# APPENDIX FOR

# FORECAST ERROR VARIANCE DECOMPOSITIONS WITH LOCAL PROJECTIONS

Yuriy Gorodnichenko
University of California – Berkeley
and NBER

Byoungchan Lee
University of California – Berkeley

## Appendix A. R2 estimator

In Section 3.1 and 3.2, we propose and study properties of the R2 estimator of forecast error variance decompositions. Here we explain how one can estimate the asymptotic variance $V_{h,R^2}$ in Proposition 1 and 2. Furthermore, we provide details of an alternative method of bias-correction briefly illustrated in footnote 5. Those methods would be useful for the researcher who wants to do inference without relying on VAR-based bootstraps.

***Estimation of $\boldsymbol{V_{h,R2}}$.*** We want to estimate $V_{h,R^2} = \Delta_{h,R^2}\left(G_{h,R^2}\right)^{-1}\Omega_{h,R^2}\left(G'_{h,R^2}\right)^{-1}\Delta'_{h,R^2}$. Let's begin with $\Delta_{h,R^2}$. A practically feasible estimator of $\theta_0 = \left(\theta'_{1,0}, \theta'_{2,0}, \theta'_{3,0}\right)'$ that we use is $\tilde{\theta} = \left(\tilde{\theta}'_1, \tilde{\theta}'_2, \tilde{\theta}'_3\right)'$, where

$$\tilde{\theta}_1 = (\boldsymbol{Z}'_h\boldsymbol{Z}_h)^{-1}\left(\boldsymbol{Z}'_h\hat{f}_h\right),\ \tilde{\theta}_2 = \frac{\boldsymbol{Z}'_h\hat{f}_h}{T_h},\ \tilde{\theta}_3 = \frac{\hat{f}'_h\hat{f}_h}{T_h},\ T_h = T - (L_{max} + h).$$

A natural estimator of $\Delta_{h,R^2} = \frac{\partial\xi(\theta_0)}{\partial\theta'} = \frac{1}{\theta_{3,0}}\left(\theta'_{2,0}, \theta'_{1,0}, -s_h\right)$ is $\widehat{\Delta}_{h,R^2} \equiv \frac{\partial\xi(\tilde{\theta})}{\partial\theta'} = \frac{1}{\tilde{\theta}_3}\left(\tilde{\theta}'_2, \tilde{\theta}'_1, -\hat{s}_h^{R2}\right)$. The last element should be a bias-corrected estimate instead of $\xi\left(\tilde{\theta}\right)$ because we find that this specification performs better in simulations.[1]

We next turn to $G_{h,R^2} = E[\nabla_\theta g_{t+h}(\theta_0)] = -diag\left(E\left[Z_t^h\left(Z_t^h\right)'\right], I_{h+2}\right)$. It can be easily estimated by $\hat{G}_{h,R^2} = -diag(\boldsymbol{Z}'_h\boldsymbol{Z}_h/T_h, I_{h+2}) = -diag\left(\sum_{t=L_{max}+1}^{T-h} Z_t^h\left(Z_t^h\right)'/T_h, I_{h+2}\right)$.

It remains to estimate $\Omega_{h,R^2} = \sum_{l=-\infty}^{\infty}\Gamma(l)$, where $\Gamma(l)$ is the autocovariance of $g_t(\theta_0)$ in equation (11) at lag $l$. Remember that

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

We pre-whiten the data following Andrews and Monahan (1992) to avoid underestimation problems of the long-run variance of $g_{t+h}(\theta_0)$. For a simple notation, we define a $2h+3$ dimensional vector $\pi_{t+h}$ as follows:

[1] Results are available upon request.

$$\pi_{t+h} \equiv \begin{pmatrix} Z_t^h \left( \hat{f}_{t+h|t-1} - \left( Z_t^h \right)' \tilde{\theta}_1 \right) \\ Z_t^h \hat{f}_{t+h|t-1} - \tilde{\theta}_2 \\ \hat{f}_{t+h|t-1}^2 - \tilde{\theta}_3 \end{pmatrix}.$$

It is worth noting that the sample average of $\pi_t$ is a zero vector, *i.e.* $\frac{1}{T_h} \sum_{t=L_{max}+1}^{T-h} \pi_t = 0$ given the definition of $\tilde{\theta}$. To whiten the series, we use a VAR(1) model that $\pi_t = A\pi_{t-1} + \eta_t$. The estimated autoregressive matrix and the residuals are denoted by $\hat{A}$ and $\hat{\eta}_t$. Then we estimate the long-run variance of $\eta_t$ by applying Newey and West (1987) estimator with the Bartlett kernel to the residuals $\{\hat{\eta}_{L_{max}+2}, \dots, \hat{\eta}_{T-h}\}$. Specifically, the estimated long-run variance is given by

$$\widehat{LRV}(\hat{\eta}_t) = \hat{\Gamma}_{\eta,0} + \frac{1}{L_{NW}+1} \left( \hat{\Gamma}_{\eta,1} + \hat{\Gamma}_{\eta,-1} \right) + \cdots + \frac{L_{NW}}{L_{NW}+1} \left( \hat{\Gamma}_{\eta,L_{NW}} + \hat{\Gamma}_{\eta,-L_{NW}} \right),$$

where $\hat{\Gamma}_{\eta,l}$ is the estimated autocovariance matrix of $\hat{\eta}_t$ at lag $l$. We use a simple rule suggested by Stock and Watson (2011) to select the number of autocovariance matrices included. Following the rule, $L_{NW} + 1$ is the closest natural number to $0.75T_h^{1/3}$. Finally, $\widehat{\Omega}_{h,R^2}$ is obtained by

$$\widehat{\Omega}_{h,R^2} = \left( I_{2h+3} - \hat{A} \right)^{-1} \widehat{LRV}(\hat{\eta}_t) \left( I_{2h+3} - \hat{A}' \right)^{-1}.$$

In sum, the asymptotic standard error of $\hat{s}_h^{R2}$ is given by

$$[s.e.(\hat{s}_h^{R2})]^2 = \frac{1}{T_h} \widehat{\Delta}_{h,R^2} \left( \hat{G}_{h,R^2} \right)^{-1} \widehat{\Omega}_{h,R^2} \left( \hat{G}'_{h,R^2} \right)^{-1} \widehat{\Delta}'_{h,R^2},$$

$$\text{where } \widehat{\Delta}_{h,R^2} = \frac{1}{\tilde{\theta}_3} \left( \tilde{\theta}'_2, \tilde{\theta}'_1, -\hat{s}_h^{R2} \right),$$

$$\hat{G}_{h,R^2} = -diag\left( \frac{\boldsymbol{Z}'_h \boldsymbol{Z}_h}{T_h}, I_{h+2} \right),$$

$$\widehat{\Omega}_{h,R^2} = \left( I_{2h+3} - \hat{A} \right)^{-1} \widehat{LRV}(\hat{\eta}_t) \left( I_{2h+3} - \hat{A}' \right)^{-1}.$$

***An alternative method for estimating biases based on the asymptotic distribution of $\tilde{\boldsymbol{\theta}}$.*** We conjecture that most of the finite sample bias in $\hat{s}_h^{R2}$ is due to the non-linear transformation $\xi(\cdot)$, not estimation of $\theta$ itself. Note that $\theta$ consists of projection coefficients of $f_{t+h|t-1}$ on $Z_t^h$, covariance between $f_{t+h|t-1}$ and $Z_t^h$, and variance of $f_{t+h|t-1}$. Estimation of all those quantities are rather standard, and significant biases of the corresponding method of moments estimators have not been reported. Below we suggest a method to capture biases originating from $\xi(\cdot)$ in small samples without relying on VAR-based bootstrap.

Because $\sqrt{T}(\hat{\theta}-\theta_0) \overset{d}{\to} \mathcal{N}\left(0, \quad G_{h,R^2}^{-1}\Omega_{h,R^2}\left(G'_{h,R^2}\right)^{-1}\right)$, we can estimate the asymptotic variance of the feasible estimator $\tilde{\theta}$ by $\frac{1}{T_h}\left(\hat{G}_{h,R^2}\right)^{-1}\hat{\Omega}_{h,R^2}\left(\hat{G}'_{h,R^2}\right)^{-1}$. We simulate $\theta^{(b)}$ for $B$ times from the following normal distribution:

$$\theta^{(b)} \sim \mathcal{N}\left(\tilde{\theta}, \qquad \frac{1}{T_h}\left(\hat{G}_{h,R^2}\right)^{-1}\hat{\Omega}_{h,R^2}\left(\hat{G}'_{h,R^2}\right)^{-1}\right).$$

We discard cases where $\theta_3^{(b)} \leq 0$ because $\theta_3 = Var\left(f_{t+h|t-1}\right) > 0$. Finally, the bias due to the non-linearity in $\xi(\cdot)$ is captured by $\frac{1}{B}\sum_{b=1}^{B}\xi\left(\theta^{(b)}\right) - \xi\left(\tilde{\theta}\right)$ and the bias-corrected estimator is given by

$$\hat{s}_h^{R2} = \xi\left(\tilde{\theta}\right) - \left(\frac{1}{B}\sum_{b=1}^{B}\xi\left(\theta^{(b)}\right) - \xi\left(\tilde{\theta}\right)\right) = 2\xi\left(\tilde{\theta}\right) - \frac{1}{B}\sum_{b=1}^{B}\xi\left(\theta^{(b)}\right).$$

## Appendix B. Other local projections estimators

As briefly discussed at the end of Section 3.1, it is possible to estimate $s_h$ by plugging estimates of $\psi_{z,i}$'s, $\sigma_z^2$, $Var(f_{t+h|t-1})$, or $Var(v_{t+h|t-1})$ into the following representations of $s_h$:

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

We call those estimators the LP-A and LP-B estimators, respectively. Below we derive their asymptotic distributions, discuss methods for joint inferences, and show how to estimate the asymptotic variances. For their performances in Monte Carlo simulations, see Appendices F and G.

***Asymptotics of the LP-A and LP-B estimators.*** Similar to Proposition 1, we begin with the case where the population forecast errors are observable, not generated. Also, using the estimated forecast errors does not change the asymptotic distributions of the LP-A and LP-B estimators.

**Proposition 4.** The local projections estimators when $f_{t+h|t-1}$ is observable have the following asymptotic distributions for some $V_{h,LPA}$ and $V_{h,LPB}$:

$$\sqrt{T}\left(\frac{\sum_{i=0}^{h}\left(\hat{\beta}_0^{i,LP}\right)^2 \hat{\sigma}_z^2}{\widehat{Var}(f_{t+h|t-1})} - s_h\right) \xrightarrow{d} \mathcal{N}\left(0, \quad V_{h,LPA}\right), \quad \text{and}$$

$$\sqrt{T}\left(\frac{\sum_{i=0}^{h}\left(\hat{\beta}_0^{i,LP}\right)^2 \hat{\sigma}_z^2}{\sum_{i=0}^{h}\left(\hat{\beta}_0^{i,LP}\right)^2 \hat{\sigma}_z^2 + \widehat{Var}\left(f_{t+h|t-1} - \sum_{i=0}^{h}\hat{\beta}_0^{i,LP}\, z_{t+h-i}\right)} - s_h\right) \xrightarrow{d} \mathcal{N}\left(0, \quad V_{h,LPB}\right)$$

*Proof.*

(i) LP-A estimator

We first derive the joint distribution of $\hat{\psi}_{z,i}\left(= \hat{\beta}_0^{i,LP}\right)$'s, $\hat{\sigma}_z^2$, and $\hat{\sigma}_{f,h}^2 \equiv \widehat{Var}\left(f_{t+h|t-1}\right)$. Then we will use the delta method to find $V_{h,LPA}$.

To begin, we describe the moment conditions for the local projections for $\hat{\psi}_{z,i}$'s. We run the following OLS regression and take the coefficient on $z_t$:

$$y_{t+i} - y_{t-1} = \beta_0^{i,LP} z_t + \cdots + \beta_{L_z}^{i,LP} z_{t-L_z} + \gamma_1^{i,LP} \Delta y_{t-1} + \cdots + \gamma_{L_y}^{i,LP} \Delta y_{t-L_y} + c_i^{LP} + r_{t+i|t-1}$$

for all $i = 0,1, \ldots, h$. In the above representation, $\beta_0^i = \psi_{z,i}$. For a simple notation, we rewrite the above equation as

$$p_{i,t} = q_t' \mathrm{B}_i + r_{t+i|t-1}, \qquad \text{where} \quad p_{i,t} = y_{t+i} - y_{t-1},$$

$$q_t = \left(z_t, \ldots, z_{t-L_z}, \Delta y_{t-1}, \ldots, \Delta y_{t-L_y}, 1\right)',$$

$$\mathrm{B}_i = \left(\beta_0^{i,LP}, \ldots, \beta_{L_z}^{i,LP}, \gamma_1^{i,LP}, \ldots, \gamma_{L_y}^{i,LP}, c_i^{LP}\right)'.$$

Then the OLS estimator $\widehat{\mathrm{B}}_i$ becomes the method of moments estimator of the following moment conditions:

$$E\left[q_t\left(p_{i,t} - q_t' \mathrm{B}_i\right)\right] = 0.$$

Also, $\widehat{\psi}_{z,i}$ is given by $\iota_1' \widehat{\mathrm{B}}_i$ where $\iota_1$ is a $L_z + L_y + 2$ dimensional vector whose first element is one and the others are zero.

To study all parameters required simultaneously, we let $\theta_0 = \left(\mathrm{B}_0', \ldots, \mathrm{B}_h', \sigma_z^2, \sigma_{f,h}^2\right)'$, where $\sigma_{f,h}^2 \equiv Var\left(f_{t+h|t-1}\right)$. We use the moment conditions that $E[g_{t+h}(\theta_0)] = 0$, where $g_{t+h}(\theta_0)$ is given as follows:

$$g_{t+h}(\theta_0) = \begin{pmatrix} q_t\left(p_{0,t} - q_t' \mathrm{B}_0\right) \\ \vdots \\ q_t\left(p_{h,t} - q_t' \mathrm{B}_h\right) \\ z_t^2 - \sigma_z^2 \\ f_{t+h|t-1}^2 - \sigma_{f,h}^2 \end{pmatrix}.$$

We define $g_{t+h}(\theta)$ similarly. It is clear that it is a just-identified system. Similar to the proof of Proposition 1, we know that

$$\sqrt{T}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} \mathcal{N}(0, \quad G^{-1}\Omega(G')^{-1}),$$

where $G = E[\nabla_\theta g_{t+h}(\theta_0)]$, $\Omega = \sum_{l=-\infty}^{\infty} \Gamma(l)$, and $\Gamma(l)$ is the autocovariance of $g_{t+h}(\theta_0)$ at lag $l$. With some algebra, we can show that

$$G = -E\begin{pmatrix} I_{h+1} \otimes q_t q_t' & 0 \\ 0 & I_2 \end{pmatrix},$$

where $\otimes$ is the Kronecker product.

A transformation $\xi$ is required to connect $\theta$ with $s_h$. We define

$$\xi(\theta_0) = s_h = \frac{\sum_{i=0}^{h} (\iota_1' \mathrm{B}_i)^2 \ \sigma_z^2}{\sigma_{f,h}^2}$$

and $\xi(\theta)$ is also defined similarly.

Regarding the delta method, we need $\Delta \equiv \frac{\partial \xi(\theta_0)}{\partial \theta'}$. We can show that

$$\Delta = \frac{1}{\sigma_{f,h}^2} \begin{pmatrix} 2\psi_{z,0}\sigma_z^2 \iota_1 \\ \vdots \\ 2\psi_{z,h}\sigma_z^2 \iota_1 \\ \sum_{i=0}^{h} \psi_{z,i}^2 \\ -s_h \end{pmatrix}'.$$

Combining the above derivations and being explicit about the fact that the moment conditions $g_{t+h}(\cdot)$ are for the LP-A approach at the horizon $h$, we have the asymptotic distribution.

$$\sqrt{T}\left(\frac{\sum_{i=0}^{h}(\hat{\psi}_{z,i})^2 \hat{\sigma}_z^2}{\widehat{Var}(f_{t+h|t-1})} - s_h\right) \xrightarrow{d} \mathcal{N}(0, \quad V_{h,LPA}),$$

where $V_{h,LPA} = \Delta_{h,LPA}(G_{h,LPA})^{-1}\Omega_{h,LPA}(G'_{h,LPA})^{-1}\Delta'_{h,LPA}$.

(ii) LP-B estimator

The joint distribution of $\hat{\psi}_{z,i}$'s is obtained similarly. To study all parameters required simultaneously, we let $\theta_0 = (\mathrm{B}'_0, \ldots, \mathrm{B}'_h, \sigma_z^2, \sigma_{v,h}^2)'$ where $\sigma_{v,h}^2 \equiv Var(f_{t+h|t-1} - \sum_{i=0}^{h} \psi_{z,i}\, z_{t+h-i})$. We use the moment conditions that $E[g_{t+h}(\theta_0)] = 0$, where $g_{t+h}(\theta_0)$ is given as follows:

$$g_{t+h}(\theta_0) = \begin{pmatrix} q_t(p_{0,t} - q'_t \mathrm{B}_0) \\ \vdots \\ q_t(p_{h,t} - q'_t \mathrm{B}_h) \\ z_t^2 - \sigma_z^2 \\ \left(f_{t+h|t-1} - \sum_{i=0}^{h} (\iota'_1 \mathrm{B}_i)\, z_{t+h-i}\right)^2 - \sigma_{v,h}^2 \end{pmatrix}.$$

We define $g_{t+h}(\theta)$ similarly. It is clear that it is a just-identified system. In such a case, the method of moments estimator $\hat{\theta}$ can be understood as a two-step estimator. It first finds $\widehat{\mathrm{B}}_i$'s using the OLS moment conditions and then plug these estimates into the remaining conditions. Then $\hat{\sigma}_z^2$ and $\hat{\sigma}_{v,h}^2$ are derived given $\widehat{\mathrm{B}}_i$'s. It is worth noting that this is the same procedure we follow when we define $\hat{s}_h^{LPB}$. The only difference is that we are using here $f_{t+h|t-1}$ instead of its estimate.

Similar to the proof of Proposition 1, we know that

$$\sqrt{T}\left(\hat{\theta}-\theta_0\right) \xrightarrow{d} \mathcal{N}(0, \quad G^{-1}\Omega(G')^{-1})$$

where $G = E[\nabla_\theta g_{t+h}(\theta_0)]$, $\Omega = \sum_{l=-\infty}^{\infty}\Gamma(l)$, and $\Gamma(l)$ is the autocovariance of $g_{t+h}(\theta_0)$ at lag $l$. With some algebra, we obtain that

$$G = -E\left(\begin{array}{ccc|c} & I_{h+1}\otimes q_t q_t' & & 0 \\ \hline 0 & \cdots & 0 & \\ 2v_{t+h|t-1}z_{t+h}\iota_1' & \cdots & 2v_{t+h|t-1}z_{t+1}\iota_1' & I_2 \end{array}\right),$$

where $\otimes$ is the Kronecker product. For the bottom left part, we use the fact that $\iota_1'\mathrm{B}_i = \psi_{z,i}$ and $v_{t+h|t-1} = f_{t+h|t-1} - \sum_{i=0}^{h}\psi_{z,i}z_{t+h-i}$ . Because $v_{t+h|t-1} = \psi_{e,0}e_{t+h} + \cdots + \left(\psi_{e,0} + \cdots + \psi_{e,h}\right)e_t$ is orthogonal to $\{z_t\}$, where $\{\psi_{e,i}\}$ and $\{e_t\}$ are defined as is the case in Section 4.1, the bottom left block of $G$ becomes a zero matrix.

A transformation $\xi$ is required to connect $\theta$ with $s_h$. We define

$$\xi(\theta_0) = s_h = \frac{\sum_{i=0}^{h}(\iota_1'\mathrm{B}_i)^2\ \sigma_z^2}{\sum_{i=0}^{h}(\iota_1'\mathrm{B}_i)^2\ \sigma_z^2 + \sigma_{v,h}^2},$$

and $\xi(\theta)$ is also defined similarly.

Regarding the delta method, we need $\Delta \equiv \frac{\partial\xi(\theta_0)}{\partial\theta'}$. For a simple notation, we write $\sigma_{f,h}^2 \equiv Var\left(f_{t+h|t-1}\right) = \sum_{i=0}^{h}\psi_{z,i}^2\,\sigma_z^2 + \sigma_{v,h}^2$. With some algebra, we can show that

$$\Delta = \frac{1-s_h}{\sigma_{f,h}^2}\begin{pmatrix} 2\psi_{z,0}\sigma_z^2\iota_1 \\ \vdots \\ 2\psi_{z,h}\sigma_z^2\iota_1 \\ \sum_{i=0}^{h}\psi_{z,i}^2 \\ -s_h/(1-s_h) \end{pmatrix}'.$$

Combining the above derivations and being explicit about the fact that the moment conditions $g_{t+h}(\cdot)$ are for the LP-B approach at the horizon $h$, we have the result:

$$\sqrt{T}\left(\frac{\sum_{i=0}^{h}\left(\hat{\beta}_0^{i,LP}\right)^2\hat{\sigma}_z^2}{\sum_{i=0}^{h}\left(\hat{\beta}_0^{i,LP}\right)^2\hat{\sigma}_z^2 + \widehat{Var}\left(f_{t+h|t-1} - \sum_{i=0}^{h}\hat{\beta}_0^{i,LP}z_{t+h-i}\right)} - s_h\right) \xrightarrow{d} \mathcal{N}\left(0, \quad V_{h,LPB}\right),$$

where $V_{h,LPB} = \Delta_{h,LPB}\left(G_{h,LPB}\right)^{-1}\Omega_{h,LPB}\left(G_{h,LPB}'\right)^{-1}\Delta_{h,LPB}'$. □

***Joint inference.*** Below we explain how one can estimate a joint distribution of the LP-B estimators $(\hat{s}_0^{LPB}, \hat{s}_1^{LPB}, \dots, \hat{s}_H^{LPB})'$. Results for the LP-A estimators can be obtained similarly.

We consider augmented moment conditions $E\left[g_{t+H}^{Joint}(\theta_0)\right] = 0$, where $\theta_0 = (\mathrm{B}_0', \dots, \mathrm{B}_H', \sigma_z^2, \sigma_{v,0}^2, \dots, \sigma_{v,H}^2)'$ is a $(H+1) * \left(L_z + L_y + 3\right) + 1$ dimensional vector, and

$$g_{t+H}^{Joint}(\theta_0) = \begin{pmatrix} q_t\left(p_{0,t} - q_t'\mathrm{B}_0\right) \\ \vdots \\ q_t\left(p_{H,t} - q_t'\mathrm{B}_H\right) \\ z_t^2 - \sigma_z^2 \\ \left(f_{t|t-1} - \sum_{i=0}^{0}(\iota_1'\mathrm{B}_i)\, z_{t-i}\right)^2 - \sigma_{v,0}^2 \\ \vdots \\ \left(f_{t+H|t-1} - \sum_{i=0}^{H}(\iota_1'\mathrm{B}_i)\, z_{t+H-i}\right)^2 - \sigma_{v,H}^2 \end{pmatrix}.$$

Then it is straightforward to extend the proof of Proposition 4 to the joint distribution of $(\hat{s}_0^{LP}, \hat{s}_1^{LP}, \dots, \hat{s}_H^{LP})'$. In practice, both $L_z$ and $L_y$ should be small not to make $(H+1) * \left(L_z + L_y + 3\right) + 1$ too large relative to $T$.

***Estimating $V_{h,LPB}$.*** Next, we explain how one can estimate $V_{h,LPB} = \Delta_{h,LPB}\left(G_{h,LPB}\right)^{-1}\Omega_{h,LPB}\left(G_{h,LPB}'\right)^{-1}\Delta_{h,LPB}'$. Again, the asymptotic variance for the LP-A estimator can be estimated similarly.

Let's begin with $\mathrm{G}_{h,LPB} = -diag(I_{h+1} \otimes E[q_t q_t'], I_2)$. It is natural to have

$$\hat{G}_{h,LPB} = -diag\left(I_{h+1} \otimes \frac{1}{T_h}\sum_{t=L_{max}+1}^{T-h} q_t {q_t}', I_2\right).$$

The feasible estimator of $\theta$ is denoted by $\tilde{\theta} \equiv \left(\widehat{\mathrm{B}}_0', \dots, \widehat{\mathrm{B}}_h', \hat{\sigma}_z^2, \hat{\sigma}_{v,h}^2\right)'$, where $\hat{\sigma}_z^2 = \frac{1}{T}\sum_{t=1}^{T} z_t^2$ and $\hat{\sigma}_{v,h}^2 = \frac{1}{T_h}\sum_{t=L_{max+1}}^{T-h}\left(\hat{f}_{t+h|t-1} - \sum_{i=0}^{h}\iota_1'\widehat{\mathrm{B}}_i\, z_{t+h-i}\right)^2$.[2] We define $(H+1) * \left(L_z + L_y + 2\right) + 2$ dimensional vector $\pi_{t+h}$ as follows:

[2] The denominator $T_h$ might be adjusted according to the degrees of freedom without affecting the asymptotics.

$$\pi_{t+h} \equiv \begin{pmatrix} q_t\left(p_{0,t} - q_t'\widehat{\mathrm{B}}_0\right) \\ \vdots \\ q_t\left(p_{h,t} - q_t'\widehat{\mathrm{B}}_h\right) \\ z_t^2 - \hat{\sigma}_z^2 \\ \left(\hat{f}_{t+h|t-1} - \sum_{i=0}^{h}\left(\iota_1'\widehat{\mathrm{B}}_i\right) z_{t+h-i}\right)^2 - \hat{\sigma}_{v,h}^2 \end{pmatrix}.$$

Then $\widehat{\Omega}_{h,LPB}$ is obtained by applying the Newey-West estimator to $\pi_{t+h}$ with pre-whitening similar to the case of $\widehat{\Omega}_{h,R^2}$.

It remains to estimate $\Delta_{h,LPB}$. It is straightforward to define

$$\widehat{\Delta}_{h.LPB} = \frac{1-\hat{s}_h^{LPB}}{\hat{\sigma}_{f,h}^2}\begin{pmatrix} 2\hat{\psi}_{z,0}\hat{\sigma}_z^2\iota_1 \\ \vdots \\ 2\hat{\psi}_{z,h}\hat{\sigma}_z^2\iota_1 \\ \sum_{i=0}^{h}\hat{\psi}_{z,i}^2 \\ -\hat{s}_h^{LPB}/(1-\hat{s}_h^{LPB}) \end{pmatrix}',$$

where $\hat{\sigma}_{f,h}^2 = \frac{1}{T_h}\sum_{t=L_{max}+1}^{T-h} \hat{f}_{t+h|t-1}^2$. We plug the bias-corrected $\hat{s}_h^{LPB}$ in the place of $s_h$.

Combining all the estimators, the standard error of $\hat{s}_h^{LPB}$ is given by

$$[s.e.(\hat{s}_h^{LPB})]^2 = \frac{1}{T_h}\widehat{\Delta}_{h,LPB}\left(\hat{G}_{h,LPB}\right)^{-1}\widehat{\Omega}_{h,LPB}\left(\hat{G}_{h,LPB}'\right)^{-1}\widehat{\Delta}_{h,LPB}'.$$

In practice, both $L_z$ and $L_y$ should be small not to make the number of moment conditions $(H+1)*\left(L_z+L_y+2\right)+2$ too large relative to the sample size $T_h$.

***An alternative method for estimating biases based on the asymptotic distribution of $\tilde{\boldsymbol{\theta}}$.*** Similar to the discussion in Appendix A, we conjecture that most of the finite sample bias is due to the non-linear transformation $\xi(\cdot)$. Here, we focus on the LP-B estimator, because one can easily apply the same procedure to the LP-A estimator.

We approximate the asymptotic variance of the feasible estimator $\tilde{\theta}$ by $\frac{1}{T_h}\left(\hat{G}_{h,LPB}\right)^{-1}\widehat{\Omega}_{h,LPB}\left(\hat{G}_{h,LPB}'\right)^{-1}$. Then we simulate $\theta^{(b)}$ for $B$ times from the following normal distribution:

$$\theta^{(b)} \sim \mathcal{N}\left(\tilde{\theta}, \quad \frac{1}{T_h}\left(\hat{G}_{h,LPB}\right)^{-1}\widehat{\Omega}_{h,LPB}\left(\hat{G}_{h,LPB}'\right)^{-1}\right).$$

We drop cases where the simulated $\hat{\sigma}_z^2$ and $\hat{\sigma}_{v,h}^2$ are negative. Then the bias is estimated by $\frac{1}{B}\sum_{b=1}^{B}\xi(\theta^{(b)}) - \xi(\tilde{\theta})$ and the bias-corrected estimator is given by

$$\hat{s}_h^{LPB} = \xi(\tilde{\theta}) - \left(\frac{1}{B}\sum_{b=1}^{B}\xi(\theta^{(b)}) - \xi(\tilde{\theta})\right) = 2\xi(\tilde{\theta}) - \frac{1}{B}\sum_{b=1}^{B}\xi(\theta^{(b)}).$$

Note that calculating $\hat{G}_{h,LPB}$ and $\hat{\Omega}_{h,LPB}$ does not require deriving the bias-corrected $\hat{s}_h^{LPB}$. Therefore, we can first estimate $\hat{G}_{h,LPB}$ and $\hat{\Omega}_{h,LPB}$, derive the bias-corrected $\hat{s}_h^{LPB}$ by using the above method, and then calculating $\hat{\Delta}_{h.LPB}$ for the (asymptotic) standard error.

## Appendix C. Finding a MA($\infty$) representation for a process driven by multiple underlying shocks

Suppose the following data generating process as in Section 4.1. In this section, we explain how an infinite-order MA representation driven by a single white noise process is obtained for the residual process $\{\Delta p_t + \Delta a_t\}$.

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

We first show why having a representation $g_y + \psi_e(L)e_t$ of $\Delta p_t + \Delta a_t$ is needed. When all three shocks are in the information set, the corresponding forecast error with $\widetilde{\Omega}_t = \{z_t, \Delta y_t, e_t^p, e_t^a, \dots\}$ is

$$\tilde{f}_{t+h|t-1} = y_{t+h} - y_{t-1} - \mathrm{P}[y_{t+h} - y_{t-1}|\widetilde{\Omega}_{t-1}] = \psi_{z,0}z_{t+h} + \cdots + \psi_{z,h}z_t$$
$$+\sigma_p e_{t+h}^p + (1+\rho_p)\sigma_p e_{t+h-1}^p + \cdots + (1 + \rho_p + \cdots + \rho_p^h)\sigma_p e_t^p$$
$$+\sigma_a e_{t+h}^a + \rho_a \sigma_a e_{t+h-1}^a + \cdots + \rho_a^h \sigma_a e_t^a.$$

Thus, the corresponding forecast error variance decomposition becomes

$$\tilde{s}_h = \frac{\left(\sum_{i=0}^h \psi_{z,i}^2\right)\sigma_z^2}{\left(\sum_{i=0}^h \psi_{z,i}^2\right)\sigma_z^2 + \sum_{i=0}^h\left(\sum_{j=0}^i \rho_p^j\right)^2 \sigma_p^2 + \sum_{i=0}^h \rho_a^{2i}\,\sigma_a^2}.$$

However, what we estimate in the simulations is $s_h$, not $\tilde{s}_h$. $s_h$ is based only on the information set $\Omega_t = \{\Delta y_t, z_t, \dots\}$, not the augmented one, $\widetilde{\Omega}_t$. Because $\Omega_t$ is coarser than $\widetilde{\Omega}_t$, $s_h \le \tilde{s}_h$ as discussed in Section 3.6. To construct the true profile of $s_h$, we need $\psi_e(L)$ and $\sigma_e$.

We use a stationary Kalman filter (Hamilton 1994, pp.391-394) to do so. We cast the above process in a state-space representation.

**State equation:**

$$S_t = \boldsymbol{F}S_{t-1} + \boldsymbol{B}\kappa_t,$$

where $S_t = (\Delta p_t - g_y, \Delta a_t, e_t^a)',$

$$\kappa_t = (e_t^p, e_t^a)' \sim (0, I),$$

$$\mathrm{F} = \begin{pmatrix} \rho_p & 0 & 0 \\ 0 & \rho_a & -\sigma_a \\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathrm{B} = \begin{pmatrix} \sigma_p & 0 \\ 0 & \sigma_a \\ 0 & 1 \end{pmatrix}.$$

**Measurement equation:** $\Delta u_t = g_y + H'S_t,$ where $H = (1,1,0)'.$

By defining $Q = BIB' = BB'$ and $R = 0$, the stationary $P$ and $K$ are obtained from the matrix equation (13.5.3) and (13.5.4) in Hamilton (1994).

$$P = F[P - PH(H'PH + R)^{-1}H'P]F' + Q,$$

$$K = FPH(H'PH + R)^{-1}.$$

The first equation is a discrete time algebraic Riccati equation for $P$ which can be solved numerically. Then deriving $K$ is straightforward from the second equation. Given $K$, it is known that

$$\Delta u_t = g_y + (I + H'(I - FL)^{-1}KL)e_t, \qquad e_t \sim WN(\sigma_e^2), \qquad \text{and} \qquad \sigma_e = \sqrt{H'PH + R}.$$

To convert $(I + H'(I - FL)^{-1}KL)$ into $\psi_e(L)$, we use the identity that $(I - FL)^{-1} = I + FL + F^2L^2 + \cdots$. Note all three eigenvalues of $F$, $\rho_p$, $\rho_a$ and 0, are less than one in absolute values.

Given the MA representation of $\Delta u_t$, we can find $s_h$ accordingly.

In Section 4.2, the Smets and Wouters (2007) model is simulated. We find $s_h$ under the assumed information set in a similar way.

## Appendix D. Unobservable Shocks and Measurement Errors

In some cases, an identified structural shock is only a part of the true shock. For example, unanticipated innovations in the federal funds rates in Romer and Romer (2004) may be a part of the entire change in monetary policy such as verbal communication, forward guidance, members of the board of governors, and regime shifts. Similarly, legislative tax changes in Romer and Romer (2010) might be a part of the whole fiscal policy shocks affecting the U.S. economy. When shocks are generated from narratives as Ramey (2011), measurement errors might be another important issue. In this section, we show that our approach can still provide interesting and meaningful quantities in such cases, because our estimates are *conservative* estimates of the 'true' value available only when all hidden confounding factors are observable.

We decompose the true shock into two parts as $z_t = z_t^o + z_t^u$. The superscripts $o$ and $u$ mean observable and unobservable components, respectively. We assume that

$$\begin{pmatrix} z_t^o \\ z_t^u \end{pmatrix} = \begin{pmatrix} \sigma_o & 0 \\ \rho_{o,u}\sigma_u & \sqrt{1-\rho_{o,u}^2}\,\sigma_u \end{pmatrix} \delta_t,$$

where $\delta_t \sim wn(I_2)$, $\sigma_o^2 = Var(z_t^o)$, $\sigma_u^2 = Var(z_t^u)$, and $\rho_{o,u}$ is the correlation between $z_t^o$ and $z_t^u$. Also, $\delta_t$ and $e_t$ are uncorrelated at all leads and lags. 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 $\rho_{o,u} < 0$. We denote the full information set with $\Omega_{t-1} = \{x_{t-1}^o, x_{t-1}^u, \Delta y_{t-1}, \ldots\}$ in this section and the econometrician's information set with $\Omega_{t-1}^e \equiv \{x_{t-1}^o, \Delta y_{t-1}, \ldots\}$. 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 $\Omega_{t-1}^e$, while the full-information forecast error $f_{t+h|t-1}$ is based on $\Omega_{t-1}$. Finally, the econometrician's regressor is denoted by $Z_t^{h,e} = (z_{t+h}^o, \ldots, z_t^o)'$.

We argue that our estimators have a negative asymptotic bias, regardless of the sign of $\rho_{o,u}$.[3] Note that $s_h = \frac{\sum_{i=0}^{h} \psi_{z,i}^2 \sigma_z^2}{Var(f_{t+h|t-1})}$ is a ratio between the amount explained by the innovations in $\{z_t\}$ and the forecast error variance. As pointed out in Proposition 3, there are (a) a positive asymptotic bias in the denominator, and (b) a negative asymptotic bias in the numerator when we

[3] Because the R2, LP-A, and LP-B estimators share the same probability limit, we here focus on the R2 estimator.

apply our methods to $\{z_t^o\}$ and ignore the existence of $\{z_t^u\}$. Therefore, (c) the estimator is downward-biased, which is conservative in favor of the null hypothesis of no effects ($s_h = 0$).

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

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

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

*Proof.*

Let's begin with (a). With the full information, we can back out the forecast error as

$$f_{t+h|t-1} = y_{t+h} - y_{t-1} - P[y_{t+h} - y_{t-1} | \Omega_{t-1}],$$

where the last term is the projection of $y_{t+h} - y_{t-1}$ on the closed subspace spanned by $\Omega_{t-1}$. However, an econometrician has only $\Omega_{t-1}^e = \{z_{t-1}^o, \Delta y_{t-1}, z_{t-2}^o, \Delta y_{t-2}, \ldots\}$. It is evident that $\Omega_t^e \subset \Omega_t$. We define the closed subspaces spanned by each information set as

$$V_t = closure(span(\Omega_t)),$$
$$V_t^e = closure(span(\Omega_t^e)).$$

Using this notation,

$$f_{t+h|t-1} = y_{t+h} - y_{t-1} - P[y_{t+h} - y_{t-1} | V_{t-1}] = P[y_{t+h} - y_{t-1} | (V_{t-1})^{\perp}],$$
$$f_{t+h|t-1}^e = y_{t+h} - y_{t-1} - P[y_{t+h} - y_{t-1} | V_{t-1}^e] = f_{t+h|t-1} + r_{t+h|t-1}^e,$$

where $V_t^{\perp}$ is the orthogonal subspace of $V_t$ and $r_{t+h|t-1}^e \equiv P[y_{t+h} - y_{t-1} | V_{t-1}] - P[y_{t+h} - y_{t-1} | V_{t-1}^e]$.

Note that $V_t^e \subset V_t$, and therefore $r_{t+h|t-1}^e \in V_{t-1}$. Because $f_{t+h|t-1} \in (V_{t-1})^{\perp}$, it follows that $f_{t+h|t-1}$ and $r_{t+h|t-1}^e$ are orthogonal. [4] Therefore, $Var(f_{t+h|t-1}^e) = Var(f_{t+h|t-1}) + Var(r_{t+h|t-1}^e) \geq Var(f_{t+h|t-1})$. Also, the equality holds only when $z_t^u$ and its lagged values have no additional power in explaining $y_{t+h}$ given $V_{t-1}^e$ implying $P[y_{t+h} - y_{t-1} | V_{t-1}] =$

[4] This result is in fact due to a decomposition of the entire vector space V into a direct sum of three mutually orthogonal closed subspaces: $V = V_{t-1}^e \oplus (V_{t-1} \cap (V_{t-1}^e)^{\perp}) \oplus (V \cap (V_{t-1})^{\perp})$, where '$\oplus$' means a direct sum. From the decomposition, it directly follows that $y_{t+h} - y_{t-1} = P[y_{t+h} - y_{t-1} | V] = P[y_{t+h} - y_{t-1} | V_{t-1}^e] + P[y_{t+h} - y_{t-1} | V_{t-1} \cap (V_{t-1}^e)^{\perp}] + P[y_{t+h} - y_{t-1} | V \cap (V_{t-1})^{\perp}] = P[y_{t+h} - y_{t-1} | V_{t-1}^e] + r_{t+h|t-1}^e + f_{t+h|t-1}$, and the last three terms are mutually orthogonal.

$P[y_{t+h}-y_{t-1}|V_{t-1}^{e}]$. This is true only for (uninteresting) special cases such as $\psi_z(L)=0$, $\rho_{o,u}=\pm 1$, and $\sigma_z^u=0$.

We next turn to (b) and show that the econometrician's numerator converges in probability to a value less than $Var(\psi_{z,0}z_{t+h}+\cdots+\psi_{z,h}z_t)=\sum_{i=0}^{h}\psi_{z,i}^2\,\sigma_z^2$. As illustrated in Equation (4),

$$Var(\psi_{z,0}z_{t+h}+\cdots+\psi_{z,h}z_t)=E(f_{t+h|t-1}\cdot Z_t^{h\prime})E(Z_t^h\,Z_t^{h\prime})^{-1}E(Z_t^h\;\cdot f_{t+h|t-1})$$
$$=\left[E(Z_t^hZ_t^{h\prime})^{-1}E(Z_t^h\;\cdot f_{t+h|t-1})\right]'\left[E(Z_t^h\,Z_t^{h\prime})\right]\left[E(Z_t^hZ_t^{h\prime})^{-1}E(Z_t^h\;\cdot f_{t+h|t-1})\right].$$

The term inside the last square brackets is a vector of population regression coefficients of $f_{t+h|t-1}$ on $Z_t^h$, which is $\Psi^h=(\psi_{z,0},\dots,\psi_{z,h})'$ by construction.

Now we investigate the econometrician's numerator $E\left(f_{t+h|t-1}^{e}\cdot Z_t^{h,e\prime}\right)E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot f_{t+h|t-1}^{e})$, where $Z_t^{h,e}=(z_{t+h}^{o},\dots,z_t^{o})'$. Note that elements of $Z_t^{h,e}$ lies in $V_{t-1}^{\perp}$, while $r_{t+h|t-1}^{e}\in V_{t-1}$. Thus,

$$E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot f_{t+h|t-1}^{e})$$
$$=E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot f_{t+h|t-1})+E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot r_{t+h|t-1}^{e})$$
$$=E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot f_{t+h|t-1})$$
$$=\left(1+\frac{Cov(z_t^o,z_t^u)}{Var(z_t^o)}\right)\Psi^h=\frac{(\sigma_o+\rho_{o,u}\cdot\sigma_u)}{\sigma_o}\Psi^h.$$

We used the fact that $f_{t+h|t-1}=\sum_{i=0}^{h}\psi_{z,i}z_{t+h-i}+\sum_{i=0}^{h}\sum_{j=0}^{i}(\psi_{e,j})\;e_{t+h-i}$ for the last line.

Finally, the econometrician's numerator becomes

$$\left[E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot f_{t+h|t-1}^{e})\right]'\left[E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)\right]\left[E\left(Z_t^{h,e}\,Z_t^{h,e\prime}\right)^{-1}E(Z_t^{h,e}\;\cdot f_{t+h|t-1}^{e})\right]$$
$$=\frac{(\sigma_o+\rho_{o,u}\cdot\sigma_u)}{\sigma_o}\Psi^{h\prime}\cdot\;\sigma_o^2 I\;\cdot\frac{(\sigma_o+\rho_{o,u}\cdot\sigma_u)}{\sigma_o}\Psi^h=\sum_{i=0}^{h}\psi_{z,i}^2\,(\sigma_o+\rho_{o,u}\cdot\sigma_u)^2.$$

Thus, any asymptotic bias in the numerator is from the differences between $\sigma_z^2$ and $(\sigma_o+\rho_{o,u}\cdot\sigma_u)^2$. Because $\sigma_z^2-(\sigma_o+\rho_{o,u}\cdot\sigma_u)^2=(1-\rho_{o,u}^2)\sigma_u^2$,

$$E(f_{t+h|t-1}\cdot Z_t^{h\prime})E(Z_t^h\,Z_t^{h\prime})^{-1}E(Z_t^h\;\cdot f_{t+h|t-1})$$

$$= E\left(f^e_{t+h|t-1} \cdot Z_t^{h,e'}\right) E\left(Z_t^{h,e}\, Z_t^{h,e'}\right)^{-1} E\left(Z_t^{h,e} \cdot f^e_{t+h|t-1}\right) + \sum_{i=0}^{h} \psi_{z,i}^2 \left(1 - \rho_{o,u}^2\right)\sigma_u^2.$$

The econometrician's numerator and denominator are asymptotically less and greater than their full-information counterparts, respectively. Therefore, we have a negative asymptotic bias as claimed in (c). Moreover, the biases become small when $\psi_{z,i}$'s are close to zero, the observed and unobserved components are highly correlated, or the variance of the unobserved part is small. In such cases, both biases in the denominator $Var\left(r^e_{t+h|t-1}\right)$ and the numerator $\sum_{i=0}^{h} \psi_{z,i}^2 \left(1 - \rho_{o,u}^2\right)\sigma_u^2$ are small.

So far, we assumed that $z_t^o$ and $z_t^u$ have the same impulse response function $\psi_z(L)$ for simplicity. However, we may consider $\psi_z^o(L) z_t^o + \psi_z^u(L) z_t^u$ instead of $\psi_z(L) z_t$. This does not change our results and the above derivations admit a straight-forward extension to this general case. In this case, the difference between two numerators becomes $\sum_{i=0}^{h} \left(\psi_{z,i}^u\right)^2 \left(1 - \rho_{o,u}^2\right)\sigma_u^2$. □

## Appendix E. Alternative specifications

This section discusses the performance of alternative specifications relative to that of our benchmark in several respects. Specifically, we consider three cases: (1) block bootstraps are used for bias-correction, (2) $L_{VAR} = L_z = L_y$ are selected by Akaike information criterion, and (3) inference is based on asymptotic standard errors.

### Appendix E1. A comparison between the block and the VAR-based bootstraps

For simulation studies in the main text, we rely on the VAR-based bootstraps to correct for biases and to construct confidence intervals. Here we compare the performances of the VAR-based bootstrap and the block bootstrap for local projections suggested by Kilian and Kim (2011).

We show results for three versions of R2 estimators: (1) R2 estimator without bias-correction, (2) R2 estimator with VAR-based bias-correction, and (3) R2 estimator with bias-correction based on block bootstraps. Similar to the main text, we experiment with DGP 1, 2, and 3, where the sample size is 160, and the replication size is 2,000. For each estimator, we calculate mean, root mean-squared error, and coverage probability of 90% confidence intervals. The VAR lag order is selected by the HQIC, and the size of each block is four following Kilian and Kim (2011). For the DGP 3, we also check the results when the VAR lag order and the size of each block are ten as in Table 4.

As illustrated in Table E1-E3, the VAR-based bootstrap performs better than the block bootstrap. The R2 estimator with VAR-based bias-correction has smaller root mean-squared errors. Furthermore, coverage rates for the estimator with VAR-bootstrap are closer to the nominal rate than the estimator based on block bootstraps.

*Table E1. Simulation results for DGP 1.*

| | Horizon $h$ | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 4 | 8 | 12 | 16 | 20 |
| **Forecast Error Variance Decomposition (VAR(HQIC), Block Size(4))** | | | | | | |
| True | 0.00 | 0.04 | 0.19 | 0.21 | 0.18 | 0.14 |
| Average estimate | | | | | | |
| R2, Without bias-correction | 0.01 | 0.06 | 0.20 | 0.25 | 0.26 | 0.27 |
| R2, VAR-based bootstrap | 0.00 | 0.02 | 0.13 | 0.16 | 0.13 | 0.11 |
| R2, Block bootstrap | 0.00 | 0.03 | 0.16 | 0.21 | 0.19 | 0.18 |
| Root mean squared error | | | | | | |
| R2, Without bias-correction | 0.01 | 0.05 | 0.11 | 0.15 | 0.19 | 0.22 |
| R2, VAR-based bootstrap | 0.01 | 0.05 | 0.12 | 0.16 | 0.17 | 0.18 |
| R2, Block bootstrap | 0.01 | 0.05 | 0.12 | 0.17 | 0.19 | 0.21 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, Without bias-correction | 0.99 | 0.81 | 0.69 | 0.65 | 0.63 | 0.61 |
| R2, VAR-based bootstrap | 0.99 | 0.95 | 0.64 | 0.64 | 0.72 | 0.81 |
| R2, Block bootstrap | 0.99 | 0.90 | 0.69 | 0.59 | 0.54 | 0.56 |

*Notes:* The results are for DGP1 in Section 4.1. The sample size is $T = 160$, and the number of simulations is 2,000. We consider three estimators in the table. 'R2, Without bias-correction' stands for $\hat{s}_h^{R2}$ without bias-correction. 'R2, VAR-based bootstrap' denotes for $\hat{s}_h^{R2} - \left(\frac{1}{B}\sum_{b=1}^{B}\hat{s}_h^{R2,(b)} - s_h^*\right)$, where $\hat{s}_h^{R2,(b)}$ is based on a simulated sample using a VAR. $s_h^*$ is the true FEVD in this DGP used for bootstrapping $\hat{s}_h^{R2,(b)}$, and $\left(\frac{1}{B}\sum_{b=1}^{B}\hat{s}_h^{R2,(b)} - s_h^*\right)$ is the estimated bias in $\hat{s}_h^{R2}$. The order of the VAR is selected by HQIC. Finally, 'R2, Block bootstrap' is for $\hat{s}_h^{R2} - \left(\frac{1}{B}\sum_{b=1}^{B}\hat{s}_h^{R2,(b)} - \hat{s}_h^{R2}\right)$, where $\hat{s}_h^{R2,(b)}$ is obtained from the block bootstrap. Similarly, $\left(\frac{1}{B}\sum_{b=1}^{B}\hat{s}_h^{R2,(b)} - \hat{s}_h^{R2}\right)$ is the estimated bias in this case. Following Kilian and Kim (2011), the size of each block is four. We rely on bootstrap estimates for constructing confidence intervals similar to Table 2-4.

*Table E2. Simulation results for DGP 2.*

| | Horizon $h$ | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 4 | 8 | 12 | 16 | 20 |
| **Forecast Error Variance Decomposition (VAR(HQIC), Block Size(4))** | | | | | | |
| True | 0.80 | 0.25 | 0.10 | 0.05 | 0.03 | 0.02 |
| Average estimate | | | | | | |
| R2, Without bias-correction | 0.79 | 0.26 | 0.15 | 0.14 | 0.15 | 0.19 |
| R2, VAR-based bootstrap | 0.81 | 0.24 | 0.09 | 0.03 | 0.01 | 0.00 |
| R2, Block bootstrap | 0.80 | 0.25 | 0.10 | 0.07 | 0.06 | 0.07 |
| Root mean squared error | | | | | | |
| R2, Without bias-correction | 0.03 | 0.11 | 0.12 | 0.14 | 0.17 | 0.21 |
| R2, VAR-based bootstrap | 0.03 | 0.10 | 0.09 | 0.09 | 0.10 | 0.12 |
| R2, Block bootstrap | 0.03 | 0.11 | 0.12 | 0.13 | 0.15 | 0.17 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, Without bias-correction | 0.90 | 0.89 | 0.89 | 0.82 | 0.73 | 0.67 |
| R2, VAR-based bootstrap | 0.92 | 0.90 | 0.97 | 0.97 | 0.95 | 0.94 |
| R2, Block bootstrap | 0.91 | 0.73 | 0.68 | 0.77 | 0.79 | 0.75 |

*Notes:* The results are for DGP2 in Section 4.1. The sample size is $T = 160$, and the number of simulations is 2,000. We consider three estimators in the table. 'R2, Without bias-correction' stands for $\hat{s}_h^{R2}$ without bias-correction. 'R2, VAR-based bootstrap' denotes for $\hat{s}_h^{R2} - \left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - s_h^*\right)$, where $\hat{s}_h^{R2,(b)}$ is based on a simulated sample using a VAR. $s_h^*$ is the true FEVD in this DGP used for bootstrapping $\hat{s}_h^{R2,(b)}$, and $\left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - s_h^*\right)$ is the estimated bias in $\hat{s}_h^{R2}$. The order of the VAR is selected by HQIC. Finally, 'R2, Block bootstrap' is for $\hat{s}_h^{R2} - \left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - \hat{s}_h^{R2}\right)$, where $\hat{s}_h^{R2,(b)}$ is obtained from the block bootstrap. Similarly, $\left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - \hat{s}_h^{R2}\right)$ is the estimated bias in this case. Following Kilian and Kim (2011), the size of each block is four. We rely on bootstrap estimates for constructing confidence intervals similar to Table 2-4.

*Table E3. Simulation results for DGP 3.*

| | Horizon $h$ | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 4 | 8 | 12 | 16 | 20 |
| **Forecast Error Variance Decomposition (VAR(HQIC), Block Size(4))** | | | | | | |
| True | 0.06 | 0.29 | 0.47 | 0.58 | 0.65 | 0.70 |
| Average estimate | | | | | | |
| R2, Without bias-correction | 0.06 | 0.22 | 0.36 | 0.45 | 0.52 | 0.57 |
| R2, VAR-based bootstrap | 0.05 | 0.21 | 0.32 | 0.40 | 0.44 | 0.46 |
| R2, Block bootstrap | 0.05 | 0.21 | 0.34 | 0.43 | 0.49 | 0.54 |
| Root mean squared error | | | | | | |
| R2, Without bias-correction | 0.03 | 0.11 | 0.17 | 0.19 | 0.20 | 0.21 |
| R2, VAR-based bootstrap | 0.04 | 0.13 | 0.20 | 0.24 | 0.27 | 0.29 |
| R2, Block bootstrap | 0.04 | 0.13 | 0.19 | 0.22 | 0.24 | 0.25 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, Without bias-correction | 0.85 | 0.76 | 0.75 | 0.78 | 0.81 | 0.84 |
| R2, VAR-based bootstrap | 0.83 | 0.72 | 0.68 | 0.68 | 0.70 | 0.71 |
| R2, Block bootstrap | 0.83 | 0.65 | 0.53 | 0.50 | 0.46 | 0.44 |
| **Forecast Error Variance Decomposition (VAR(10), Block Size(10))** | | | | | | |
| True | 0.06 | 0.29 | 0.47 | 0.58 | 0.65 | 0.70 |
| Average estimate | | | | | | |
| R2, Without bias-correction | 0.06 | 0.25 | 0.39 | 0.48 | 0.54 | 0.57 |
| R2, VAR-based bootstrap | 0.07 | 0.29 | 0.46 | 0.56 | 0.62 | 0.65 |
| R2, Block bootstrap | 0.06 | 0.27 | 0.43 | 0.53 | 0.59 | 0.62 |
| Root mean squared error | | | | | | |
| R2, Without bias-correction | 0.04 | 0.11 | 0.16 | 0.19 | 0.21 | 0.22 |
| R2, VAR-based bootstrap | 0.05 | 0.12 | 0.16 | 0.19 | 0.21 | 0.23 |
| R2, Block bootstrap | 0.04 | 0.13 | 0.18 | 0.21 | 0.23 | 0.26 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, Without bias-correction | 0.81 | 0.83 | 0.81 | 0.83 | 0.83 | 0.83 |
| R2, VAR-based bootstrap | 0.74 | 0.78 | 0.79 | 0.80 | 0.82 | 0.82 |
| R2, Block bootstrap | 0.76 | 0.72 | 0.67 | 0.65 | 0.63 | 0.61 |

*Notes:* The results are for DGP3 in Section 4.1. The sample size is $T = 160$, and the number of simulations is 2,000. We consider three estimators in the table. ‘R2, Without bias-correction’ stands for $\hat{s}_h^{R2}$ without bias-correction. ‘R2, VAR-based bootstrap’ denotes for $\hat{s}_h^{R2} - \left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - s_h^*\right)$, where $\hat{s}_h^{R2,(b)}$ is based on a simulated sample using a VAR. $s_h^*$ is the true FEVD in this DGP used for bootstrapping $\hat{s}_h^{R2,(b)}$, and $\left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - s_h^*\right)$ is the estimated bias in $\hat{s}_h^{R2}$. The order of the VAR is either selected by HQIC or ten. Finally, ‘R2, Block bootstrap’ is for $\hat{s}_h^{R2} - \left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - \hat{s}_h^{R2}\right)$, where $\hat{s}_h^{R2,(b)}$ is obtained from the block bootstrap. Similarly, $\left(\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{R2,(b)} - \hat{s}_h^{R2}\right)$ is the estimated bias in this case. The size of each block is either four or ten. We rely on bootstrap estimates for constructing confidence intervals similar to Table 2-4.

## Appendix E2. VAR with Akaike information criterion

This section considers the Akaike information criterion (AIC) when we select the lag order $L_{VAR} = L_z = L_y$ instead of the HQIC. It is well known that the AIC suggests higher-order models than the HQIC or the Bayesian information criterion. Thus, we can check how sensitive our new estimators of FEVDs are to the lag order of VAR models that are used to correct for the biases. Also, we can compare the performance of the higher-order VARs with our estimators when estimating FEVDs.

Tables E4-E6 illustrate results for DGP 1-3. The lag orders $L_{VAR} = L_z = L_y$ are selected by the AIC, while all the other details are similar to the simulations in Section 4.1. It is clear from the tables that the FEVDs based on bivariate VARs are strongly biased for DGP 1 and 3. Note that DGP 2 admits a VAR(1) representation, and this is the only case in our simulations when VARs perform better than our suggested estimators. Furthermore, the R2 estimator is slightly more efficient when the VAR lag order is selected via HQIC than AIC. This can be seen from Table 2(3) and Table E4(E5) for DGP 1(2).

*Table E4. 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 | 1.40 | 3.01 | 2.03 | 0.83 | 0.27 |
| VAR(AIC) | 0.00 | 0.81 | 1.41 | 1.40 | 1.37 | 1.38 |
| **Forecast Error Variance Decomposition** | | | | | | |
| True | 0.00 | 0.04 | 0.19 | 0.21 | 0.18 | 0.14 |
| Average estimate | | | | | | |
| R2 | 0.01 | 0.06 | 0.21 | 0.26 | 0.26 | 0.27 |
| LP A | 0.01 | 0.05 | 0.19 | 0.24 | 0.23 | 0.24 |
| LP B | 0.01 | 0.05 | 0.19 | 0.23 | 0.22 | 0.22 |
| VAR(AIC) | 0.01 | 0.04 | 0.10 | 0.12 | 0.13 | 0.13 |
| Root mean squared error | | | | | | |
| R2 | 0.01 | 0.06 | 0.12 | 0.16 | 0.19 | 0.22 |
| LP A | 0.01 | 0.05 | 0.13 | 0.16 | 0.19 | 0.21 |
| LP B | 0.01 | 0.05 | 0.12 | 0.15 | 0.16 | 0.16 |
| VAR(AIC) | 0.01 | 0.04 | 0.15 | 0.18 | 0.17 | 0.17 |
| Coverage (90 % level) (asymptotic) | | | | | | |
| R2 | 0.98 | 0.85 | 0.80 | 0.78 | 0.75 | 0.72 |
| LP A | 0.99 | 0.90 | 0.66 | 0.70 | 0.80 | 0.79 |
| LP B | 0.99 | 0.89 | 0.64 | 0.67 | 0.78 | 0.77 |
| VAR(AIC) | 0.99 | 0.85 | 0.38 | 0.37 | 0.38 | 0.39 |
| **Forecast Error Variance Decomposition (bias-corrected, VAR(AIC))** | | | | | | |
| True | 0.00 | 0.04 | 0.19 | 0.21 | 0.18 | 0.14 |
| Average estimate | | | | | | |
| R2 | 0.00 | 0.03 | 0.15 | 0.18 | 0.16 | 0.14 |
| LP A | 0.00 | 0.03 | 0.15 | 0.17 | 0.15 | 0.13 |
| LP B | 0.00 | 0.03 | 0.16 | 0.18 | 0.17 | 0.15 |
| VAR(AIC) | 0.00 | 0.02 | 0.09 | 0.11 | 0.12 | 0.13 |
| Root mean squared error | | | | | | |
| R2 | 0.01 | 0.06 | 0.14 | 0.17 | 0.18 | 0.19 |
| LP A | 0.01 | 0.05 | 0.12 | 0.16 | 0.17 | 0.18 |
| LP B | 0.01 | 0.05 | 0.13 | 0.16 | 0.17 | 0.17 |
| VAR(AIC) | 0.01 | 0.05 | 0.16 | 0.19 | 0.18 | 0.18 |
| Coverage (90 % level) | | | | | | |
| R2 | 0.98 | 0.94 | 0.68 | 0.72 | 0.79 | 0.86 |
| LP A | 1.00 | 0.90 | 0.62 | 0.66 | 0.77 | 0.89 |
| LP B | 1.00 | 0.90 | 0.58 | 0.60 | 0.72 | 0.84 |
| VAR(AIC) | 0.99 | 0.73 | 0.36 | 0.35 | 0.36 | 0.36 |

*Notes:* The table reports the performances of various 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 number of lags is selected by the Akaike information criterion (AIC). 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. The average lag order $L_{VAR}$ is 3.05.

*Table E5. 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 | 1.89 | 1.14 | 0.64 | 0.30 | 0.08 |
| VAR(AIC) | 2.95 | 1.95 | 1.36 | 1.00 | 0.77 | 0.62 |
| **Forecast Error Variance Decomposition** | | | | | | |
| True | 0.80 | 0.25 | 0.10 | 0.05 | 0.03 | 0.02 |
| Average estimate | | | | | | |
| R2 | 0.78 | 0.27 | 0.15 | 0.14 | 0.16 | 0.18 |
| LP A | 0.80 | 0.27 | 0.13 | 0.09 | 0.09 | 0.10 |
| LP B | 0.79 | 0.26 | 0.13 | 0.09 | 0.09 | 0.09 |
| VAR(AIC) | 0.80 | 0.27 | 0.13 | 0.08 | 0.06 | 0.05 |
| Root mean squared error | | | | | | |
| R2 | 0.04 | 0.10 | 0.11 | 0.14 | 0.17 | 0.21 |
| LP A | 0.03 | 0.09 | 0.08 | 0.08 | 0.10 | 0.12 |
| LP B | 0.03 | 0.08 | 0.07 | 0.08 | 0.09 | 0.11 |
| VAR(AIC) | 0.03 | 0.08 | 0.07 | 0.06 | 0.06 | 0.05 |
| Coverage (90 % level) (asymptotic) | | | | | | |
| R2 | 0.88 | 0.88 | 0.90 | 0.82 | 0.73 | 0.67 |
| LP A | 0.93 | 0.90 | 0.92 | 0.87 | 0.82 | 0.77 |
| LP B | 0.87 | 0.89 | 0.91 | 0.86 | 0.80 | 0.75 |
| VAR(AIC) | 0.90 | 0.89 | 0.92 | 0.96 | 0.97 | 0.97 |
| **Forecast Error Variance Decomposition (bias-corrected, VAR(AIC))** | | | | | | |
| True | 0.80 | 0.25 | 0.10 | 0.05 | 0.03 | 0.02 |
| Average estimate | | | | | | |
| R2 | 0.81 | 0.25 | 0.09 | 0.04 | 0.02 | 0.00 |
| LP A | 0.79 | 0.25 | 0.10 | 0.05 | 0.03 | 0.02 |
| LP B | 0.81 | 0.25 | 0.10 | 0.05 | 0.03 | 0.02 |
| VAR(AIC) | 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 | 0.11 | 0.13 |
| LP A | 0.04 | 0.08 | 0.07 | 0.07 | 0.08 | 0.09 |
| LP B | 0.03 | 0.08 | 0.07 | 0.07 | 0.07 | 0.08 |
| VAR(AIC) | 0.03 | 0.08 | 0.06 | 0.05 | 0.05 | 0.04 |
| Coverage (90 % level) | | | | | | |
| R2 | 0.93 | 0.91 | 0.96 | 0.96 | 0.94 | 0.93 |
| LP A | 0.93 | 0.90 | 0.92 | 0.95 | 0.93 | 0.91 |
| LP B | 0.90 | 0.86 | 0.89 | 0.93 | 0.90 | 0.90 |
| VAR(AIC) | 0.90 | 0.89 | 0.91 | 0.96 | 0.99 | 0.99 |

*Notes:* The table reports the performances of various 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 number of lags is selected by the Akaike information criterion (AIC). 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. The average lag order $L_{VAR}$ is 1.25.

*Table E6. Simulation results for DGP 3.*

| | 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 | 0.97 | 3.84 | 5.61 | 6.70 | 7.28 | 7.56 |
| VAR(AIC) | 0.95 | 3.09 | 3.55 | 3.66 | 3.69 | 3.71 |
| **Forecast Error Variance Decomposition** | | | | | | |
| True | 0.06 | 0.29 | 0.47 | 0.58 | 0.65 | 0.70 |
| Average estimate | | | | | | |
| R2 | 0.06 | 0.24 | 0.37 | 0.47 | 0.53 | 0.58 |
| LP A | 0.06 | 0.25 | 0.40 | 0.51 | 0.60 | 0.67 |
| LP B | 0.06 | 0.24 | 0.37 | 0.46 | 0.52 | 0.56 |
| VAR(AIC) | 0.06 | 0.19 | 0.23 | 0.25 | 0.26 | 0.26 |
| Root mean squared error | | | | | | |
| R2 | 0.04 | 0.11 | 0.17 | 0.19 | 0.20 | 0.21 |
| LP A | 0.04 | 0.11 | 0.17 | 0.20 | 0.23 | 0.26 |
| LP B | 0.04 | 0.11 | 0.16 | 0.18 | 0.20 | 0.20 |
| VAR(AIC) | 0.04 | 0.14 | 0.27 | 0.37 | 0.43 | 0.47 |
| Coverage (90 % level) (asymptotic) | | | | | | |
| R2 | 0.85 | 0.77 | 0.77 | 0.80 | 0.83 | 0.86 |
| LP A | 0.88 | 0.81 | 0.77 | 0.79 | 0.81 | 0.82 |
| LP B | 0.87 | 0.78 | 0.75 | 0.76 | 0.79 | 0.81 |
| VAR(AIC) | 0.87 | 0.56 | 0.30 | 0.18 | 0.12 | 0.10 |
| **Forecast Error Variance Decomposition (bias-corrected, VAR(AIC))** | | | | | | |
| True | 0.06 | 0.29 | 0.47 | 0.58 | 0.65 | 0.70 |
| Average estimate | | | | | | |
| R2 | 0.06 | 0.23 | 0.36 | 0.44 | 0.48 | 0.51 |
| LP A | 0.06 | 0.23 | 0.36 | 0.45 | 0.52 | 0.57 |
| LP B | 0.06 | 0.23 | 0.36 | 0.44 | 0.49 | 0.52 |
| VAR(AIC) | 0.06 | 0.18 | 0.23 | 0.24 | 0.25 | 0.26 |
| Root mean squared error | | | | | | |
| R2 | 0.04 | 0.12 | 0.19 | 0.22 | 0.25 | 0.27 |
| LP A | 0.04 | 0.12 | 0.18 | 0.22 | 0.25 | 0.28 |
| LP B | 0.04 | 0.12 | 0.18 | 0.21 | 0.23 | 0.25 |
| VAR(AIC) | 0.04 | 0.15 | 0.28 | 0.37 | 0.43 | 0.47 |
| Coverage (90 % level) | | | | | | |
| R2 | 0.82 | 0.75 | 0.72 | 0.73 | 0.74 | 0.74 |
| LP A | 0.86 | 0.77 | 0.73 | 0.75 | 0.78 | 0.79 |
| LP B | 0.85 | 0.74 | 0.69 | 0.70 | 0.72 | 0.73 |
| VAR(AIC) | 0.85 | 0.54 | 0.29 | 0.17 | 0.13 | 0.10 |

*Notes:* The table reports the performances of various 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. The number of lags is selected by the Akaike information criterion (AIC). 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. The average lag order $L_{VAR}$ is 2.40.

## Appendix E3. Asymptotic vs. bootstrap standard error

As illustrated in Section 3.5, our benchmark method for inference is based on the bootstrap standard errors and the distribution of $\hat{s}_h^{R2,(b)}$ across $b \leq B$. On the other hand, one may estimate the asymptotic standard deviation of $\hat{s}_h^{R2}$ in Proposition 1 and 2 directly to compute the standard errors. This section compares the performance of those different approaches of inference in small samples. We consider DGP 1-3 in Section 4.1 and the Smets and Wouters (2007) model in Section 4.2. For the Smets and Wouters model, we investigate how either economic output or price inflation responds to monetary policy shocks where the information set includes output growth rate, price inflation, monetary policy rate, and monetary policy shock. The results are shown in Tables E7-E11.

In each table, coverage rates under the name of "R2, bootstrap [P5, P95]" means the probability of $s_h$ being in the interval $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC},\ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right]$ as discussed in Section 3.5, where $\alpha = 0.1$. In this case, confidence intervals are obtained from the bootstrap distribution of $\hat{s}_h^{R2,(b)}$. Instead, "R2, bootstrap s.e." implies that the confidence interval is constructed in a symmetric way with the bootstrap standard error, which is the standard deviation of $\hat{s}_h^{R2,(b)}$ across $b$. The critical values are obtained from a standard normal distribution. When $\alpha = 0.1$, it is 1.65. Symmetric confidence bands based on the asymptotic s.e., not the bootstrap s.e., is denoted by "R2, asymptotic s.e." Finally, we also present the results for the VAR-based FEVDs. Confidence bands are constructed using bootstrap samples as is the case in "R2, bootstrap [P5, P95]."

When we compare the coverage rates of confidence intervals around bias-corrected estimators of FEVDs, it becomes clear that asymptotic standard errors do not perform strictly better than our benchmark method. While the coverage rates of the symmetric confidence intervals based on the asymptotic standard error are closer to the nominal rate of 90% for DGP1, the opposite is true for DGP 3. Furthermore, the asymptotic standard error is rather "spiky" across horizons. Combined with non-smooth $\hat{s}_h$, erratic standard errors will produce figures with choppy confidence bands.

*Table E7. 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, bootstrap [P5, P95] | 0.99 | 0.81 | 0.69 | 0.65 | 0.63 | 0.61 |
| R2, bootstrap s.e. | 1.00 | 0.93 | 0.72 | 0.71 | 0.74 | 0.72 |
| R2, asymptotic s.e. | 1.00 | 0.91 | 0.80 | 0.79 | 0.78 | 0.77 |
| VAR(HQIC), bootstrap [P5, P95] | 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, bootstrap [P5, P95] | 0.99 | 0.95 | 0.64 | 0.64 | 0.72 | 0.81 |
| R2, bootstrap s.e. | 1.00 | 0.92 | 0.60 | 0.61 | 0.67 | 0.77 |
| R2, asymptotic s.e. | 0.78 | 0.72 | 0.72 | 0.72 | 0.71 | 0.69 |
| VAR(HQIC), bootstrap [P5, P95] | 1.00 | 0.53 | 0.06 | 0.05 | 0.07 | 0.09 |

*Notes:* The table reports the performances of various 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 number of lags is selected by the Hannan-Quinn information criterion (HQIC). We consider three different methods of constructing confidence intervals. First, we can use $\alpha/2$ and $1-\alpha/2$ quantiles of the bootstrap estimates $\hat{s}_h^{R2,(b)}$ as discussed in Section 3.5, where $\alpha = 0.1$. The two other methods build symmetric confidence intervals with critical values from a standard normal distribution. We investigate both bootstrap standard error and asymptotic standard error. Confidence intervals for the other estimators are constructed similarly. For the asymptotic s.e., we use the asymptotic variance in Proposition 1 and follow the implementation details in Section A.

*Table E8. 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, bootstrap [P5, P95] | 0.90 | 0.89 | 0.89 | 0.82 | 0.73 | 0.67 |
| R2, bootstrap s.e. | 0.91 | 0.87 | 0.91 | 0.89 | 0.84 | 0.79 |
| R2, asymptotic s.e. | 0.85 | 0.83 | 0.82 | 0.90 | 0.86 | 0.80 |
| VAR(HQIC), bootstrap [P5, P95] | 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, bootstrap [P5, P95] | 0.92 | 0.90 | 0.97 | 0.97 | 0.95 | 0.94 |
| R2, bootstrap s.e. | 0.90 | 0.89 | 0.96 | 0.98 | 0.98 | 0.97 |
| R2, asymptotic s.e. | 0.83 | 0.86 | 0.83 | 0.85 | 0.86 | 0.85 |
| VAR(HQIC), bootstrap [P5, P95] | 0.88 | 0.89 | 0.91 | 0.96 | 0.99 | 0.99 |

*Notes:* The table reports the performances of various 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 number of lags is selected by the Hannan-Quinn information criterion (HQIC). We consider three different methods of constructing confidence intervals. First, we can use $\alpha/2$ and $1-\alpha/2$ quantiles of the bootstrap estimates $\hat{s}_h^{R2,(b)}$ as discussed in Section 3.5, where $\alpha = 0.1$. The two other methods build symmetric confidence intervals with critical values from a standard normal distribution. We investigate both bootstrap standard error and asymptotic standard error. Confidence intervals for the other estimators are constructed similarly. For the asymptotic s.e., we use the asymptotic variance in Proposition 1 and follow the implementation details in Section A.

*Table E9. Simulation results for DGP 3.*

| | 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 | 0.95 | 3.80 | 5.55 | 6.57 | 7.15 | 7.43 |
| VAR(HQIC) | 0.94 | 2.46 | 2.58 | 2.61 | 2.62 | 2.62 |
| **Forecast Error Variance Ddecomposition** | | | | | | |
| True | 0.06 | 0.29 | 0.47 | 0.58 | 0.65 | 0.70 |
| Average estimate | | | | | | |
| R2 | 0.06 | 0.22 | 0.36 | 0.45 | 0.52 | 0.57 |
| VAR(HQIC) | 0.06 | 0.14 | 0.15 | 0.16 | 0.16 | 0.16 |
| Root mean squared error | | | | | | |
| R2 | 0.03 | 0.11 | 0.17 | 0.19 | 0.20 | 0.21 |
| VAR(HQIC) | 0.04 | 0.17 | 0.33 | 0.43 | 0.50 | 0.54 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, bootstrap [P5, P95] | 0.85 | 0.76 | 0.75 | 0.78 | 0.81 | 0.84 |
| R2, bootstrap s.e. | 0.83 | 0.70 | 0.68 | 0.71 | 0.74 | 0.79 |
| R2, asymptotic s.e. | 0.82 | 0.74 | 0.71 | 0.70 | 0.72 | 0.71 |
| VAR(HQIC), bootstrap [P5, P95] | 0.86 | 0.31 | 0.08 | 0.03 | 0.02 | 0.01 |
| **Forecast Error Variance Decomposition (bias-corrected, VAR(HQIC))** | | | | | | |
| True | 0.06 | 0.29 | 0.47 | 0.58 | 0.65 | 0.70 |
| Average estimate | | | | | | |
| R2 | 0.05 | 0.21 | 0.32 | 0.40 | 0.44 | 0.46 |
| VAR(HQIC) | 0.05 | 0.13 | 0.15 | 0.15 | 0.15 | 0.16 |
| Root mean squared error | | | | | | |
| R2 | 0.04 | 0.13 | 0.20 | 0.24 | 0.27 | 0.29 |
| VAR(HQIC) | 0.04 | 0.18 | 0.34 | 0.44 | 0.51 | 0.55 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, bootstrap [P5, P95] | 0.83 | 0.72 | 0.68 | 0.68 | 0.70 | 0.71 |
| R2, bootstrap s.e. | 0.80 | 0.67 | 0.58 | 0.61 | 0.63 | 0.63 |
| R2, asymptotic s.e. | 0.79 | 0.69 | 0.62 | 0.57 | 0.53 | 0.46 |
| VAR(HQIC), bootstrap [P5, P95] | 0.84 | 0.29 | 0.08 | 0.03 | 0.02 | 0.01 |

*Notes:* The table reports the performances of various 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. The number of lags is selected by the Hannan-Quinn information criterion (HQIC). We consider three different methods of constructing confidence intervals. First, we can use $\alpha/2$ and $1 - \alpha/2$ quantiles of the bootstrap estimates $\hat{s}_h^{R2,(b)}$ as discussed in Section 3.5, where $\alpha = 0.1$. The two other methods build symmetric confidence intervals with critical values from a standard normal distribution. We investigate both bootstrap standard error and asymptotic standard error. Confidence intervals for the other estimators are constructed similarly. For the asymptotic s.e., we use the asymptotic variance in Proposition 1 and follow the implementation details in Section A.

*Table E10. Simulation results for the Smets and Wouters (2007) model, output.*

| | Horizon $h$ | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 4 | 8 | 12 | 16 | 20 |
| **Impulse Response** | | | | | | |
| True | -0.18 | -0.31 | -0.18 | -0.07 | -0.02 | 0.00 |
| Local projections | -0.18 | -0.30 | -0.16 | -0.07 | -0.02 | 0.00 |
| VAR(HQIC) | -0.18 | -0.34 | -0.34 | -0.33 | -0.33 | -0.33 |
| **Forecast Error Variance Decomposition** | | | | | | |
| True | 0.05 | 0.08 | 0.06 | 0.05 | 0.04 | 0.03 |
| Average estimate | | | | | | |
| R2 | 0.05 | 0.09 | 0.10 | 0.12 | 0.15 | 0.19 |
| VAR(HQIC) | 0.05 | 0.09 | 0.10 | 0.09 | 0.09 | 0.09 |
| Root mean squared error | | | | | | |
| R2 | 0.03 | 0.07 | 0.10 | 0.13 | 0.16 | 0.20 |
| VAR(HQIC) | 0.03 | 0.06 | 0.08 | 0.09 | 0.09 | 0.09 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, bootstrap [P5, P95] | 0.88 | 0.91 | 0.90 | 0.81 | 0.73 | 0.67 |
| R2, bootstrap s.e. | 0.85 | 0.89 | 0.94 | 0.88 | 0.85 | 0.80 |
| R2, asymptotic s.e. | 0.84 | 0.80 | 0.88 | 0.90 | 0.86 | 0.81 |
| VAR(HQIC), bootstrap [P5, P95] | 0.90 | 0.86 | 0.86 | 0.89 | 0.87 | 0.87 |
| **Forecast Error Variance Decomposition (bias-corrected, VAR(HQIC))** | | | | | | |
| True | 0.05 | 0.08 | 0.06 | 0.05 | 0.04 | 0.03 |
| Average estimate | | | | | | |
| R2 | 0.05 | 0.06 | 0.05 | 0.04 | 0.04 | 0.04 |
| VAR(HQIC) | 0.05 | 0.08 | 0.08 | 0.08 | 0.08 | 0.07 |
| Root mean squared error | | | | | | |
| R2 | 0.03 | 0.07 | 0.09 | 0.10 | 0.12 | 0.14 |
| VAR(HQIC) | 0.03 | 0.06 | 0.08 | 0.09 | 0.09 | 0.09 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, bootstrap [P5, P95] | 0.85 | 0.87 | 0.95 | 0.93 | 0.91 | 0.90 |
| R2, bootstrap s.e. | 0.82 | 0.83 | 0.94 | 0.96 | 0.95 | 0.94 |
| R2, asymptotic s.e. | 0.82 | 0.72 | 0.70 | 0.74 | 0.77 | 0.78 |
| VAR(HQIC), bootstrap [P5, P95] | 0.87 | 0.83 | 0.83 | 0.84 | 0.87 | 0.88 |

*Notes:* The table reports the performance of various estimators introduced in Section 3 for the Smets and Wouters (2007) model. The dependent variable is output, and we consider monetary policy shocks as inputs. The information set further includes price inflation and the monetary policy rate. The sample size is $T = 160$, and the number of simulations is 2,000. We consider three different methods of constructing confidence intervals. First, we can use $\alpha/2$ and $1-\alpha/2$ quantiles of the bootstrap estimates $\hat{s}_h^{R2,(b)}$ as discussed in Section 3.5, where $\alpha = 0.1$. The two other methods build symmetric confidence intervals with critical values from a standard normal distribution. We investigate both bootstrap standard error and asymptotic standard error. Confidence intervals for the other estimators are constructed similarly. For the asymptotic s.e., we use the asymptotic variance in Proposition 1 and follow the implementation details in Section A.

*Table E11. Simulation results for the Smets and Wouters (2007) model, price inflation.*

| | Horizon $h$ | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 4 | 8 | 12 | 16 | 20 |
| **Impulse Response** | | | | | | |
| True | -0.04 | -0.04 | -0.02 | -0.01 | 0.00 | 0.00 |
| Local projections | -0.04 | -0.04 | -0.02 | 0.00 | 0.00 | 0.00 |
| VAR(HQIC) | -0.04 | -0.02 | -0.01 | 0.00 | 0.00 | 0.00 |
| **Forecast Error Variance Decomposition** | | | | | | |
| True | 0.02 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 |
| Average estimate | | | | | | |
| R2 | 0.02 | 0.07 | 0.10 | 0.13 | 0.16 | 0.19 |
| VAR(HQIC) | 0.02 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 |
| Root mean squared error | | | | | | |
| R2 | 0.02 | 0.07 | 0.09 | 0.11 | 0.14 | 0.17 |
| VAR(HQIC) | 0.02 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, bootstrap [P5, P95] | 0.91 | 0.80 | 0.73 | 0.64 | 0.54 | 0.42 |
| R2, bootstrap s.e. | 0.88 | 0.89 | 0.82 | 0.74 | 0.65 | 0.53 |
| R2, asymptotic s.e. | 0.82 | 0.86 | 0.89 | 0.86 | 0.81 | 0.71 |
| VAR(HQIC), bootstrap [P5, P95] | 0.94 | 0.92 | 0.93 | 0.93 | 0.94 | 0.94 |
| **Forecast Error Variance Decomposition (bias-corrected, VAR(HQIC))** | | | | | | |
| True | 0.02 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 |
| Average estimate | | | | | | |
| R2 | 0.02 | 0.05 | 0.05 | 0.05 | 0.05 | 0.06 |
| VAR(HQIC) | 0.02 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| Root mean squared error | | | | | | |
| R2 | 0.02 | 0.06 | 0.08 | 0.09 | 0.09 | 0.10 |
| VAR(HQIC) | 0.02 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 |
| Coverage (90 % level, asymptotic) | | | | | | |
| R2, bootstrap [P5, P95] | 0.85 | 0.80 | 0.83 | 0.85 | 0.85 | 0.85 |
| R2, bootstrap s.e. | 0.81 | 0.79 | 0.82 | 0.84 | 0.85 | 0.85 |
| R2, asymptotic s.e. | 0.77 | 0.74 | 0.76 | 0.80 | 0.83 | 0.84 |
| VAR(HQIC), bootstrap [P5, P95] | 0.87 | 0.82 | 0.80 | 0.81 | 0.81 | 0.82 |

*Notes:* The table reports the performances of various estimators introduced in Section 3 for the Smets and Wouters (2007) model. The dependent variable is price inflation, and we consider monetary policy shocks as inputs. The information set further includes price inflation and the monetary policy rate. The sample size is $T = 160$, and the number of simulations is 2,000 We consider three different methods of constructing confidence intervals. First, we can use $\alpha/2$ and $1-\alpha/2$ quantiles of the bootstrap estimates $\hat{s}_h^{R2,(b)}$ as discussed in Section 3.5, where $\alpha = 0.1$. The two other methods build symmetric confidence intervals with critical values from a standard normal distribution. We investigate both bootstrap standard error and asymptotic standard error. Confidence intervals for the other estimators are constructed similarly. For the asymptotic s.e., we use the asymptotic variance in Proposition 1 and follow the implementation details in Section A.

## Appendix F. Supplementary Figures for Bivariate Simulations

This section provides additional results for bivariate DGPs in Section 4.1. We include results for the LP-A and LP-B estimators. We further consider large-sample performances of the estimators.

For bias-correction, we apply our estimators to bootstrap samples and obtain $\hat{s}_h^{R2,(b)}$, $\hat{s}_h^{LPA,(b)}$, $\hat{s}_h^{LPB,(b)}$, and $\hat{s}_h^{VAR,(b)}$ for $b = 1, \dots, 2000$. The biases for each estimator are calculated by

$$\frac{1}{B}\sum_{b=1}^{B} \hat{s}_h^{m,(b)} - s_h^*$$

for $m = R2, LPA, LPB$, and VAR, where $s_h^*$ denotes the true contribution of $z$ to the forecast error variance of $y$ at horizon $h$ for the DGP used for bootstrap. For other details, see Section 3.4 and Section 4.1.

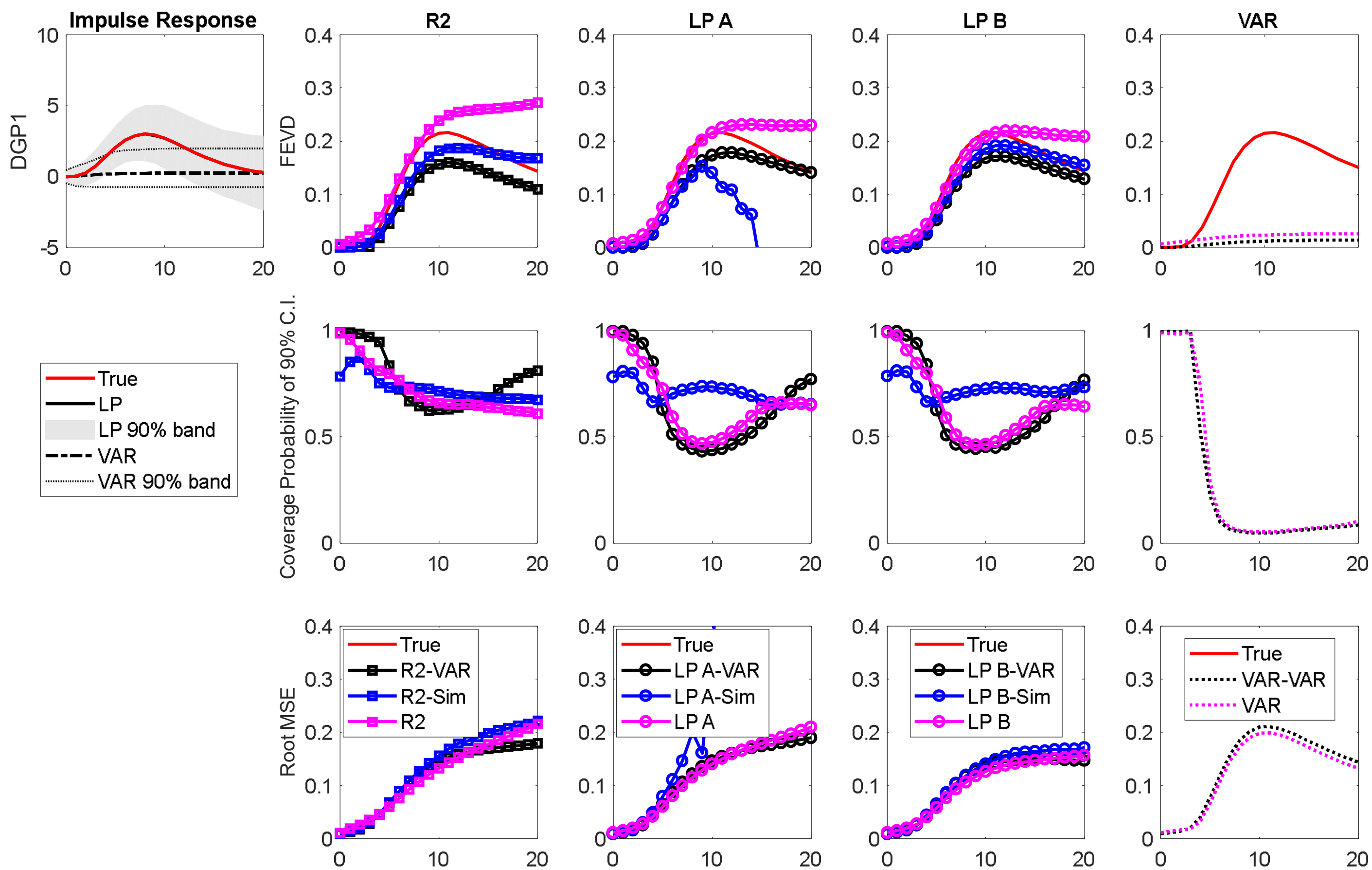
How to read the figure legend in the following pages:

1. Impulse responses
   - The 90% bands are based on the 5th and 95th percentiles of the estimates across 2,000 replications.

2. FEVD, Coverage probability, and Root MSE

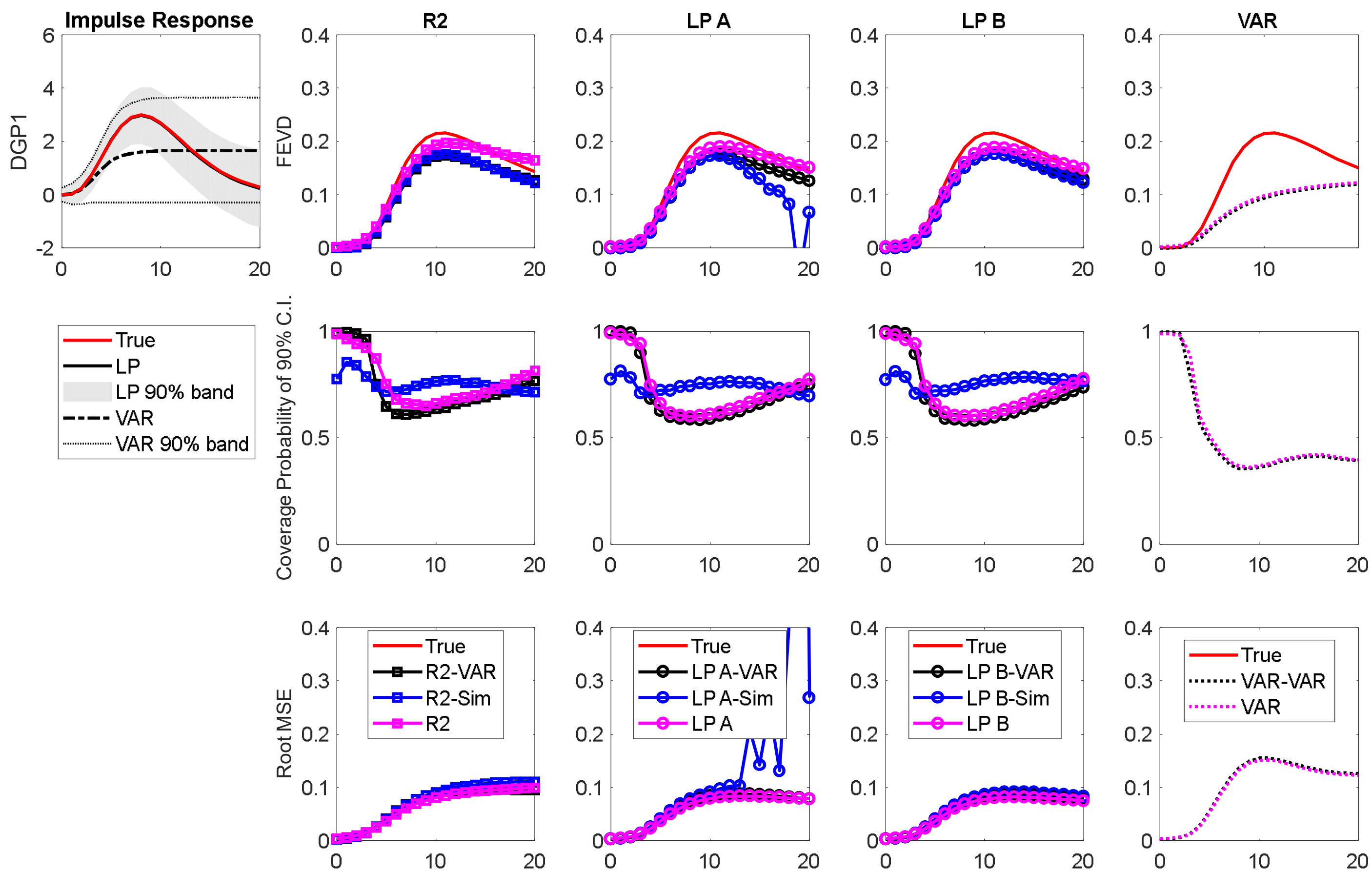
- 'R2-VAR' means the bias-corrected R2 estimator with a VAR-bootstrap. For the coverage rates, we construct the confidence intervals $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2,BC},\ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2,BC}\right]$ as discussed in Section 3.5, where $\alpha = 0.1$.
- 'R2-Sim' uses the method in Appendix A, which does not rely on VAR-based bootstraps. The coverage probability is based on the asymptotic standard error with pre-whitening as discussed in Appendix A. That is, we consider a symmetric confidence interval.
- 'R2' denotes for the estimator without any finite-sample correction. Its coverage rate is calculated in a way similar to 'R2-VAR': $\left[\hat{q}_{h,\alpha/2}^{R2} + \hat{s}_h^{R2},\ \hat{q}_{h,1-\alpha/2}^{R2} + \hat{s}_h^{R2}\right]$. Note that the interval is centered around the estimate without bias-correction.
- 'LP A/B-VAR', 'LP A/B-Sim', 'LP A/B', 'VAR-VAR', and 'VAR' are defined similarly.

DGP 1, T = 160.

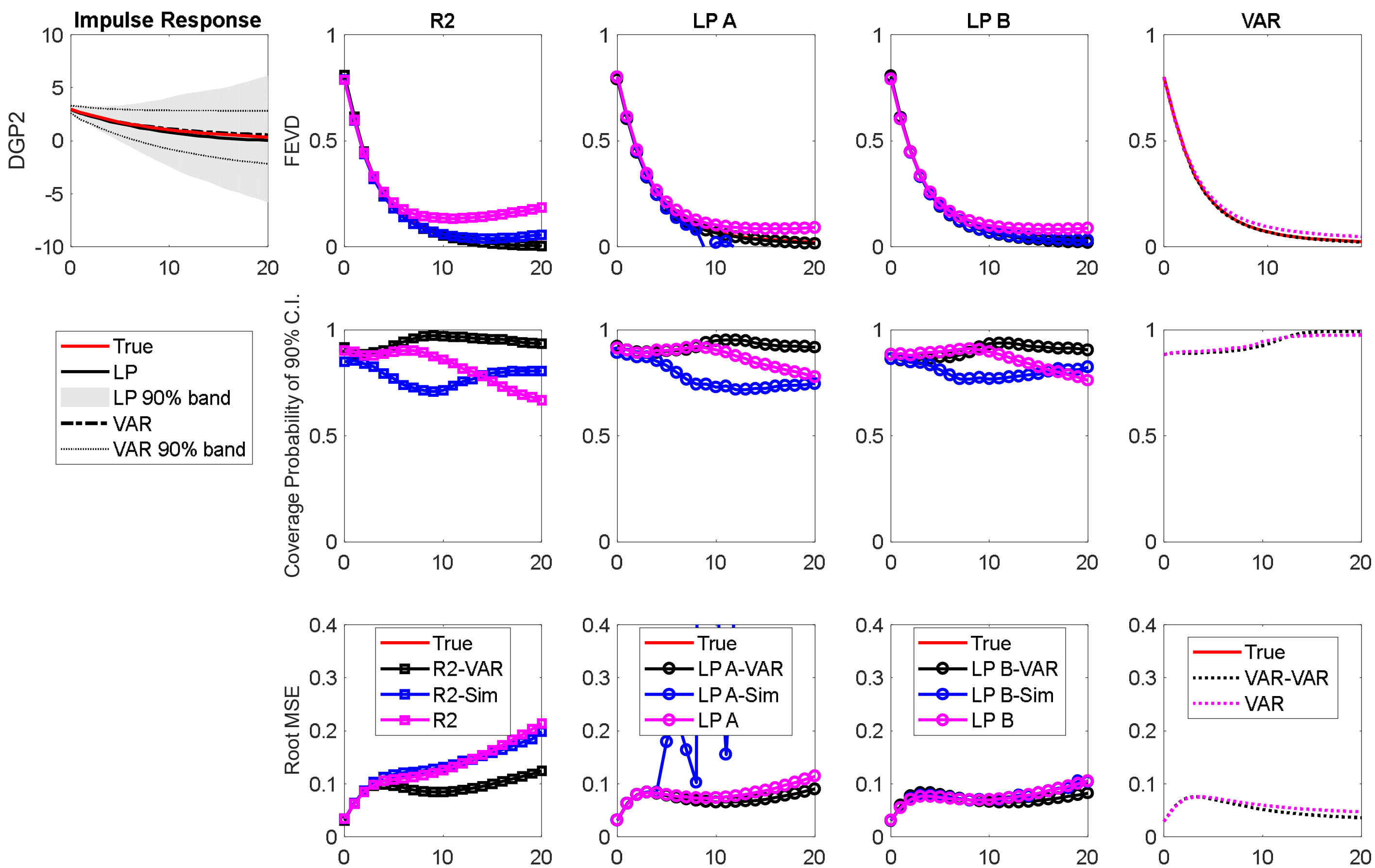


Impulse Response
R2
LP A
LP B
VAR
DGP1
FEVD
Coverage Probability of 90% C.I.
Root MSE
True
LP
LP 90% band
VAR
VAR 90% band
True
R2-VAR
R2-Sim
R2
True
LP A-VAR
LP A-Sim
LP A
True
LP B-VAR
LP B-Sim
LP B
True
VAR-VAR
VAR

DGP1, T = 500.

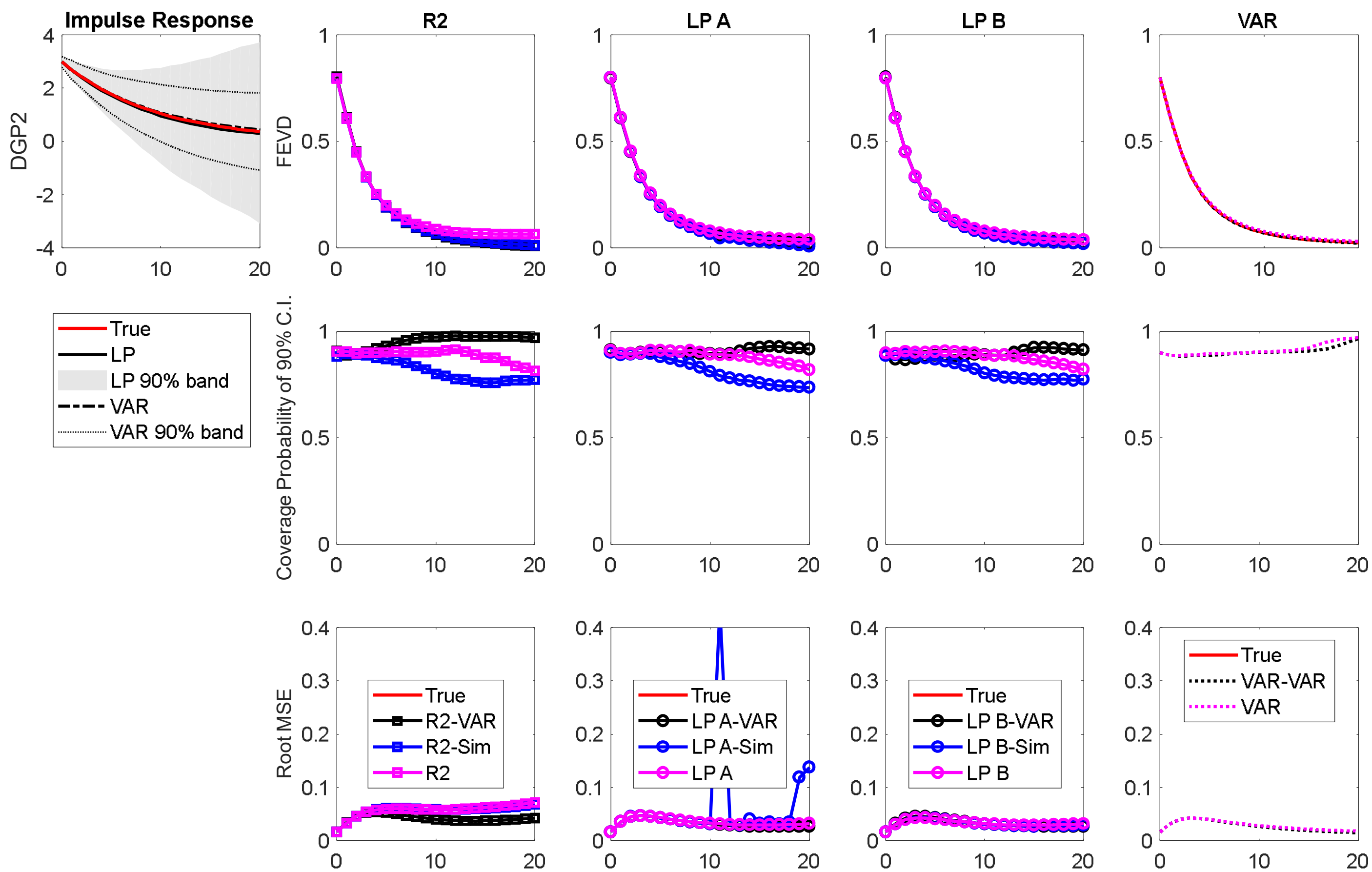

Impulse Response
R2
LP A
LP B
VAR
DGP1
FEVD
Coverage Probability of 90% C.I.
Root MSE
True
LP
LP 90% band
VAR
VAR 90% band
True
R2-VAR
R2-Sim
R2
True
LP A-VAR
LP A-Sim
LP A
True
LP B-VAR
LP B-Sim
LP B
True
VAR-VAR
VAR

DGP2, T = 160.

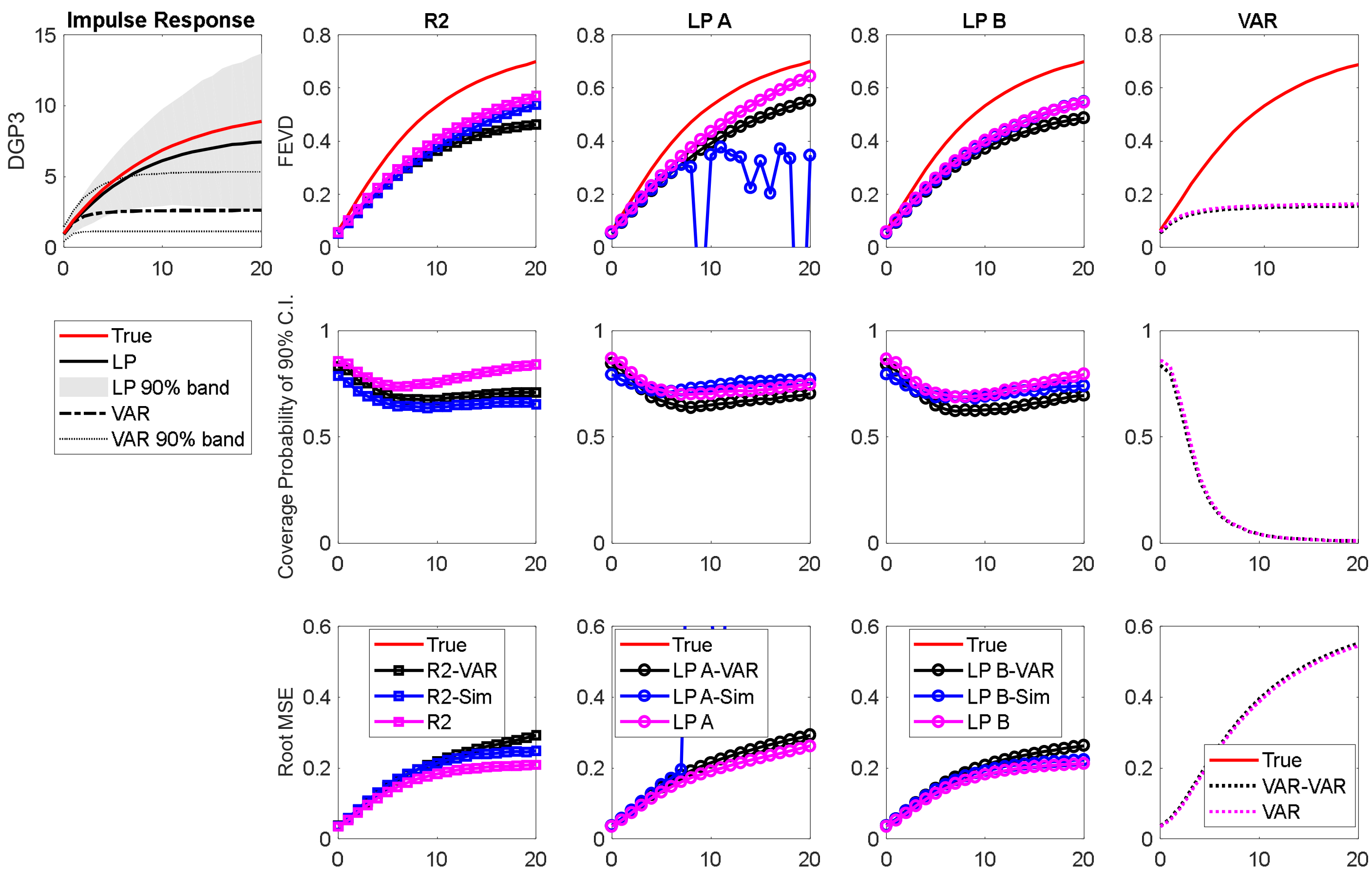


Impulse Response
R2
LP A
LP B
VAR
DGP2
FEVD
Coverage Probability of 90% C.I.
Root MSE
True
LP
LP 90% band
VAR
VAR 90% band
True
R2-VAR
R2-Sim
R2
True
LP A-VAR
LP A-Sim
LP A
True
LP B-VAR
LP B-Sim
LP B
True
VAR-VAR
VAR

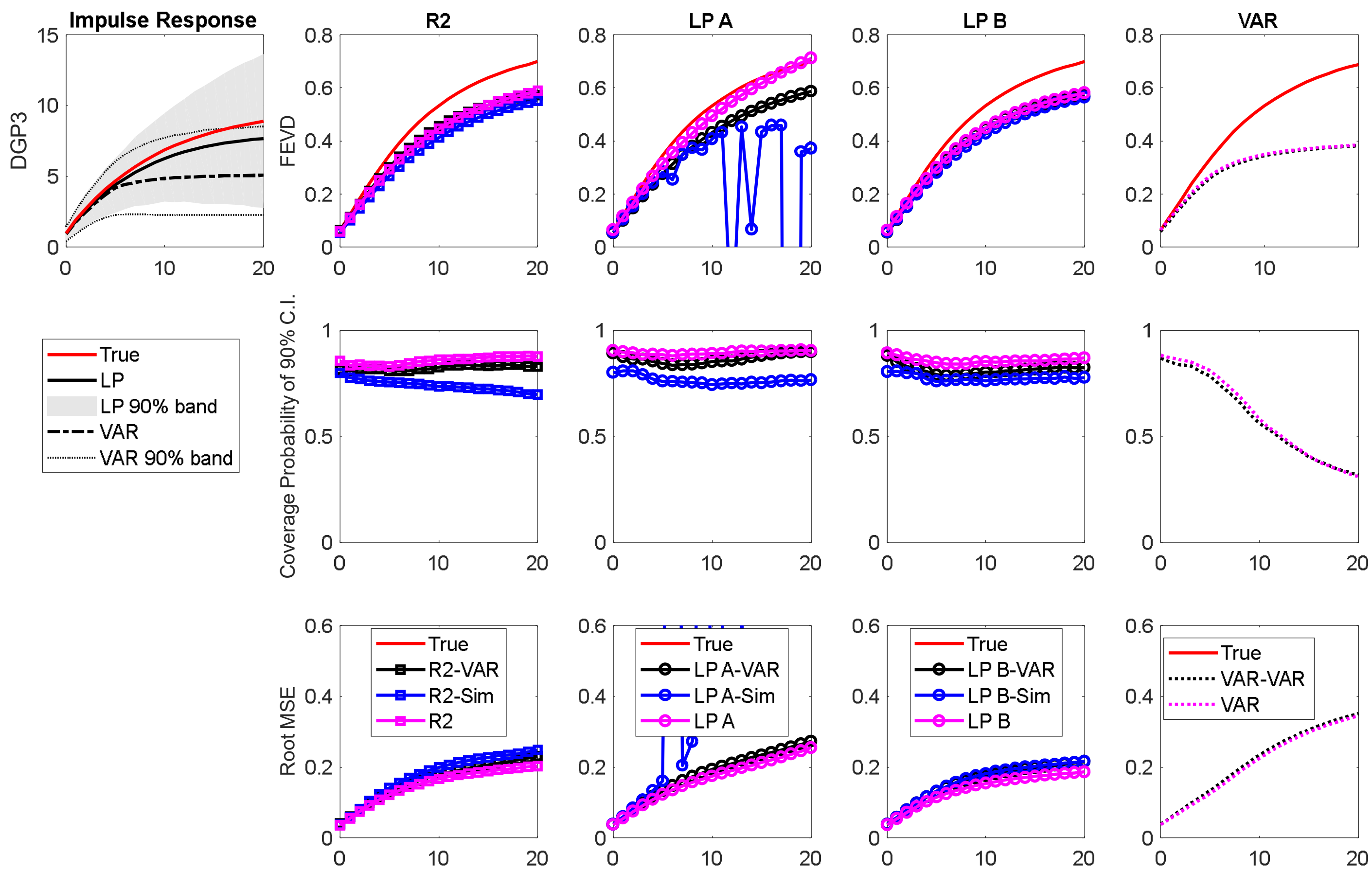
DGP2, T= 500.


Impulse Response
R2
LP A
LP B
VAR
DGP2
FEVD
Coverage Probability of 90% C.I.
Root MSE
True
LP
LP 90% band
VAR
VAR 90% band
True
R2-VAR
R2-Sim
R2
True
LP A-VAR
LP A-Sim
LP A
True
LP B-VAR
LP B-Sim
LP B
True
VAR-VAR
VAR

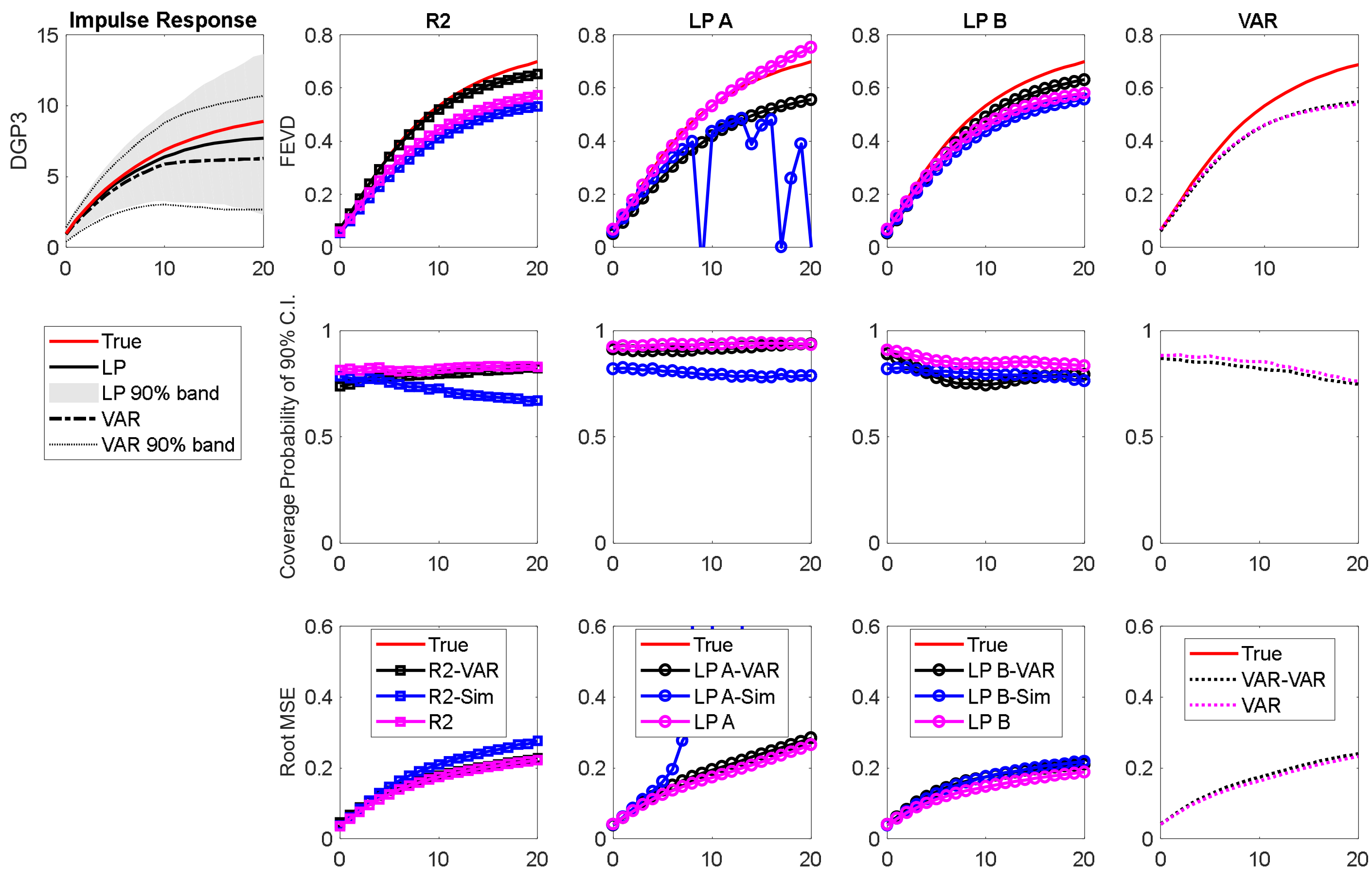
DGP3, T = 160, VAR(HQIC).



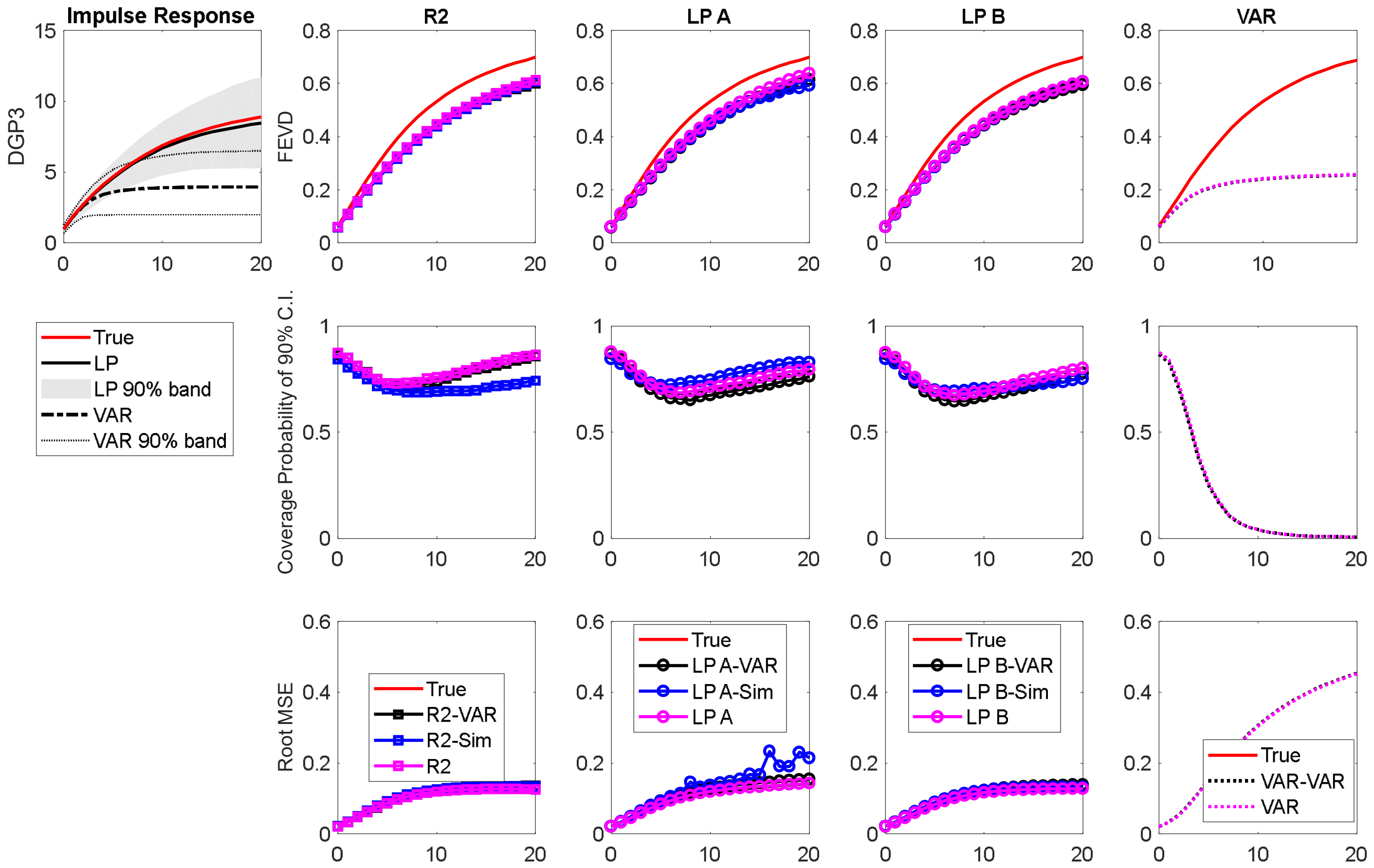
DGP3, T = 160, VAR(5).



DGP3, T = 160, VAR(10).



DGP3, T = 500, VAR(HQIC).



## Appendix G. Supplementary Figures for the Smets and Wouters model

We presents results for the LP-A and LP-B estimators in this section. We further consider large-sample performances of the estimators. For details of the simulations, see Section 4.2.

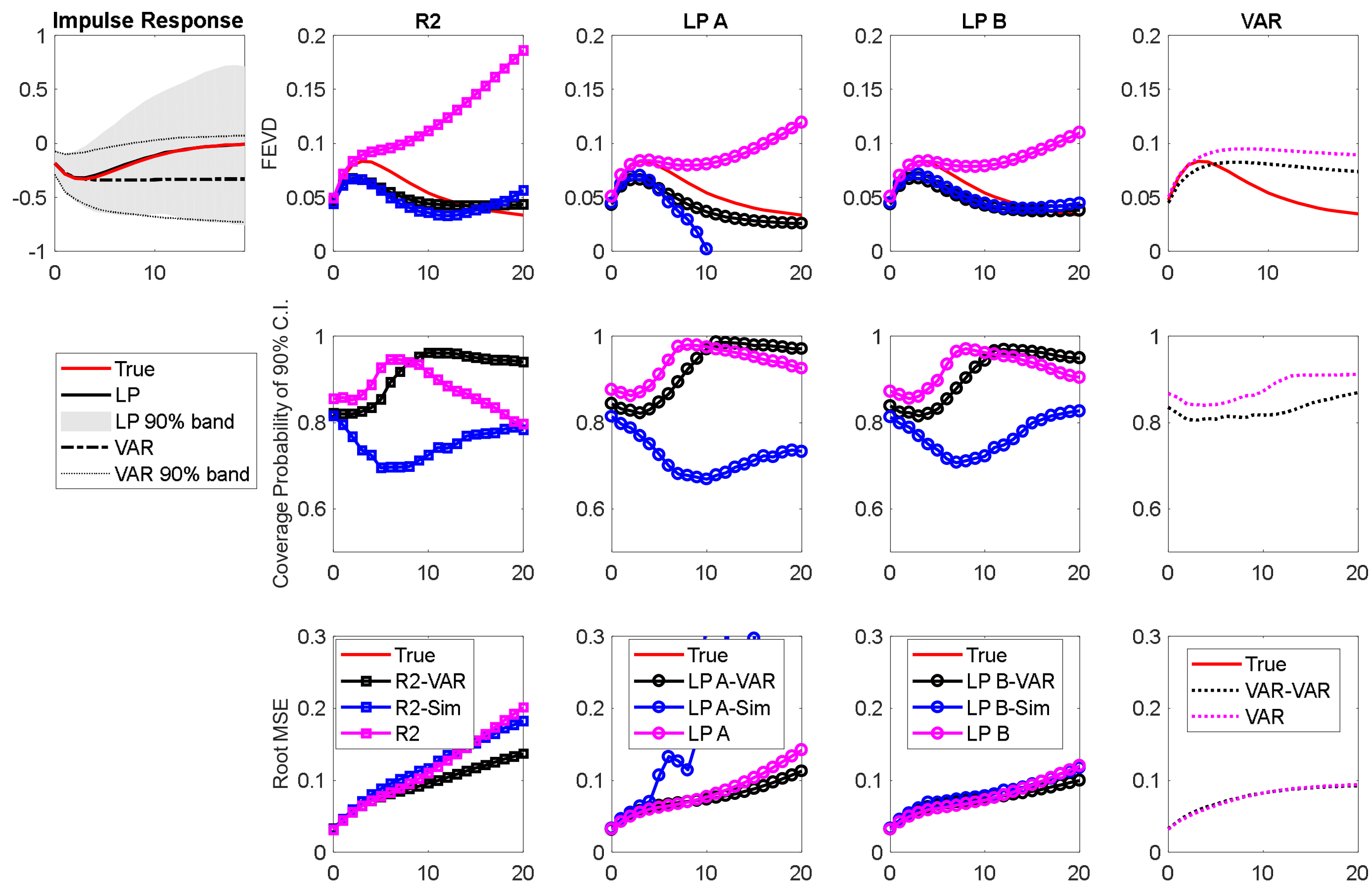
How to read the figure legend in the following pages:

1. Impulse responses
   - The 90% bands are based on the 5th and 95th percentiles of the estimates across 2,000 replications.

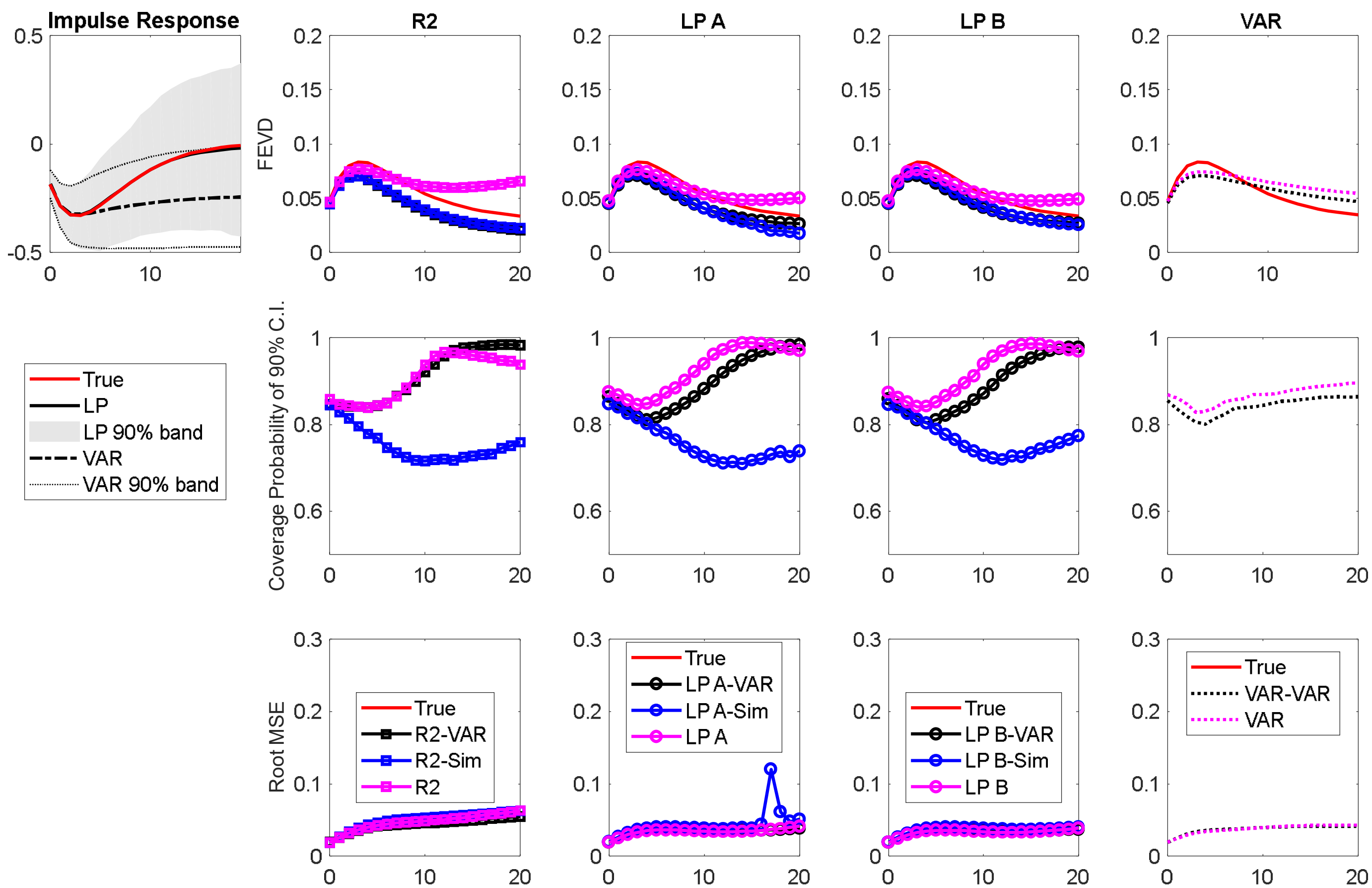
2. FEVD, Coverage probability, and Root MSE

- 'R2-VAR' means the bias-corrected R2 estimator with a VAR-bootstrap. For the coverage rates, we construct the confidence interval $\left[\hat{q}^{R2}_{h,\alpha/2} + \hat{s}^{R2,BC}_{h},\ \hat{q}^{R2}_{h,1-\alpha/2} + \hat{s}^{R2,BC}_{h}\right]$ as discussed in Section 3.5, where $\alpha = 0.1$.
- 'R2-Sim' uses the method in Appendix A, which does not rely on VAR-based bootstraps. The coverage probability is based on the asymptotic standard error with pre-whitening as discussed in Appendix A. That is, we consider a symmetric confidence interval.
- 'R2' denotes for the estimator without any finite-sample correction. Its coverage rate is calculated in a way similar to 'R2-VAR': $\left[\hat{q}^{R2}_{h,\alpha/2} + \hat{s}^{R2}_{h},\ \hat{q}^{R2}_{h,1-\alpha/2} + \hat{s}^{R2}_{h}\right]$. Note that the interval is centered around the estimate without bias-correction.
- 'LP A/B-VAR', 'LP A/B-Sim', 'LP A/B', 'VAR-VAR', and 'VAR' are defined similarly.

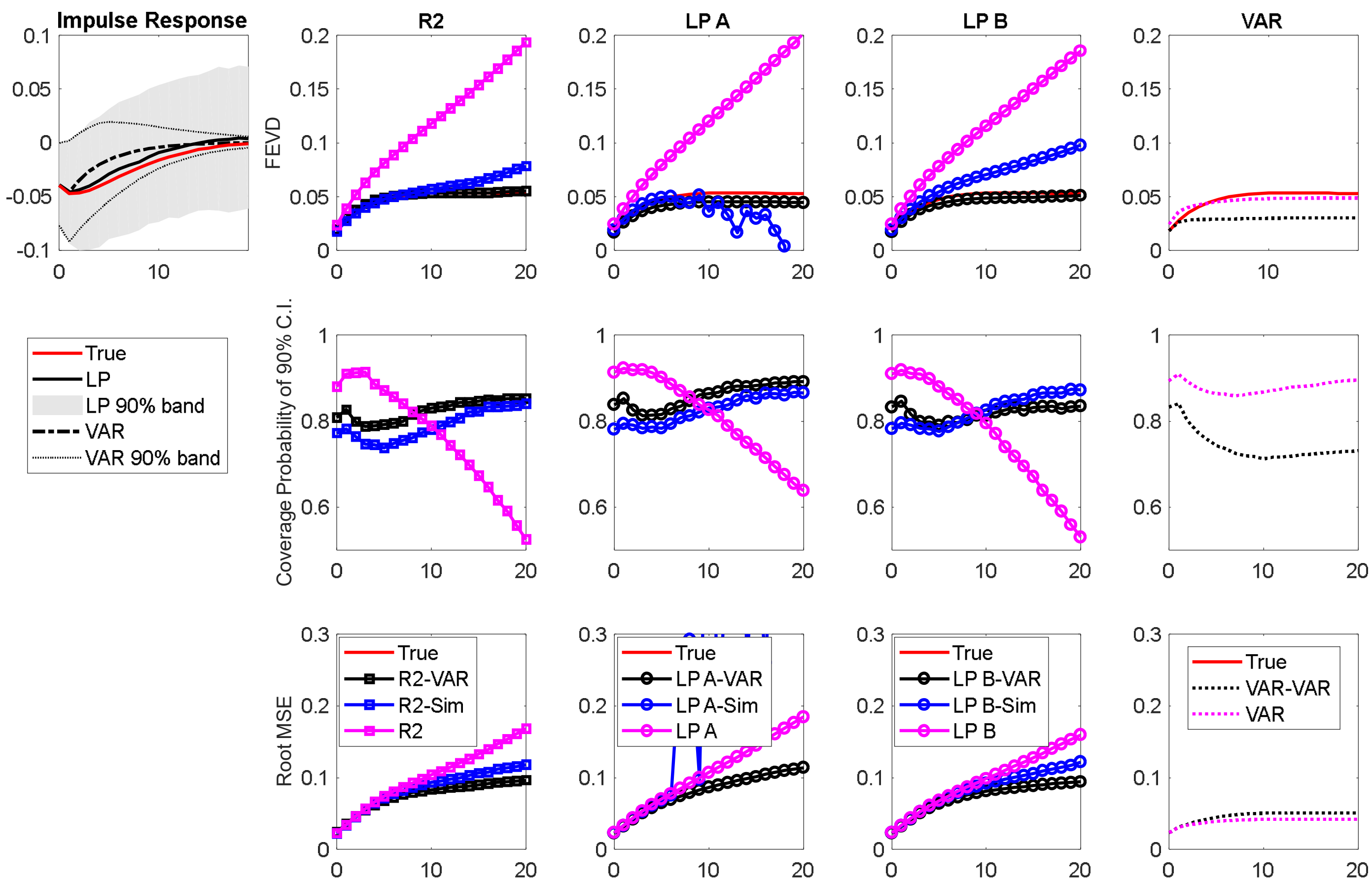
Real GDP and monetary policy shock, T = 160



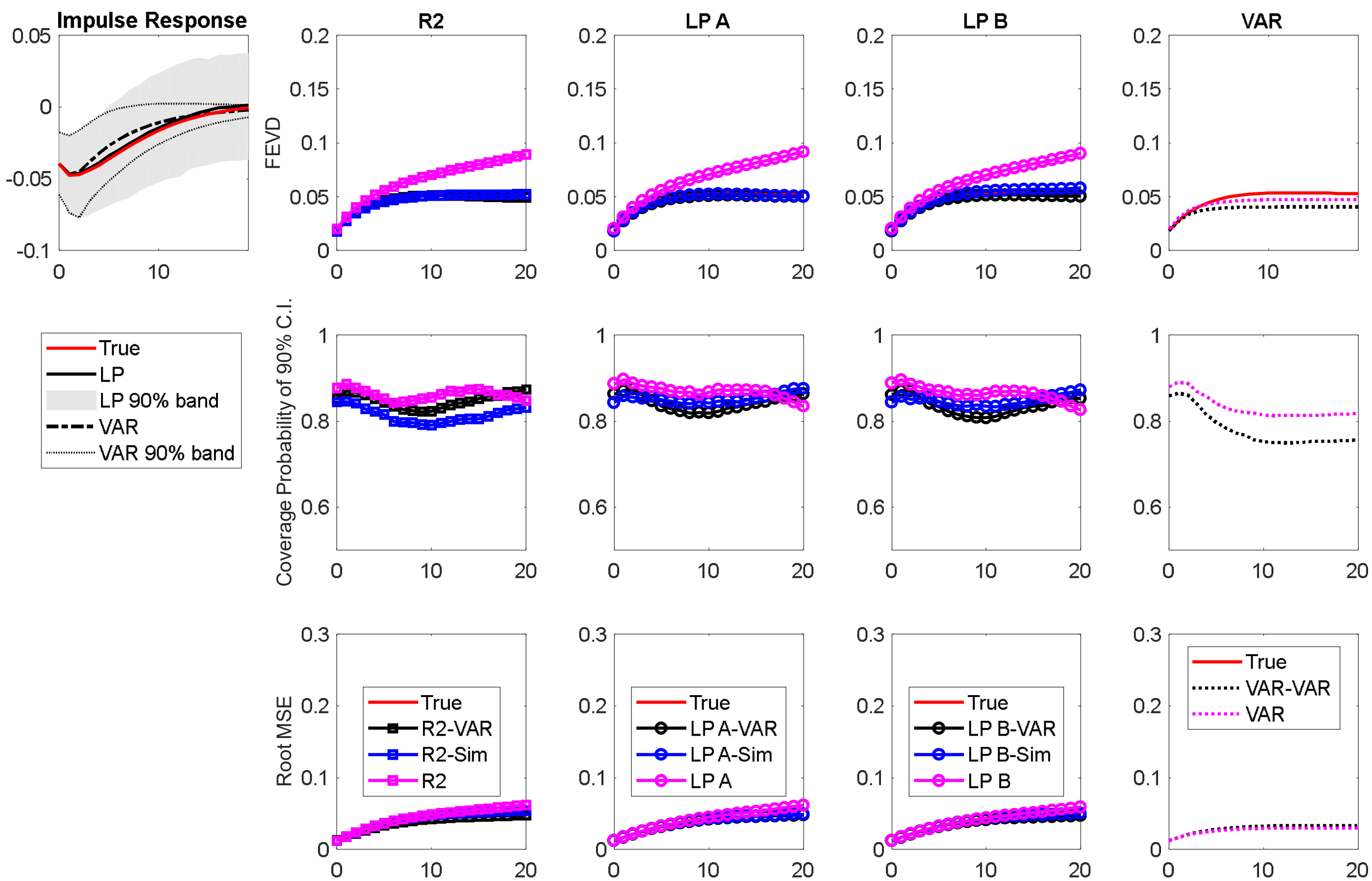
Real GDP and monetary policy shock, T = 500.



Price inflation and monetary policy shock, T = 160



Price inflation and monetary policy shock, T = 500.



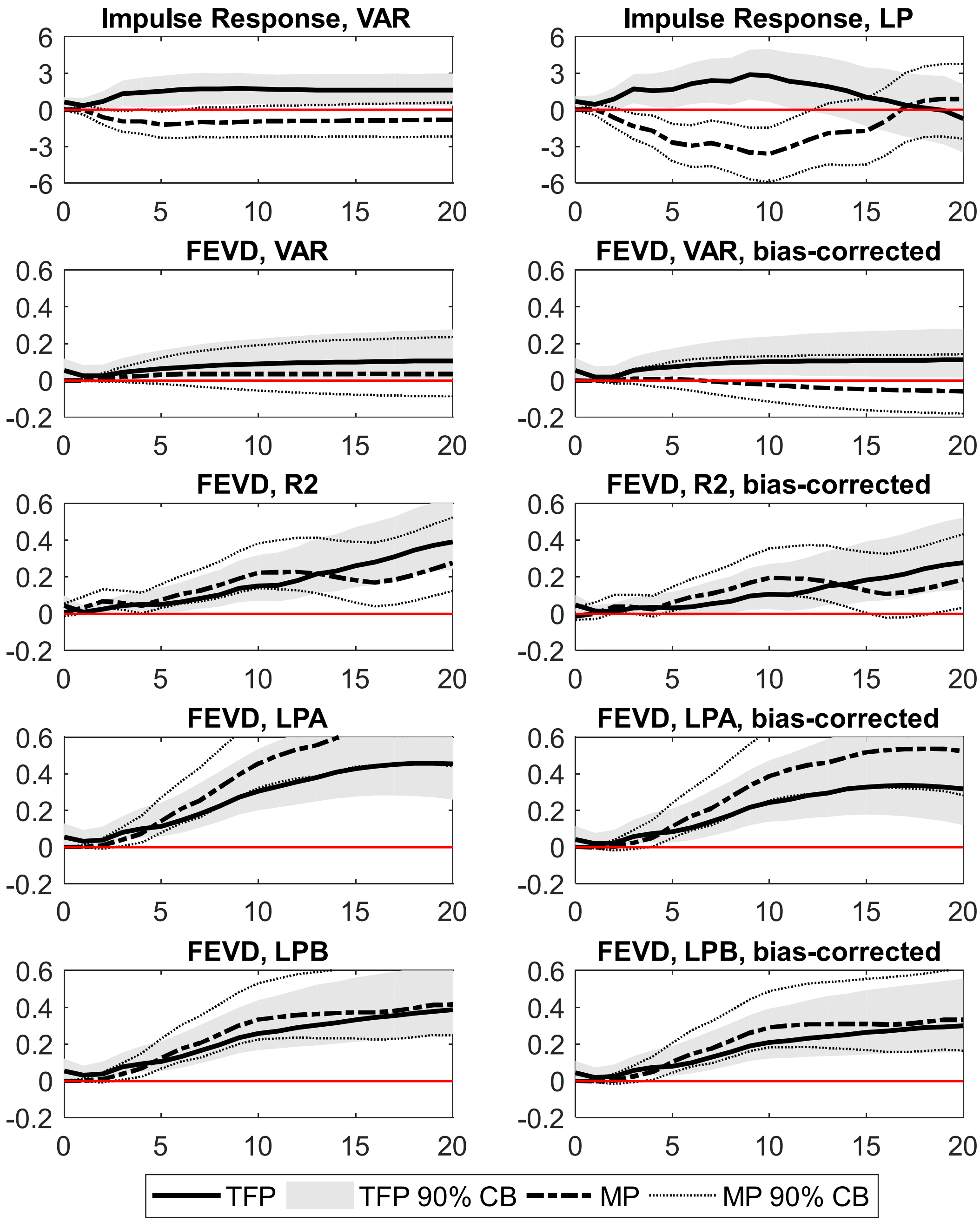
# Appendix H. Applications

This section covers additional results on applications. Appendix H1 includes results for the total factor productivity and monetary policy shocks when we employ either LP-A or LP-B estimator. The figures in Appendix H2 depicts the contribution of the military news shock series constructed by Ramey and Zubairy (2018) on output and price inflation.

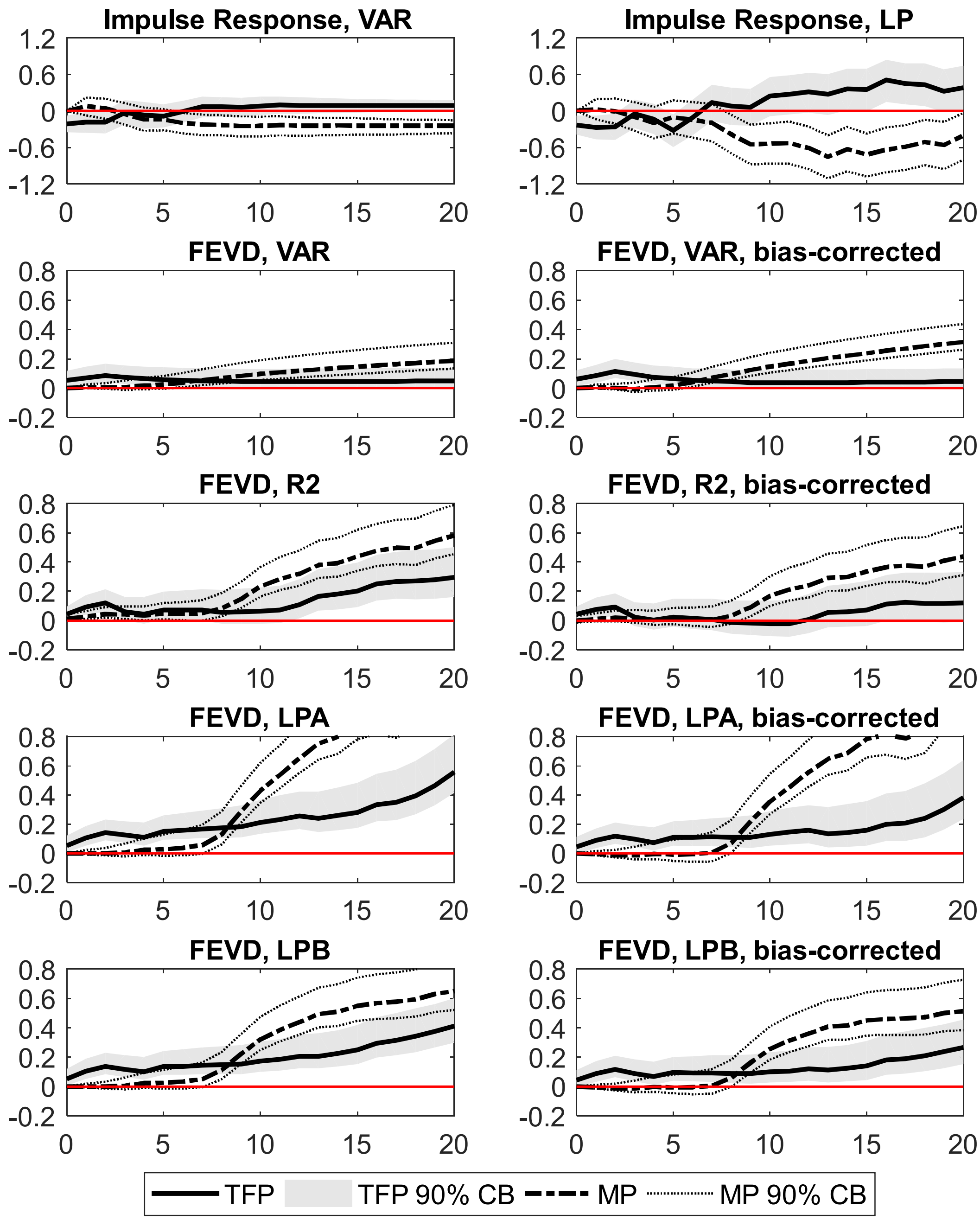
## Appendix H1. Supplementary figures to Figures 3 and 4

Below we show figures similar to Figures 3 and 4. While other things are the same, we add the estimates and 90% confidence intervals for the LP-A and LP-B estimators discussed in Appendix B. Biases are corrected with VAR-bootstraps, and confidence intervals are based on the P5 and P95 of the bootstrap estimates as discussed in Section 3.5.

1969:Q1-2008:Q4, Real GDP.



1969:Q1-2008:Q4, Inflation.

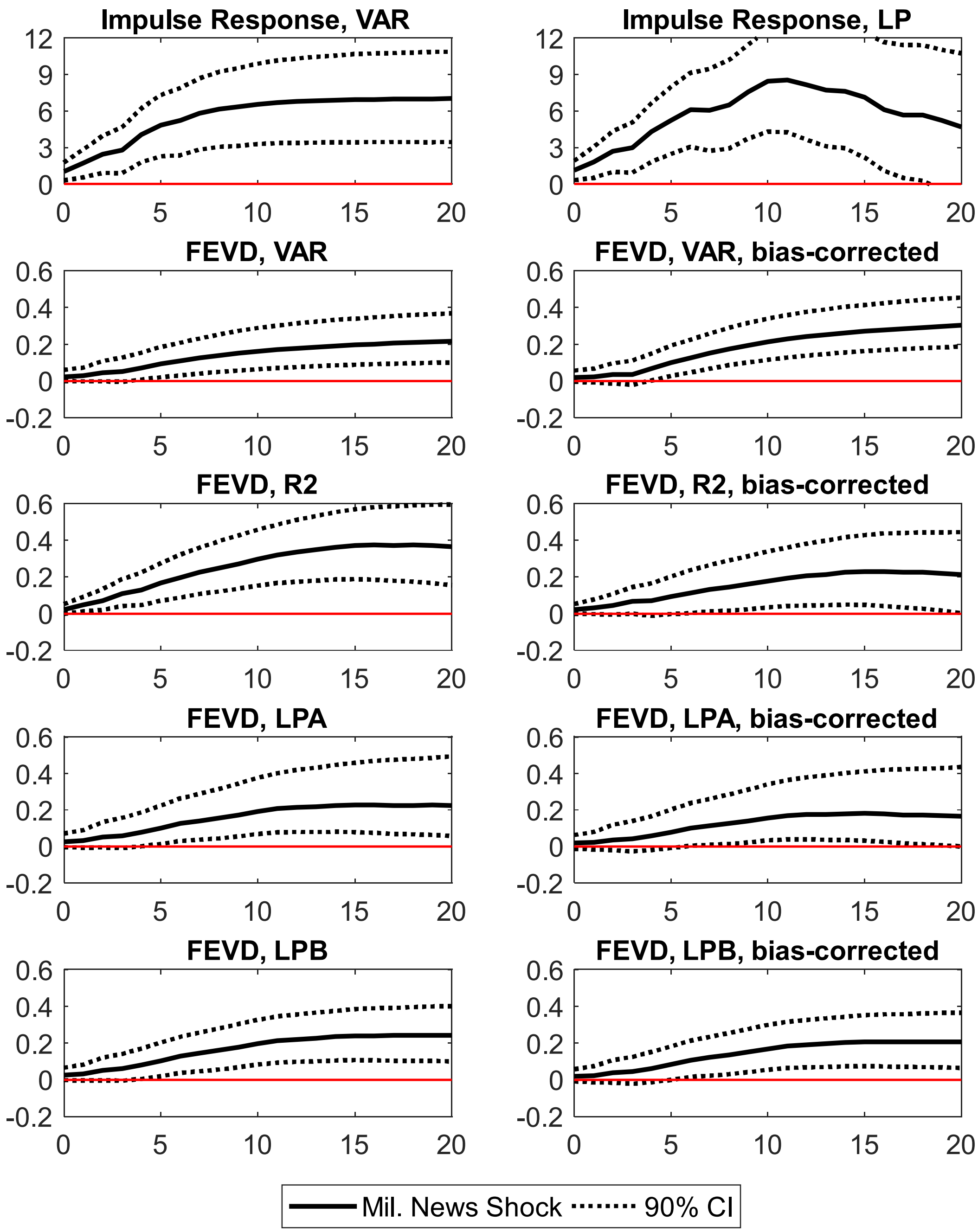


## Appendix H2. Military news shocks

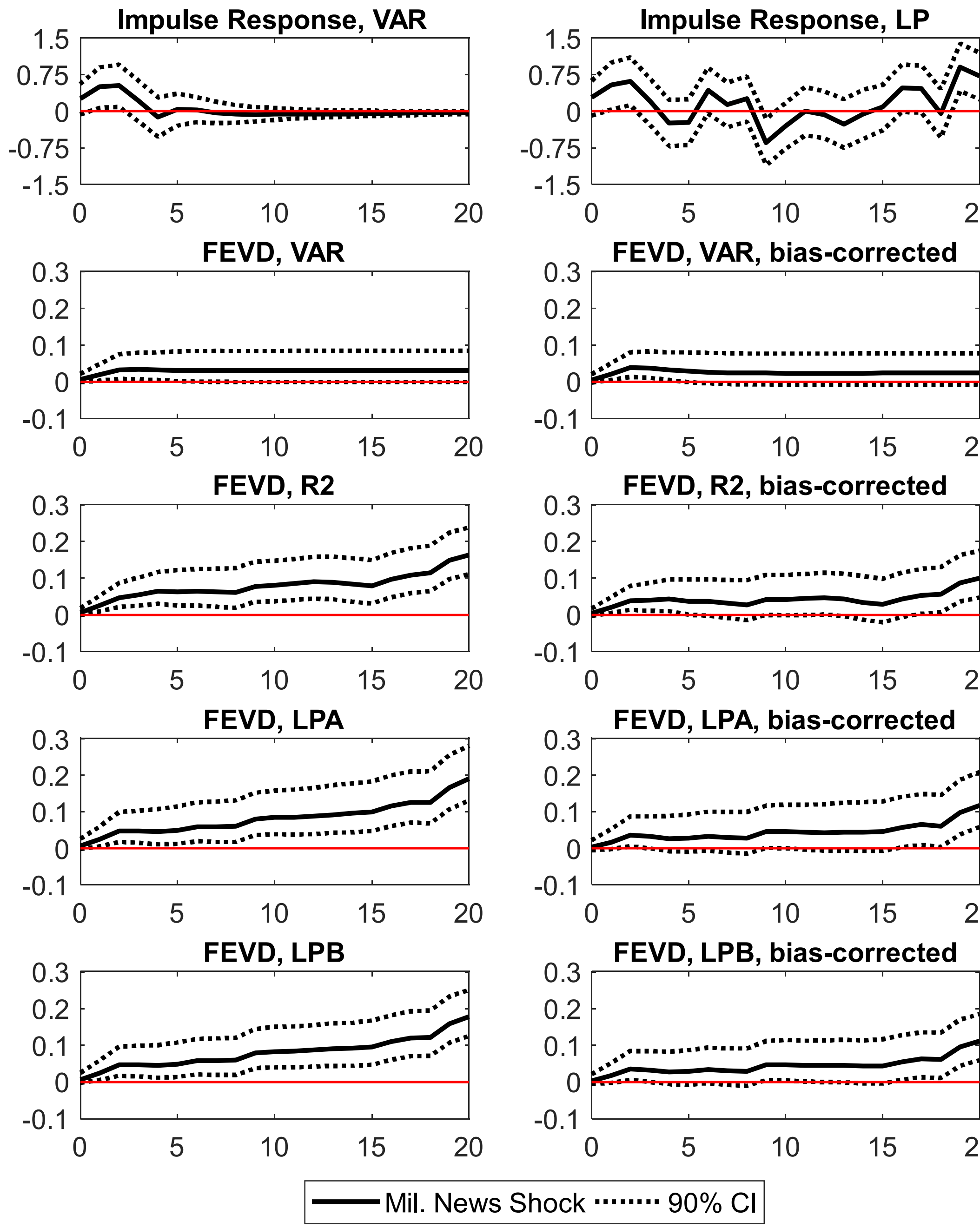
We study the effects and contribution of military news shocks on real GDP and price inflation in the U.S. economy. Similar to Section 5, our information set includes the military news shock constructed by Ramey and Zubairy (2018), output growth rate, inflation, and 3-month Treasury bill rate in a secondary market. This ordering is also used for the VAR analysis. The sample period is from 1920Q1 to 2015Q4. During the sample period, one standard deviation military news shock amounts to 6.4%. This implies that the sum of the present discounted values of future increases in military expenditure corresponds to 6.4% of the current trend GDP.

Military news shocks have both statistically and economically significant effects on real GDP. The estimated FEVD is slightly less than 20 percent at the 5-year horizon. However, its contribution to inflation is negligible. The estimated FEVD is around 5 percent in the medium-run.

1920:Q1-2015:Q4, Real GDP.



1920:Q1-2015:Q4, Inflation.



# References


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

Hamilton, James D., 1994. *Time series analysis*. Princeton: Princeton university press.

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.

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.

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 H., and Mark W. Watson, 2011. *Introduction to Econometrics*, 3/E. Boston: Addison Wesley.

Whitney K. Newey and Kenneth D. West, 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55(3): 703-708.