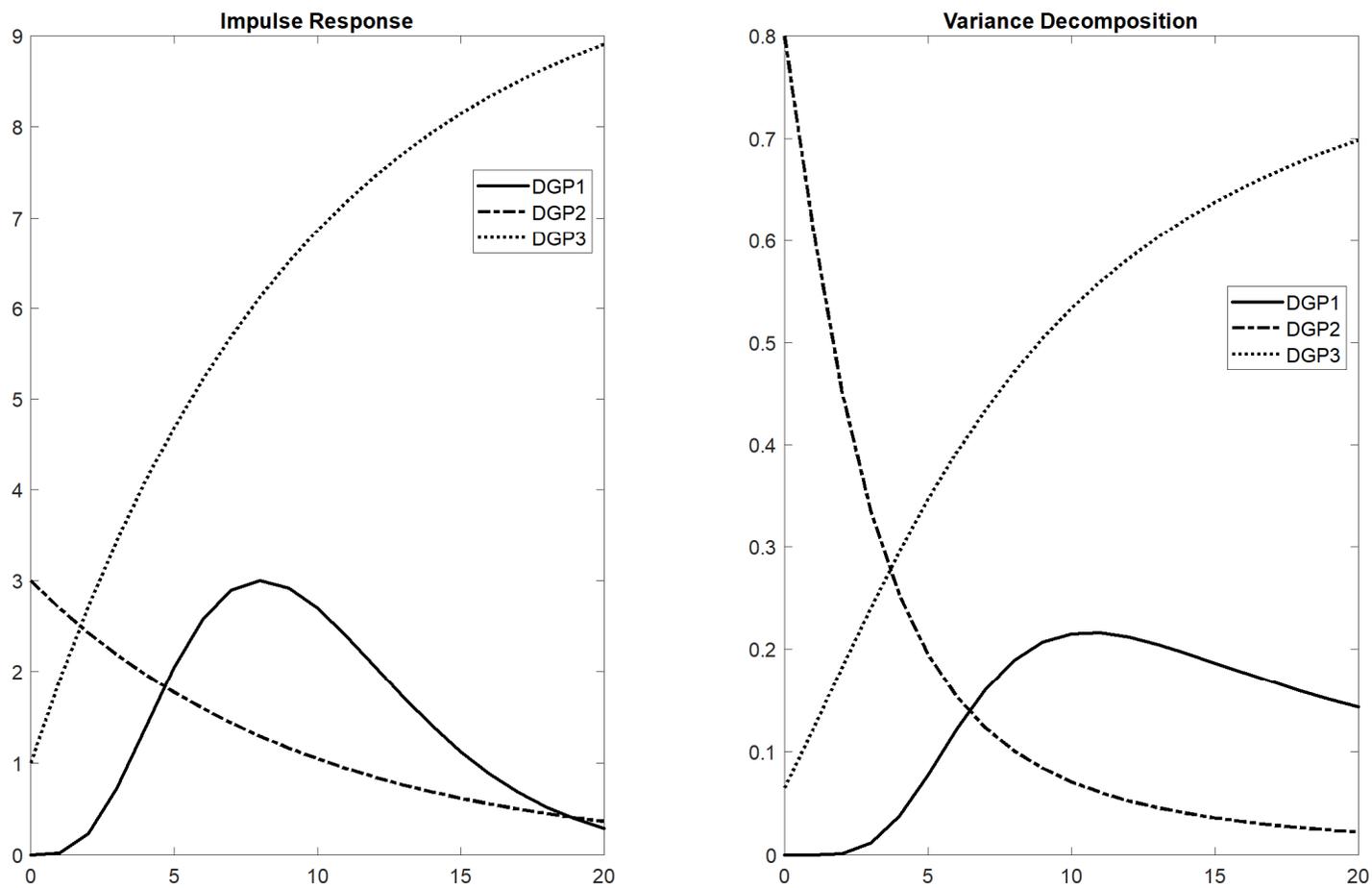
Notes: The left panel shows the impulse response functions for three bivariate data generating processes (DGPs) in Section 4.1. The right panel shows the contribution of the structural shocks to the forecast error variances of an outcome variable for the DGPs.

Figure 2. Smets and Wouters (2007) model, real GDP and monetary policy shock, T = 160.
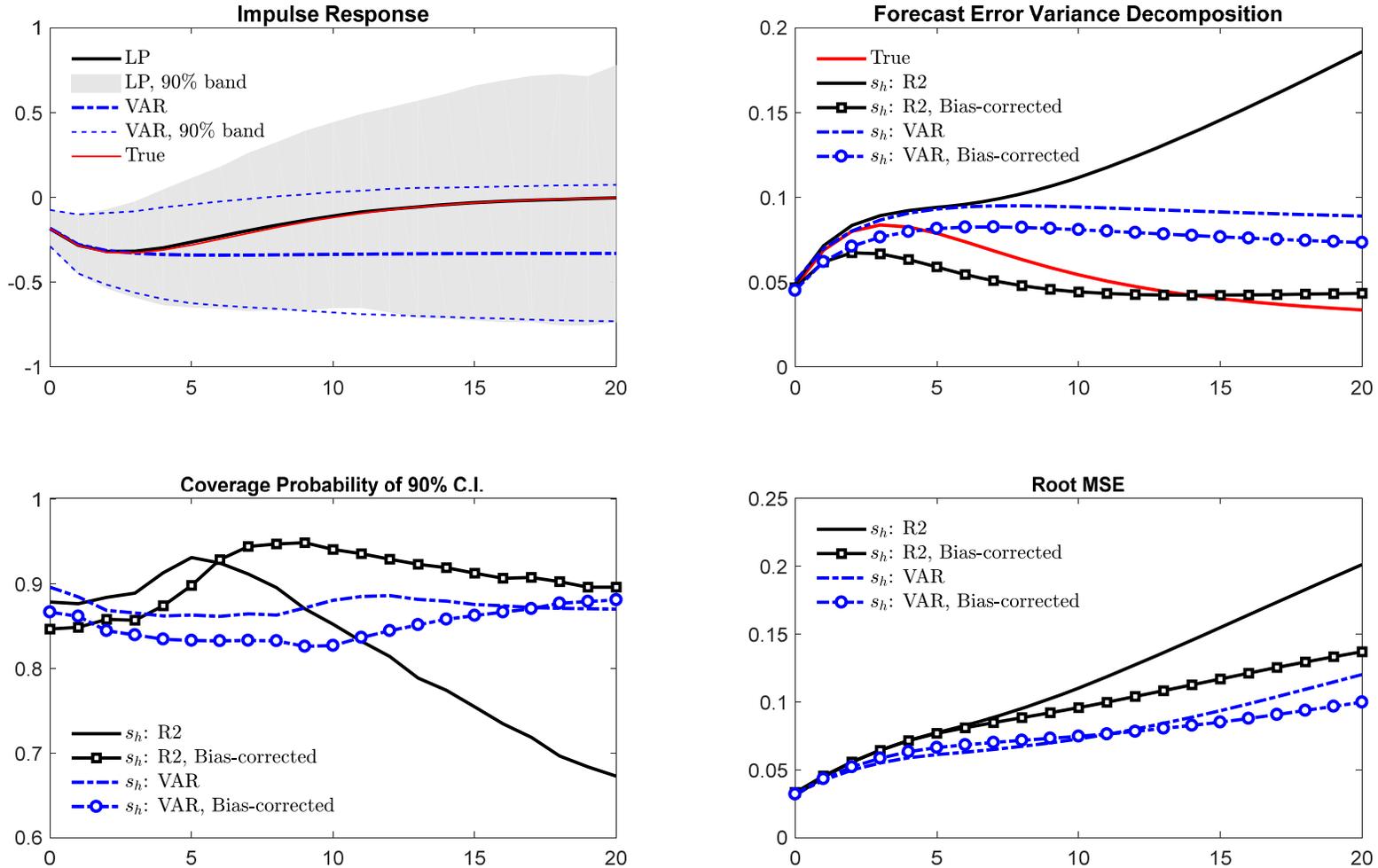
Notes: We simulate the Smets and Wouters (2007) model to evaluate the performance of our estimators as discussed in Section 4.2. The top-left panel covers local projections (LP) and VAR estimators of impulse responses where lag lengths are determined with the Hannan-Quinn information criterion (HQIC). The shaded area and the dashed lines represent the 5th and 95th percentiles of the simulated LP and VAR estimates, respectively. For forecast error variance decompositions (FEVDs), $E[\hat{s}_h^{R2}]$ and $E[\hat{s}_h^{VAR}]$ with or without bias-correction can be found in the top-right panel. Coverage probabilities of 90% confidence intervals around $\hat{s}_h^{R2}$ and $\hat{s}_h^{VAR}$ with or without bias-correction are shown in the bottom-left panel. We construct the confidence intervals using $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ percentiles of bootstrapped estimates (see Section 3.5). The bottom-right panel illustrates root mean squared errors of the estimators of the FEVDs.

Figure 3. Real GDP. Sample: 1969:Q1-2008:Q4.
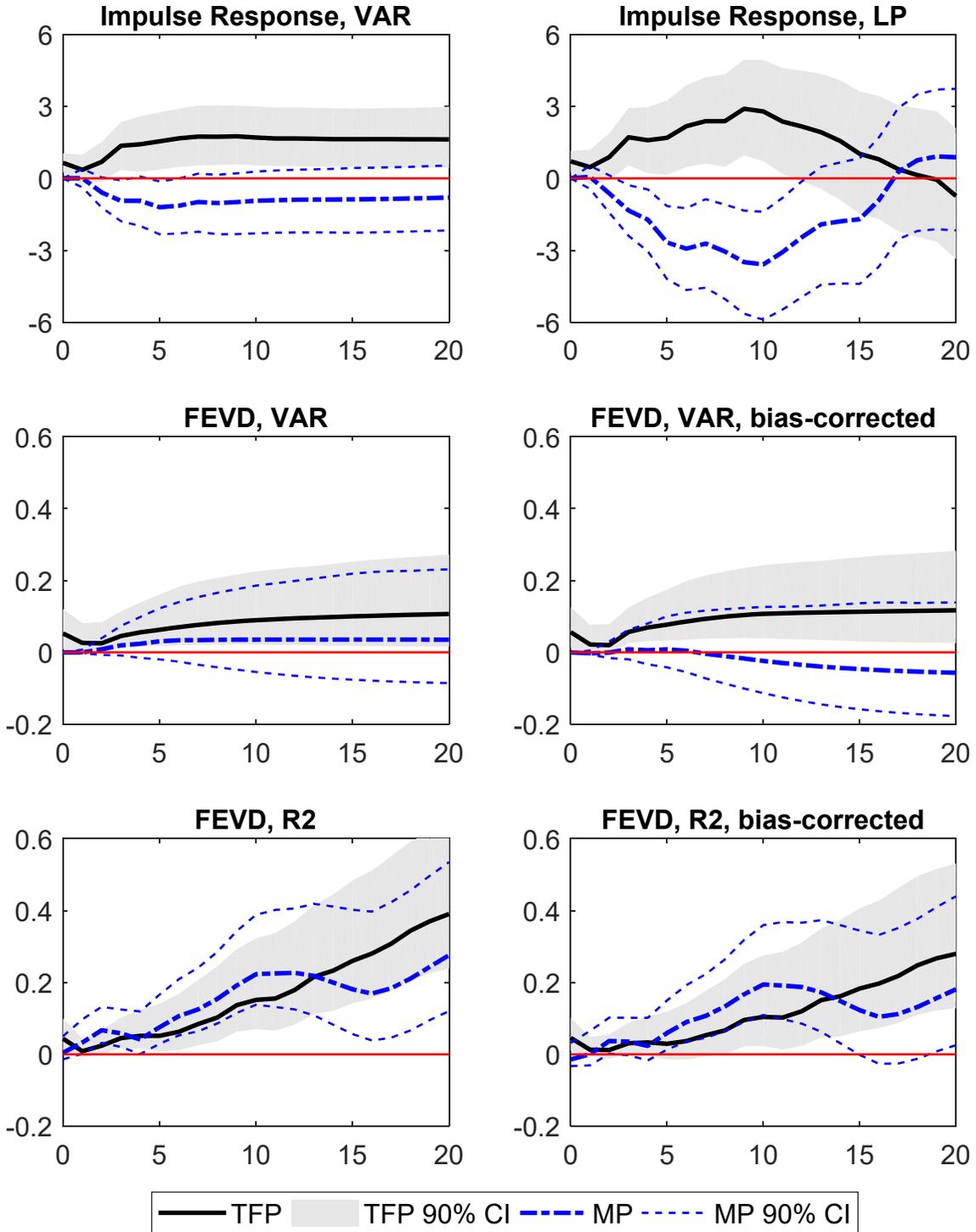
Notes: We estimate impulse responses and forecast error variance decompositions (FEVDs) of real GDP in Section 5. We focus on total factor productivity (TFP) shocks identified as in Fernald (2014) and monetary policy (MP) shocks of Romer and Romer (2004) extended by Coibion et al. (2017). The first row covers the estimated impulse responses and 90% bootstrap confidence intervals in response to a one standard deviation shock to TFP and MP. We depict the results for VARs (top-left panel) and local projections (LP, top-right panel). The unit of the y-axis is annualized percent. The second row shows $\hat{s}_h^{VAR}$ and 90% bootstrap confidence intervals with and without bias-correction. The last row is for $\hat{s}_h^{R2}$ and 90% bootstrap confidence intervals with and without bias-correction.

Figure 4. Inflation. Sample:1969:Q1-2008:Q4.
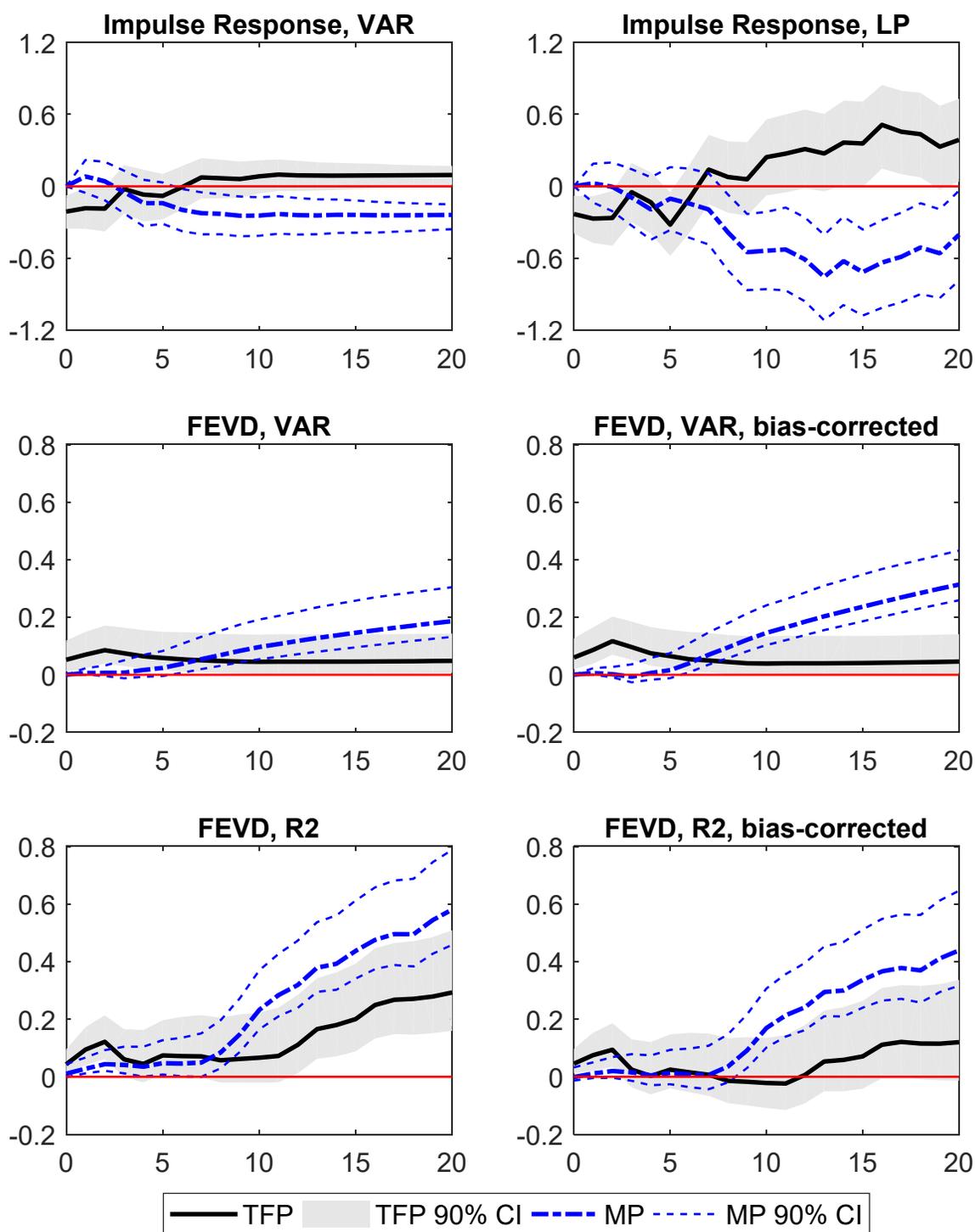
Notes: We estimate impulse responses and forecast error variance decompositions (FEVDs) of inflation in Section 5. We focus on total factor productivity (TFP) shocks identified as in Fernald (2014) and monetary policy (MP) shocks of Romer and Romer (2004) extended by Coibion et al. (2017). The first row covers the estimated impulse responses and 90% bootstrap confidence intervals in response to a one standard deviation shock to TFP and MP. We depict the results for VARs (top-left panel) and local projections (LP, top-right panel). The unit of the y-axis is annualized percentage points. The second row shows $\hat{s}_h^{VAR}$ and 90% bootstrap confidence intervals with and without bias-correction. The last row is for $\hat{s}_h^{R2}$ and 90% bootstrap confidence intervals with and without bias-correction.