Journal of Econometrics 192 (2016) 421–432

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Joint confidence sets for structural impulse responses[☆]

Atsushi Inoue [a], Lutz Kilian [b,*]

[a] Department of Economics, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, United States
[b] University of Michigan, Department of Economics, 611 Tappan Street, Ann Arbor, MI 48109-1220, United States

ARTICLE INFO

Article history:
Available online 11 February 2016

JEL classification:
C32
C52
C53

Keywords:
Joint inference
Shotgun plots
Confidence bands
Impulse response shapes
Bootstrap
Degenerate limiting distribution

ABSTRACT

Many questions of economic interest in structural VAR analysis involve estimates of multiple impulse response functions. Other questions relate to the shape of a given impulse response function. Answering these questions requires joint inference about sets of structural impulse responses, allowing for dependencies across time as well as across response functions. Such joint inference is complicated by the fact that the joint distribution of the structural impulse response estimators becomes degenerate when the number of structural impulse responses of interest exceeds the number of model parameters, as is often the case in applied work. This degeneracy may be overcome by transforming the estimator appropriately. We show that the joint Wald test is invariant to this transformation and converges to a nonstandard distribution, which can be approximated by the bootstrap, allowing the construction of asymptotically valid joint confidence sets for any subset of structural impulse responses, regardless of whether the joint distribution of the structural impulse responses is degenerate or not. We propose to represent the joint confidence sets in the form of "shotgun plots" rather than joint confidence bands for impulse response functions. Several empirical examples demonstrate that this approach not only conveys the same information as confidence bands about the statistical significance of response functions, but may be used to provide economically relevant additional information about the shape of and comovement across response functions that is lost when reducing the joint confidence set to two-dimensional bands.

1. Introduction

It is well known that impulse response estimates from structural vector autoregressive (VAR) models tend to be imprecisely estimated, given the short samples typical of applied work. This fact makes it important to assess the reliability of these estimates by constructing confidence sets. It has become standard in the literature to evaluate the statistical significance of the estimated structural impulse response functions using pointwise confidence intervals (see, e.g. Lütkepohl, 1990; Kilian, 1998). This approach is questionable because in practice many of these pointwise intervals are evaluated at the same time and pointwise intervals ignore the fact that structural impulse response estimators tend to be dependent both across horizons and across impulse response functions. As a result, confidence bands obtained by connecting pointwise confidence intervals tend to be too narrow and lack coverage accuracy, resulting in spurious findings of statistical significance. This problem has been recognized for a long time, but there is no consensus on how to overcome these distortions.

Analogous problems also arise in Bayesian inference. In related work, Sims and Zha (1999) and Inoue and Kilian (2013) have discussed possible solutions to this problem from a Bayesian point of view. The current paper, in contrast, takes a frequentist perspective. To the extent that the problem of joint impulse response confidence sets has been discussed in the frequentist VAR literature, it has often been reduced to a problem of conducting joint inference across a range of horizons for a given impulse response function. For example, Jordà (2009) proposes one solution to this problem and Lütkepohl et al. (2015) several alternatives. Simulation evidence on the finite-sample accuracy of these confidence bands is discussed in Kilian and Kim (2011) and Lütkepohl et al. (2015).

It is important to stress that these approaches, while representing an important step forward, are too restrictive for applied work. Many users of structural VAR models are interested in conducting inference about multiple impulse response functions at the same time. For example, a macroeconomist may be interested in whether an oil price shock creates stagflation in the domestic

☆ We thank Helmut Lütkepohl, Sophocles Mavroeidis, Peter C.B. Phillips, the editor, and three anonymous referees for helpful comments.
* Corresponding author.
E-mail addresses: atsushi.inoue@vanderbilt.edu (A. Inoue), lkilian@umich.edu (L. Kilian).

http://dx.doi.org/10.1016/j.jeconom.2016.02.008
0304-4076/

economy, which by necessity involves studying the responses of inflation as well as real output. The same would be true if we studied the effect of a U.S. monetary policy shock because the loss function of the Federal Reserve depends on both real output and inflation. It is also common for researchers to be interested in assessing the implications of economic theory for a range of different impulse response functions simultaneously. For example, Blanchard (1989) uses a macroeconomic VAR model to evaluate the implication of standard Keynesian models that (1) positive demand innovations increase output and decrease unemployment persistently, and (2) that a favorable supply shock triggers an increase in unemployment without a decrease in output. This example involves inference about four impulse response functions simultaneously. There even are cases in which users of structural VAR models are interested in studying the responses of all model variables to all structural shocks simultaneously. A good example is recent structural VAR models of the global market for industrial commodities such as crude oil (e.g., Kilian, 2009).

A proper solution to this problem requires taking account of the dependence of all structural impulse responses of interest, not just of the responses in a given impulse response function. This is the objective of the current article. We propose a novel approach to constructing asymptotically valid joint confidence sets for any subset of the structural impulse responses of interest based on inverting a Wald test statistic. One difficulty in this context is that in many situations the standard asymptotic results for the joint distribution of the structural impulse responses do not hold even in stationary vector autoregressions. Specifically, when the number of structural impulse responses exceeds the number of model parameters, the joint asymptotic distribution is degenerate, and the distribution of the Wald test statistic is nonstandard. This problem has also been noted in Lütkepohl and Poskitt (1991, p. 493), for example.

This degeneracy may be overcome by transforming the estimator appropriately. We show that the joint Wald test statistic is invariant to this transformation and converges to a nonstandard distribution, which can be approximated by the bootstrap, thus providing a theoretical justification for the use of the bootstrap Wald test statistic in constructing joint confidence sets for structural impulse responses even in the absence of joint asymptotic normality. This result greatly extends the range of problems that can be addressed with this bootstrap method.[1] Although the current paper is concerned with structural impulse response analysis, our approach of addressing the potential degeneracy of the joint asymptotic distribution of the estimator of the structural impulse responses is of more general interest and may be adapted to other inference problems.[2]

We show that in idealized settings the joint Wald confidence region is expected to be smaller than alternative confidence sets such as the Bonferroni set. We also analyze the coverage accuracy of these confidence sets in finite samples by simulation. Our simulation design focuses on data generating processes with many lags, large roots, and realistic sample sizes. We find that the bootstrap Wald confidence set to be reasonably accurate even in large-dimensional and highly persistent VAR models, while the Bonferroni approach is conservative. The latter result is not unexpected because the number of structural impulse responses of interest in these models is large.

A closely related approach has been proposed independently in Lütkepohl et al. (2014). One difference is that we focus on the Wald test statistic for the vector of structural impulse responses, γ, whereas Lütkepohl et al. (2014) utilize a Wald test statistic for the parameters of the VAR model, θ. They then infer the confidence region for γ from the mapping $\gamma = \gamma(\theta)$. We contrast these two approaches and note that the Wald test statistic in Lütkepohl et al. (2014) has certain theoretical advantages in constructing joint confidence regions for impulse responses compared with our approach. Our simulation study, however, suggests that the differences in coverage accuracy tend to be small. In the few cases in which there is a larger difference in coverage accuracy, the Wald test statistic for γ yields more accurate confidence regions.

The second difference between our analysis and Lütkepohl et al.'s is that we propose to plot the sets of impulse responses associated with each bootstrap draw that is contained in the joint Wald confidence set, resulting in plots for each impulse response function that resemble a shotgun trajectory chart ("shotgun plot"). In contrast, Lütkepohl et al. (2014) connect the pointwise maxima of the impulse responses in the joint set to form the upper bound of a confidence band and similarly connect the pointwise minima of the impulse responses in the joint set to form the lower bound of a confidence band. This approach results in a loss of information compared with our shotgun plots because one is no longer able to discern the evolution over time of the impulse response functions associated with any one structural model estimate in the joint confidence set.

It is precisely this evolution of the response function that users of structural VAR models typically are interested in (see, e.g. Cochrane, 1994). For example, many macroeconomists have abandoned frictionless neoclassical models and adopted models with nominal or real rigidities based on VAR evidence of sluggish or delayed responses of inflation and output (see, e.g. Woodford, 2003). This is also true for other applications. Whereas macroeconomists are interested in whether a response function for real output is hump-shaped or not, for example, users of structural VAR models in international economics often are interested in whether there is delayed overshooting in the response of the exchange rate to monetary policy shocks. It is difficult to answer such questions about the shape of a given impulse response function based on two-dimensional joint confidence bands in general because such bands are consistent with a multitude of different response patterns. These difficulties are compounded when considering the analysis of more than one impulse response function at a time, as is common in applied work.

We illustrate these points based on two empirical examples. Our first empirical example is a semi-structural model of U.S. monetary policy; the second example is a semi-structural model of the response of the U.S. economy to oil price shocks. We use these examples to illustrate that in some situations, the use of shotgun plots and of joint confidence bands will yield the same answer by construction. For example, if we are interested only in one impulse response function at a time, the lower bound of the joint confidence band will include zero at some horizon, if and only if some of the impulse response functions in the shotgun plot crosses zero. In other situations, the shotgun plots implied by joint Wald confidence sets may reveal features of the data that are obscured by more traditional pointwise confidence intervals or by two-dimensional joint confidence bands. For example, the shotgun plot

[1] A similar econometric problem has been described in a different context by Andrews (1987). Andrews provides a sufficient condition that allows standard inference based on asymptotic χ^2-critical values even when the covariance matrix used in constructing the test statistic is asymptotically singular. This condition does not apply in our setting, because the bootstrap covariance matrix is almost always positive definite. Its eigenvalues are positive in finite samples and equal to zero only in the limit. Thus, our approach is of independent interest. An illustrative example is provided in the working paper version of this article.

[2] Related work also includes Dufour and Valery (2015) who propose regularized Wald tests in the presence of singular covariance matrices. Their approach is more general than ours in some dimensions, including the fact that it allows estimators of parameters to have non-Gaussian asymptotic distributions. Our analysis, in contrast, is specifically geared to structural impulse response estimators and exploits the asymptotic normality of such estimators. The advantage of our approach is that it does not require the choice of a tuning parameter.