Vanderbilt University Department of Economics Working Paper Series
Nashville, TN 37235-1819

Vanderbilt University Department of Economics Working Papers 19-00001

The uniform validity of impulse response inference in autoregressions

Atsushi Inoue
Vanderbilt University

Lutz Kilian
University of Michigan

Abstract

Existing proofs of the asymptotic validity of conventional methods of impulse response inference based on higher-order autoregressions are pointwise only. In this paper, we establish the uniform asymptotic validity of conventional asymptotic and bootstrap inference about individual impulse responses and vectors of impulse responses at fixed horizons. For inference about vectors of impulse responses based on Wald test statistics to be uniformly valid, lag-augmented autoregressions are required, whereas inference about individual impulse responses is uniformly valid under weak conditions even without lag augmentation. We introduce a new rank condition that ensures the uniform validity of inference on impulse responses and show that this condition holds under weak conditions. Simulations show that the highest finite-sample accuracy is achieved when bootstrapping the lag-augmented autoregression using the bias adjustments of Kilian (1999). The resulting confidence intervals remain accurate even at long horizons. We provide a formal asymptotic justification for this result.

We thank Xu Cheng, Nikolay Gospodinov, Helmut Lutkepohl, Mikkel Plagborg-Moller, Barbara Rossi, Katsumi Shimotsu and Xiaoqing Zhou for helpful discussions.
Citation: Atsushi Inoue and Lutz Kilian, (2019) "The uniform validity of impulse response inference in autoregressions", *Vanderbilt University Department of Economics Working Papers*, VUECON-19-00001.
Contact: Atsushi Inoue - atsushi.inoue@vanderbilt.edu, Lutz Kilian - lkilian@umich.edu.
Submitted: January 31, 2019. **Published:** March 25, 2019.
URL:http://www.accessecon.com/Pubs/VUECON/VUECON-19-00001.pdf

The Uniform Validity of Impulse Response Inference in Autoregressions

Atsushi Inoue*
Vanderbilt University

Lutz Kilian†
University of Michigan
CEPR

First draft: September 10, 2018. This version: January 27, 2019.

Abstract: Existing proofs of the asymptotic validity of conventional methods of impulse response inference based on higher-order autoregressions are pointwise only. In this paper, we establish the uniform asymptotic validity of conventional asymptotic and bootstrap inference about individual impulse responses and vectors of impulse responses at fixed horizons. For inference about vectors of impulse responses based on Wald test statistics to be uniformly valid, lag-augmented autoregressions are required, whereas inference about individual impulse responses is uniformly valid under weak conditions even without lag augmentation. We introduce a new rank condition that ensures the uniform validity of inference on impulse responses and show that this condition holds under weak conditions. Simulations show that the highest finite-sample accuracy is achieved when bootstrapping the lag-augmented autoregression using the bias adjustments of Kilian (1999). The resulting confidence intervals remain accurate even at long horizons. We provide a formal asymptotic justification for this result.

JEL code: C22, C52

Key words: Impulse response, autoregression, lag augmentation, asymptotic normality, bootstrap, uniform inference.

Acknowledgments: We thank Xu Cheng, Nikolay Gospodinov, Helmut Lütkepohl, Mikkel Plagborg-Møller, Barbara Rossi, Katsumi Shimotsu and Xiaoqing Zhou for helpful discussions.

*Department of Economics, Vanderbilt University, Nashville, TN 37235-1819. E-mail: atsushi.inoue@vanderbilt.edu

†University of Michigan, Department of Economics, 238 Lorch Hall, Ann Arbor, MI 48109-1220. E-mail: lkilian@umich.edu.

1 Introduction

Impulse response analysis based on autoregressions plays a central role in quantitative economics (see Kilian and Lütkepohl 2017). Many researchers have cautioned against relying on pre-tests to diagnose and remove apparent unit roots in autoregressive processes (e.g., Elliott 1998; Rossi and Pesavento 2007; Gospodinov, Herrera and Pesavento 2011). As result, autoregressions are often estimated based on highly persistent data. A long-standing question has been how to assess the uncertainty about response estimates when the dominant autoregressive root may be close to unity. The asymptotic validity of conventional methods of asymptotic and bootstrap inference for impulse responses for stationary processes has been established in Lütkepohl (1990) and Gonçalves and Kilian (2004). Extensions to possibly integrated autoregressive processes are provided in Inoue and Kilian (2002, 2003), building on Park and Phillips (1989) and Sims, Stock and Watson (1990). Kilian and Lütkepohl (2017) note that the assumptions underlying the analysis of higher-order autoregressions in Inoue and Kilian (2002) may be relaxed further by fitting lag-augmented autoregressions, as proposed by Dolado and Lütkepohl (1996) and Toda and Yamamoto (1995).[1] All these asymptotic justifications, however, rely on pointwise convergence results. It is unclear whether they are valid uniformly across the parameter space.

In many econometric applications the distinction between pointwise and uniform validity, as discussed in Giraitis and Phillips (2006), Mikusheva (2007a), Andrews and Guggenberger (2009), and Kasy (2018), among others, is of no practical importance. This distinction matters, however, when the distribution of the statistic of interest changes with the value of the population parameter to be estimated, as would be the case in the AR(1) model when the autoregressive root approaches unity. The concern is that for a $1-\alpha$ confidence interval C to be asymptotically valid we need to show that

$$\lim_{T\to\infty} \inf_{\rho} P_\rho(\rho \in C) \geq 1-\alpha,$$

where ρ denotes the AR(1) slope parameter and the infimum is taken over the parameter space of ρ. This means that there exists a sample size that guarantees the coverage accuracy of the interval for any parameter value ρ. In contrast, under the pointwise approximation, the actual coverage accuracy is not known and may become arbitrarily low, since the true value of ρ is not known.[2]

[1] Related work also includes Kurozumi and Yamamoto (2000) and Bauer and Maynard (2012).

[2] For example, Mikusheva (2007a) demonstrates that the conventional asymptotic normal approximation for the slope parameter in the AR(1) model with near unit roots is pointwise correct, but not uniformly correct, which helps explain the poor coverage accuracy of conventional confidence intervals in this model, as ρ approaches unity (see Nankervis and Savin 1988; Hansen 1999; Kilian 1999).

Concern over the reliability of conventional methods of inference when applied to highly persistent autoregressive processes has subsequently motivated the development of nonstandard asymptotic approximations based on local-to-unity processes. For example, Stock (1991) proposes constructing confidence intervals for the dominant autoregressive root by inverting unit root tests. Phillips (2014), however, proves that inference about the AR(1) slope parameter based on Stock's (1991) confidence interval, while asymptotically valid when the root is local to unity, has zero coverage asymptotically when the root is far enough from unity.[3]

The lack of a uniform asymptotic approximation across the parameter space has undermined the profession's confidence in the accuracy of either of these confidence intervals in applied work and has created interest in confidence intervals that remain asymptotically valid whether the AR(1) slope parameter is unity, close to unity or far from unity. For example, under weak conditions, the grid bootstrap of Hansen (1999) can be shown to provide a uniformly asymptotically valid approximation to the distribution of the AR(1) slope parameter under both stationary and local-to-unity asymptotics (see Mikusheva 2007a).

While the AR(1) process has been studied extensively in the literature, there has been much less work on the problem of uniform inference in higher-order autoregressions, which are the workhorse model in applied work. Allowing for additional lags turns out to change the properties of the estimator of the autoregressive model dramatically. Our analysis shows that the lack of uniform validity of the conventional Gaussian asymptotic approximation does not extend to inference on individual slope parameters in higher-order autoregressive models. In the latter case, asymptotic normality holds uniformly across the parameter space. This result has important implications for inference on smooth functions of autoregressive slope coefficients such as impulse responses in autoregressions.

Our contribution to this literature is fourfold. First, we show that conventional asymptotic and bootstrap confidence intervals for individual impulse responses remain uniformly asymptotically valid, as long as the horizon of the impulse response remains fixed with respect to the sample size, generalizing the pointwise asymptotic results in Park and Phillips (1989), Sims, Stock and Watson (1990) and Inoue and Kilian (2002). Our analysis covers both higher-order autoregressions and lag-augmented autoregressions. We provide a suitable rank condition that ensures that inference on impulse responses is uniformly valid. We show that lag-augmented autoregressions based on

[3] Phillips' conclusion is consistent with simulation evidence in Hansen (1999), which illustrates the comparatively poor coverage accuracy of Stock's method in the stationary region.

stationary, unit root or local-to-unity AR(p) processes always satisfy this rank condition at horizons $h \leq p$.[4]

Second, we establish the uniform asymptotic validity of Gaussian inference on the vectors of autoregressive slope parameters and vectors of impulse responses. The joint asymptotic normality of the estimator of autoregressive slope parameters has been postulated as a high-level assumption in a range of studies including Montiel Olea, Stock and Watson (2016), Guerron-Quintana, Inoue and Kilian (2017), and Gafarov, Meier and Montiel Olea (2018). Our analysis establishes the unifom joint asymptotic normality of the lag-augmented estimator of the autoregressive slope parameters under conditions not requiring the process to be stationary. We furthermore establish, under the same conditions, the uniform joint asymptotic normality of the impulse response estimator postulated by Granziera, Moon and Schorfheide (2018).

The latter result is also central for the construction of joint impulse response confidence intervals based on Wald test statistics. Joint inference on impulse response functions has become increasingly recognized as essential for practitioners interested in understanding the true extent of the uncertainty about estimates of impulse response functions (e.g., Jordà 2009; Lütkepohl; Staszewska-Bystrova and Winker 2015a,b,c; Inoue and Kilian 2016; Kilian and Lütkepohl 2017; Bruder and Wolf 2018; Montiel Olea and Plagborg-Møller 2018). Our analysis shows that the use of lag-augmented autoregressions is required for inference about impulse response functions based on Wald test statistics to be uniformly asymptotically valid, when the dominant autoregressive root may be arbitrarily close to unity.[5]

Third, a simulation study involving univariate autoregressions with varying degrees of persistence confirms that the conventional asymptotic approximation based on fixed impulse response horizons remains accurate even uniformly, as long as the horizon is reasonably small relative to the sample size. We find that impulse response confidence intervals based on lag-augmented autoregressions are considerably more accurate in small samples than confidence intervals based on the original autoregression. Substantial further improvements in coverage accuracy may be achieved by bootstrapping the lag-augmented autoregression. The reason that delta method confidence intervals tend to be less accurate than suitably constructed bootstrap confidence intervals is that, even

[4] In related work, Mikusheva (2012) established the uniform validity of one-dimensional impulse response inference for a specific form of the grid bootstrap applied to autoregressions. Her uniformity results, however, do not apply to the conventional delta method and bootstrap confidence intervals considered in our analysis.

[5] In contrast, the use of conservative sup-t or Bonferroni bounds, as discussed in Montiel Olea and Plagborg-Møller(2018), only requires the marginal distributions of the impulse responses to be uniformly asymptotically normal with strictly positive variances.

for stationary processes, the finite-sample distribution of impulse responses is far from Gaussian. The longer the horizon, the worse the normal approximation becomes. One potential remedy is the use of the Hall percentile interval, which allows the distribution of the impulse response estimator to be non-Gaussian. We find, however, that the use of the Hall percentile interval yields at best modest improvements in practice and cannot be recommended.

An alternative is the bias-adjusted bootstrap method of Kilian (1999), which was designed to improve the small-sample accuracy of impulse response confidence intervals in stationary autoregressions. This method yields consistently high uniform coverage accuracy when applied to lag-augmented autoregressions. For example, for $T = 240$, uniform coverage rates range from 87% to 89% and for $T = 480$ from 89% to 90%, for horizons between 1 and 12 periods, which is a substantial improvement relative to the delta method for the same model. In contrast, without lag augmentation, both delta method and bootstrap confidence intervals are much less accurate, consistent with earlier simulation evidence in the literature. These results suggest that highly persistent autoregressions in applied work should be routinely lag-augmented when conducting impulse response inference. While lag-augmenting the autoregression will cause an increase in the average width of the interval, we show by simulation that this loss in efficiency may for all practical purposes be ignored, when the original autoregressive lag order is already large, as is typically the case in applied work.

Fourth, we establish the asymptotic validity of the Efron percentile interval at long horizons based on the lag-augmented autoregression within the local-to-unity framework. Although impulse response inference is not uniformly valid at long horizons, these results explain the excellent coverage accuracy of this interval for persistent autoregressive processes at horizons as long as 60 periods. Our simulation evidence suggests that there is little need for nonstandard interval estimators based on long-horizon asymptotics in many applications of impulse response analysis. This result is in stark contrast to earlier theoretical and bootstrap simulation results based on autoregressions that were not lag augmented (see Phillips 1998; Kilian and Chang 2000). We also formally show that other bootstrap confidence intervals for impulse responses based on lag-augmented autoregressions such as Hall's percentile interval or, for that matter, the delta method are not asymptotically valid at long horizons. Likewise, equal-tailed and symmetric percentile-t intervals are not asymptotically valid. This is the first example to our knowledge of a situation in which Efron's percentile interval is asymptotically valid for impulse response inference, but other intervals are not. Our results also provide a formal justification for conducting long-horizon inference based on autoregressions in levels rather than in differences.

In related work, Mikusheva (2012) proposed a generalization of the grid bootstrap of Hansen (1999) for autoregressions that allows uniformly asymptotically valid inference on individual impulse responses. The advantage of Mikusheva's procedure is that it nests as special cases the conventional normal approximation for short-horizon impulse response estimators and the nonstandard asymptotic approximation for long-horizon impulse response estimators, as discussed in Phillips (1998), Wright (2000), Gospodinov (2004) and Pesavento and Rossi (2006). Its disadvantage is that it tends to be computationally costly. Our simulations show that for roots arbitrarily close to or equal to unity the coverage accuracy of our computationally much less costly bias-adjusted bootstrap method based on lag-augmented autoregressions is close to nominal coverage for reasonably large sample sizes. It matches or surpasses the coverage accuracy of the grid bootstrap interval for comparable horizons, as reported in Mikusheva (2012).

The remainder of the paper is organized as follows. In section 2, we establish notation and state our assumptions about the data generating process and the estimated model. Section 3 contains the derivation of the uniform validity of the conventional asymptotic Gaussian approximation. We consider inference on individual impulse responses as well as vectors of impulse responses. In section 4, we establish the uniform asymptotic validity of inference based on the recursive-design bootstrap for autoregressions. In section 5, we examine the practical relevance of our asymptotic analysis in finite samples. In section 6, we provide long-horizon asymptotics based on the local-to-unity framework for impulse responses estimated from lag-augmented autoregressions. Section 7 contains the concluding remarks. Details of the proofs can be found in the appendix.

2 Notation and Assumptions

Consider a scalar autoregressive process of known order $p > 1$:

$$
\begin{aligned}
y_t &= d_t^\dagger + y_t^\dagger, \\
y_t^\dagger &= \phi_1 y_{t-1}^\dagger + \phi_2 y_{t-2}^\dagger + \cdots + \phi_p y_{t-p}^\dagger + u_t,
\end{aligned}
$$

where $d_t^\dagger$ is a deterministic function of time, u_t is iid with zero mean and variance σ^2 and $\Delta y_0^\dagger = \cdots = \Delta y_{1-p}^\dagger = 0$. Without loss of generality, we will focus on linear time trends, i.e., $d_t^\dagger = \delta_0^\dagger + \delta_1^\dagger (t/T)$.

This process has an augmented Dickey-Fuller representation:

$$\begin{aligned}\Delta y_t &= \delta_0 + \delta_1 \frac{t}{T} + \pi y_{t-1} + \gamma_1 \Delta y_{t-1} + \cdots + \gamma_{p-1}\Delta y_{t-p+1} + u_t \\ &= \beta' x_t + u_t, \end{aligned} \tag{1}$$

where $\delta_0 = \phi(1)\delta_0^\dagger + \delta_1^\dagger(\phi_1 + 2\phi_2 + \cdots + p\phi_p)$, $\delta_1 = \phi(1)\delta_1^\dagger$, $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\pi = -\phi(1) = \sum_{j=1}^p \phi_j - 1$, $\gamma_j = -(\phi_{j+1} + \cdots + \phi_p)$, $u_t \overset{iid}{\sim} (0, \sigma^2)$, $\beta = [\delta_0\ \delta_1\ \pi\ \gamma_1\ \cdots\ \gamma_{p-1}]'$, and $x_t = [1\ t/T\ \ y_{t-1}\ \Delta y_{t-1}\ \cdots\ \Delta y_{t-p+1}]'$. When $\phi_p = 0$, the AR(p) process underlying equation (1) without loss of generality may be reinterpreted as a lag-augmented autoregression of order $p-1$.

The autoregressive lag-order polynomial may equivalently be expressed as $\phi(L) = \Pi_{i=1}^p (1-\rho_i L)$, where $|\rho_1| \leq |\rho_1| \leq ... \leq |\rho_{p-1}| \leq |\rho_p|$ are the p autoregressive roots. Then $\omega = \sigma^2/(\phi^\dagger(1))^2$ where $\phi^\dagger(L) = \phi(L)/(1-\rho_p L)$ and ρ_p is the largest root. Let $\theta = [\beta', \sigma^2]'$. Let the parameter space $\Theta \subset \Re^{d_\theta}$ denote the set of θ where $d_\theta = p+3$. Finally, let $J_c(r)$ denote an Ornstein-Uhlenbeck process such that $J_c(r) = \int_0^r e^{c(r-s)} dW(s)$, where $W(s)$ is a standard Brownian motion defined on $[0,1]$ and $c = T\log(|\rho_p|)$. The model is estimated by least squares, yielding

$$\begin{aligned}\widehat{\beta}_T &= (\sum_{t=p+1}^T x_t x_t')^{-1} \sum_{t=p+1}^T x_t y_t, \\ \widehat{u}_t &= y_t - \widehat{\beta}_T' x_t, \\ \widehat{\sigma}_T^2 &= \frac{1}{T-p} \sum_{t=p+1}^T \widehat{u}_t^2, \\ \widehat{\sigma}_{4,T} &= \frac{1}{T-p} \sum_{t=1}^T (\widehat{u}_t^2 - \widehat{\sigma}_T^2)^2, \\ \widehat{\Sigma}_T &= \begin{bmatrix} \widehat{\sigma}_T^2 \otimes (\sum_{t=p+1}^T x_t x_t')^{-1} & 0_{(2+p)\times 1} \\ 0_{1\times(2+p)} & T^{-1}\widehat{\sigma}_{4,T} \end{bmatrix}. \end{aligned}$$

Assumption A. The data generating process satisfies:

(i) There are constants $\overline{\rho}$ and $\underline{\rho}$ in $(0,1)$ such that $|\rho_{p-1}| \leq \overline{\rho}$ and either $|\rho_p| \leq \overline{\rho}$ or $\underline{\rho} \leq \rho_p \leq 1$.

(ii) $\{u_t\}_{t=1}^T$ is a sequence of iid random variables with $E(u_t) = 0$ and

$$E\left\{ \begin{bmatrix} u_t \\ u_t^2 - \sigma^2 \end{bmatrix} \begin{bmatrix} u_t \\ u_t^2 - \sigma^2 \end{bmatrix}' \right\} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma_4 \end{bmatrix}.$$

(iii) There are constants κ and K, where $0 < \kappa < K < \infty$, that do not depend on the data generating process such that $\kappa \leq \sigma^2 \leq K$, $\kappa \leq \sigma_4 \leq K$, $\kappa \leq E(u_t^6) \leq K$, $-K \leq \delta_0^\dagger \leq K$, and $-K \leq \delta_1^\dagger \leq K$.

Remarks.

1. Assumption (i) implies that the roots of $|\phi(z)| = 0$ are either all outside the unit circle in modulus or that $\phi(z) = 0$ has at most one unit root and all the other roots are outside the unit circle in modulus. We rule out the possibility that the data are generated by an I(2) process or that the process is explosive. This assumption is standard in the literature (e.g., Mikusheva 2012). Assumption (i) also rules out complex near unit roots and roots near -1.

2. When the model is augmented with one lag, the population coefficient on that lag is known to be zero. Although the augmented lag parameter is estimated, the uniform coverage rate is defined as the limit of the infimum of the coverage probabilities with respect to the other parameters, with the augmented lag parameter fixed at zero. This coverage rate is greater than or equal to the uniform coverage rate in which the infimum is taken with respect to *all* parameters. Thus, without loss of generality, we focus on the latter.

3. We abstract from the complications introduced by conditional heteroskedasticity in the error term (see Gonçalves and Kilian 2004; Andrews and Guggenberger 2009).

4. We deliberately abstract from the lag order selection problem. As discussed in Kilian and Lütkepohl (2017), conditioning on estimates of the lag order invalidates the asymptotic validity of inference on the autoregressive parameters. One way of circumventing this problem is to set p equal to a conservative upper bound on the lag order, not unlike the upper bound that users of infomation criteria already have to provide when estimating the lag order.

3 Asymptotic Results for the Delta Method

Let $\widehat{\theta}_T$ denote the least-squares estimator of θ.

Proposition 1: Suppose that Θ satisfies Assumption A. Then

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} \sup_{x\in\Re^{d_\theta}} \left| P(\widehat{\Sigma}_T^{-\frac{1}{2}}(\widehat{\theta}_T - \theta) \leq x) - P(\eta(\theta, T) \leq x) \right| = 0, \tag{2}$$

where $\eta(\theta, T)$ is

$$\begin{cases} \left[\left(\Upsilon^{-1}(c) \begin{bmatrix} 1 & \frac{1}{2} & \omega \int_0^1 J_c(r)dr & 0 \\ \frac{1}{2} & \frac{1}{3} & \omega \int_0^1 rJ_c(r)dr & 0 \\ \omega \int_0^1 J_c(r)dr & \omega \int_0^1 rJ_c(r)dr & \omega^2 \int_0^1 J_c(r)^2 dr & 0 \\ 0 & 0 & 0 & M \end{bmatrix} \Upsilon^{-1}(c) \right)^{-\frac{1}{2}} \Upsilon^{-1}(c) \begin{bmatrix} N_1 \\ N_2 \\ \sigma\omega \int_0^1 J_c(r)dW(r) \\ N_3 \end{bmatrix} \right] & \text{if } |\rho_p| < 1, \\ \quad N_4 & \\ N & \text{if } \rho_p = 1, \end{cases} \tag{3}$$

$\Upsilon(c) = \text{diag}(1, 1, \frac{1}{\sqrt{-2c}}, I_{p+1}, 1)$, $c = T\log(|\rho_p|)$, ρ_p is the largest root of $\phi(z) = 0$, $[N_1\ N_2]'$, N_3 and N_4 are independent normal random vectors with zero means and covariance matrices given by

$$\sigma^2 \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix}, \ \sigma^2 M \ \equiv \ \sigma^2 E \left(\begin{bmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p+1} \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p+1} \end{bmatrix}' \right) \text{ and } \sigma_4, \tag{4}$$

respectively, and N is the standard normal random vector that is the limit of $\eta(\theta, T)$ as $\rho_p \uparrow 1$.

Express the $d_\psi \times 1$ vector of impulse responses ψ as a function of θ:

$$\psi \ = \ f(\theta), \tag{5}$$

where $f : X \to \Re^{d_\psi}$ and $\Theta \subset X \subset \Re^{p+3}$. Our goal is to provide methods for uniform inference on impulse responses in Θ.

Assumption B. Suppose that $f : X \ \to \Re^{d_\psi}$ is continuously differentiable, that ψ does not depend on $\delta_0^\dagger$ and $\delta_1^\dagger$, and that the rank of

$$Df(\theta)diag(I_2, \sqrt{-2\log(|\rho_p|)}, I_{p+1}) \tag{6}$$

is d_ψ for all $\theta \in \Theta$ where $Df(\theta) = \partial f(\theta)/\partial\theta'$ and ρ_p is the largest autoregressive root.

Remarks.

1. A violation of Assumption B would occur, for example, if the autoregressice parameters were zero in population and the impulse response horizon $h > p$ (see Benkwitz, Lütkepohl and Neumann 2000). We abstract from this well-known problem, as is standard in the literature, since we are concerned with inference about impulse responses estimated from persistent time

series processes.

2. Likewise, assumption B may fail for some exact unit root processes. Specifically, standard delta method and bootstrap inference fails when the first-order linear approximation to ψ is proportionate to ρ. In that case, the limiting variance of $\sqrt{T}(\widehat{\psi} - \psi)$ is zero, as discussed in Kilian and Lütkepohl (2017), and the rank condition fails. Lag augmenting the autoregression helps rule out this singularity in the asymptotic variance of $f(\theta)$.

3. To appreciate the usefulness of the rank condition in Assumption B, consider the AR(2) process

$$\Delta y_t \;=\; \pi y_{t-1} + \gamma \Delta y_{t-1} + u_t,$$

where $\pi = \phi_1 + \phi_2 - 1$ and $\gamma = -\phi_2$, and the first two impulse responses are ϕ_1 and $\phi_1^2 + \phi_2$. Then

$$Df(\theta) \;=\; \begin{bmatrix} 1 & 1 \\ 2(\pi+\gamma+1) & 2(\pi+\gamma+1)-1 \end{bmatrix}$$

Note that the rank condition for the delta method is always satisfied. The matrix (6) can be written as

$$T^{\frac{1}{2}} Df(\theta)\Upsilon_T^{-1}(c) \;=\; \begin{bmatrix} \sqrt{-2\log(|\rho_2|)} & 1 \\ 2\sqrt{-2\log(|\rho_2|)}(\pi+\gamma+1) & 2(\pi+\gamma+1)-1 \end{bmatrix},$$

where the first two columns are omitted because there is no deterministic component. This expression shows that if one is interested in inference on the first impulse response, the rank condition is always satisfied. As $\phi_1, \phi_2 \to 1/2$, however, the second row approaches zeros, so inference about the second impulse response is not possible. Even if the conventional rank condition is satisfied, we can conduct uniform inference only on one of the parameters of interest. Replacing the normalization matrix $\Upsilon_T(c)$ so that the scaled Jacobian matrix has full rank does not solve the problem because the joint asymptotic normality is likely to be lost under a different normalization.

In contrast, when fitting an AR(3) model to the data generated by the AR(2) DGP, the rank of the matrix (6),

$$\begin{bmatrix} \sqrt{-2\log(|\rho_3|)} & 1 & 0 \\ 2\sqrt{-2\log(|\rho_3|)}(\pi+\gamma_1+1) & 2(\pi+\gamma_1+1)-1 & 1 \end{bmatrix},$$

is always 2. Thus, lag augmentation allows inference about the second impulse response as well as joint inference about both of the impulse responses.

4. This example may be generalized. It can be shown that the rank condition in Assumption B is satisfied for the first p impulse responses for all θ in the parameter space specified in Assumption A and for all $p = 1, 2, ...,$ when the autoregression is augmented by one lag. This is true even for the processes described in Benkwitz et al. (2000). Suppose that y_t follows an $AR(p)$ process. The companion matrix for the first p coefficients of the lag-augmented model is given by

$$F = \begin{bmatrix} \pi+\gamma_1+1 & \gamma_2-\gamma_1 & \gamma_3-\gamma_2 & \cdots & \gamma_p-\gamma_{p-1} & \gamma_{p+1}-\gamma_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \quad (7)$$

Then the h-step-ahead impulse response is given by the (1,1) element of F^h. The responses are given by

$$\begin{aligned} &\pi+\gamma_1+1, \\ &(\pi+\gamma_1+1)^2+\gamma_2-\gamma_1, \\ &(\pi+\gamma_1+1)^3+2(\pi+\gamma_1+1)(\gamma_2-\gamma_1)+(\gamma_3-\gamma_2), \\ &\vdots \\ &(\pi+\gamma_1+1)^p+p(\gamma_{p+1}-\gamma_p)^{p-1}(\gamma_2-\gamma_1)+\cdots+\gamma_{p+1}-\gamma_p, \end{aligned}$$

for $h = 1, 2, 3, ..., p$, respectively. We are concerned about uniform joint inference about the first p impulse responses. Our claim is that the submatrix M_p obtained from eliminating the first three columns of the $p \times (p+3)$ matrix in Assumption B has rank p. Note that the first $p-1$ impulse responses are identical to those from an $AR(p)$ model with $\gamma_p = 0$. Thus, the $(p-1) \times (p-1)$ upper-left submatrix of M_p matches the corresponding $(p-1) \times (p-1)$ submatrix for the $AR(p)$ model that is obtained from lag-augmenting an $AR(p-1)$ model. We prove this claim by mathematical induction. When $p = 1$ (i.e., the DGP is an AR(1) process and an AR(2) model is fitted), the 1×2 matrix in Assumption B always has rank 1 satisfying Assumption B. Suppose that the rank condition is satisfied for $p = k$. That is, the $k \times k$ submatrix has rank k. Denote that matrix by M_k. Then the $(k+1) \times (k+1)$ submatrix for the $AR(k+1)$ model can be written as

$$\begin{bmatrix} M_k & 0_{k\times 1} \\ 0_{1\times k} & 0 \end{bmatrix} + \begin{bmatrix} 0_{k\times k} & 0_{k\times 1} \\ 0_{1\times(k-1)} \; -1 & 1 \end{bmatrix}. \tag{8}$$

Because M_k has rank k, this matrix has rank $k+1$ anywhere in the parameter space specified in Assumption A. Thus, the claim holds for $p = k+1$. Since this result holds for the first p impulse responses jointly, it holds also for individual elements in this vector.

Proposition 2. Under Assumptions A and B,

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} \left| P\left((Df_T(\widehat{\theta}_T)\widehat{\Sigma}_T Df_T(\widehat{\theta}_T)')^{-\frac{1}{2}}(f(\widehat{\theta}_T) - f(\theta)) \leq x \right) - \Phi(x) \right| = 0, \tag{9}$$

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} \left| P\left((f(\widehat{\theta}_T) - f(\theta))'(Df(\widehat{\theta}_T)\widehat{\Sigma}_T Df_T(\widehat{\theta}_T)')^{-1}(f(\widehat{\theta}_T) - f(\theta)) \leq x \right) - F_{\chi^2_{d_\psi}}(x) \right| = 0, \tag{10}$$

where $F_{\chi^2_{d_\psi}}(\cdot)$ is the cdf of the chi-square distribution with d_ψ degrees of freedom.

It follows from Lemma 2 of Kasy (2018) that confidence sets constructed from quantiles of the standard normal and chi-square distributions have confidence level $1-\alpha$ uniformly on Θ.

4 Asymptotic Results for Bootstrap Inference

Bootstrap approximations of the asymptotic distribution of the impulse response estimators may be generated by standard recursive residual-based bootstrap algorithms for autoregressions (see Kilian and Lütkepohl 2017). Let $\widehat{\theta}^*_T$ denote the bootstrap estimator of $\widehat{\theta}_T$, constructed by bootstrapping the original or the lag-augmented autoregressive model. Similarly, let P^* denote the bootstrap analogue of P.

Proposition 3. Under Assumptions A and B,

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} |P^*(\Upsilon_T(c^*)(\widehat{\theta}^*_T - \widehat{\theta}_T) \leq x) - P^*(\eta^*(\widehat{\theta}_T, T) \leq x)| = 0, \tag{11}$$

almost surely conditional on the data, where $c^* = c + \int_0^1 J_c^\tau(r)dW(r)/\int_0^1 (J_c^\tau(x))^2 dr$, $J_c^\tau(r) = J_c(r) - \int_0^1 (4-6s)J_c(s)ds - r\int_0^1 (12s-6)J_c(s)ds$ and $\eta^*(\cdot,\cdot)$ is $\eta(\cdot,\cdot)$ with c replaced by c^*.

Proposition 4. Under Assumptions A and B,

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} |P^*([(Df(\widehat{\theta}_T^*)\widehat{\Sigma}_T^* Df(\widehat{\theta}_T^*)')^{-\frac{1}{2}} f(\widehat{\theta}_T^*) - f(\widehat{\theta})] \le x) - \Phi(x)| \;=\; 0, \tag{12}$$

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} |P((f(\widehat{\theta}_T^*) - f(\widehat{\theta}_T))(Df(\widehat{\theta}_T^*)\widehat{\Sigma}_T^* Df_T(\widehat{\theta}_T^*)')^{-1}(f(\widehat{\theta}_T^*) - f(\widehat{\theta}_T)) \le x)) - F_{\chi^2}(x)| \;=\; 0, \tag{13}$$

almost surely conditional on the data.

To summarize, Propositions 1 through 4, extend the pointwise asymptotic results in Park and Phillips (1989), Sims, Stock and Watson (1990), and Inoue and Kilian (2002) by establishing the uniform validity of asymptotic and bootstrap inference about individual slope parameters and impulse responses based on higher-order autoregressions. They also establish the corresponding results for asymptotic and bootstrap inference based on lag-augmented autoregressions. Finally, they establish the uniform validity of asymptotic and bootstrap inference based on lag-augmented autoregressions about vectors of impulse responses.

5 Simulation Evidence

In this section, we demonstrate that our asymptotic analysis helps understand the finite-sample accuracy of delta method and bootstrap confidence intervals for impulse responses. Without loss of generality, we generate 5,000 samples of $\{y_t\}_{t=1}^T$ from the data generating process $y_t = \rho y_{t-1} + u_t$, $u_t \overset{iid}{\sim} N(0,1)$, where $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$, $T \in \{80, 120, 240, 480, 600, 6000, 12000, 24000\}$, and $y_0 = 0$. We focus on roots exceeding 0.95 because for smaller roots conventional bootstrap approximations are known to work well (see Kilian 1999).[6] For each sample of length T, we fit an AR(p) model, $p \in \{2,4,6\}$, with intercept and construct the implied responses to a unit shock at horizons $h \in \{1,...,60\}$. Lag-augmented autoregressions include an additional lag, but the impulse responses are based on the estimates of the first p slope coefficients only. We do not

[6] Alternative specifications of the data generating process with standardized t_4 or standardized χ_3^2 errors, as in Kilian (1998a), yield results very similar to the baseline specification with $N(0,1)$ errors and, hence, are not shown to conserve space. These specific distributions were chosen because their moments resemble those of residual distributions often encountered in applied work (see Kilian 1998a). Although the standardized t_4 distribution does not satisfy our sufficient condition A(iii), the simulation results are robust to this violation.

include a linear time trend in the fitted model because the inclusion of deterministic time trends is rare in applied work.[7]

Since the results are not sensitive to the lag order p, the tables shown in this section concentrate on the case of $p = 4$. Our analysis focuses on confidence intervals for individual impulse responses. The nominal confidence level is 90%. In constructing the uniform coverage rates as the infimum of the coverage rates for a given impulse response across $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$ for given T, one has to account for the bias caused by data mining across ρ. The reason is that the coverage rates in the simulation study (like any estimate of a proportion) are subject to estimation error. They have an approximate Gaussian distribution. Thus, even if the estimate of the coverage rate were centered on 0.90 for each ρ, there would be sampling variation in the simulated coverage rates. It can be shown by that under the null hypothesis that the coverage is truly 0.90, the lowest coverage rate across all ρ would be 0.894, which biases downward our estimate of the uniform coverage rate. This means that we need to adjust upward the infimum across ρ obtained in the simulation by 0.006 to control for data mining. This adjustment is independent of the sample size because it only reflects the Monte Carlo simulation error. Details of the rationale of this adjustment can be found in the appendix.[8]

The delta method intervals are based on closed-form solutions for the impulse-response standard error, as discussed in Lütkepohl (1990). Table 1 shows that the uniform coverage rates of the delta method interval converge to 0.90, as $T \to \infty$, as predicted by asymptotic theory, whether inference is based on the AR(4) model or the lag-augmented AR(5) model. There is strong evidence that delta method intervals based on the lag-augmented AR(5) model are considerably more accurate in small samples than delta method intervals based on the AR(4) model. For example, for $T = 480$, the uniform coverage accuracy at horizon 12 is 86% for the lag-augmented model compared with only 63% for the original model. These differences are not predicted by our asymptotic analysis in section 3. For large T, as expected, there is nothing to choose between these approaches.

Not surprisingly, the coverage accuracy is excellent at short horizons, but deteriorates as h increases, except when T is large. This finding mirrors the conclusions of Kilian and Chang (2000) and Phillips (1998) that the conventional asymptotic approximation remains accurate, as long as

[7] Likewise, we do not consider autoregressions excluding an intercept. Normal asymptotic appoximations tend to work better when the regression model does not include an intercept because the exclusion of deterministic regressors reduces the small-sample bias of the least-squares estimator. This regression specification is hardly ever used in applied work, however.

[8] An alternative approach would have been to view ρ as local to unity and to report results for the implied ρ, given T and a grid of Pitman drifts. Since our asynptotic results do not hinge on this particular asymptotic thought experiment, it is more natural to focus on the grid of possible ρ values in the simulation study.

the horizon is small relative to the sample size. Only when the horizon of the impulse response is allowed to grow with the sample size, conventional asymptotic approximations for impulse response estimators become asymptotically invalid in a local-to-unity setting. Table 1 illustrates that even for horizons as large as $h = 12$, for moderately large samples, the conventional asymptotic Gaussian approximation remains reasonably accurate. For example, for $T = 480$, coverage rates for the lag-augmented model range from 90% at short horizons to 86% at horizon 12.

An important question is whether the accuracy of impulse response inference may be improved by bootstrapping the impulse responses. In Table 2, we examine the coverage accuracy of several commonly used bootstrap confidence intervals.[9] The distribution of impulse response estimators is known to be non-normal in small samples (see Kilian 1999). The first two panels in Table 2 show results based on the Hall percentile interval which accounts for small-sample bias in the impulse response estimator and which does not require normality to hold (see Hall 1992). The bootstrap data are generated based on a recursive design-bootstrap, as discussed in Kilian and Lütkepohl (2017). All results are based on 1,000 bootstrap replications. Table 2 shows that bootstrap confidence intervals greatly improves the accuracy of inference based on the AR(4) model. For example, the uniform coverage rates for $T = 480$ range from 90% at horizon 1 to 85% at horizon 12. In contrast, for the lag-augmented model bootstrap inference does not yield improved uniform coverage accuracy. Thus, overall, the Hall percentile interval cannot be recommended.

An alternative approach that has been shown to work well in bootstrapping stationary autoregressions is the bias-adjusted bootstrap of Kilian (1999), which replaces the least-squares estimates of the slope parameters by first-order mean bias-adjusted estimates when implementing the bootstrap. Impulse response intervals are based on the standard Efron percentile interval (see Efron 1979). Table 2 shows that this method greatly improves the uniform coverage accuracy of the bootstrap confidence intervals, whether the model is lag-augmented or not, but by far the most accurate coverage rates are obtained based on the lag-augmented model. For $T = 80$, the coverage rates are between 80% and 87%, depending on the horizon. For $T = 120$, the coverage accuracy improves to between 83% and 88%. For $T = 240$, they are at least 87% and for $T = 480$ and $T = 600$ at least 89%. This evidence suggests that the conventional asymptotic approximation remains accurate at longer horizons than previously thought possible. Performance deteriorates, when the autoregression is not lag-augmented, whether inference is based on the delta method or

[9] A general introduction to bootstrap methods for autoregressions and details of the construction of each of these intervals can be found in Kilian and Lütkepohl (2017).

the bootstrap.

Our coverage results suggest that highly persistent autoregressions in applied work should be routinely lag-augmented. While lag-augmenting the autoregression will cause an increase in the average width of the interval, a reasonable conjecture is that this loss in efficiency may for all practical purposes be ignored, when the original autoregressive lag order is already large, as is typically the case in applied work. Since lag augmentation is necessary to control coverage accuracy in finite samples, it is difficult in general to compare the average width of intervals based on the original and on the lag-augmented model. Some insight may be gained, however, by comparing the average width of the Efron percentile interval at short horizons, because at these horizons both the intervals based on the bias-adjusted lag-augmented autoregression and the intervals based on the bias-adjusted original autoregression are about equally accurate (see Table 2). Table 3 shows the percentage increase in the average interval width at these horizons, computed as the average of the percentage increases in average interval width obtained for each ρ. We find that the loss in power from lag augmentation tends to be negligible. Even for $T = 80$ and $T = 120$ the average interval width increases by only 1.5% and 1%, respectively, when a fifth autoregressive lag is added. For $T = 240$, that increase drops to 0.5% and for larger samples sizes the increase further reduces to 0.2%, consistent with our conjecture.

An important question is how quickly the accuracy of our asymptotic approximation deteriorates with the impulse response horizon. Table 4 shows that the coverage accuracy of the Efron percentile interval based on the bias-adjusted lag-augmented autoregression is preserved even at much longer horizons. For example, for $T = 240$ uniform coverage accuracy of the bootstrap confidence interval for the lag-augmented model is at least 88% at every impulse response horizon from 12 to 36. For $T = 480$ the lowest uniform coverage rate at these horizons is 89% and for $T = 600$ it is 90%. Even for horizons as long as 60, the coverage accuracy remains excellent.[10]

These results suggest that for many applications of impulse response analysis there is no need to rely on nonstandard interval estimators based on long-horizon asymptotics for impulse responses, as long as we apply the bias-adjusted bootstrap method to the lag-augmented autoregression. The superior accuracy of this method at longer horizons is not explained by our fixed-horizon

[10]Our analysis in Tables 1 through 4 focused on nominal 90% intervals. Confidence levels of 90% or smaller are conventional for element-wise impulse response inference. In constructing joint confidence intervals based on the Bonferroni method, it is customary to rely on much higher element-wise confidence levels (see Lütkepohl et al. 2015, 2018). It is therefore useful to note that, for $T = 240$ or larger, nominal 99% impulse response confidence intervals based on our preferred method have effective uniform coverage rates of between 98% and 99% at all horizons considered.

asymptotics in sections 3 and 4, however. In the next section, we formally establish the asymptotic validity of this method (and this method alone) under the assumption that the impulse response horizon increases linearly with the sample size, which helps explain its greater robustness to the impulse response horizon.

6 Impulse Responses at Long Horizons

In this section, we show that Efron's percentile interval bootstrap is asymptotically valid for long-horizon impulse response inference when there is a near unit root. We first consider the AR(1) model for illustration:

$$y_t \;=\; \phi_1 y_{t-1} + u_t, \tag{14}$$

where $\phi_1 = e^{c/T}$ for some constant $c \leq 0$, $y_0 = 0$ and $u_t \overset{iid}{\sim} (0, \sigma^2)$. As is well known, the estimator $\widehat{\phi}_1$ in this model has a nonstandard distribution, as does the $[\lambda T]$-step-ahead impulse response $\widehat{\phi}_1^{[\lambda T]}$. In contrast, in the lag-augmented model

$$y_t \;=\; \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t, \tag{15}$$

which may equivalently be expressed as

$$\Delta y_t \;=\; \pi y_{t-1} + \gamma \Delta y_{t-2} + u_t, \tag{16}$$

$\sqrt{T}(\widehat{\phi}_{1,T} - \phi_1) \;=\; \sqrt{T}(\widehat{\pi}_T + \widehat{\gamma}_T - \pi - \gamma)$ is asymptotically normally distributed.

Let z and z^* be random variables such that $\sqrt{T}(\widehat{\gamma} - \gamma) \overset{d}{\to} z$ and $\sqrt{T}(\widehat{\gamma}^* - \widehat{\gamma}) \overset{d}{\to} z^*$, where * denotes random variables defined under the bootstrap probability measure. The linear approximation to the limiting distribution of the impulse response depends on the limit of the Jacobian. The Jacobian of $\widehat{\phi}_1^{[\lambda T]}$, $[\lambda T](\widehat{\phi}_1)^{[\lambda T]-1}$, converges to zero when z is negative and diverges to infinity when z is positive. In contrast, the bootstrap version of the Jacobian $[\lambda T](\widehat{\phi}_1^*)^{[\lambda T]-1}$ converges to zero when $z + z^*$ is negative and diverges to infinity when $z + z^*$ is positive. Thus, Hall's percentile interval fails in the lag-augmented model.

Moreover, the delta method interval fails because it is based on the t-statistic

$$\frac{\widehat{\phi}_1^{[\lambda T]} - \phi_1^{[\lambda T]}}{[\lambda T]\widehat{\phi}_1^{[\lambda T]-1}\widehat{ASE}(\widehat{\phi}_1)} \tag{17}$$

where $\widehat{ASE}(\widehat{\phi}_1)$ is the estimate of the asymptotic standard error of $\widehat{\phi}_1$. Expressing the numerator of the t-statistic as

$$(\widehat{\phi}_1^{[\lambda T]-1} + \widehat{\phi}_1^{[\lambda T]-2}\phi_1 + \cdots + \phi_1^{[\lambda T]-1})(\widehat{\phi}_1 - \phi_1), \tag{18}$$

the t-statistic can be expressed as

$$\frac{(\widehat{\phi}_1^{[\lambda T]-1} + \widehat{\phi}_1^{[\lambda T]-2}\phi_1 + \cdots + \phi_1^{[\lambda T]-1})}{[\lambda T]\widehat{\phi}_1^{[\lambda T]-1}} \cdot \frac{\widehat{\phi}_1 - \phi_1}{\widehat{ASE}(\widehat{\phi}_1)}. \tag{19}$$

Note that the second component of (19) is the t-statistic for ϕ_1 and is asymptotically normally distributed in the lag-augmented model. The first component of (19) can be written as

$$\begin{aligned} & \frac{1}{[\lambda T]}\left(1 + \frac{\phi_1}{\widehat{\phi}_1} + \left(\frac{\phi_1}{\widehat{\phi}_1}\right)^2 + \cdots + \left(\frac{\phi_1}{\widehat{\phi}_1}\right)^{[\lambda T]-1}\right) \\ = & \frac{1}{[\lambda T]}\frac{1 - (\phi_1/\widehat{\phi}_1)^{[\lambda T]}}{1 - (\phi_1/\widehat{\phi}_1)} \\ = & \frac{1 - (\phi_1/\widehat{\phi}_1)^{[\lambda T]}}{[\lambda T]/\sqrt{T}} \cdot \frac{\widehat{\phi}_1}{\sqrt{T}(\widehat{\phi}_1 - \phi_1)}. \end{aligned} \tag{20}$$

Because

$$\frac{\phi_1}{\widehat{\phi}_1} = \frac{1}{1 + \frac{1}{\sqrt{T}}\frac{\sqrt{T}(\widehat{\phi}_1 - \phi_1)}{\phi_1}} \tag{21}$$

it follows that

$$\frac{1}{[\lambda T]/\sqrt{T}}\left(1 - \left(\frac{\phi_1}{\widehat{\phi}_1}\right)^{[\lambda T]}\right) = -\frac{1}{[\lambda T]/\sqrt{T}}\left(1 + \frac{1}{\sqrt{T}}\frac{\sqrt{T}(\widehat{\phi}_1 - \phi_1)}{\phi_1}\right)^{[\lambda T]} + o(1).$$

Thus, depending on the sign of the limit z, the first component either converges to zero or diverges to infinity. The delta method fails because the t-statistic (17) does not converge to a standard normal distribution.

Similarly, intervals based on bootstrapping the t-statistic fail. The bootstrap-t statistic can be written as

$$
\begin{aligned}
&\frac{1}{[\lambda T]/\sqrt{T}}\left(1+\frac{1}{\sqrt{T}}\frac{\sqrt{T}(\widehat{\phi}_1^*-\widehat{\phi}_1)}{\widehat{\phi}_1}\right)^{[\lambda T]}\\
&\times\frac{\widehat{\phi}_1^*}{\sqrt{T}(\widehat{\phi}_1^*-\widehat{\phi}_1)}\\
&\times\frac{\widehat{\phi}_1^*-\widehat{\phi}_1}{\widehat{ASE}^*(\widehat{\phi}_1^*)}.
\end{aligned}
\tag{22}
$$

The percentile-t bootstrap interval fails because the original t-statistic either converges to zero or diverges to infinity, whereas the limiting value of the bootstrap-t statistic depends on the sign of z^*, which may differ from the sign of z.

In contrast, impulse response inference based on Efron's percentile interval remains asymptotically valid at long horizons, because, as the horizon lengthens, the impulse response can be expressed as a monotonic function of the asymptotically normal estimator $\widehat{\phi}_{1,T}$. Unlike other confidence intervals, the Efron percentile interval is transformation-respecting. In other words, the interval for a given monotonic transformation of the original parameter may be obtained by transforming the interval endpoints obtained for the original parameter using the same function (see Efron 1979). The implications of this point for long-horizon impulse response inference are formalized in the following proposition.

Assumption C. The data generating process satisfies:

(i) $|\rho_j| < 1$ for $j = 1, ..., p-1$ and $\rho_p = e^{c/T}$ for some constant $c \leq 0$.

(ii) $\{u_t\}_{t=1}^T$ is a sequence of iid random variables with $E(u_t) = 0$ and $E(u_t^2) < \infty$.

Proposition 5: Suppose that the DGP is an AR(p) model and that Assumption C is satisfied. The AR(p) model is augmented by one lag. Let $\widehat{F}$ and $\widehat{F}^*$ denote the companion matrices for the first p coefficients of the estimated lag-augmented model and its bootstrap analogue,

$$
\widehat{F} = \begin{bmatrix} \widehat{\phi}_1 & \widehat{\phi}_2 & \cdots & \widehat{\phi}_{p-1} & \widehat{\phi}_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad \widehat{F}^* = \begin{bmatrix} \widehat{\phi}_1^* & \widehat{\phi}_2^* & \cdots & \widehat{\phi}_{p-1}^* & \widehat{\phi}_p^* \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},
$$

respectively. The h-step-ahead impulse response estimate and bootstrap estimate are the (1,1) elements of $\widehat{F}^h$ and $\widehat{F}^{*h}$, respectively.

F can be written as

$$F = PJP^{-1}, \tag{23}$$

where J is the Jordan normal form of F and P consists of eigenvectors and generalized eigenvectors of P. Thus

$$F^h = PJ^hP^{-1} \tag{24}$$

The hth power of the Jordan normal form is given by

$$J^h = \begin{bmatrix} \rho_p^h & 0 & \cdots & 0 \\ 0 & J_{m_2}^h(\rho_{p-1}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_q}^h(\rho_1) \end{bmatrix}, \tag{25}$$

where m_j is the multiplicity of the jth largest root with $m_1 = 1$ such that $\sum_{j=1}^q m_j = p$, and

$$J_{m_j}^h(\rho_k) = \begin{bmatrix} \rho_k^h & \binom{h}{1}\rho_k^{h-1} & \binom{h}{2}\rho_k^{h-2} & \cdots & \binom{h}{m_j-1}\rho_k^{h-m_j+1} \\ 0 & \rho_k^h & \binom{h}{1}\rho_k^{h-1} & \cdots & \binom{h}{m_j-2}\rho_k^{h-m_j+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \rho_k^h \end{bmatrix} \tag{26}$$

for $j = 1, 2, ..., q$. Because there is one and only one local-to-unity root (ρ_p), the (1,1) element of $F^{[\lambda T]}$ can be approximated by

$$P_{11}\rho_p^{[\lambda T]}P^{11} + o(1), \tag{27}$$

where P_{11} is the (1,1) element of P and P^{11} is the (1,1) element of P^{-1}. Similarly, the (1,1) element of $\widehat{F}^{[\lambda T]}$ and that of $\widehat{F}^{*[\lambda T]}$ can be approximated by

$$\widehat{P}_{11}\widehat{\rho}_p^{[\lambda T]}\widehat{P}^{11} + o_p(1), \tag{28}$$

$$\widehat{P}_{11}^*\widehat{\rho}_p^{*[\lambda T]}\widehat{P}^{11*} + o_p^*(1), \tag{29}$$

respectively. Taking the log on both sides

$$\log(\widehat{P}_{11}) + \log(\widehat{P}^{11}) + [\lambda T] \log(\widehat{\rho}_p) + o_p(1), \tag{30}$$

$$\log(\widehat{P}^*_{11}) + \log(\widehat{P}^{11*}) + [\lambda T] \log(\widehat{\rho}^*_p) + o^*_p(1). \tag{31}$$

Because there is one and only one local-to-unity root (ρ_p), $\widehat{\rho}_p$ and $\widehat{\rho}^*_p$ are continuously differentiable in $(\widehat{\phi}_1, ..., \widehat{\phi}_p)$ and $(\widehat{\phi}^*_1, ..., \widehat{\phi}^*_p)$, respectively. Thus, $\widehat{\rho}_p$ is asymptotically normally distributed in the lag-augmented model. Because Efron's percentile bootstrap method is transformation-respecting and because $\widehat{\rho}_p$ is asymptotically normally distributed, the percentile bootstrap remains asymptotically valid, when other intervals fail. Since bias adjustments of the slope parameters are of order T, this argument remains valid when using Efron's interval in conjunction with bias adjustments (see Kilian 1998b).

It is useful to illustrate these points by simulation. Table 5 restricts the impulse response horizon to be a fraction $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ of the sample size T. We focus on impulse response inference based on the lag-augmented autoregression, since impulse response inference based on the original autoregression fails at long horizons, regardless of how the interval is constructed. This fact follows from Phillips (1998) and may easily be verified by simulation. The more interesting question is how well our asymptotic approximation works for the lag-augmented AR(4) model. As predicted by our theoretical analysis, conventional delta method inference breaks down at long horizons, even when working with the lag-augmented model. Even for $T = 600$, the coverage rates of the nominal 90% delta method interval remain between 49% and 68%, depending on λ. Similarly, the coverage rates of the Hall percentile interval range from 43% to 52% for $T = 600$, illustrating the failure of this method. Likewise, equal-tailed and symmetric percentile-t impulse response intervals show no tendency to converge to their nominal probability content (results not shown to conserve space). In contrast, even without bias adjustments, the coverage rates of Efron's percentile interval improve range about 85% for $T = 240$ to 89% for $T = 600$. Applying bias adjustments for the slope parameters, further increases the finite-sample accuracy. The coverage accuracy increases to about 89% for $T = 240$ and 90% for $T = 600$, regardless of λ.

Remarks.

1. Proposition 5 establishes the asymptotic validity of impulse response inference based on Efron's percentile interval within the local-to-unity framework. It does not establish its

asymptotic validity in the stationary region. Thus, long-horizon inference is not uniformly valid in the parameter space. It is valid only for roots close to unity. Further simulations (not shown to conserve space) suggest that for reasonably large samples our asymptotic approximation is excellent for roots of 0.8 or larger. An immediate implication is that impulse response inference is likely to be more reliable at long horizons when persistent autoregressive processes are expressed in levels rather than in differences.

2. Proposition 5 may be generalized to vector autoregressive processes, as long as there is only one large root, as is commonly assumed in related studies (see, e.g., Pesavento and Rossi 2006; Mikusheva 2012). For a potential alternative approach that allows for multiple large roots see Phillips and Lee (2016) and the references therein.

3. In related work, Mikusheva (2012) proposes a generalization of the grid bootstrap of Hansen (1999) for autoregressions that allows inference on individual impulse responses that, like our approach, is uniformly asymptotically valid in the parameter space. The advantage of Mikusheva's asymptotic approximation is that it nests as special cases the conventional normal approximation for short-horizon impulse response estimators and the nonstandard asymptotic approximation for long-horizon impulse response estimators, as proposed by Phillips (1998), Wright (2000), Gospodinov (2004) and Pesavento and Rossi (2006). The disadvantage of Mikusheva's procedure is that its computational cost tends to be prohibitive for all but the simplest autoregressive processes.[11] Our approach provides a computationally less costly alternative to Mikusheva's grid bootstrap in many applied settings for short as well as long horizons. For example, at horizons up to 12 periods, even for $T = 240$, the infimum of the impulse response coverage rates based on our conventional bootstrap asymptotics for lag-augmented autoregressive models ranges from 87% to 89%. For $T = 480$, the coverage rates reach 89% to 90%, depending on the horizon. The latter coverage rates are at least as accurate as the grid bootstrap coverage rates for $T = 500$ reported in Mikusheva (2012), which range from 87% to 92% at similar horizons.

7 Concluding Remarks

Although impulse response inference has played an important role in macroeconometrics since the

[11] The largest process considered by Mikusheva (2012) in her simulation analysis is an AR(2) process. Her method does not appear to have been applied to autoregressions with more lags.

1980s, all existing proofs of the asymptotic validity of conventional delta method and bootstrap confidence intervals are based on pointwise Gaussian asymptotic approximations. As has been shown for the AR(1) model, this approximation may fail when the dominant root of the autoregressive process is near unity, resulting in intervals with poor coverage accuracy. In this paper, we showed that the failure of conventional confidence intervals, as the dominant autoregressive roots approaches unity, does not extend to higher-order autoregressions. We established the uniform asymptotic validity of conventional asymptotic and bootstrap inference about individual impulse responses and vectors of impulse responses at fixed horizons. We showed that for inference about vectors of impulse responses based on Wald test statistics to be uniformly valid in the parameter space, autoregressions must be lag augmented. Inference about individual impulse responses, in contrast, under weak conditions is uniformly valid even without lag augmentation.

We further documented that the conventional asymptotic approximation works well in moderately large samples, as long as the impulse response horizon remains reasonably small relative to the sample size. The highest small-sample accuracy is achieved when bootstrapping the lag-augmented autoregressive model using the bias-adjusted bootstrap method of Kilian (1999). For horizons between 1 and 12 periods, for example, this approach achieves uniform coverage rates of between 87% and 90% for $T = 240$ and between 89% and 90% for $T = 480$ and $T = 600$, which is a substantial improvement relative to the delta method for the same model. In contrast, without the lag augmentation, both delta method and bootstrap confidence intervals are much less accurate, consistent with earlier simulation evidence in the literature. We provided formal asymptotic arguments why our preferred method of inference based on lag-augmented autoregressions retains its accuracy even at very long impulse response horizons, when other methods do not. Although the latter result does not hold uniformly across the parameter space, it does hold in a local-to-unity setting.

These results suggest that highly persistent autoregressions in applied work should be routinely lag-augmented when conducting impulse response analysis. While lag-augmenting the autoregression will cause an increase in the average width of the interval, we showed that the loss in efficiency caused by including one extra lag may for all practical purposes be ignored, when the original autoregressive lag order is already large, as is common in applied work.

Our approach provides a highly accurate alternative to the much more computationally costly grid bootstrap of Mikusheva (2012) in many applied settings. In fact, the coverage accuracy of the bias-adjusted bootstrap remains excellent at horizons as long as 60 periods for sample sizes as small as $T = 240$. Thus, we overturn the standard finding that conventional bootstrap confidence intervals become increasingly inaccurate at longer horizons, even when using bias adjustments. For example, Kilian and Chang (2000) documented by simulation that the coverage accuracy of conventional methods of impulse response inference is inadequate for all but the shortest horizons. The reason for this difference in results was our use of the lag-augmented autoregressive model, which is new in impulse response inference.

Appendix A

To prove Proposition 1, we follow the steps taken in Mikusheva (2007a,b). First we show that the estimation uncertainty about the asymptotic covariance matrix is asymptotically negligible (Lemma B1). Next, we show that the distribution of the least-squares estimator can be uniformly approximated by that based on Gaussian autoregressive processes (Lemma B2). Third, we show that the latter can be uniformly approximated by the local-to-unity asymptotic distribution (Lemma B3). Proposition 2 follows from Proposition 1 and the rank condition in Assumption B.

As in Mikusheva (2007a, 2012), we split the parameter space into two overlapping parts:

$$\begin{aligned}\mathcal{A}_T &= \{\theta \in \Theta : |1-\rho_p| < T^{1-\alpha}\}, \\ \mathcal{B}_T &= \{\theta \in \Theta : |1-\rho_p| > T^{1-\alpha}\},\end{aligned}$$

for some $0 < \alpha < 1$.

Proof of Proposition 1. First, it follows from Lemma B1 that

$$\lim_{T\to\infty} \sup_{\theta\in\Theta} \sup_{x\in\Re^{d_\theta}} \left| P(\widehat{\Sigma}_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta) \leq x) - P(\Sigma_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta) \leq x) \right| = 0, \tag{A.1}$$

where

$$\widehat{\Sigma}_T = \begin{bmatrix} \widehat{\sigma}_T^2(\sum_{t=p+1}^T x_t x_t')^{-1} & 0_{(p+2)\times 1} \\ 0_{1\times(p+2)} & \widehat{\sigma}_{4,T} \end{bmatrix}, \quad \Sigma_T = \begin{bmatrix} \sigma^2(\sum_{t=p+1}^T x_t x_t')^{-1} & 0_{(p+2)\times 1} \\ 0_{1\times(p+2)} & \sigma_4 \end{bmatrix}. \tag{A.2}$$

Next, it follows from Lemma B2 that the distribution of $\Sigma_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta)$ based on $\{y_t\}_{t=1}^T$ can be uniformly approximated by that based on $\{\bar{y}_t\}_{t=1}^T$ where

$$\begin{aligned}\Delta\bar{y}_t &= \delta_0 + \delta_1\left(\frac{t}{T}\right) + \Pi\bar{y}_{t-1} + \Gamma_1\Delta\bar{y}_{t-1} + \cdots + \Gamma_{p-1}\Delta\bar{y}_{t-p+1} + \bar{u}_t \\ &= \beta\bar{x}_t + \bar{u}_t, \end{aligned} \tag{A.3}$$

and $\bar{u}_t \overset{iid}{\sim} N(0,\sigma^2)$. It follows from Lemma 5 of Mikusheva (2007a) and Lemma B3 that

$$\Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}\bar{x}_t\bar{x}_t'\Upsilon_T^{-1}(c) - \begin{bmatrix} 1 & \frac{1}{2} & \omega(c)\int_0^1 J_c(r)dr & 0 \\ \frac{1}{2} & \frac{1}{3} & \omega\int_0^1 rJ_c(r)dr & 0 \\ \omega\int_0^1 J_c(r)dr & \omega\int_0^1 rJ_c(r)dr & \omega^2\int_0^1 J_c(r)^2dr & 0 \\ 0 & 0 & 0 & M \end{bmatrix} = o_p(1) \quad (A.4)$$

and

$$\Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}\bar{x}_t\bar{e}_t - \begin{bmatrix} N_1 \\ N_2 \\ \sigma\omega(c)\int_0^1 J_c(r)dW(r) \\ N_3 \end{bmatrix} = o_p(1), \quad (A.5)$$

uniformly over $\mathcal{A}_T$, where $c = T\log(|\rho_p|)$.

Third, it follows from Lemma 12(a) and (b) of Mikusheva (2007b), (A.4) and (A.5) that

$$\begin{aligned} T^{\frac{1}{2}}(\widehat{\sigma}_T^2-\sigma^2) &= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}(\widehat{u}_t^2-\sigma^2) \\ &= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}((u_t-(\widehat{\beta}_T-\beta)'x_t)^2-\sigma^2) \\ &= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}(u_t^2-\sigma^2-2(\widehat{\beta}_T-\beta)'x_tu_t+(\widehat{\beta}_T-\beta)'x_tx_t'(\widehat{\beta}_T-\beta)) \\ &= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}(u_t^2-\sigma^2-\frac{2\sqrt{T}(\widehat{\beta}_T-\beta)'\Upsilon_T(c)}{T}\Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}x_tu_t \\ &\quad +\frac{1}{T}(\widehat{\beta}_T-\beta)'\Upsilon_T(c)\Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}x_tx_t'\Upsilon_T^{-1}(c)\Upsilon_T(c)(\widehat{\beta}_T-\beta) \\ &= \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}(u_t^2-\sigma^2)+O_p(T^{-\frac{1}{2}}) \\ &\overset{d}{\to} N_4, \end{aligned} \quad (A.6)$$

uniformly on Θ.

It follows from (A.4), (A.5) and (A.6) that

$$\lim_{T\to\infty}\sup_{\theta\in\mathcal{A}_T}\sup_{x\in\Re^{d_\theta}}\left|P(\Sigma_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta)\le x)-P(\eta(\theta,T)\le x)\right| = 0. \quad (A.7)$$

Using

$$\sqrt{-2c}\int_0^1 J_c(r)dW(r) \;\; \xrightarrow{d} \;\; N(0,1), \tag{A.8}$$

$$(-2c)\int_0^1 J_c^2(r)dx \;\; \xrightarrow{p} \;\; 1, \tag{A.9}$$

as in Phillips (1987), it follows from Lemma 12(a) and (b) of Mikusheva (2007b) and (A.6) that

$$\lim_{T\to\infty}\sup_{\theta\in\mathcal{B}_T}\sup_{x\in\Re^{d_\theta}}\left|P(\Sigma_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta)\leq x)-P(\eta(\theta,T)\leq x)\right| \;=\; 0. \tag{A.10}$$

Therefore, Proposition 1 follows from (A.7) and (A.10). ■

Proof of Proposition 2. Because $f(\cdot)$ is continuously differentiable,

$$\begin{aligned}
T^{\frac{1}{2}}(f(\widehat{\theta}_T)-f(\theta)) &= T^{\frac{1}{2}}Df(\bar{\theta}_T)(\widehat{\theta}_T-\theta) \\
&= T^{\frac{1}{2}}Df(\bar{\theta}_T)\Upsilon_T^{-1}(c)\Upsilon_T(c)(\widehat{\theta}_T-\theta) \\
&= T^{\frac{1}{2}}Df(\bar{\theta}_T)\Sigma_T^{\frac{1}{2}}\Sigma_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta) \\
&= T^{\frac{1}{2}}Df(\bar{\theta}_T)\Upsilon_T^{-1}(c)(\Upsilon_T(c)\Sigma_T\Upsilon_T(c))^{\frac{1}{2}} \\
&\quad \times\Sigma_T^{-\frac{1}{2}}(\widehat{\theta}_T-\theta),
\end{aligned} \tag{A.11}$$

where $\bar{\theta}_T$ is a point between $\widehat{\theta}_T$ and θ.

Note that $T^{\frac{1}{2}}F(\theta)\Upsilon_T(c)$ equals (6), that the first two columns consist of zeros by Assumption B and that in the nonstationary region $\mathcal{A}_T$, the elements of the third column converge to zero. Thus, it follows from (A.4), (A.5) and (A.11) that

$$(Df(\widehat{\theta}_T)\widehat{\Sigma}_T Df(\widehat{\theta}_T)')^{-\frac{1}{2}}T^{\frac{1}{2}}(f(\widehat{\theta}_T)-f(\theta)) \tag{A.12}$$

converges in distribution to the standard normal random vector uniformly on $\mathcal{A}_T$. In the stationary region $\mathcal{B}$, it follows from Lemma 12(a) and (b) of Mikusheva (2007b) and (A.6) that (A.12) converges in distribution to the standard normal random vector uniformly on $\mathcal{B}_T$. Thus, (9) follows. (10) follows from the second remark about Theorem 1 in Kasy (2018). ■

Proof of Proposition 3. Because we assume that the variance is uniformly bounded away from zero and uniformly bounded from above in Assumption A(iii), the arguments in the proof of Lemma 6 of

Mikusheva (2007a) carry through after scaling the residual in her proof by its standard deviation. Thus, the empirical distribution of the scaled residuals belongs to the $\mathcal{L}_r(K, M, \theta)$ class[12], and the Skorohod representation result in Lemma 12 of Mikusheva (2007a) applies. In other words, for any realization of the disturbance term, there exists $K > 0$, $M > 0$ and θ such that the empirical distribution function of the residual-based bootstrap, $\widehat{F}_T$, belongs to $\mathcal{L}_r(K, M, \theta)$ for all $\theta \in \Theta$. Thus, there is an almost sure approximation of the partial sum process by Brownian motions: For any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\lim_{T\to\infty} \sup_{F_T \in \mathcal{L}_r(K,M,\theta)} P^* \left(\sup_{0\le s\le 1} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} u_t^* - \sigma W(s) \right| > \varepsilon T^{-\delta} \right) \; = \; 0. \tag{A.13}$$

Moreover, by Lemma B2 with c, u_t, x_t, y_t replaced by c^*, u_t^*, x_t^*, y_t^*, respectively, the relevant bootstrap sample moments can be approximated by those generated from a Gaussian autoregressive process with $\beta = \widehat{\beta}_T$ almost surely conditional on the data. A bootstrap version of Lemma B1 may be constructed by replacing $\widehat{\sigma}_{4,T}$ and $\sigma_{4,T}$ by $\widehat{\sigma}^*_{4,T}$ and $\widehat{\sigma}_{4,T}$, respectively. Repeating the arguments in the proof of Lemma 5 of Mikusheva (2007a) yields a bootstrap version of Lemma B3 in which c is replaced by c^* from which we obtain the desired result. ■

Proof of Proposition 4. The proof of Proposition 4 is analogous to that of Proposition 2. ■

Appendix B

Throughout Appendix B, suppose that Assumptions A and B are satisfied.

The following lemma builds on Lemma 3 of Mikusheva (2007a):

Lemma B1.

$$\widehat{\sigma}_{4,T} \; = \; \sigma_4 + o_p(1) \tag{B.1}$$

[12]Mikusheva (2007a) defines this class to be the class of sequences of distributions F_T such that the mean is zero, the variance σ_T^2 satisfies $|\sigma_T^2| \le MT^{-\theta}$ and the supremum of the rth moment with respect to T is less than K.

uniformly on Θ.

Proof of Lemma B1. $\widehat{\sigma}_{4,T}$ can be approximated by

$$\begin{aligned}
\widehat{\sigma}_{4,T} &= \frac{1}{T-p}\sum_{t=p+1}^{T}\widehat{u}_t^4 - \widehat{\sigma}_T^4 \\
&= \frac{1}{T-p}\sum_{t=p+1}^{T}\widehat{u}_t^4 - \sigma^4 + o_p(1) \\
&= \frac{1}{T-p}\sum_{t=p+1}^{T}(u_t - (\widehat{\beta}_T-\beta)'x_t)^4 - \sigma^4 + o_p(1) \\
&= \frac{1}{T-p}\Bigg[\sum_{t=p+1}^{T}u_t^4 - 4((\widehat{\beta}_T-\beta)'x_t)u_t^3 \\
&\quad +6((\widehat{\beta}_T-\beta)'x_t)^2u_t^2 - 4((\widehat{\beta}_T-\beta)'x_t)^3u_t + ((\widehat{\beta}_T-\beta)'x_t)^4\Bigg] \\
&= \frac{1}{T-p}\left[\sum_{t=p+1}^{T}u_t^4 - 4(\zeta_T'\tilde{x}_t)u_t^3 + 6(\zeta_T'\tilde{x}_t)^2u_t^2 - 4(\zeta_T'\tilde{x}_t)^3u_t + (\zeta_T'x_t)^4\right] \qquad \text{(B.2)}
\end{aligned}$$

where the second equality follows from Lemma 3 of Mikusheva (2007a), $\zeta_T = (\sum_{t=p+1}^{T}x_tx_t')^{\frac{1}{2}}(\widehat{\beta}_T - \beta)$ and $\tilde{x}_t = (\sum_{t=p+1}^{T}x_tx_t')^{-\frac{1}{2}}x_t$. As shown in the proof of Proposition 1, $\zeta_T = O_p(1)$ uniformly on Θ. Because $\sum_{t=p+1}^{T}\tilde{x}_t\tilde{x}_t' = I_{p+2}$,

$$\sum_{t=p+1}^{T}\sum_{j=1}^{p+2}\tilde{x}_t^k \;\leq\; 1, \qquad \text{(B.3)}$$

for $k = 4, 6, 8$. Thus, it follows from the Cauchy-Schwarz inequality that

$$\begin{aligned}
\left|\frac{1}{T-p}\sum_{t=p+1}^{T}(\zeta_T'\tilde{x}_t)u_t^3\right| &\leq \left(\frac{1}{T-p}\sum_{t=p+1}^{T}(\zeta_T'\tilde{x}_t\tilde{x}_t'\zeta_T)\right)^{\frac{1}{2}}\left(\frac{1}{T-p}\sum_{t=p+1}^{T}u_t^6\right)^{\frac{1}{2}} \\
&= O_p(T^{-\frac{1}{2}}), \qquad \text{(B.4)} \\
\left|\frac{1}{T-p}\sum_{t=p+1}^{T}(\zeta_T'\tilde{x}_t)^2u_t^2\right| &\leq \left(\frac{1}{T-p}\sum_{t=p+1}^{T}(\zeta_T'\tilde{x}_t)^4\right)^{\frac{1}{2}}\left(\frac{1}{T-p}\sum_{t=p+1}^{T}u_t^4\right)^{\frac{1}{2}} \\
&\leq C\left(\frac{\|\zeta_T\|^4}{T-p}\sum_{t=p+1}^{T}\sum_{j=1}^{p+2}\tilde{x}_{j,t}^4\right)^{\frac{1}{2}}\left(\frac{1}{T-p}\sum_{t=p+1}^{T}u_t^4\right)^{\frac{1}{2}} \\
&= O_p(T^{-\frac{1}{2}}), \qquad \text{(B.5)}
\end{aligned}$$

$$\left|\frac{1}{T-p}\sum_{t=p+1}^{T}(\zeta_T'\tilde{x}_t)^3 u_t\right| = O_p(T^{-\frac{1}{2}}), \tag{B.6}$$

$$\left|\frac{1}{T-p}\sum_{t=p+1}^{T}(\zeta_T'\tilde{x}_t)^4\right| = O_p(T^{-\frac{1}{2}}), \tag{B.7}$$

where the last two results follow from arguments similar to the one used in the second result and the $O_p(T^{-\frac{1}{2}})$ terms are uniformly on Θ. Thus, (B.1) follows. ■

The next lemma is a slight extension of Lemma 11 of Mikusheva (2007b) which we present for completeness.

Lemma B2. Suppose that $\bar{y}_t$ follows

$$\begin{aligned}\Delta\bar{y}_t &= c + d(t/T) + \pi\bar{y}_{t-1} + \gamma_1\Delta\bar{y}_{t-1} + \cdots + \gamma_{p-1}\Delta\bar{y}_{t-p+1} + \bar{u}_t \\ &= \beta\bar{x}_t + \bar{u}_t, \end{aligned} \tag{B.8}$$

where $\bar{u}_t \overset{iid}{\sim} N(0,\sigma^2)$. Then thre exists a completion of the initial probability space and the realization of $\bar{y}_t$ on this probability space such that

$$\sup_{\theta\in\mathcal{A}_T}\sup_{t=1,\ldots,T}\left\|\frac{y_t}{\sqrt{T}} - \frac{\bar{y}_t}{\sqrt{T}}\right\| = o(1) \text{ a.s.}, \tag{B.9}$$

$$\sup_{\theta\in\mathcal{A}_T}\sup_{t=1,\ldots,T}\left\|\frac{y_t}{\sqrt{T}}\right\| = O(1) \text{ a.s.}, \tag{B.10}$$

$$\sup_{\theta\in\mathcal{A}_T}\left\|\Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}x_t u_t - \Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}\bar{x}_t\bar{u}_t\right\| = o(1), \tag{B.11}$$

$$\sup_{\theta\in\mathcal{A}_T}\left\|\Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}x_t x_t'\Upsilon_T^{-1}(c) - \Upsilon_T^{-1}(c)\sum_{t=p+1}^{T}\bar{x}_t\bar{x}_t'\Upsilon_T^{-1}(c)\right\| = o(1), \tag{B.12}$$

$$\sup_{\theta\in\mathcal{A}_T}\left\|(\sum_{t=p+1}^{T}x_t x_t')^{-\frac{1}{2}}\sum_{t=p+1}^{T}x_t u_t' - (\sum_{t=p+1}^{T}\bar{x}_t\bar{x}_t')^{-\frac{1}{2}}\sum_{t=p+1}^{T}\bar{x}_t e_t'\right\| = o(1). \tag{B.13}$$

Proof of Lemma B2.

Because $\kappa \leq \sigma^2 \leq K$, (B.9) and (B.10) follow from Lemma 11(a) and (b), respectively, of Mikusheva (2007b) who normalizes σ^2 to one. Similarly, (B.11) follows from Lemma 11(c), (d), (e), and (f) of Mikusheva (2007b) and (B.12) from her Lemma 11(g), (h), and (i). (B.13) follows from (B.11) and (B.12). ■

The next two results follow from the arguments used in the proof of Lemma 5 of Mikusheva (2007a).

Lemma B3.

$$\lim_{T\to\infty}\sup_{\theta\in\mathcal{A}_T} E\left\|\text{vech}\left(\Upsilon_T^{-1}\sum_{t=p+1}^{T}\bar{x}_t\bar{x}_t\Upsilon_T^{-1} - \begin{bmatrix} 1 & \frac{1}{2} & \omega\int_0^1 J_c(r)dr & 0 \\ \frac{1}{2} & \frac{2}{3} & \omega\int_0^1 rJ_c(r)dr & 0 \\ \omega\int_0^1 J_c(r)dr & \omega\int_0^1 rJ_c(r)dr & \omega^2\int_0^1 J_c(r)^2dr & 0 \\ 0 & 0 & 0 & M \end{bmatrix}\right)\right\|^2 = 0, \tag{B.14}$$

$$\lim_{T\to\infty}\sup_{\theta\in\mathcal{A}_T} E\left\|\Upsilon_T^{-1}\sum_{t=p+1}^{T}\bar{x}_t\bar{u}_t - \begin{bmatrix} N_1 \\ N_2 \\ \sigma\omega\int_0^1 J_c(r)dW(r) \\ N_3 \end{bmatrix}\right\|^2 = 0. \tag{B.15}$$

The following result is a slight modification of equation (21) of Inoue and Kilian (2002):

Lemma B4.

Let $\delta_{0,T} = \phi_T(1)\delta_{0,T}^{\dagger} + \delta_{1,T}^{\dagger}(\phi_{T,1} + 2\phi_{T,2} + \cdots + p\phi_{T,p})$, $\delta_{0,T}^{\dagger} = \delta_0^{\dagger} + \xi_c T^{-1/2} + o(T^{-1/2})$, $\delta_{1,T} = \phi_T(1)\delta_{1,T}^{\dagger}$, $\delta_{1,T}^{\dagger} = \delta_1^{\dagger} + \xi_d T^{-1/2} + o(T^{-1/2})$, $\phi_T(L) = 1 - \phi_{T,1}L - \cdots - \phi_{T,p}L^p$, $\pi_T = -\phi_T(1) = \sum_{j=1}^{p}\phi_{T,j} - 1 = \xi_0 T^{-1} + o(T^{-1})$, $\gamma_{T,j} = -(\phi_{T,j+1} + \cdots + \phi_{T,p}) = \gamma_j + \xi_j T^{-1/2} + o(T^{-1/2})$, $u_{T,t} \overset{iid}{\sim} (0_{n\times 1}, \sigma_T^2)$, and $\sigma_T^2 = \sigma^2 + \xi_{\sigma^2}T^{-1/2} + o(T^{-1/2})$ for some $[\xi_c\ \xi_d\ \xi_0, \xi_1\ \cdots\ \xi_{p-1}\ \xi_{\sigma^2}]'w$. Define a triangular array

$$\begin{aligned} \Delta y_{T,t} &= \delta_{0,T} + \delta_{1,T}\left(\frac{t}{T}\right) + \pi_T y_{T,t-1} + \gamma_{T,1}\Delta y_{T,t-1} + \cdots + \gamma_{T,p-1}\Delta y_{T,t-p+1} + u_{T,t} \\ &= \beta_T' x_{T,t} + u_{T,t}, \end{aligned} \tag{B.16}$$

where $\beta_T = [\delta_{0,T}\ \delta_{1,T}\ \pi_T\ \gamma_{T,1}\ \cdots\ \gamma_{T,p}]'$, and $x_t = [1\ t/T\ \ y_{t-1}\ \Delta y_{t-1}\ \cdots\ \Delta y_{t-p+1}]'$. Then

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}\sup_{x\in\Re^{d_\theta}}\left|P(\Upsilon_T(c_T)(\widehat{\theta}_T - \theta) \le x) - P(\eta_T(\theta, T) \le x)\right| = 0, \tag{B.17}$$

where $c_T = T\log(1+\rho_{p,T})$.

The proof of Lemma B4 is based on the same arguments already used in the proof of Proposition 1 and is omitted. ■

Appendix C

We want to evaluate the $1-\alpha$ level confidence set

$$\min_{i\in\{1,2,\ldots,7\}} P_{\rho_i}(\psi_i \in C_i), \tag{B.18}$$

where ψ_i and C_i are the true parameter values and ρ_i is the ith value of ρ.

Because $P_{\rho_i}(\psi_i \in C_i)$ is not analytically tractable in finite samples, it is approximated by simulation:

$$\widehat{P}_{\rho_i}(\psi_i \in C_i) = \frac{1}{M}\sum_{j=1}^{M} I(\psi_i \in C_i^{(j)}) \tag{B.19}$$

where M is the number of Monte Carlo simulations and $C_i^{(j)}$ is a level $1-\alpha$ confidence set for the jth Monte Carlo iteration.

The problem is that even when the coverage rate is uniform in finite samples the estimate (B.19) may have downward bias due to "data mining".

To estimate the bias, express the Monte Carlo estimate of the coverage rate as:

$$X_i = \frac{1}{5000}\sum_{j=1}^{M} d_{ij}, \tag{B.20}$$

where d_{ij}'s are iid Bernoulli random variables with parameter $1-\alpha$ for $i=1,\ldots,7$. That is, $5000X_i$ is a binomial random variable with parameters 5000 and $1-\alpha$. Then the expectation of

$$Y = \min_{i\in\{1,2,\ldots,7\}} X_i \tag{B.21}$$

minus $(1-\alpha)$ is the data mining bias. Thus, the expectation of Y can be estimated from

$$\frac{1}{n}\sum_{j=1}^{n} \min_{i\in\{1,2,\ldots,7\}} X_i^{(j)} \tag{B.22}$$

When $\alpha = 0.1$, $d = 7$, and $M = 5000$, as in our simulation study, the normal approximation yields 0.894, implying a data mining bias of 0.006. The same answer is obtained when simulating this bias rather than relying on the normal approximation.

References

1. Andrews, D.W.K., and P. Guggenberger (2009), "Hybrid and Size-Corrected Subsample Methods," *Econometrica*, 77, 721-762.

2. Bauer, D., and A. Maynard (2012), "Persistence-Robust Surplus-Lag Granger Causality Testing," *Journal of Econometrics,* 169, 293–300.

3. Benkwitz, A., Lütkepohl, H., and M.H. Neumann (2000), "Problems Related to Confidence Intervals for Impulse Responses of Autoregressive Processes," *Econometric Reviews*, 19, 69-103.

4. Bruder, S., and M. Wolf (2018), "Balanced Bootstrap Joint Confidence Bands for Structural Impulse Response Functions,"*Journal of Time Series Analysis*, 39, 641-664.

5. Dolado, J., and H. Lütkepohl (1996), "Making Wald Tests Work for Cointegrated VAR Systems," *Econometric Reviews,* 15, 369–386.

6. Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.

7. Elliott, G. (1998), "On the Robustness of Cointegration Methods when Regressors Almost Have Unit Roots," *Econometrica*, 66, 149–158.

8. Gafarov, B., Meier, M., and J.L. Montiel Olea (2018), "Delta Method Inference for a Class of Set-Identified SVARs," *Journal of Econometrics*, 203, 316-327.

9. Giraitis, L., and P.C.B. Phillips (2006), "Uniform Limit Theory for Stationary Autoregression," *Journal of Time Series Analysis*, 27, 51–60.

10. Gonçalves, S., and L. Kilian (2004), "Bootstrapping Autoregressions in the Presence of Conditional Heteroskedasticity of Unknown Form," *Journal of Econometrics*, 123, 89-120.

11. Gospodinov, N. (2004), "Asymptotic Confidence Intervals for Impulse Responses of Near-Integrated Processes," *Econometrics Journal*, 7, 505-527.

12. Gospodinov, N., Herrera, A.M., and E. Pesavento (2011), "Unit Roots, Cointegration and Pre-Testing in VAR Models," *Advances in Econometrics*, 31, 81-115.

13. Granziera, E., Moon, H.R., and F. Schorfheide (2018), "Inference for VARs Identified with Sign Restrictions," *Quantitative Economics*, 9, 1087-1121.

14. Guerron-Quintana, P., Inoue, A., and L. Kilian (2017), "Impulse Response Matching Estimators for DSGE Models," *Journal of Econometrics*, 196, 144-155.

15. Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York.

16. Hansen, B.E. (1999), "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics*, 81, 594-607.

17. Inoue, A., and L. Kilian (2002), "Bootstrapping Autoregressive Processes with Possible Unit Roots," *Econometrica*, 70, 377-391.

18. Inoue, A., and L. Kilian (2003), "The Continuity of the Limit Distribution in the Parameter of Interest is not Essential for the Validity of the Bootstrap," *Econometric Theory*, 19, 944-961.

19. Inoue, A., and L. Kilian (2016), "Joint Confidence Sets for Structural Impulse Responses," *Journal of Econometrics*, 192, 421-432.

20. Jordà, Ò. (2009), "Simultaneous Confidence Regions for Impulse Responses,"*Review of Economics and Statistics*, 91, 629–647.

21. Kasy, M. (2018), "Uniformity and the Delta Method," *Journal of Econometric Methods,* forthcoming.

22. Kilian, L. (1998a), "Confidence Intervals for Impulse Responses Under Departures from Normality," *Econometric Reviews*, 17, 1-29.

23. Kilian, L. (1998b), "Small-Sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics*, 80, 218-230.

24. Kilian, L. (1999), "Finite-Sample Properties of Percentile and Percentile-t Bootstrap Confidence Intervals for Impulse Responses," *Review of Economics and Statistics*, 81, 652-660.

25. Kilian, L, and P.L. Chang (2000), "How Accurate are Confidence Intervals for Impulse Responses in Large VAR Models?", *Economics Letters*, 69, 299-307.

26. Kilian, L., and H. Lütkepohl (2017), *Structural Vector Autoregressive Analysis*, Cambridge University Press, New York.

27. Kurozumi, E., and T. Yamamoto (2000), "Modified Lag Augmented Vector Autoregressions," *Econometric Reviews*, 19, 207–231.

28. Lütkepohl, H. (1990), "Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models," *Review of Economics and Statistics*, 72, 116–125.

29. Lütkepohl, H., Staszewska-Bystrova, A., and P. Winker (2015a), "Confidence Bands for Impulse Responses: Bonferroni vs. Wald," *Oxford Bulletin of Economics and Statistics*, 77, 800-821.

30. Lütkepohl, H., Staszewska-Bystrova, A., and P. Winker (2015b), "Comparison of Methods for Constructing Joint Confidence Bands for Impulse Response Functions," *International Journal of Forecasting*, 31, 782-798.

31. Lütkepohl, H., Staszewska-Bystrova, A., and P. Winker (2018), "Calculating Joint Confidence Bands for Impulse Response Functions Using Highest Density Regions," *Empirical Economics*, 55, 1389-1411.

32. Mikusheva, A. (2007a), "Uniform Inference in Autoregressive Models," *Econometrica*, 75, 1411–1452.

33. Mikusheva, A. (2007b), "Uniform Inference in Autoregressive Models: Supplementary Appendix," https://economics.mit.edu/files/9467.

34. Mikusheva, A. (2012), "One-Dimensional Inference in Autoregressive Models with the Potential Presence of a Unit Root," *Econometrica*, 80, 173–212.

35. Montiel Olea, J.L., and M. Plagborg-Møller (2018), "Simultaneous Confidence Bands: Theory, Implementation, and an Application to SVARs," *Journal of Applied Econometrics*, forthcoming.

36. Montiel Olea, J.L., Stock, J.H., and M.W. Watson (2016), "Inference in SVARs Identified with an External Instrument," manuscript, Columbia University.

37. Nankervis J.C., and N.E. Savin (1988), "The Student's t Approximation in a Stationary First Order Autoregressive Model," *Econometrica*, 56, 119-145.

38. Park, J.Y., and P.C.B. Phillips (1989), "Statistical Inference in Regressions with Integrated Processes: Part 2," *Econometric Theory* 5, 95–131.

39. Pesavento, E., and B. Rossi (2006), "Small-Sample Confidence Intervals for Multivariate Impulse Response Functions at Long Horizons," *Journal of Applied Econometrics*, 21, 1135-1155.

40. Pesavento, E., and B. Rossi (2007), "Impulse Response Confidence Intervals for Persistent Data: What Have We Learned?," *Journal of Economic Dynamics and Control*, 31, 2398-2412.

41. Phillips, P.C.B. (1987), "Toward a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535–547.

42. Phillips, P.C.B. (1998), "Impulse Response and Forecast Error Variance Asymptotics in Non-stationary VARs," *Journal of Econometrics*, 83, 21–56.

43. Phillips, P.C.B. (2014), "On Confidence Intervals for Autoregressive Roots and Predictive Regression," *Econometrica*, 82, 1177-1195.

44. Phillips, P.C.B., and J.H. Lee (2016), "Robust Econometric Inference with Mixed Integrated and Mildly Explosive Regressors," *Journal of Econometrics,* 192, 433–450.

45. Sims, C.A., J.H. Stock and M.W. Watson (1990), "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica* 58, 113–144.

46. Toda, H., and T., Yamamoto (1995), "Statistical Inference in Vector Autoregressions with Possibly Integrated Processes," *Journal of Econometrics,* 66, 225–250.

47. Wright, J.H. (2000), "Confidence Intervals for Univariate Impulse Responses," *Journal of Business and Economic Statistics*, 18, 368-373.

Table 1: Uniform coverage rates of nominal 90% impulse response confidence intervals based on the delta method

	Impulse response horizon											
T	1	2	3	4	5	6	7	8	9	10	11	12
						AR model						
80	0.861	0.809	0.761	0.665	0.582	0.540	0.504	0.474	0.450	0.428	0.408	0.391
120	0.875	0.841	0.804	0.725	0.646	0.602	0.569	0.540	0.509	0.485	0.467	0.450
240	0.892	0.872	0.851	0.792	0.738	0.701	0.664	0.636	0.604	0.576	0.556	0.536
480	0.899	0.886	0.878	0.840	0.805	0.774	0.745	0.717	0.698	0.672	0.651	0.633
600	0.896	0.883	0.882	0.850	0.812	0.786	0.759	0.735	0.715	0.691	0.671	0.655
6000	0.906	0.902	0.903	0.902	0.899	0.896	0.891	0.888	0.881	0.876	0.869	0.862
12000	0.910	0.908	0.902	0.902	0.899	0.899	0.897	0.892	0.891	0.888	0.884	0.880
24000	0.912	0.907	0.905	0.904	0.901	0.901	0.899	0.899	0.897	0.894	0.891	0.889
						Lag-augmented AR model						
80	0.864	0.806	0.766	0.727	0.732	0.744	0.739	0.737	0.723	0.709	0.688	0.671
120	0.873	0.844	0.808	0.774	0.773	0.781	0.778	0.774	0.767	0.757	0.744	0.729
240	0.890	0.872	0.850	0.824	0.831	0.837	0.840	0.838	0.835	0.829	0.821	0.811
480	0.899	0.886	0.879	0.867	0.870	0.872	0.874	0.876	0.872	0.869	0.865	0.860
600	0.896	0.884	0.882	0.874	0.879	0.885	0.883	0.881	0.881	0.880	0.878	0.875
6000	0.906	0.901	0.904	0.898	0.899	0.900	0.899	0.901	0.901	0.900	0.897	0.897
12000	0.908	0.909	0.903	0.899	0.898	0.899	0.902	0.901	0.901	0.900	0.901	0.900
24000	0.912	0.907	0.905	0.901	0.901	0.899	0.898	0.899	0.899	0.900	0.898	0.899

Notes: The data are generated from $y_t = \rho y_{t-1} + u_t$, $u_t \overset{iid}{\sim} N(0,1)$, $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$. The fitted model is an AR(4) with intercept. The lag-augmented fitted model is an AR(5) model with intercept. Inference on the individual impulse responses is based on the delta method based on closed-form solutions for the impulse response standard error in Lütkepohl (1990). The uniform coverage rates are computed as the infimum across the ρ values, adjusted for data mining bias.

Table 2: Uniform coverage rates of nominal 90% impulse response confidence intervals based on the bootstrap

	Impulse response horizon											
T	1	2	3	4	5	6	7	8	9	10	11	12
AR model: Hall percentile interval												
80	0.888	0.854	0.823	0.793	0.776	0.745	0.700	0.656	0.610	0.557	0.516	0.475
120	0.981	0.873	0.852	0.828	0.820	0.813	0.796	0.763	0.726	0.682	0.650	0.612
240	0.898	0.888	0.879	0.861	0.862	0.862	0.863	0.859	0.846	0.828	0.807	0.788
480	0.989	0.892	0.890	0.880	0.876	0.879	0.886	0.889	0.878	0.869	0.857	0.845
600	0.897	0.895	0.892	0.880	0.875	0.872	0.875	0.880	0.882	0.879	0.872	0.863
Lag-augmented AR model: Hall percentile interval												
80	0.887	0.858	0.827	0.796	0.767	0.739	0.718	0.686	0.650	0.610	0.568	0.528
120	0.892	0.870	0.852	0.836	0.816	0.794	0.775	0.752	0.725	0.691	0.662	0.620
240	0.897	0.887	0.881	0.871	0.864	0.852	0.847	0.838	0.822	0.797	0.770	0.745
480	0.898	0.895	0.892	0.889	0.885	0.879	0.879	0.873	0.867	0.855	0.841	0.825
600	0.897	0.893	0.892	0.892	0.891	0.889	0.887	0.885	0.878	0.869	0.860	0.843
AR model with bias adjustment: Efron percentile interval												
80	0.868	0.853	0.815	0.794	0.753	0.751	0.733	0.727	0.712	0.708	0.701	0.697
120	0.882	0.869	0.847	0.828	0.798	0.797	0.784	0.773	0.761	0.753	0.745	0.742
240	0.897	0.890	0.872	0.862	0.839	0.835	0.822	0.812	0.804	0.798	0.791	0.787
480	0.900	0.891	0.888	0.882	0.868	0.867	0.861	0.854	0.861	0.836	0.829	0.822
600	0.897	0.892	0.890	0.883	0.873	0.869	0.860	0.853	0.843	0.834	0.828	0.821
Lag-augmented AR model with bias adjustment: Efron percentile interval												
80	0.866	0.848	0.812	0.801	0.802	0.819	0.819	0.823	0.830	0.835	0.839	0.842
120	0.877	0.869	0.842	0.836	0.830	0.840	0.840	0.844	0.847	0.849	0.853	0.854
240	0.894	0.888	0.873	0.868	0.866	0.873	0.875	0.878	0.880	0.880	0.880	0.881
480	0.901	0.893	0.888	0.887	0.887	0.890	0.892	0.892	0.891	0.892	0.893	0.893
600	0.896	0.891	0.889	0.892	0.895	0.897	0.894	0.895	0.895	0.896	0.896	0.896

Notes: The data are generated from $y_t = \rho y_{t-1} + u_t$, $u_t \overset{iid}{\sim} N(0,1)$, $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$. The fitted model is an AR(4) with intercept. The lag-augmented fitted model is an AR(5) model with intercept. The bootstrap data are generated using the recursive-design bootstrap method, as discussed in Kilian and Lütkepohl (2017). The uniform coverage rates are computed as the infimum across the ρ values, adjusted for data mining bias.

Table 3: Percentage increase in average width of nominal 90% Efron percentile intervals based on bias-adjusted lag-augmented autoregression relative to the same interval based on the bias-adjusted original autoregression

	Impulse response horizon		
T	1	2	3
80	1.43	1.53	1.53
120	0.94	0.98	1.03
240	0.47	0.47	0.46
480	0.21	0.22	0.25
600	0.16	0.19	0.17

Notes: The data are generated from $y_t = \rho y_{t-1} + u_t$, $u_t \overset{iid}{\sim} N(0,1)$, $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$. The lag-augmented fitted model is an AR(5) model with intercept. The bootstrap data are generated using the recursive-design bootstrap method, as discussed in Kilian and Lütkepohl (2017). The percentage increase in the average interval width is based on the average of the percentage increases in average interval width obtained for each ρ.

Table 4: Uniform coverage rates of nominal 90% impulse response confidence intervals based on the bootstrap

	Impulse response horizon								
T	12	18	24	30	36	42	48	54	60
	Lag-augmented AR model with bias adjustment: Efron percentile interval								
80	0.842	0.903	0.867	0.908	0.896	0.899	0.906	0.902	0.909
120	0.853	0.890	0.870	0.877	0.886	0.880	0.892	0.889	0.888
240	0.879	0.883	0.883	0.885	0.887	0.888	0.888	0.889	0.889
480	0.891	0.893	0.893	0.893	0.895	0.895	0.895	0.895	0.895
600	0.897	0.898	0.898	0.898	0.899	0.899	0.899	0.899	0.900

Notes: The data are generated from $y_t = \rho y_{t-1} + u_t$, $u_t \overset{iid}{\sim} N(0,1)$, $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$. The lag-augmented fitted model is an AR(5) model with intercept. The bootstrap data are generated using the recursive-design bootstrap method, as discussed in Kilian and Lütkepohl (2017). The uniform coverage rates are computed as the infimum across the ρ values, adjusted for data mining bias.

Table 5: Uniform coverage rates of nominal 90% impulse response confidence intervals based on lag-augmented autoregression

	$\lambda = h/T$				
T	0.1	0.3	0.5	0.7	0.9
			Delta method interval		
120	0.729	0.546	0.487	0.454	0.433
240	0.704	0.554	0.477	0.477	0.464
480	0.697	0.568	0.499	0.499	0.488
600	0.682	0.552	0.494	0.494	0.484
			Hall percentile interval		
120	0.621	0.401	0.376	0.363	0.357
240	0.543	0.429	0.407	0.399	0.393
480	0.542	0.458	0.440	0.433	0.429
600	0.524	0.457	0.443	0.435	0.432
			Efron percentile interval		
120	0.745	0.797	0.795	0.805	0.804
240	0.840	0.847	0.850	0.848	0.850
480	0.878	0.881	0.882	0.882	0.882
600	0.887	0.887	0.888	0.888	0.887
			Efron percentile interval after bias adjustment		
120	0.855	0.887	0.888	0.891	0.890
240	0.884	0.887	0.889	0.889	0.891
480	0.896	0.888	0.898	0.888	0.896
600	0.901	0.900	0.898	0.900	0.900

Notes: The data are generated from $y_t = \rho y_{t-1} + u_t$, $u_t \overset{iid}{\sim} N(0,1)$, $\rho \in \{0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1\}$. The lag-augmented fitted model is an AR(5) model with intercept. The bootstrap data are generated using the recursive-design bootstrap method, as discussed in Kilian and Lütkepohl (2017). The uniform coverage rates are computed as the infimum across the ρ values, adjusted for data mining bias.