Predictive Accuracy of Impulse Responses Estimated Using Local Projections and Vector Autoregressions*

Zacharias Psaradakis[†]
Birkbeck, University of London, U.K.
`z.psaradakis@bbk.ac.uk`

Martin Sola
Universidad Torcuato di Tella, Argentina
`msola@utdt.edu`

Nicola Spagnolo
Brunel University London, U.K.
`nicola.spagnolo@brunel.ac.uk`

Patricio Yunis
University of Pennsylvania, U.S.A.
`pyunis@sas.upenn.edu`

July 22, 2025

Abstract

We examine the finite-sample accuracy of impulse responses obtained using local projections (LP) and vector autoregressive (VAR) models. In view of the fact that impulse responses are differences between multistep predictors, we propose to assess the relative performance of impulse-response estimators using tests for equal predictive accuracy. In our Monte Carlo experiments, LP-based and VAR-based estimators are found to be equally accurate in large samples under a mean-squared-error risk function. VAR-based estimators tend to have an advantage over LP-based estimators in small and moderately sized samples, particularly at long horizons.

Keywords: Local projections; Predictive accuracy; VAR models.
JEL Classification: C32; C53.

---

* The authors thank Martin Gonzalez-Rozada, Constantino Hevia, Fabio Spagnolo and two anonymous referees for helpful comments and/or discussions. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

[†] Corresponding author. E-mail: `z.psaradakis@bbk.ac.uk`; Address: Birkbeck Business School, Birkbeck, University of London, Malet Street, London WC1E 7HX, United Kingdom.

1 Introduction

Estimating impulse responses by the method of local projections (LP), as proposed by Jordà (2005), has become an increasingly popular alternative to the conventional approach based on vector autoregressive (VAR) models. For a wide range of empirically relevant data-generating processes (DGPs), the two approaches are known to agree approximately, both in population and in large samples, especially at impulse-response horizons of short length (see Plagborg-Møller and Wolf (2021)). Nevertheless, there is ongoing debate as to whether LP-based estimators fare better or worse than VAR-based estimators, in terms of bias and efficiency, in samples of relatively small or intermediate sizes like those that are often used in empirical work (see Brugnolini (2018), Herbst and Johannsen (2021) and Li et al. (2024), inter alia).

This paper contributes to the expanding literature on the relative merits of LP-based and VAR-based approaches to impulse-response estimation by considering finite-sample accuracy from a new perspective. Specifically, motivated by the observation that impulse responses are forecasts conditional on specific innovations, we propose to assess the relative performance of LP-based and VAR-based estimators by means of tests for equal predictive accuracy (cf. Diebold and Mariano (1995), Harvey et al. (1997)). This way of evaluating the accuracy of the two approaches, based on their relative performance with respect to a mean-squared-error risk function, allows us to draw conclusions about the statistical significance, or otherwise, of the differences between the estimation errors associated with the two competing methods of impulse-response estimation. The approach is, to our knowledge, novel and a useful addition to the more familiar bias/efficiency comparisons that are found in the literature.

By means of Monte Carlo experiments, we demonstrate that, under a finite-order VAR DGP, there is not much to choose between the two estimators of impulse responses in large samples in terms of predictive ability (evaluated with a quadratic loss function). By contrast, in small and moderately sized samples, VAR-based estimators outperform LP-based ones, particularly at long horizons where LP-based estimators of impulse responses tend to be significantly biased.

The organization of the paper is as follows. Section 2 introduces the impulse-response estimators of interest. Section 3 presents and discusses the results of a simulation study that compares LP-based and VAR-based estimators of impulse responses in terms of finite-sample bias and predictive accuracy. Section 4 summarizes and concludes.

2 VAR-Based and LP-Based Impulse-Response Estimators

Consider the VAR model of order p for the n-dimensional time series $y_t = (Y_{1,t}, \ldots, Y_{n,t})'$ given by

$$y_t = \mu + \sum_{k=1}^{p} \Phi_k y_{t-k} + \varepsilon_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where μ is an n-dimensional vector of constants, $\Phi_1, \ldots, \Phi_p$ are $n \times n$ parameter matrices, and $\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{n,t})'$ are n-dimensional random vectors such that $\mathsf{E}[\varepsilon_t] = 0$, $\mathsf{E}[\varepsilon_t \varepsilon_r'] = 0$ for $t \neq r$ and $\mathsf{E}[\varepsilon_t \varepsilon_t'] = \Sigma$, with Σ being a positive definite matrix. The model is assumed to be stable, in the sense that the $np \times np$ matrix

$$F = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{bmatrix}$$

has spectral radius strictly less than 1, I_n being the identity matrix of order n.

For any non-negative integer τ, the impulse response of y_t at horizon τ to a shock of size $\delta = (\delta_1, \ldots, \delta_n)'$ at date t may be defined in general as the difference between two conditional expectations of $y_{t+\tau}$, one which is conditional on the information set $\mathbb{Y}_{t-1} = \{y_{t-1}, y_{t-2}, \ldots\}$ and on a shock of size δ to ε_t and one which is conditional on $\mathbb{Y}_{t-1}$ alone (Koop et al. (1996)). More formally, the impulse response function of y_t to a shock of size δ is

$$R(\tau, \delta) = \mathsf{E}[y_{t+\tau} | \varepsilon_t = \delta, \mathbb{Y}_{t-1}] - \mathsf{E}[y_{t+\tau} | \varepsilon_t = 0, \mathbb{Y}_{t-1}], \quad \tau = 0, 1, 2, \ldots. \tag{2}$$

When the DGP is the model in (1), it can be easily verified that

$$R(\tau, \delta) = F_n^{\tau} \delta, \tag{3}$$

where F_n^τ is the $n \times n$ leading principal submatrix of the τ-th power F^τ of F (with $F_n^0 = I_n$). Evidently, $R(\tau, \delta)$ is invariant with respect to $\mathbb{Y}_{t-1}$ in this case and tends to the zero vector as τ tends to infinity.

Given any horizon $\tau \geq 0$, the conventional approach to estimating impulse responses (under a maintained VAR model of order p) amounts to replacing F_n^τ in (3) with the $n \times n$ leading principal submatrix $\widehat{F}_n^\tau$ of the τ-th power of the least-squares estimator of F (with $\widehat{F}_n^0 = I_n$) to obtain the VAR-based estimator

$$\widehat{R}_{\mathrm{VAR}}(\tau, \delta) = \widehat{F}_n^\tau \delta.$$

An alternative approach, put forward by Jordà (2005), relies on the linear projection model

$$y_{t+\tau} = c_\tau + B_\tau y_t + \sum_{k=1}^{p-1} A_{\tau,k} y_{t-k} + u_{\tau,t+\tau}, \quad t = 1, 2, \ldots, T - \tau, \tag{4}$$

where, for each horizon $\tau \geq 1$, c_τ is an n-dimensional vector of constants, $B_\tau, A_{\tau,1} \ldots, A_{\tau,p-1}$ are $n \times n$ parameter matrices, and $u_{\tau,t+\tau}$ is a n-dimensional projection error that is autocorrelated in general (for instance, $u_{\tau,t+\tau} = \sum_{k=0}^{\tau-1} F_n^k \varepsilon_{t+\tau-k}$ when y_t satisfies (1)). The standard LP estimator of $R(\tau, \delta)$ is

$$\widehat{R}_{\mathrm{LP}}(\tau, \delta) = \widehat{B}_\tau \delta,$$

where $\widehat{B}_\tau$ is the least-squares estimator of B_τ in (4) (with $\widehat{B}_0 = I_n$). Note that, unless it is maintained that the DGP is the VAR model in (1), LP impulse response estimates may be obtained from linear projections of $y_{t+\tau}$ onto a constant and $(y_t, y_{t-1}, \ldots, y_{t-m})$ for some positive integer m that need not be the same for different horizons τ.

3 Bias and Predictive Accuracy of Impulse-Response Estimators

This section presents and discusses the results of Monte Carlo experiments designed to shed light on the comparative performance of VAR-based and LP-based impulse responses in terms of bias and predictive accuracy. In order to focus on these two aspects of impulse responses, we abstract from issues associated with structural identification of shocks by considering a setting in which the innovations in the Wold representation of y_t have a diagonal covariance matrix. We note that popular recursive and non-recursive VAR-based identification schemes, which identify as structural shocks particular linear combinations of the Wold innovations, can equivalently be implemented using LP, with the implied impulse responses of the two approaches being approximately equal (up to scale) under general conditions (see Plagborg-Møller and Wolf (2021)). When impulse responses with respect to structural shocks $\xi_t = \Gamma\varepsilon_t$ are of interest, where Γ is an $n \times n$ parameter matrix of full rank (satisfying suitable identification restrictions), both VAR-based and LP-based estimates of impulse responses typically involve an adjustment based on the same consistent estimate of Γ. In the setting of our experiments, such an adjustment would not have any material effect on the relative merits of VAR-based and LP-based estimators of impulse responses and would not alter the conclusions.

The DGPs used in the experiments are bivariate versions ($n = 2$) of the VAR model in (1), with $\mu = (0.4, 0.4)'$ and ε_t being normally distributed with identity covariance matrix. We consider three parameter configurations:

$$\text{DGP-1 } (p = 1)\text{: } \Phi_1 = \begin{bmatrix} 0.49 & 0.36 \\ 0.36 & 0.49 \end{bmatrix};$$

$$\text{DGP-2 } (p = 1)\text{: } \Phi_1 = \begin{bmatrix} 0.80 & 0.10 \\ 0.75 & 0.40 \end{bmatrix};$$

$$\text{DGP-3 } (p = 2)\text{: } \Phi_1 = \begin{bmatrix} 0.45 & 0.16 \\ 0.53 & 0.42 \end{bmatrix}, \quad \Phi_2 = \begin{bmatrix} 0.15 & 0.15 \\ -0.28 & 0.19 \end{bmatrix}.$$

Letting $e_1 = (1, 0)'$ and $e_2 = (0, 1)'$, the objects of interest are the horizon-τ impulse responses $R_i(\tau, e_j) = e_i' R(\tau, e_j) = e_i' F_2^{\tau} e_j$ $(i, j = 1, 2)$ and their VAR-based and LP-based estimators $\widehat{R}_{\mathrm{VAR},i}(\tau, e_j) = e_i' \widehat{R}_{\mathrm{VAR}}(\tau, e_j)$ and $\widehat{R}_{\mathrm{LP},i}(\tau, e_j) = e_i' \widehat{R}_{\mathrm{LP}}(\tau, e_j)$, respectively. All reported simulation results are obtained from 1000 artificial samples.

3.1 Bias

Monte Carlo estimates of the bias of VAR-based and LP-based estimators of impulse responses at each horizon $\tau = 1, 2, \ldots, 15$ are shown in Figures 1 and 2 for the value of p that matches the order of the DGP and for sample size $T = 100, 400, 1600$. These are computed as the average over Monte Carlo replications of the differences between the estimated horizon-τ impulse responses $\widehat{R}_{\mathrm{VAR},i}(\tau, e_j)$ and $\widehat{R}_{\mathrm{LP},i}(\tau, e_j)$ and the true impulse response $R_i(\tau, e_j) = e_i' F_2^{\tau} e_j$.

The clear pattern that emerges from these plots is that the finite-sample bias of both estimators is a decreasing function of the sample size and becomes negligible for $T = 1600$. A notable difference between the two estimators of impulse responses is that the bias of VAR-based estimators tends to decrease in general (albeit not necessarily monotonically) as τ increases. By contrast, the bias of LP-based estimators has a tendency to increase as the horizon length τ becomes longer, especially when the sample size is relatively small. The VAR-based estimator generally has a bias advantage for all three DGPs, particularly for small sample sizes and long horizon lengths.

In an additional set of experiments, we assess the implications of using a VAR the order of which does not necessarily coincide with that of the DGP. Specifically, to reflect empirical practice, VAR-based impulse responses are obtained from a model the order p of which is selected from the set $\{1, 2, 3\}$ so as to minimize Akaike's information criterion $\mathrm{AIC}(p) = T \ln |\widehat{\Sigma}| + 2n^2 p$ or the Bayesian information criterion $\mathrm{BIC}(p) = T \ln |\widehat{\Sigma}| + n^2 p \ln T$, where $\widehat{\Sigma}$ is the estimated covariance matrix of the innovations.

The percentage of Monte Carlo replications in which each of the three possible values of p is selected by the two information criteria when artificial data are generated according to DGP-3 are shown in Table 1. AIC is more successful than BIC in selecting the correct order $p = 2$ when $T \leq 200$, while the reverse is true when $T \geq 400$. The known consistency of BIC is reflected in the observation that the true order is always selected when $T \geq 800$.

The average estimation errors of VAR-based and LP-based estimators of impulse responses under DGP-3 when the order of the VAR is chosen by AIC and BIC are shown in Figures 3 and 4, respectively. In this case, LP-based estimators are given an (arguably unfair) advantage by specifying a lag structure for the linear projection model which matches the order of the DGP, i.e., $p = 2$ in (4). As in previous experiments, substantial differences between the two estimators are observed only for the smallest of the sample sizes under consideration. The general pattern of the bias appears to be roughly the same when AIC and BIC are used, especially for $T \geq 400$, a finding which perhaps is not very surprising in light of the results reported in Table 1.

Overall, the differences in the observed estimation errors committed by the two estimators of impulse responses appear to favour the VAR-based approach. Whether these differences are significant enough for the VAR-based approach to be considered to outperform the LP-based approach is a question to which we turn our attention next.

3.2 Predictive Accuracy

Since horizon-τ impulse responses are in essence differences between τ-step-ahead predictors, the relative accuracy of VAR-based and LP-based estimators may be assessed by comparing their predictive accuracy using the methodology of Diebold and Mariano (1995), or its small-sample modification discussed in Harvey et al. (1997). Specifically, let

$$d_{ij}(\tau) = \{R_i(\tau, e_j) - \widehat{R}_{\mathrm{LP},i}(\tau, e_j)\}^2 - \{R_i(\tau, e_j) - \widehat{R}_{\mathrm{VAR},i}(\tau, e_j)\}^2, \quad i, j = 1, 2,$$

be the difference between the squared errors associated with the LP-based and VAR-based estimators of the horizon-τ impulse response of the i-th variable $Y_{i,t}$ to the j-th shock $\varepsilon_{j,t}$. The hypothesis of equal accuracy, with respect to the mean-squared-error risk function, may be formulated as $\mathsf{E}[d_{ij}(\tau)] = 0$, whereas $\mathsf{E}[d_{ij}(\tau)] > 0$ (< 0) implies that VAR-based impulse responses are more (less) accurate than LP-based responses.

Following Harvey et al. (1997), when N horizon-τ differentials $d_{ij,s}(\tau)$, $s = 1, 2, \ldots, N$ $(N \geq \tau)$, are available, with arithmetic mean $\bar{d}_{ij}(\tau) = N^{-1} \sum_{s=1}^{N} d_{ij,s}(\tau)$, a test for equal accuracy may be based on the statistic

$$\mathcal{D}_{ij}(\tau) = \left\{ \frac{N + 1 - 2\tau + N^{-1}\tau(\tau - 1)}{\widehat{\omega}_{ij}^2(\tau)} \right\}^{1/2} \bar{d}_{ij}(\tau), \tag{5}$$

where $\widehat{\omega}_{ij}^2(\tau)$ is the estimator of the asymptotic variance of $N^{1/2}\bar{d}_{ij}(\tau)$ given by

$$\begin{aligned}\widehat{\omega}_{ij}^2(\tau) = N^{-1} &\sum_{s=1}^{N} \{d_{ij,s}(\tau) - \bar{d}_{ij}(\tau)\}^2 \\ &+ 2N^{-1} \sum_{k=1}^{\tau-1} \sum_{s=k+1}^{N} \{d_{ij,s}(\tau) - \bar{d}_{ij}(\tau)\}\{d_{ij,s-k}(\tau) - \bar{d}_{ij}(\tau)\}.\end{aligned}$$

The null hypothesis of equal accuracy is rejected at level α, in favour of the one-sided alternative that VAR-based responses are superior, if $\mathcal{D}_{ij}(\tau)$ exceeds the $100(1-\alpha)$-th percentile of the Student-t distribution with $N-1$ degrees of freedom.

Differentials $d_{ij,s}(\tau)$, $s = 1, \ldots, N$, are constructed in our analysis using a recursive scheme. More specifically, for each artificial time series of length $T = 100, 200, 400, 800, 1600$ from the DGP of interest and each horizon $\tau = 1, 2, \ldots, 15$, impulse responses $\widehat{R}_{\mathrm{VAR},i}(\tau, e_j)$ and $\widehat{R}_{\mathrm{LP},i}(\tau, e_j)$ are computed using the first $T_0 = 60$ data points to obtain $\widehat{F}_2^{\tau}$ and $\widehat{B}_{\tau}$. This is then repeated for the first $T_0+1, T_0+2, \ldots, T$ data points to obtain $N = T - T_0 + 1$ differentials $d_{ij,s}(\tau)$ in total.

Tables 2, 3 and 4 show the percentage of artificial samples in which VAR-based impulse responses are found to be more accurate than LP-based responses, according to a 5%-level test based on the statistic in (5), when the VAR and LP lag structures coincide with that of the DGP. The results suggest that the use of VAR-based impulse responses is preferable in samples comprising 200 observations or less, especially so at long horizons; at such horizons, VAR-based responses have a clear advantage over LP-based responses under all three DGPs. For medium-sized samples of 400 observations, there is not much to choose between the two competing estimators, although VAR-based impulse responses tend to have a slight advantage the longer the impulse horizon is. Finally, for samples of 800 observations or more, the hypothesis of equal accuracy is very rarely rejected.

Similar conclusions are reached when the order of the VAR model is determined by AIC or BIC instead of being fixed at $p = 2$. The percentage rejections of the test of equal accuracy reported in Tables 5 and 6 reveal that VAR-based impulse responses retain their advantage over LP-based responses in samples of 200 or less observations, especially at longer horizons. This is the case even though the LP lag structure is always fixed at the value $p = 2$ implied by the DGP, whereas that of the VAR model is selected in each replication by AIC or BIC and is, therefore, potentially incorrect. As the sample size increases, differences between the squared errors of the two estimators become statistically insignificant.

Simulation results (not shown) for a 5%-level test of the hypothesis of equal accuracy

($\mathsf{E}[d_{ij}(\tau)] = 0$) against the one-sided alternative that VAR-based impulse responses are less accurate than LP-based responses ($\mathsf{E}[d_{ij}(\tau)] < 0$) reveal that the null hypothesis is very rarely rejected when the lag structures used coincide with that of the DGP. The highest rejection rates (ranging between 8% and 13%) are obtained for $R_1(\tau, e_1)$ and $R_2(\tau, e_1)$ under DGP-2, when $\tau \geq 2$ and $T = 100$. When the order of the VAR model is selected by AIC or BIC, rejection rates are, once again, very small for $T \geq 200$. LP-based responses (obtained with the correct number of lags $p = 2$) have an advantage only in samples of 100 observations and for relatively short horizons. In these cases, percentage rejections are higher than those expected for a 5%-level test, the highest rejection rates (ranging from 11.8% to 41.2%) being obtained for $R_1(\tau, e_2)$ and $R_2(\tau, e_2)$, with $\tau \leq 8$, using BIC.

We note that the results based on predictive accuracy tests are generally in agreement with the findings in the simulation exercises of Brugnolini (2018) and Herbst and Johannsen (2021), who also report that VAR-based estimators of impulse responses have superior performance when using well-specified models that contain the DGP. For DGPs that do not admit a finite-order VAR representation, the results in Li et al. (2024) suggest that while LP-based and VAR-based estimators tend to perform similarly at short horizons, they behave differently at intermediate and long horizons, with VAR-based estimators being a more attractive choice in terms of mean-squared-error loss.

4 Conclusion

This paper has investigated the comparative performance of VAR-based and LP-based estimators of impulse responses in terms of finite-sample predictive accuracy. Motivated by the observation that impulse responses are differences between optimal predictors, we have examined whether VAR-based and LP-based estimators of impulse responses have equal predictive ability (with respect to the mean-squared-error risk function) using a modification of the Diebold–Mariano approach. This is a novel way of assessing the statistical significance of the differences in the estimation errors associated with the two competing approaches.

Although one should exercise caution when drawing conclusions from Monte Carlo experiments based on a small number of DGPs, the results of our simulation study have shown that there is not much to choose between the two estimators of impulse responses in large samples. However, in small and moderately sizes samples similar to those commonly used in applied macroeconomics, VAR-based estimators have an advantage over LP-based ones, particularly at long horizons where LP-based estimators tend to be significantly biased. This advantage is observed when the order of the VAR model coincides with that of the DGP, as well as in the case where the order of the model is chosen by AIC or BIC and, consequently, is potentially incorrect.

References

Brugnolini, L. (2018). About local projection impulse response function reliability. CEIS Working Paper No. 440, University of Rome Tor Vergata.

Diebold, F. X. and R. S. Mariano (1995). Testing the equality of prediction mean squared errors. *Journal of Business & Economic Statistics 13*(3), 253–263.

Harvey, D., S. Leybourne, and P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting 13*(2), 281–291.

Herbst, E. and B. K. Johannsen (2021). Bias in local projections. FEDS Working Paper No. 2020-010R1, Board of Governors of the Federal Reserve System.

Jordà, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review 95*(1), 161–182.

Koop, G., H. M. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*(1), 119–147.

Li, D., M. Plagborg-Møller, and C. K. Wolf (2024). Local projections vs. VARs: Lessons from thousands of DGPs. *Journal of Econometrics 244*(2), Article 105722.

Plagborg-Møller, M. and C. K. Wolf (2021). Local projections and VARs estimate the same impulse responses. *Econometrica 89*(2), 955–980.

Table 1: Frequencies of Selected VAR Orders (in Percentage)

	$T = 100$		$T = 200$		$T = 400$		$T = 800$		$T = 1600$	
Order	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1	17.5	64.5	0.50	19.8	0.00	0.30	0.00	0.00	0.00	0.00
2	75.7	35.5	93.7	80.2	94.2	99.7	94.1	100	94.8	100
3	6.80	0.00	5.80	0.00	5.80	0.00	5.90	0.00	5.20	0.00

Table 2: Percentage Rejections of the Hypothesis of Equal Accuracy Under DGP-1 ($p = 1$)

$R_1(\tau, e_1)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	24.5	26.5	30.1	32.5	34.9	38.3	39.2	42.5	46.7	51.4	54.2	60.5	66.1	69.6
200	0.0	12.1	12.4	15.0	14.6	16.3	17.8	17.9	19.9	22.7	24.6	26.9	29.3	32.6	34.5
400	0.0	5.7	5.2	6.8	6.8	6.5	7.0	6.6	7.9	9.4	10.9	13.9	13.6	14.0	17.4
800	0.0	2.0	2.4	3.3	3.3	1.9	3.7	4.0	5.7	4.9	6.0	6.7	6.2	6.7	7.5
1600	0.0	1.0	1.0	0.7	0.8	1.0	1.7	2.2	2.2	2.6	2.6	2.2	2.1	2.9	3.1

$R_1(\tau, e_2)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	20.1	22.7	24.5	27.8	29.5	36.0	38.9	42.4	45.4	48.4	55.1	61.8	65.4	70.9
200	0.0	8.9	9.6	11.4	13.4	14.5	16.1	18.2	19.8	23.1	26.3	30.2	32.0	31.7	36.9
400	0.0	4.1	5.2	5.0	4.9	6.1	7.2	7.8	9.5	9.2	11.3	11.3	12.8	14.0	16.2
800	0.0	1.9	2.0	2.3	2.3	2.7	3.1	3.4	3.0	4.5	4.2	4.3	5.4	5.6	6.5
1600	0.0	0.7	1.2	0.8	1.3	1.5	2.2	1.0	1.5	1.6	2.2	2.2	3.2	3.2	3.3

$R_2(\tau, e_1)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	22.0	25.3	25.5	27.6	33.5	36.2	38.9	40.4	47.5	52.4	57.2	60.9	67.3	70.0
200	0.0	8.4	11.4	12.4	13.8	16.2	16.7	18.6	18.2	21.6	24.5	27.6	29.1	32.4	34.9
400	0.0	4.5	6.4	5.2	5.3	6.1	5.6	7.1	7.9	10.3	12.4	11.8	14.5	14.2	16.1
800	0.0	2.2	2.8	2.4	2.8	3.0	3.8	3.8	5.0	4.8	6.1	6.1	5.7	6.2	6.9
1600	0.0	0.7	0.8	1.2	0.8	1.4	1.6	1.6	2.3	1.9	2.0	2.2	2.8	2.3	3.4

$R_2(\tau, e_2)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	25.1	25.0	25.3	30.8	34.3	36.5	39.9	43.2	46.0	49.6	54.3	61.6	64.5	69.4
200	0.0	12.3	11.4	13.6	14.6	17.0	18.3	17.9	18.8	23.1	27.7	29.0	31.7	32.2	35.8
400	0.0	4.9	5.1	5.2	6.1	6.1	6.7	7.9	8.9	10.3	12.3	12.6	13.3	13.5	15.4
800	0.0	2.3	2.3	2.3	2.7	2.8	3.8	3.5	3.7	4.1	4.8	4.2	5.9	6.0	6.7
1600	0.0	0.5	0.4	1.6	0.9	1.4	1.1	2.0	1.6	2.1	2.8	2.5	2.7	2.8	2.1

.

Table 3: Percentage Rejections of the Hypothesis of Equal Accuracy Under DGP-2 ($p = 1$)

$R_1(\tau, e_1)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	21.0	22.1	21.9	25.0	27.4	29.1	30.8	32.2	35.6	39.4	42.3	45.6	49.6	51.5
200	0.0	7.8	9.4	10.7	10.2	10.6	10.8	11.8	11.4	13.4	15.2	17.3	19.2	21.5	23.7
400	0.0	3.1	3.9	4.2	4.3	3.4	3.7	4.4	4.4	5.6	5.2	6.6	7.2	8.3	8.9
800	0.0	1.7	1.1	1.8	1.2	1.1	1.4	1.4	1.4	1.6	1.4	1.5	2.2	2.0	2.7
1600	0.0	0.6	0.7	0.5	0.5	0.4	0.4	0.4	0.5	0.4	0.4	0.3	0.2	0.4	0.5

$R_1(\tau, e_2)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	19.4	22.3	24.1	26.7	29.8	34.4	38.5	42.5	48.0	50.4	55.4	61.1	63.2	68.9
200	0.0	8.7	10.7	12.2	12.1	14.0	15.3	16.7	18.6	21.4	23.3	27.3	28.9	30.7	34.3
400	0.0	3.5	3.7	4.9	4.6	4.9	6.5	6.4	7.4	7.0	8.4	10.0	10.5	12.0	13.3
800	0.0	2.5	1.3	1.9	2.0	2.3	3.3	2.8	3.2	3.2	3.3	3.4	3.4	3.9	4.3
1600	0.0	1.2	0.6	0.9	1.5	1.5	1.8	1.6	1.8	2.1	2.0	1.8	1.4	2.0	2.5

$R_2(\tau, e_1)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	17.9	17.0	18.4	21.1	22.1	24.6	26.5	29.3	32.8	33.9	39.9	42.9	43.9	49.5
200	0.0	8.1	7.5	7.9	9.0	8.2	9.4	8.9	10.5	12.4	13.2	15.7	17.6	16.9	20.3
400	0.0	3.4	3.5	3.8	3.6	2.6	2.4	3.0	3.9	5.0	5.9	6.2	7.2	7.9	7.3
800	0.0	1.8	2.1	1.4	1.8	1.3	1.3	1.3	1.8	1.3	1.1	1.6	2.1	2.3	2.2
1600	0.0	0.3	0.6	0.7	0.6	0.6	0.4	0.4	0.9	0.7	0.6	0.6	0.6	0.7	0.2

$R_2(\tau, e_2)$

τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	23.7	23.0	25.2	27.7	31.2	34.4	38.3	39.6	47.2	49.9	55.2	59.9	63.9	66.8
200	0.0	10.4	11.4	12.5	13.9	13.9	16.2	16.5	17.0	20.1	23.8	26.7	30.4	31.2	34.6
400	0.0	3.6	5.5	5.2	5.1	5.7	6.6	6.9	6.4	8.3	7.6	10.0	10.7	13.4	13.6
800	0.0	2.0	2.3	2.0	2.3	2.0	2.7	2.6	3.5	3.9	3.5	4.1	3.5	3.9	4.5
1600	0.0	0.5	0.8	1.7	1.4	1.3	1.5	1.3	1.5	1.7	1.4	2.1	1.0	1.7	1.7

Table 4: Percentage Rejections of the Hypothesis of Equal Accuracy Under DGP–3 ($p = 2$)

	$R_1(\tau, e_1)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	18.8	20.2	22.7	24.3	25.4	29.6	33.3	33.7	38.2	44.5	42.2	51.0	54.0	59.2
200	0.0	10.2	9.3	8.4	10.7	11.3	13.1	13.4	13.5	16.0	19.6	17.6	23.1	25.6	27.1
400	0.0	3.9	4.8	4.1	5.3	4.3	5.2	4.4	5.3	5.6	6.9	7.5	9.3	9.4	12.9
800	0.0	1.2	2.3	2.0	1.5	1.6	1.7	1.2	2.5	2.5	3.0	2.8	4.0	3.5	4.9
1600	0.0	0.8	1.3	0.7	0.5	0.5	0.5	0.4	0.7	0.7	1.7	1.3	1.0	0.8	1.9

	$R_1(\tau, e_2)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.1	15.3	19.7	18.6	20.9	22.9	27.5	29.9	33.7	37.1	39.0	41.6	49.3	52.7	55.9
200	0.0	7.3	8.3	9.4	7.8	10.0	12.3	12.9	14.8	15.5	18.0	20.5	24.3	25.2	28.3
400	0.0	3.1	2.8	4.1	3.6	4.5	5.1	5.2	6.8	6.1	7.7	8.3	11.4	10.8	12.6
800	0.0	1.6	1.7	1.9	1.6	2.1	2.1	2.7	2.4	2.7	3.2	3.5	4.3	4.8	5.0
1600	0.0	0.4	1.1	1.0	1.3	0.8	0.7	1.0	0.7	1.1	1.4	1.5	1.4	1.5	1.9

	$R_2(\tau, e_1)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	17.6	19.6	20.1	22.1	27.7	27.6	35.8	35.4	39.2	43.8	47.8	52.7	58.5	61.2
200	0.0	9.1	9.4	9.3	9.2	10.3	11.5	13.2	14.5	16.0	18.3	20.9	23.3	27.3	28.2
400	0.0	3.8	4.9	4.2	4.2	4.4	3.9	5.4	4.0	5.4	7.6	8.7	9.1	11.5	10.8
800	0.0	1.5	1.4	1.4	2.5	1.4	1.9	3.0	2.7	1.9	3.4	4.1	3.3	3.7	4.1
1600	0.0	0.5	0.5	0.5	0.7	0.9	0.9	1.0	1.4	0.8	1.3	1.5	2.0	0.8	1.7

	$R_2(\tau, e_2)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	0.0	19.2	20.4	23.0	26.0	27.1	30.3	34.1	36.7	38.0	41.4	44.5	51.1	55.9	58.5
200	0.0	8.1	10.1	9.6	12.0	12.2	13.2	15.1	13.9	16.8	18.0	19.8	24.8	24.9	29.4
400	0.0	3.5	3.8	3.6	4.3	4.9	4.3	5.8	5.3	7.5	8.2	8.8	10.3	12.0	14.6
800	0.0	1.1	1.6	2.2	1.5	2.4	2.4	2.8	2.3	3.5	3.6	3.4	3.6	5.7	5.8
1600	0.0	0.4	0.3	0.8	0.7	0.7	1.1	2.2	1.0	1.6	1.5	2.3	1.8	2.6	2.3

Table 5: Percentage Rejections of the Hypothesis of Equal Accuracy Under DGP-3 (AIC-Selected VAR Order)

	$R_1(\tau, e_1)$														
T \ τ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	2.8	16.8	18.7	21.3	22.0	25.1	27.6	30.8	35.1	34.9	39.9	45.8	44.5	49.8	55.6
200	0.5	5.8	8.2	10.8	8.9	9.8	10.5	13.5	13.9	14.6	14.1	20.8	21.9	22.6	27.7
400	0.1	3.1	2.7	4.0	3.2	4.1	4.5	5.0	5.6	4.9	6.6	8.4	10.0	9.5	11.6
800	0.0	1.1	1.0	1.6	1.7	1.5	1.5	2.0	2.0	2.3	1.8	2.2	2.5	3.0	2.8
1600	0.0	0.1	0.8	0.8	0.8	0.7	0.7	0.9	1.1	0.4	0.8	0.6	1.1	1.1	1.0

	$R_1(\tau, e_2)$														
T \ τ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	1.9	14.2	17.3	18.3	16.4	20.5	24.1	26.5	31.3	33.2	39.1	40.9	47.1	50.6	54.1
200	0.7	6.1	8.6	8.2	9.7	10.6	10.2	12.9	14.2	15.0	16.4	19.1	21.5	22.8	26.4
400	0.2	2.6	3.6	3.4	3.8	4.8	4.8	5.2	5.4	5.1	6.7	7.3	8.3	9.4	10.3
800	0.1	1.0	1.3	1.4	1.5	1.2	2.3	2.3	2.9	2.8	3.9	3.6	3.7	3.6	4.5
1600	0.0	0.6	0.9	0.4	0.9	0.5	0.7	1.0	0.9	1.6	1.0	1.8	1.4	1.7	1.9

	$R_2(\tau, e_1)$														
T \ τ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	3.3	12.8	18.3	18.3	21.5	22.9	26.6	30.1	33.7	37.8	39.6	44.3	52.5	53.6	60.2
200	0.6	7.3	9.6	8.4	9.0	10.1	10.2	11.8	13.4	13.5	18.1	20.7	21.1	22.5	27.3
400	0.4	2.9	4.3	4.2	3.9	3.7	4.4	4.5	4.8	5.9	7.5	7.7	8.1	8.8	10.2
800	0.1	1.5	2.1	0.9	1.9	1.6	1.3	2.3	2.1	1.9	2.8	3.7	2.6	3.9	5.0
1600	0.1	0.2	0.8	0.6	0.7	1.3	0.7	0.7	0.7	1.0	1.2	0.9	0.6	1.6	2.3

	$R_2(\tau, e_2)$														
T \ τ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	1.7	14.2	16.7	21.1	23.3	23.9	29.3	31.6	34.0	37.0	42.7	45.8	48.2	54.6	58.5
200	0.5	6.7	8.8	10.5	11.2	11.6	12.6	14.6	14.2	17.9	18.3	20.5	23.2	25.5	31.2
400	0.0	2.5	4.0	4.2	4.7	4.8	5.4	5.4	7.1	6.9	7.8	8.1	10.6	9.9	13.8
800	0.2	0.8	1.3	1.4	2.1	2.9	2.0	1.7	3.5	3.2	3.1	3.8	4.5	3.9	5.8
1600	0.1	0.3	0.4	0.9	0.7	1.2	1.1	1.1	1.3	1.5	1.6	1.3	1.6	1.7	2.2

Table 6: Percentage Rejections of the Hypothesis of Equal Accuracy Under DGP-3 (BIC-Selected VAR order)

	$R_1(\tau, e_1)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	10.2	18.7	18.6	21.4	22.0	25.4	27.4	30.0	36.2	35.5	39.8	45.8	44.7	49.5	56.0
200	0.9	6.1	8.5	10.4	8.4	9.5	10.3	13.3	13.3	14.6	13.9	21.0	21.6	22.3	28.2
400	0.0	3.3	2.8	4.0	3.1	4.0	4.6	5.1	5.7	4.9	6.7	8.5	10.1	9.5	11.6
800	0.0	1.2	1.3	1.6	1.7	1.5	1.7	2.0	2.1	2.3	1.8	2.3	2.5	3.0	2.8
1600	0.0	0.4	0.8	0.8	0.8	0.7	0.6	0.8	1.2	0.4	0.8	0.6	1.1	1.1	1.0

	$R_1(\tau, e_2)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	5.3	15.5	15.2	13.6	13.4	16.0	19.4	22.7	28.0	30.5	35.8	37.2	44.7	48.1	52.7
200	0.6	6.7	8.4	8.2	8.1	9.5	9.4	11.7	13.7	14.3	15.8	17.3	20.2	22.1	26.4
400	0.0	3.3	4.0	3.8	4.0	4.7	4.6	5.1	5.5	5.2	6.6	7.2	8.2	9.2	10.5
800	0.0	1.3	1.1	1.6	1.4	1.2	2.4	2.3	2.8	2.8	3.9	3.7	3.6	3.6	4.4
1600	0.0	0.6	0.9	0.4	0.9	0.5	0.7	1.0	0.9	1.6	1.0	1.7	1.4	1.7	1.9

	$R_2(\tau, e_1)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	6.0	9.1	19.5	19.3	22.3	24.9	27.9	31.0	35.0	39.5	40.5	45.0	53.7	53.9	61.1
200	0.7	5.8	9.4	8.3	9.2	10.1	10.0	12.2	13.8	13.5	18.5	21.1	20.7	22.8	27.6
400	0.0	3.3	4.6	4.2	4.1	3.7	4.4	4.4	4.8	5.9	7.7	7.9	8.1	8.8	10.0
800	0.0	1.8	2.0	1.0	1.9	1.7	1.4	2.2	2.1	1.9	2.9	3.7	2.6	3.9	5.0
1600	0.1	0.3	0.9	0.6	0.7	1.3	0.7	0.7	0.7	1.0	1.2	0.9	0.6	1.7	2.3

	$R_2(\tau, e_2)$														
τ T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
100	6.3	8.4	11.2	17.3	17.9	20.3	25.2	29.2	32.8	34.5	40.2	43.8	46.7	53.0	57.1
200	1.2	6.6	7.0	10.1	9.9	10.1	11.7	13.5	13.3	16.9	16.7	19.5	22.8	24.7	30.3
400	0.0	3.0	4.3	4.6	4.7	5.0	5.5	5.3	7.1	7.0	7.8	8.0	10.6	9.9	13.9
800	0.0	1.0	1.3	1.6	2.1	2.9	2.1	1.7	3.3	3.2	3.1	3.8	4.4	4.0	5.8
1600	0.0	0.3	0.5	0.8	0.7	1.1	1.1	1.1	1.3	1.5	1.6	1.3	1.6	1.7	2.2

Figure 1: Simulation-Estimated Bias of LP-Based and VAR-Based Estimators Under DGP-1 and DGP-2 ($p = 1$)

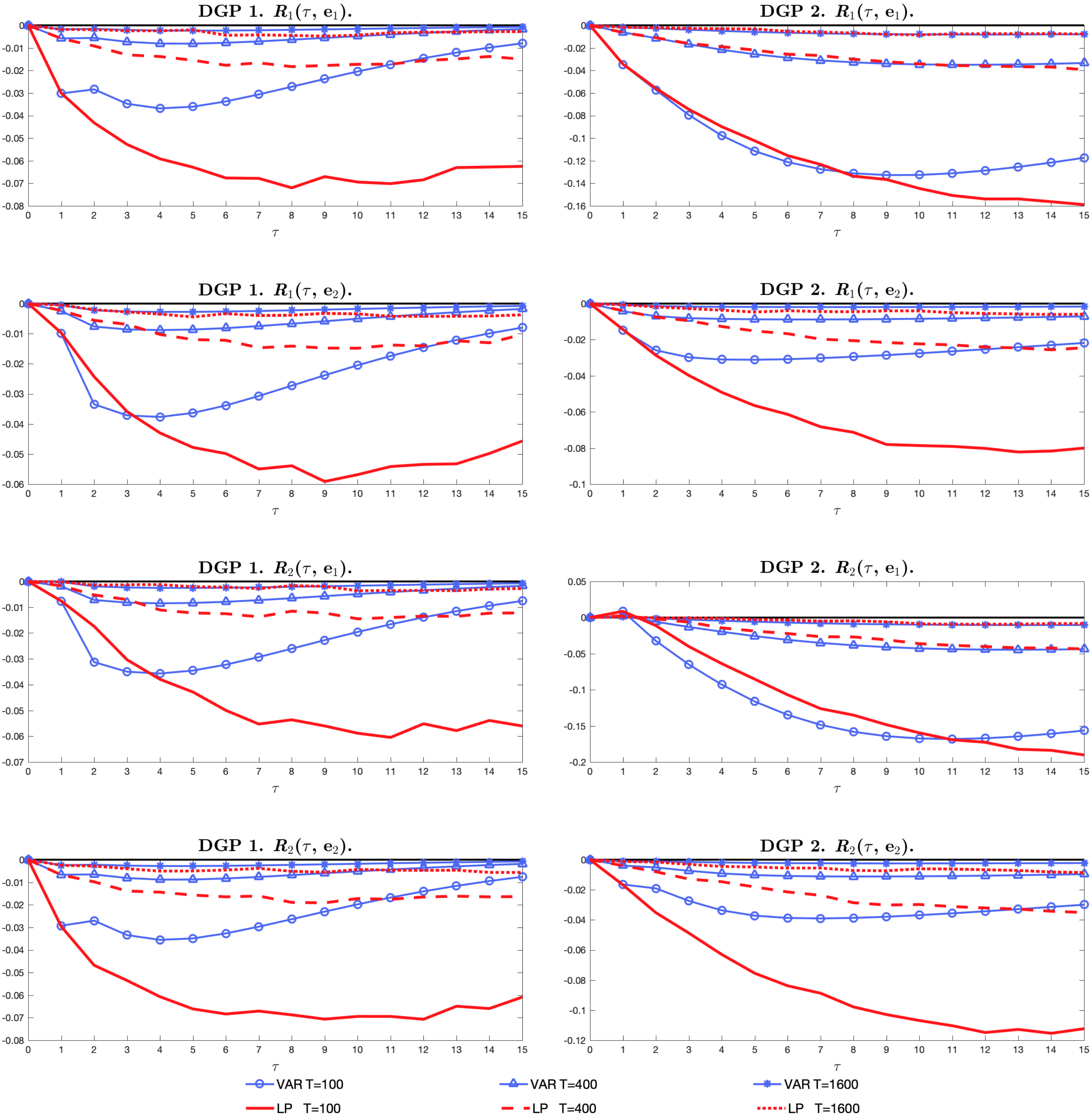


Figure 2: Simulation-Estimated Bias of LP-Based and VAR-Based Estimators Under DGP-3 ($p = 2$)

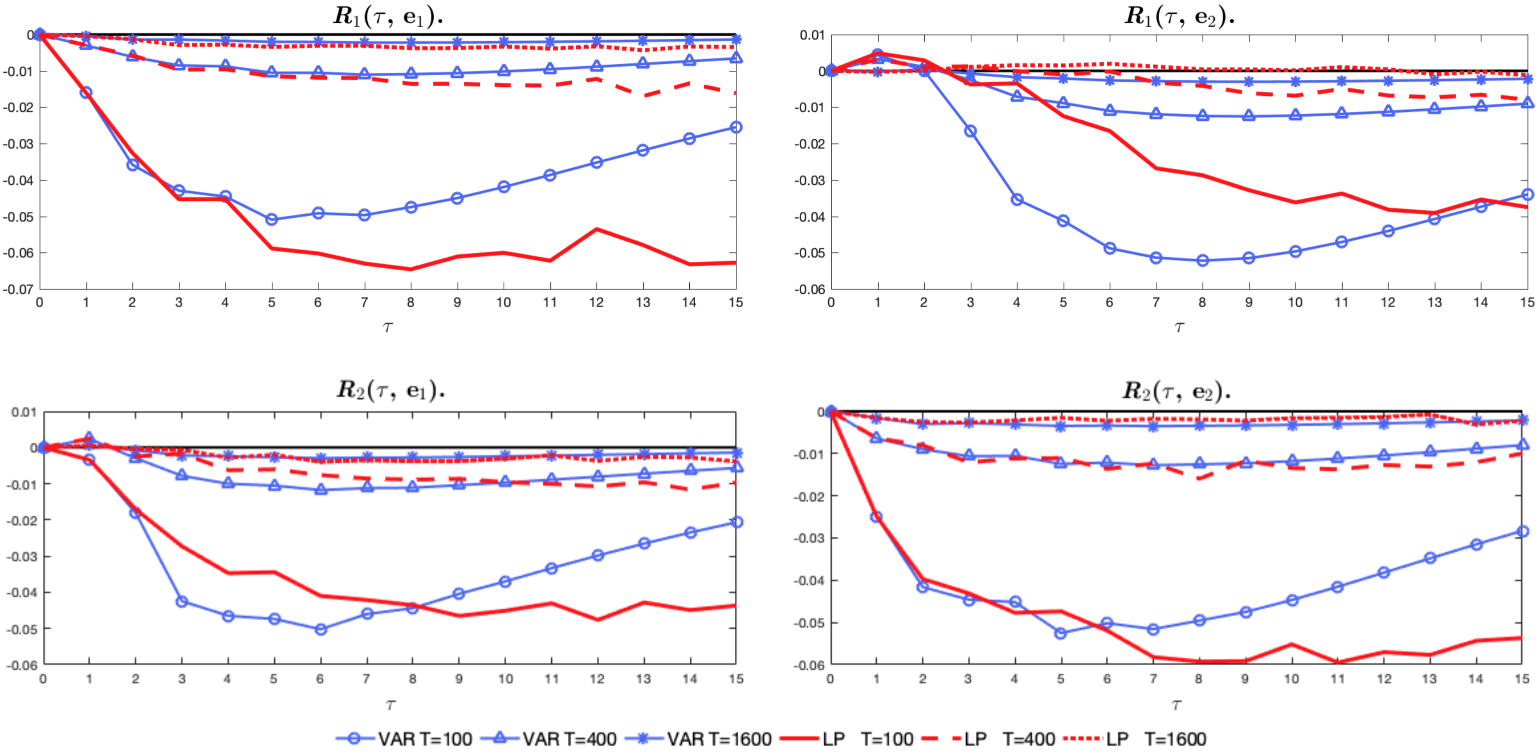
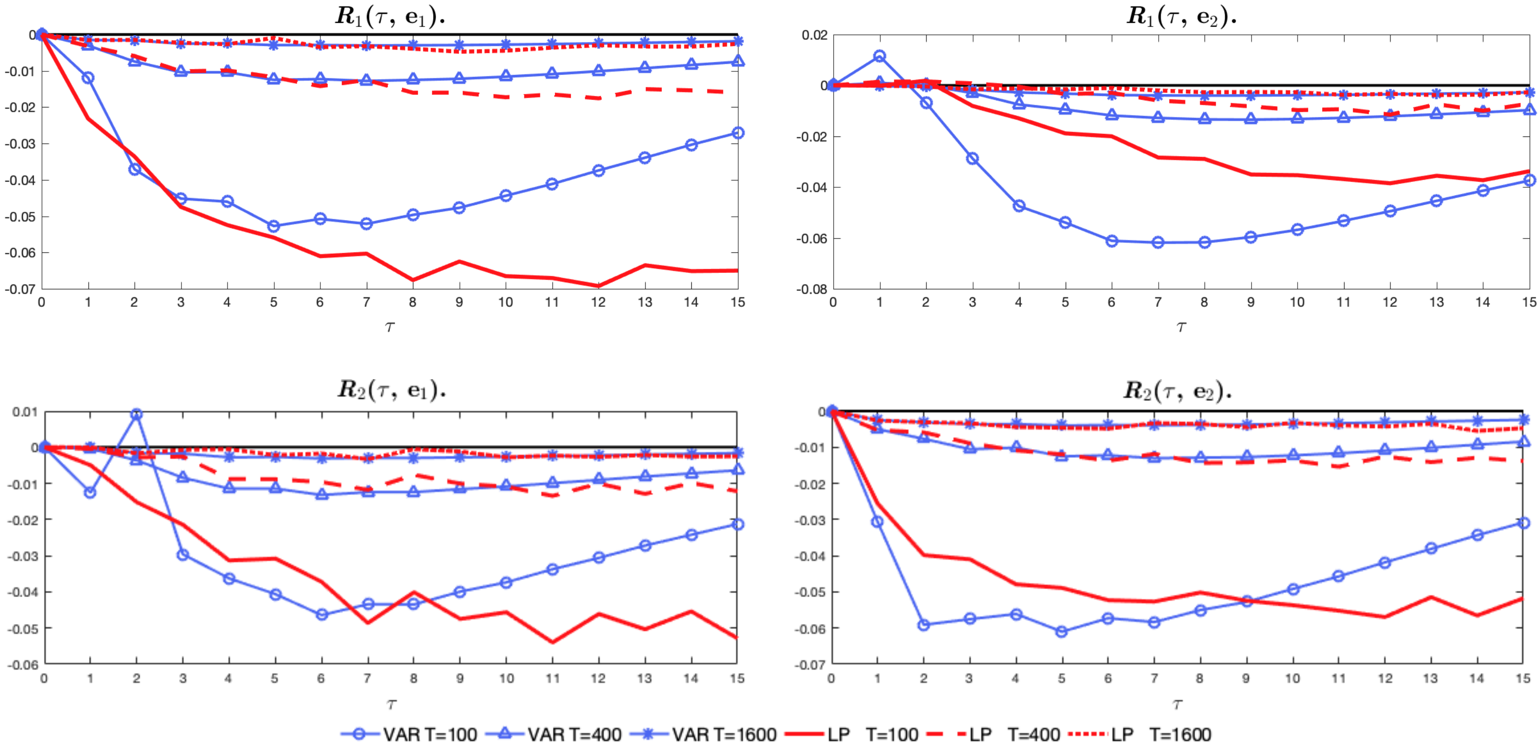


Figure 3: Simulation-Estimated Bias of LP-Based and VAR-Based Estimators Under DGP-3 (AIC-Selected VAR Order)



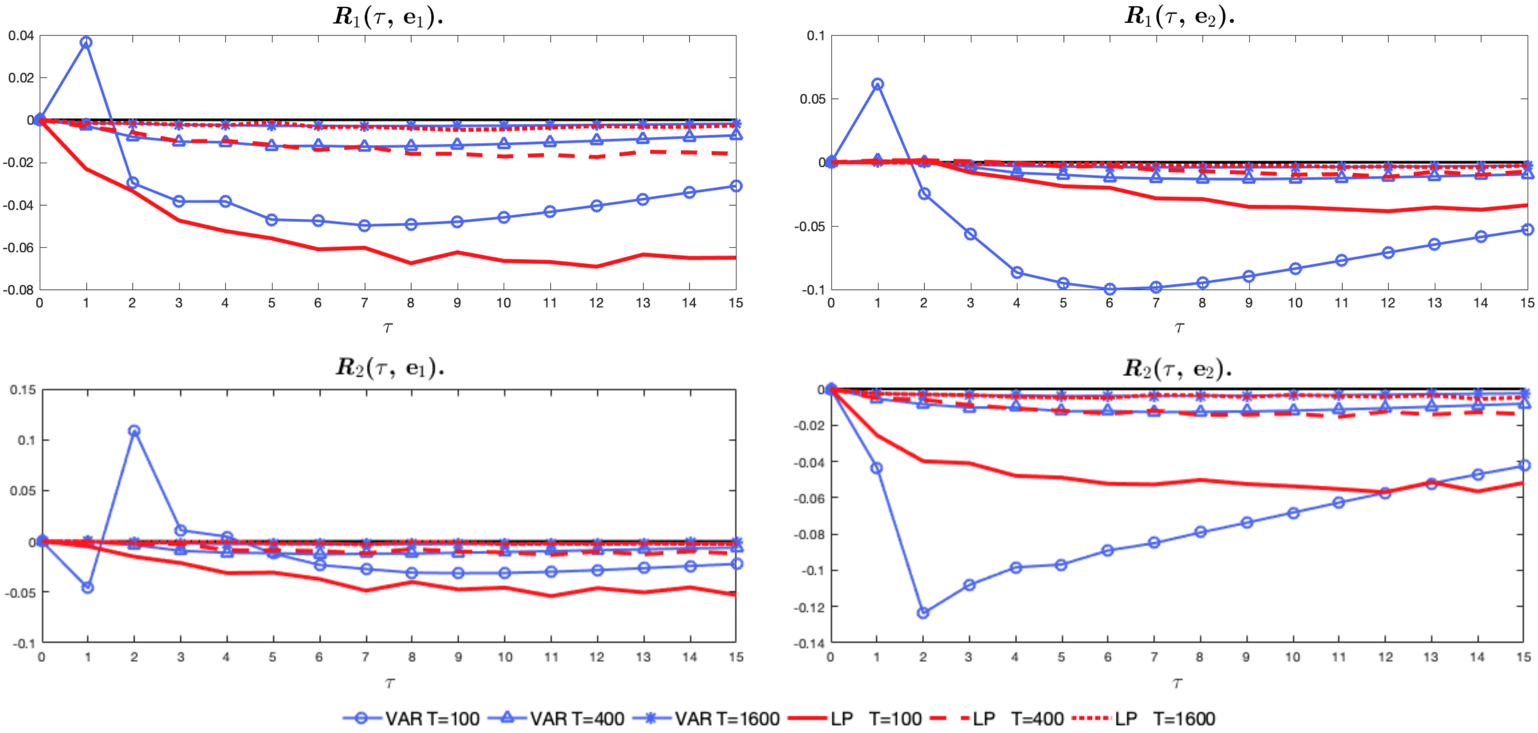
$\boldsymbol{R}_1(\tau,\ \mathbf{e}_1)$.

$\boldsymbol{R}_1(\tau,\ \mathbf{e}_2)$.

$\boldsymbol{R}_2(\tau,\ \mathbf{e}_1)$.

$\boldsymbol{R}_2(\tau,\ \mathbf{e}_2)$.

Figure 4: Simulation-Estimated Bias of LP-Based and VAR-Based Estimators Under DGP-3 (BIC-Selected VAR Order)



$\boldsymbol{R}_1(\tau,\, \mathbf{e}_1)$.

$\boldsymbol{R}_1(\tau,\, \mathbf{e}_2)$.

$\boldsymbol{R}_2(\tau,\, \mathbf{e}_1)$.

$\boldsymbol{R}_2(\tau,\, \mathbf{e}_2)$.