

Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 682

First Version: September 2000, This version: August 2001

EXACT CONFIDENCE INTERVALS FOR IMPULSE RESPONSES IN A GAUSSIAN VECTOR
AUTOREGRESSION

Jonathan H. Wright

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors. Recent IFDPs are available on the Web at www.bog.frb.fed.us.

EXACT CONFIDENCE INTERVALS FOR IMPULSE RESPONSES IN A GAUSSIAN VECTOR AUTOREGRESSION

Jonathan H. Wright*

Abstract: Many techniques have been proposed for forming confidence intervals for the impulse responses in a vector autoregression. However, numerous Monte-Carlo simulations have shown that all of these methods often have coverage well below the nominal level. This paper proposes a new approach to constructing confidence intervals for impulse responses in a vector autoregression, making the additional assumption of Gaussianity. These confidence intervals are conservative in all sample sizes; by construction they have coverage that must be greater than or equal to the nominal level.

Keywords: Confidence Intervals; Vector Autoregressions; Impulse Responses; Bootstrap.

* Wright is an economist in the International Finance Division of the Federal Reserve Board. I am grateful to Jon Faust and Lutz Kilian for extremely helpful comments. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any person associated with the Federal Reserve System.

1. Introduction

Ever since the seminal paper of Sims (1980), the analysis of impulse response functions in a vector autoregression (VAR) has been a major topic in empirical macroeconomics. There are many techniques in use for constructing confidence intervals for the impulse response function in a VAR. These include the delta method (Lütkepohl (1990)), the bootstrap (Runkle (1987)), the bias-adjusted bootstrap (Kilian (1998a)) and the Bayesian Monte-Carlo integration method, interpreted as a classical confidence interval (Doan (1990) and Sims and Zha (1999)). A great many papers have simulated the coverage of these confidence intervals and have found that all of them can have coverage that is substantially below the nominal level, in the sample sizes that are relevant to macroeconomic researchers (see, for example, Griffiths and Lütkepohl (1993), Fachin and Bravetti (1996), Kilian (1998a, 1999), Lütkepohl (1996), Benkwitz, Lütkepohl and Neumann (2000) and the simulations below). The bias-adjusted bootstrap and Bayesian methods often do better in terms of coverage than the delta method or the simple bootstrap, but can still have coverage that is far below the nominal level. Problems with coverage arise when the vector autoregression is close to being nonstationary, but are by no means limited to this situation (see, for example, Benkwitz, Lütkepohl and Neumann).

This paper proposes a new approach to inference for impulse response functions in VARs. The idea is to perform a size-adjustment of the standard delta method

such that the coverage of the confidence interval is at or above the nominal level, uniformly in a plausible region of the parameter space. More precisely, I propose forming S_θ , an exact confidence set for the vector of VAR parameters, and then simulating the minimum size adjustment of a delta method confidence interval for the impulse response function that is required to ensure that the coverage is at least equal to the nominal level for all VAR parameter vectors in S_θ . I appeal to the Bonferroni inequality to argue that this parametric bootstrap algorithm must give coverage that is greater than or equal to some specified level. A related algorithm was proposed in Wright (2000), which relied on specifying that the model contained roots local to unity. This other algorithm works well in a univariate autoregression, but its extension to a general VAR requires the researcher to make assumptions about the number and location of large roots. The algorithm proposed in this paper does not rely on any special asymptotic nesting and does not require the researcher to make any assumptions beyond Gaussianity and the lag order.

The plan of the remainder of this paper is as follows. In section 2, I describe the model and outline the existing approaches to inference for impulse responses. In section 3, I describe the approach proposed in this paper. In section 4, the coverage and mean width of the proposed confidence intervals and of the leading existing alternatives are examined in a Monte-Carlo simulation of a simple bivariate VAR. Section 5 concludes. Throughout the paper, I focus on confidence intervals for im-

pulse responses, but all the same algorithms can of course be used to form confidence intervals for any function of the VAR model parameters, such as variance decompositions. Also, throughout the paper, as in virtually all of the related literature (e.g. Runkle 1987; Lütkepohl 1990; Kilian 1998a), I construct pointwise confidence intervals, i.e. confidence intervals for the impulse response at a single lead time.

2. The Model and the Existing Methods

Consider the Gaussian VAR

$$A(L)y_t = \varepsilon_t$$

where $A(L) = I - \sum_{k=0}^p A_k L^k$, L denotes the lag operator, ε_t is iidN(0,Ω), y_t and ε_t are $m \times 1$ vectors and the sample size is T . The observed time series, y_t , could be augmented by a constant, or by a constant and a trend. The lag order p is assumed to be finite and known. Let $\theta = (\theta'_A, \theta'_\Omega)'$ denote the vector of model parameters, where $\theta_A = (\text{vec}(A_1)', \text{vec}(A_2)', \dots, \text{vec}(A_p)')'$ and $\theta_\Omega = \text{vech}(\Omega)$ ¹. The VAR may be written in companion form as $\xi_t = A\xi_{t-1} + \zeta_t$, where $\xi_t = (y'_t, y'_{t-1}, \dots, y'_{t-p})'$, $\zeta_t = (\varepsilon'_t, 0', \dots, 0)'$ and

$$A = \begin{pmatrix} A_1 & A_2 & \dots & A_p \\ & I_{m(p-1)} & & 0_{m(p-1) \times m} \end{pmatrix}.$$

¹Throughout this paper, $\text{vec}(\cdot)$ denotes the column stacking operator and $\text{vech}(\cdot)$ denotes the column stacking operator that omits all elements above the main diagonal.

Let $u_t = P^{-1}\varepsilon_t$ be an $m \times 1$ iidN(0, I_m) vector of structural shocks², where P denotes the lower triangular Cholesky factor of Ω , so that $\Omega = PP'$. Let $h_{ij}^l(\theta)$ denote the effect of a unit increase in the i th element of u_t on the j th element of y_{t+l} , or the $i - j$ impulse response at lead time l . This is simply the j th element of $B_l P$ where $B(L) = I + \sum_{k=1}^{\infty} B_k L^k = A(L)^{-1}$; B_k is the matrix given by the first m rows and columns of A^k . No stationarity assumption is being made in this general model; accordingly the model is initialized by the condition $y_t = 0$ for $t \leq 0$.

Let $\hat{\theta} = (\hat{\theta}'_A, \hat{\theta}'_{\Omega})'$ denote the usual least squares estimator of θ . The standard estimator of the impulse response $h_{ij}^l(\theta)$ is simply $h_{ij}^l(\hat{\theta})$. This point estimate is of little value without a reliable measure of sampling error. There are several widely used approaches to forming confidence intervals for $h_{ij}^l(\theta)$.

The first approach is the delta method (Runkle (1987) and Lütkepohl (1990)).

If y_t is stationary and $\Gamma = E(\xi_t \xi_t')$ is of full rank³, then

$$\sqrt{T}(\hat{\theta} - \theta) \rightarrow_d N(0, W(\theta)),$$

²More generally, u_t could be defined as a vector of independent standard normal structural shocks, such that $u_t = R^{-1}\varepsilon_t$, where R is a matrix such that $RR' = \Omega$. This does not uniquely define R ; some other identifying assumptions are required. In this paper, I use the most common identification scheme (a causal ordering), by setting $R = P$, but all the methods discussed in this paper would work with any other choice of R . The Cholesky factor P is unique only up to postmultiplication by any diagonal matrix with elements 1 and -1 on the diagonal. The extra identifying assumptions that are required to circumvent this lack of uniqueness are not contentious, as they only restrict the sign of the impulse responses.

³Note that Γ is a closed-form function of θ , as $vec(\Gamma) = (I - A \otimes A)^{-1}vec(\Sigma_{\zeta})$ where $\Sigma_{\zeta} = E(\zeta_t \zeta_t') = \begin{pmatrix} \Omega & 0 \\ 0 & 0 \end{pmatrix}$.

and so, if $\frac{dh_{ij}^l(\theta)}{d\theta'} \neq 0$,

$$\sqrt{T}(h_{ij}^l(\hat{\theta}) - h_{ij}^l(\theta)) \rightarrow_d N(0, \frac{dh_{ij}^l(\theta)}{d\theta'} W(\theta) \frac{dh_{ij}^l(\theta)'}{d\theta'}),$$

where

$$W(\theta) = \begin{pmatrix} \Gamma^{-1} \otimes \Omega & 0 \\ 0 & 2(D'_m D_m)^{-1} D'_m (\Omega \otimes \Omega) D_m (D'_m D_m)^{-1} \end{pmatrix},$$

and D_m is the duplication matrix, defined such that $vec(F) = D_m vech(F)$ for any symmetric $m \times m$ matrix F . Accordingly, a 95% confidence interval for the impulse response $h_{ij}^l(\theta)$ is simply given by

$$h_{ij}^l(\hat{\theta}) \pm \frac{1.96}{\sqrt{T}} \sqrt{\frac{dh_{ij}^l(\hat{\theta})}{d\theta'} W(\hat{\theta}) \frac{dh_{ij}^l(\hat{\theta})'}{d\theta'}}.$$

Lütkepohl provides a rather complicated explicit expression for $\frac{dh_{ij}^l(\theta)}{d\theta'}$. This confidence interval may not have the correct coverage asymptotically if y_t is not stationary⁴ or if Γ is singular⁵ (see Benkwitz, Lütkepohl and Neumann (2000)).

The second approach is the bootstrap (Runkle (1987)). Let $\{\hat{\varepsilon}_t\}_{t=1}^T$ denote the residuals corresponding to $\hat{\theta}_A$. Resampling from these residuals with replacement, bootstrap samples may be constructed, and $h_{ij}^l(\theta)$ may be estimated from these bootstrap samples. The percentiles of the distribution of the estimated impulse response

⁴The delta method is asymptotically correct under nonstationarity in most VARs, but not in a VAR(1) (see Phillips (1998)).

⁵A univariate AR(1), where the true value of the AR coefficient is zero, is an example of a situation in which Γ will be singular.

across the bootstrap replications may be used to construct a confidence interval for $h_{ij}^l(\theta)$ ⁶.

The third approach is the bias-adjusted bootstrap of Kilian (1998a), which can generate substantial improvements in effective coverage. This involves using the bootstrap to estimate the bias in $\hat{\theta}_A$ and then anchoring the bootstrap at a bias-adjusted estimate of θ_A , rather than at $\hat{\theta}_A$.

The final commonly used existing method for forming impulse response confidence intervals is the Bayesian Monte-Carlo integration method (Doan (1990) and Sims and Zha (1999)). This involves constructing error bands based on the posterior mean and posterior standard deviation of $h_{ij}^l(\theta)$, using a diffuse prior. These error bands can then be interpreted as a classical confidence interval. They can easily be obtained by simulation, since the posterior of θ is normal-inverse Wishart.

In this paper, I assume that the VAR innovations are Gaussian. This assumption is not necessary for any of the above procedures, except for the Bayesian procedure. However, I adopt the assumption because it is necessary for the approach to inference that I propose in this paper. In most of the literature that has found that the effective coverage of confidence intervals for impulse responses is often too low

⁶In the terminology of Hall (1992), this is the “other percentile” bootstrap confidence interval. Kilian (1999) argues that this bootstrap confidence interval works best in empirically relevant sample sizes, despite the higher order asymptotic refinements that are available if the “symmetric percentile-t” confidence interval is used instead.

in Monte-Carlo simulations, the focus has been on Gaussian VARs. Kilian (1998b) discusses the impact of nonnormality on inference for impulse responses.

3. The Proposed Approach to Inference

The methods for constructing confidence intervals for $h_{ij}^l(\theta)$ discussed in section 2 have been evaluated in numerous Monte-Carlo simulations. Despite their asymptotic justifications, they can all have coverage that is well below the nominal level in empirically relevant sample sizes, as discussed above. This motivates consideration of an alternative algorithm for constructing impulse response confidence intervals. The basic idea is to form S_θ , an exact confidence set for θ , and to simulate the size adjustment of a delta method confidence interval for the impulse response function that is required to ensure that the coverage is at least equal to the nominal level for all θ in S_θ . I appeal to the Bonferroni inequality to argue that this parametric bootstrap algorithm must give coverage that is greater than or equal to some specified level.

Suppose that θ_0 is the true parameter value and let $F(\theta_0)$ denote the F statistic testing the hypothesis that $\theta = \theta_0$, i.e.

$$F(\theta_0) = T(\hat{\theta} - \theta_0)'W(\hat{\theta})^{-1}(\hat{\theta} - \theta_0).$$

Since the model is completely specified when the parameter vector is known, it is possible to simulate the distribution of $F(\theta_0)$, to an arbitrary degree of accuracy, and

so to solve for the critical values $\lambda_T(\theta_0, \alpha_1)$ such that $P[F(\theta_0) \geq \lambda_T(\theta_0, \alpha_1)] = \alpha_1$. An exact test of the hypothesis that $\theta = \theta_0$ can thus be conducted by comparing F with the correct critical value. Let $S_\theta(\alpha_1)$ denote the $100(1-\alpha_1)\%$ confidence set for θ obtained by inverting the acceptance region of this test, i.e. $S_\theta(\alpha_1) = \{\theta : F(\theta) \leq \lambda_T(\theta, \alpha_1)\}$. This confidence set for θ is essentially a multivariate extension of the grid bootstrap confidence set proposed by Hansen (1999), except that the critical values are simulated from the actual parametric model rather than from any residual-based bootstrap. It can also be thought of as a multivariate extension of the exact confidence set for the coefficient of a univariate Gaussian AR(1) model, proposed by Andrews (1993). While $F(\theta_0)$ may have an asymptotic χ^2 distribution, the confidence set $S_\theta(\alpha_1)$ is valid whether this is true or not and has the correct coverage not just asymptotically, but also in small samples.

Continue to suppose that θ_0 is the true parameter value. Since the model is completely specified when the parameter vector is known, it is possible to solve for critical values $\kappa_T(i, j, l, \theta_0, \alpha_2)$ such that

$$P\left[h_{ij}^l(\hat{\theta}) - \frac{\kappa_T(i, j, l, \theta_0, \alpha_2)}{\sqrt{T}} \sqrt{\frac{dh_{ij}^l(\hat{\theta})}{d\theta'}} W(\hat{\theta}) \frac{dh_{ij}^l(\hat{\theta})'}{d\theta'}} \leq h_{ij}^l(\theta_0) \leq h_{ij}^l(\hat{\theta}) + \frac{\kappa_T(i, j, l, \theta_0, \alpha_2)}{\sqrt{T}} \sqrt{\frac{dh_{ij}^l(\hat{\theta})}{d\theta'}} W(\hat{\theta}) \frac{dh_{ij}^l(\hat{\theta})'}{d\theta'}}\right] = \alpha_2,$$

by simulation, to an arbitrary degree of accuracy. A different simulation is required for each impulse response and for each θ_0 . Of course, the true value of the parameter

vector is unknown, but $S_\theta(\alpha_1)$ is an exact confidence set for it. So, by the Bonferroni inequality, the confidence interval

$$h_{ij}^l(\hat{\theta}) \pm \frac{\kappa_T^*(i, j, l, \alpha_1, \alpha_2)}{\sqrt{T}} \sqrt{\frac{dh_{ij}^l(\hat{\theta})}{d\theta'} W \frac{dh_{ij}^l(\hat{\theta})'}{d\theta'}} \quad (3.1)$$

has coverage of at least $100(1-\alpha_1 - \alpha_2)\%$, where

$$\kappa_T^*(i, j, l, \alpha_1, \alpha_2) = \sup_{\theta \in S_\theta(\alpha_1)} \kappa_T(i, j, l, \theta, \alpha_2) = \sup_{\theta: F(\theta) \leq \lambda(\theta, \alpha_1)} \kappa_T(i, j, l, \theta, \alpha_2). \quad (3.2)$$

This statement is true in small samples, and requires no assumptions that $\hat{\theta}$ is root-T consistent or asymptotically normal. The critical values $\kappa_T^*(i, j, l, \alpha_1, \alpha_2)$ can be obtained by standard numerical optimization methods, although this is computationally intensive, since simulation methods must be used to obtain the function evaluations⁷. The confidence intervals in equation 3.1 are those proposed in this paper⁸. Loosely speaking, they consist of using a parametric bootstrap to size-adjust the delta method confidence intervals. In the same way, it would be possible to size-adjust the other approaches to inference for confidence intervals (the regular bootstrap, the bias-adjusted

⁷In the optimisation in equation 3.2, the parameter space for θ_A is unrestricted, while the parameter space for θ_Ω is restricted only by the requirement that Ω must be positive definite.

⁸Another possible conservative confidence interval for the $i - j$ impulse response at lead time l is $(\inf_{\theta \in S_\theta(\alpha_1)} h_{ij}^l(\theta), \sup_{\theta \in S_\theta(\alpha_1)} h_{ij}^l(\theta))$, which has coverage of at least $100(1-\alpha_1)\%$. But simulations indicate that this confidence interval is far wider than the conservative confidence intervals proposed here.

bootstrap or the Bayesian method). But this would be still more computationally intensive.

A large part of what makes inference about impulse responses difficult is that the t-statistics associated with these impulse responses are far from being pivotal in empirically relevant sample sizes, as discussed in particular by Kilian (1999). This lack of pivotalness can arise both in nonstationary (or nearly nonstationary) VARs and in VARs that are clearly stationary. Meanwhile, the bootstrap is often motivated based on the assumption of asymptotic pivotalness (Hall (1992)). If these t-statistics were exactly pivotal, then $\kappa_T(i, j, l, \theta, \alpha_2)$ would not depend on θ and would equal $\kappa_T^*(i, j, l, \alpha_1, \alpha_2)$. The confidence intervals in 3.1 would then have an effective coverage of exactly $100(1-\alpha_2)\%$. Setting $\alpha_1 = 0$ (effectively dispensing with the $F(\theta_0)$ test), the confidence intervals in equation 3.1 would then be exact. But the t-statistics associated with the impulse response are typically far from being pivotal. The proposed confidence intervals can be thought of as circumventing this problem by calculating an upper bound on their lack of pivotalness.

The proposed confidence intervals must by construction have an effective coverage rate greater than or equal to the nominal level, in all sample sizes. They are exactly conservative confidence intervals⁹. While the formal justification of these con-

⁹In a slight abuse of language, following for example Campbell and Dufour (1997), I refer to these confidence intervals as “exact” in the title of the paper.

confidence intervals requires Gaussianity (as the model must be completely specified in order to simulate $\lambda(\theta_0, \alpha_1)$ and $\kappa_T(i, j, l, \theta, \alpha_2)$), I conjecture that they are also conservative in many reasonable nonnormal VARs. The fact that the coverage of these confidence intervals is typically above the nominal level is not a problem in itself. If a researcher is concerned that the coverage of a confidence interval is “too high”, then it is easy to fix this alleged problem by simply specifying that the confidence interval is empty with some specified probability, thereby raising the probability of making a Type I error. But conservative confidence intervals can be expected to be relatively wide. The effective coverage and mean width of the alternative confidence intervals will be explored by Monte-Carlo simulations in section 4.

4. Monte-Carlo Results

In this section, I report the results of Monte-Carlo simulations, assessing the properties of the proposed confidence intervals and comparing them with those of the delta method, bootstrap, bias-adjusted bootstrap and Bayesian confidence intervals. The simulation design follows Griffiths and Lütkepohl (1993), Fachin and Bravetti (1996) and Kilian (1998a, 1999). The population model is a bivariate VAR(1) of the form

$$y_t = \begin{pmatrix} \beta & 0 \\ 0.5 & 0.5 \end{pmatrix} y_{t-1} + \varepsilon_t,$$

where ε_t is iidN($0, \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix}$). The model is estimated either with just an intercept, or with an intercept and a trend. The true values of the intercept/trend are equal to zero.

The construction of the proposed confidence intervals is considerably more computationally demanding than any of the other algorithms. Most notably, a new set of simulations is required for *each* impulse response at *each* lead time. It takes about 15 minutes with a 800Mhz. processor to construct the proposed confidence interval for one impulse response at one particular lead time in this simple bivariate VAR. While this is not a problem for any one replication, the computational burden in a Monte-Carlo context is very considerable. I therefore restrict myself to evaluating the confidence intervals for the 1-2 impulse response, h_{12}^l , with $l=5,10,15$ and $\beta=0.5,0.9,0.97,1$, both with and without a time trend. The sample size is 100. These are parameter configurations which generated some of the most interesting results in Kilian (1998a)¹⁰. I use 500 replications in all the simulations to approximate $\lambda(\theta_0, \alpha_1)$ and $\kappa_T(i, j, l, \theta, \alpha_2)$. For simulating the properties of the proposed confidence intervals, I conduct 500 replications of each of the Monte-Carlo experiments (total CPU

¹⁰These are also similar to the parameter configurations in the results reported by Fachin and Bravetti (1996), who likewise considered only the 1-2 impulse responses, with $l=0,1,2,3$ and $\beta=-0.5$ and 0.9. However, I think that it is important to consider impulse responses at higher lead times because they are of interest to applied researchers and because the coverage problems generally get more serious as the lead time rises.

time in each of the 24 experiments is about 5 days). This is a small number of replications, but it reflects the high computational cost. Besides, the differences that I find between the properties of the alternative methods for forming confidence intervals are so large that they are clearly significant, even with this small number of replications.

The effective coverage of the various alternative confidence intervals and their mean width are reported in Tables 1 and 2, respectively. These confidence intervals all have nominal coverage of 68%. I follow Sims and Zha (1999) in considering confidence intervals of 68% coverage, instead of more conventional 90% or 95% confidence intervals. The delta method, bootstrap (regular and bias-adjusted) and Bayesian confidence intervals all have coverage that can be substantially too low. The problem gets worse as the lead time l increases and as β gets closer to 1. In these simulations, the bias-adjusted bootstrap generally has the highest coverage, with the Bayesian, delta method and bootstrap confidence intervals following in that order. The bias-adjusted bootstrap does well in terms of coverage provided that β is 0.97 or lower, but it too breaks down in the case of an exact unit root. In the model with a time trend, if $\beta = 1$, none of these methods achieves an effective coverage above 31%, for any lead time. On the other hand, the proposed confidence intervals have coverage that is consistently around 90%. This is not materially affected by β or l . The difference in effective coverage rates across the alternative methods is unambiguous, notwithstanding the small number of replications.

The proposed confidence intervals are conservative in all sample sizes, by construction, and so it is not surprising that they have an effective coverage well above the nominal level. The mean width of the alternative confidence intervals is also of interest. The proposed confidence intervals are much wider than confidence intervals constructed by any of the other methods: they are generally between 2 and 6 times wider than the bias-adjusted bootstrap confidence intervals. This is the price that we pay for controlling coverage. Considering how poor the coverage of the other confidence intervals can be, it may be a price worth paying. The conventional frequentist approach to forming confidence intervals specifies lexicographic preferences, as it minimizes width subject to the constraint that coverage must be controlled. By this criterion, the proposed confidence intervals must be preferred to all the alternatives, since none of these alternatives controls coverage uniformly in the parameter space. Even taking a less rigid approach to the tradeoff between coverage and width, it still seems helpful to have an algorithm for constructing impulse response confidence intervals that is guaranteed to control coverage. It means that the researcher can separate those situations in which the significance of an impulse response is beyond question (provided that the model is correctly specified) from those in which it may be an artifact of the serious econometric difficulties associated with delta method, bootstrap (regular or bias-adjusted) or Bayesian confidence intervals. In the bivariate VAR considered in these simulations, if $\beta = 1$, then the true 1-2 impulse

response at an infinite lead time is 1. Meanwhile, if $\beta < 1$, then it is 0. When the true value of β is close to 1, it is not possible reliably to infer from a sample whether $\beta = 1$ or β is slightly below 1. So a confidence interval for the impulse response at a long lead time should have a width of at least 1: otherwise it cannot hope to control coverage uniformly in β . The delta method, bootstrap, bias-adjusted bootstrap and Bayesian confidence intervals all have mean width in the range 0.25-0.45, while the proposed confidence intervals have mean width in the range 0.9-4. In the light of this intuition for what the width of a properly constructed confidence interval should be, it seems that the width of the proposed confidence intervals is mainly an appropriate statement of our uncertainty about impulse responses.

5. Conclusion

In this paper, I have proposed a new method for forming confidence intervals for impulse responses in Gaussian VARs. The method has the advantage that it is exactly conservative; the resulting confidence intervals have coverage that is greater than or equal to the nominal level in all sample sizes. By contrast, all of the existing methods for forming confidence intervals for impulse responses can easily have effective coverage far below the nominal level in empirically relevant models and sample sizes.

The proposed confidence intervals are relatively wide. The conventional frequentist approach to forming confidence intervals specifies lexicographic preferences, as it minimizes width subject to the constraint that coverage must be controlled. By this criterion, the proposed confidence intervals must be preferred to all the alternatives, since none of these alternatives controls coverage. Even taking a less rigid approach to the tradeoff between coverage and width, it still seems helpful to have an algorithm for constructing impulse response confidence intervals that is guaranteed to control coverage. Indeed, I have argued that the width of the proposed confidence intervals is mainly an appropriate statement of our uncertainty about impulse responses in VARs that are on the stationary-nonstationary borderline.

The algorithm for constructing these confidence intervals is computationally intensive, but is still workable in some simple VARs. As computational power increases, the method proposed in this paper will become practical in larger VARs. Perhaps in future research some other algorithm will be proposed that gives shorter confidence intervals that still control coverage. The algorithm in this paper can be thought of as a size-adjustment to the delta method: conceptually exactly the same size-adjustment could be applied to the bias-adjusted bootstrap. This might give a shorter confidence interval while still controlling coverage, because of its superior location, but the computational cost of this is at present entirely prohibitive.

References

- Andrews, D.W.K. (1993): Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models, *Econometrica*, 61, pp.139-166.
- Benkwitz, A., H. Lütkepohl and M.H. Neumann (2000): Problems Related to Confidence Intervals for Impulse Responses of Autoregressive Processes, *Econometric Reviews*, 19, pp.69-103.
- Campbell, B. and J.M. Dufour (1997): Exact Nonparametric Tests of Orthogonality and Random Walk in the Presence of a Drift Parameter, *International Economic Review*, 38, pp.151-173.
- Fachin, S. and L. Bravetti (1996): Asymptotic Normal and Bootstrap Inference in Structural VAR Analysis, *Journal of Forecasting*, 15, pp.329-341.
- Doan, T.A. (1990): "RATS User's Manual", VAR Econometrics, Evanston.
- Griffiths, W. and H. Lütkepohl (1993): Confidence Intervals for Impulse Responses from VAR Models: A Comparison of Asymptotic Theory and Simulation Approaches. Mimeo.
- Hall, P. (1992): "The Bootstrap and the Edgeworth Expansion", Springer-Verlag, New York.
- Hansen, B.E. (1999): The Grid Bootstrap and the Autoregressive Model, *Review of Economics and Statistics*, 81, pp.594-607.
- Kilian, L. (1998a): Small-Sample Confidence Intervals for Impulse Response Func-

- tions, *Review of Economics and Statistics*, 80, pp.218-230.
- Kilian, L. (1998b): Confidence Intervals for Impulse Responses Under Departures from Normality, *Econometric Reviews*, 17, pp.1-29.
- Kilian, L. (1999): Finite Sample Properties of Percentile and Percentile-t Bootstrap Confidence Intervals for Impulse Responses, *Review of Economics and Statistics*, 81, pp.652-660.
- Lütkepohl, H. (1990): Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models, *Review of Economics and Statistics*, 72, pp.116-125.
- Lütkepohl, H. (1996): Testing for Nonzero Impulse Responses in Vector Autoregressive Processes, *Journal of Statistical Planning and Inference*, 50, pp.1-20.
- Phillips, P.C.B. (1998), Impulse Response and Forecast Variance Asymptotics in Non-stationary VARs, *Journal of Econometrics*, 83, 21-56.
- Runkle, D.E. (1987): Vector Autoregressions and Reality, *Journal of Business and Economic Statistics*, 5, pp.437-442.
- Sims, C.A. (1980): Macroeconomics and Reality, *Econometrica*, 48, pp.1.48.
- Sims, C.A. and T. Zha (1999): Error Bands for Impulse Responses, *Econometrica*, 67, pp.1113-1155.
- Wright, J.H. (2000): Confidence Intervals for Univariate Impulse Responses with a Near Unit Root, *Journal of Business and Economic Statistics*, 18, pp.368-373.

Table 1: Effective Coverage of Alternative Confidence Intervals for the 1-2 Impulse Response
(68% Nominal Coverage)

| | | Delta Method | Bootstrap | B.A. Bootstrap | Bayesian | Proposed |
|---------------------------------|----------|--------------|-----------|----------------|----------|----------|
| <i>Model with no time trend</i> | | | | | | |
| $\beta = 0.5$ | $l = 5$ | 57.8 | 53.0 | 65.4 | 68.2 | 85.6 |
| | $l = 10$ | 57.8 | 58.0 | 64.0 | 89.2 | 84.2 |
| | $l = 15$ | 66.4 | 71.0 | 77.6 | 98.0 | 85.6 |
| $\beta = 0.9$ | $l = 5$ | 53.8 | 38.2 | 68.4 | 59.4 | 85.2 |
| | $l = 10$ | 51.2 | 33.4 | 65.4 | 62.6 | 86.6 |
| | $l = 15$ | 48.2 | 34.0 | 64.0 | 65.2 | 87.6 |
| $\beta = 0.97$ | $l = 5$ | 43.4 | 25.0 | 67.0 | 50.0 | 92.0 |
| | $l = 10$ | 35.4 | 10.6 | 75.0 | 48.6 | 88.0 |
| | $l = 15$ | 34.6 | 6.8 | 76.2 | 50.8 | 85.0 |
| $\beta = 1$ | $l = 5$ | 36.0 | 15.0 | 53.2 | 42.2 | 94.6 |
| | $l = 10$ | 20.0 | 4.0 | 50.8 | 30.8 | 91.8 |
| | $l = 15$ | 14.8 | 2.0 | 52.2 | 27.8 | 88.8 |
| <i>Model with time trend</i> | | | | | | |
| $\beta = 0.5$ | $l = 5$ | 58.4 | 41.6 | 71.4 | 67.8 | 89.6 |
| | $l = 10$ | 52.6 | 47.6 | 68.6 | 90.0 | 87.2 |
| | $l = 15$ | 61.6 | 66.0 | 78.8 | 98.6 | 88.2 |
| $\beta = 0.9$ | $l = 5$ | 40.0 | 14.6 | 74.6 | 48.8 | 90.4 |
| | $l = 10$ | 34.6 | 5.0 | 75.4 | 46.4 | 88.6 |
| | $l = 15$ | 32.2 | 5.6 | 75.6 | 53.0 | 91.2 |
| $\beta = 0.97$ | $l = 5$ | 27.0 | 5.2 | 60.4 | 33.2 | 93.8 |
| | $l = 10$ | 15.8 | 0.0 | 64.2 | 25.2 | 89.2 |
| | $l = 15$ | 13.0 | 0.0 | 65.6 | 26.2 | 90.4 |
| $\beta = 1$ | $l = 5$ | 14.2 | 2.2 | 30.4 | 19.2 | 89.8 |
| | $l = 10$ | 3.8 | 0.0 | 26.2 | 9.4 | 88.2 |
| | $l = 15$ | 2.4 | 0.0 | 26.8 | 8.2 | 85.8 |

Notes: The different methods of constructing the confidence intervals are described in the text; B.A. bootstrap refers to the bias-adjusted bootstrap. The sample size in these simulations is 100.

Table 2: Mean Width of Alternative Confidence Intervals for the 1-2 Impulse Response
(68% Nominal Coverage)

| | | Delta Method | Bootstrap | B.A. Bootstrap | Bayesian | Proposed |
|---------------------------------|----------|--------------|-----------|----------------|----------|----------|
| <i>Model with no time trend</i> | | | | | | |
| $\beta = 0.5$ | $l = 5$ | 0.16 | 0.14 | 0.19 | 0.19 | 0.33 |
| | $l = 10$ | 0.14 | 0.04 | 0.03 | 0.06 | 0.07 |
| | $l = 15$ | 0.01 | 0.01 | 0.02 | 0.03 | 0.10 |
| $\beta = 0.9$ | $l = 5$ | 0.28 | 0.29 | 0.35 | 0.31 | 0.88 |
| | $l = 10$ | 0.30 | 0.26 | 0.47 | 0.37 | 1.28 |
| | $l = 15$ | 0.24 | 0.18 | 0.49 | 0.38 | 1.29 |
| $\beta = 0.97$ | $l = 5$ | 0.27 | 0.30 | 0.35 | 0.30 | 0.93 |
| | $l = 10$ | 0.35 | 0.35 | 0.56 | 0.42 | 1.57 |
| | $l = 15$ | 0.36 | 0.31 | 0.71 | 0.50 | 1.95 |
| $\beta = 1$ | $l = 5$ | 0.26 | 0.29 | 0.34 | 0.28 | 0.90 |
| | $l = 10$ | 0.34 | 0.38 | 0.54 | 0.40 | 1.56 |
| | $l = 15$ | 0.39 | 0.38 | 0.71 | 0.51 | 2.12 |
| <i>Model with time trend</i> | | | | | | |
| $\beta = 0.5$ | $l = 5$ | 0.15 | 0.13 | 0.20 | 0.18 | 0.39 |
| | $l = 10$ | 0.03 | 0.02 | 0.07 | 0.06 | 0.18 |
| | $l = 15$ | 0.01 | 0.00 | 0.02 | 0.03 | 0.16 |
| $\beta = 0.9$ | $l = 5$ | 0.27 | 0.28 | 0.39 | 0.30 | 1.22 |
| | $l = 10$ | 0.26 | 0.20 | 0.50 | 0.33 | 1.90 |
| | $l = 15$ | 0.19 | 0.10 | 0.52 | 0.31 | 2.44 |
| $\beta = 0.97$ | $l = 5$ | 0.27 | 0.30 | 0.39 | 0.29 | 1.25 |
| | $l = 10$ | 0.32 | 0.27 | 0.59 | 0.38 | 2.45 |
| | $l = 15$ | 0.29 | 0.18 | 0.69 | 0.42 | 3.97 |
| $\beta = 1$ | $l = 5$ | 0.26 | 0.29 | 0.37 | 0.28 | 1.21 |
| | $l = 10$ | 0.32 | 0.28 | 0.55 | 0.38 | 2.46 |
| | $l = 15$ | 0.31 | 0.21 | 0.64 | 0.43 | 4.30 |

Notes: The different methods of constructing the confidence intervals are described in the text; B.A. bootstrap refers to the bias-adjusted bootstrap. The sample size in these simulations is 100.