Bootstrapping Structural VARs:

Avoiding a Potential Bias in Confidence Intervals for Impulse Response Functions[1]

Kerk L. Phillips
Department of Economics
P.O. Box 22363
Brigham Young University
Provo, UT 84602-2363
phone: (801) 422-5928
fax: (801) 422-2844
email: kerk_phillips@byu.edu

David E. Spencer
Department of Economics
P.O. Box 22363
Brigham Young University
Provo, UT 84602-2363
phone: (801) 422-7277
fax: (801) 422-2844
email: david_spencer@byu.edu

June 2008

JEL Codes: C150, C320, E370

Keywords: impulse response function, structural VAR, bias, bootstrap

Abstract

Constructing bootstrap confidence intervals for impulse response functions (IRFs) from structural vector autoregression (SVAR) models has become standard practice in empirical macroeconomic research. The accuracy of such confidence intervals can deteriorate severely, however, if the bootstrap IRFs are biased, resulting in inappropriate inference. In this paper, we document an apparently common source of bias in the estimation of the VAR error covariance matrix. The bias is easily corrected with a straightforward scale adjustment. This bias is often unrecognized because it only affects the bootstrap estimates of the error variance, not the original OLS estimates. Nevertheless, as we illustrate here, analytically, with sampling experiments, and in an example from the literature, the bootstrap error variance bias can have significant distorting effects on bootstrap IRF confidence intervals even if the original IRF estimate relies on unbiased parameter estimates.

[1] We are grateful to Lutz Kilian and James B. McDonald for their comments on previous versions of this paper.

I. Introduction

In this paper we identify and document a commonly occurring, but easily corrected, source of bias in bootstrap estimates of impulse response functions (IRFs) from structural vector autoregression (SVAR) models. Given the extensive use of SVARs in macroeconomics and the pervasiveness of the techniques that lead to this bias, it has important implications.

IRFs from SVAR models are widely employed to investigate the response of key macroeconomic variables to identified structural shocks. In order to assess uncertainty and draw inferences, it is important to construct confidence intervals around each point of the estimated IRF. These intervals can be based on asymptotic theory; see Lütkepohl (1990) and Mittnik and Zadrosny (1993). Alternatively, confidence intervals can be constructed using bootstrap techniques; see Runkle (1987), and Berkowitz and Kilian (2000). Indeed, use of bootstrap techniques has become increasingly popular.

The accuracy of such bootstrap confidence intervals depends on the bias of the bootstrap IRFs. That traditional bootstrap techniques tend to be biased has been discussed by Kilian (1998) and Sims and Zha (1999). As noted in those papers, such bias can lead to significantly distorted bootstrap confidence intervals. Kilian (1998, p. 220) indicates that this bias "explains the curious result … that the standard bootstrap [confidence] interval often tends to lie almost entirely below (or above) the initial OLS point estimate." Sims and Zha (1999, p. 1125, fn 13) note that some studies have found it necessary to "use a modification of [the bootstrap confidence interval] that makes *ad hoc* adjustments to prevent the computed bands from failing to include the point estimates."

This bias-caused distortion can also be seen in the results reported by Blanchard and Quah (1989); see especially their Figures 3 and 5 in which one side of the one standard deviation bounds appears to actually coincide with the original estimated IRF over some horizon intervals. The fact that the IRF does not actually cross the bounds is due to the way they compute the bounds.[2]

[2] Blanchard and Quah compute their bootstrap one standard deviation bounds as follows. They obtain 1000 bootstrap IRFs which, for each horizon, they separate into those above and those below the original IRF. They then compute the standard deviation for each class to obtain the asymmetric one standard deviation bounds. This

If researchers do not allow for asymmetric confidence intervals and simply plot error bands that are the estimated IRFs plus or minus one or two standard deviations, bias is completely invisible and the reported error bands are incorrect.

When researchers do allow for asymmetric confidence intervals, bias is often misinterpreted as skewness in the distribution of the IRF. The misinterpreted bias, however, distorts the bootstrap IRFs and therefore the bootstrap confidence intervals for the original estimated IRF in important ways.

Not all researchers attribute this odd behavior of IRFs completely to skewness. Christiano, Eichebaum and Vigfusson (2006), for example, note that, in their case, the mean value of the bootstrapped IRFs is not the same as the IRFs from the original estimation. They plot both of these along with confidence intervals and note that the asymmetric percentile confidence intervals reflect a downward bias.

As emphasized in Section III below, there are several potential sources of bias in the estimation of bootstrap confidence intervals for IRFs. This is a consequence of the fact that the IRF for an identified SVAR is a highly nonlinear function of the OLS estimates of both the slope *and* error variance parameters in the underlying reduced form vector autoregression (VAR). Furthermore, the *bootstrap* IRF is the same nonlinear function of the *bootstrap* estimates of these same parameters. If these estimates are systematically biased for their population (or, in the case of the bootstrap, pseudo-population) counterparts, the resulting IRFs and the implied bootstrap confidence intervals are likely to be distorted in potentially complicated and significant ways as well.

Though there are several sources of bias, we focus on the source arising from systematically biased bootstrap estimates of the reduced form VAR error covariance matrix. This bias is apparently common[3] but easily corrected. It is not corrected in practice because it is generally unrecognized as it only affects the *bootstrap* estimates of the error variance, not the

---

procedure assures that the IRF will not “cross” the bounds. A bound that is coincident with the original IRF indicates that, at that horizon, *none* of the bootstrap IRFs were above (or below) the original IRF.

[3] Of course, we have not documented this for all or even most SVAR studies. We have, however, examined programs that authors have been willing to post on web sites. In *none* of the cases was the appropriate bias-adjusted bootstrap error covariance estimator used. We therefore conclude that this is a common practice. We prefer not to mention names, especially since one of the authors of this paper is also guilty of this common practice in his previously published papers which estimate IRFs for SVARs.

original OLS estimates. But, since the bootstrap IRFs and, thus, the bootstrap confidence intervals, depend on these bootstrap error covariance matrix estimates, the bootstrap IRF confidence intervals can be significantly distorted even if the original IRF estimate relies on unbiased parameter estimates.

In the next section, we illustrate the specific source of this bias in the bootstrap estimate of error variances in the context of simple models and confirm its impact. In Section III we show how this bias in bootstrap error variance estimates effects the bootstrap IRFs and thus the bootstrap confidence intervals for the original IRF. In Section IV we illustrate how correcting for this bias affects the IRF confidence intervals obtained in a widely-cited previous study. The final section offers a conclusion.

II. The Source of Bias

A. The Standard Linear Regression Model

The simplest way to illustrate the bias under investigation is to examine a standard linear regression model with nonstochastic regressors. We first consider a univariate regression model represented by

(1) $y = X\beta + u$

where *y* is an $T \times 1$ vector of observations on a dependent variable, *X* is an $T \times R$ matrix of observations on *R* nonstochastic regressors (perhaps including a constant), β is an $R \times 1$ vector of regression coefficients, and *u* is an $T \times 1$ vector of errors. We assume that $E(u) = 0$ and $E(uu') = \sigma^2 I_T$. Applying ordinary least squares (OLS), we obtain coefficient and error variance estimates: $\hat{\beta} = (X'X)^{-1}X'y$, $\hat{\sigma}^2 = \frac{1}{T-R}(\hat{u}'\hat{u})$, where $\hat{u} = y - X\hat{\beta}$. The indicated degrees of freedom correction makes $\hat{\sigma}^2$ is an unbiased estimator for σ^2.

To help us understand the key argument to follow, it is useful to interpret the degrees of freedom adjustment from the perspective that it is necessary to compensate for the fact that the OLS residuals tend to be smaller than the error terms. Note that the expected value of the average squared error is σ^2; i.e., $E\left(\frac{u'u}{T}\right) = \sigma^2$. On the other hand,

$E\left(\frac{\hat{u}'\hat{u}}{T}\right)=\left(\frac{T-R}{T}\right)E\left(\frac{u'u}{T}\right)$, which reflects that, on average, the squared residuals are $\left((T-R)/T\right)$ times as large as the squared errors[4]. I.e., the estimated OLS residuals tend to be "deflated" relative to the regression errors resulting in the average squared residual giving a biased estimate of σ^2. To obtain an unbiased estimate, we must "re-inflate" the residuals. There are two ways to accommodate this re-inflation:

(a) Rescale each residual by $\left(\frac{T}{T-R}\right)^{1/2}$ and then compute the average squared *rescaled* residual.

(b) Compute the average squared residual, $\left(\frac{\hat{u}'\hat{u}}{T}\right)$, and then rescale by $\left(\frac{T}{T-R}\right)$.

Either procedure results in the usual unbiased estimate for σ^2, $\hat{\sigma}^2=\frac{1}{T-R}(\hat{u}'\hat{u})$.

Now, suppose we obtain bootstrap estimates of the error variance as follows. For bootstrap replications b=1,…, B, generate

(2) $y_b = X\hat{\beta}+u_b^*$

where the elements of u_b^* are drawn with replacement from the OLS residuals, $\hat{u}$. Then, apply OLS to equation (2) to get bootstrap estimates of $\hat{\beta}$, which we denote $\tilde{\beta}_b$, and bootstrap residuals, $\tilde{u}_b$, from which we calculate the bootstrap "error" variance estimate $\tilde{\sigma}_b^2$. There at least two candidates we might consider for $\tilde{\sigma}_b^2$. Perhaps the natural choice follows from the original OLS estimator:

(3) $\tilde{\sigma}_{b(1)}^2=\frac{1}{T-R}(\tilde{u}_b{}'\tilde{u}_b)$

An alternative and asymptotically equivalent choice is

(4) $\tilde{\sigma}_{b(2)}^2=\frac{T}{(T-R)^2}(\tilde{u}_b{}'\tilde{u}_b)=\frac{T}{T-R}\tilde{\sigma}_{b(1)}^2$

Obviously, the choice makes no difference asymptotically but there will be a difference in small samples.

[4] See Davidson and MacKinnon (1993), pp. 69-70.

To see that the estimate given by (4) is likely to be superior, note that, given the analogy on which the bootstrap technique is based, $\tilde{\sigma}_b^2$ is an estimate of $\hat{\sigma}^2$, the "population" error variance in the pseudo-population given by the OLS residuals. I.e., in the case of the error variance estimator, the basic idea of the bootstrap is that the bootstrap estimate $\tilde{\sigma}_b^2$ is expected to relate to the OLS estimate $\hat{\sigma}^2$ as $\hat{\sigma}^2$ relates to the unobservable population parameter σ^2. Since $\hat{\sigma}^2$ is an unbiased estimate of σ^2, the bootstrap only works if $\tilde{\sigma}_b^2$ is an unbiased estimate of $\hat{\sigma}^2$.

Proceeding as in our discussion above, we note that $E\left(\frac{\hat{u}'\hat{u}}{T-R}\right)=\sigma^2$. On the other hand, $E\left(\frac{\tilde{u}_b{}'\tilde{u}_b}{T-R}\right)=\left(\frac{T-R}{T}\right)E\left(\frac{u_b^{*\prime}u_b^*}{T-R}\right)=\left(\frac{T-R}{T}\right)\sigma^2$ since the elements of u_b^* are drawn randomly from $\hat{u}$. This relationship reflects that, on average, the squared bootstrap residuals are $\left((T-R)/T\right)$ times as large as the squared OLS residuals which have become the pseudo-population errors. Consequently, $\tilde{\sigma}_{b(1)}^2$ given by equation (3) yields a biased estimate of $\hat{\sigma}^2$ which, in turn, is an unbiased estimate of σ^2. We can obtain an unbiased bootstrap estimate of $\hat{\sigma}^2$ by re-inflating the residuals in one of two equivalent ways:

(a) Rescale each bootstrap residual by $\left(\frac{T}{T-R}\right)^{1/2}$ and then compute the degrees of freedom adjusted average squared *rescaled* residual; i.e., divide the sum of squared rescaled residuals by $(T-R)$.

(b) Compute the usual degrees of freedom adjusted average squared bootstrap residual, $\left(\frac{\tilde{u}_b{}'\tilde{u}_b}{T-R}\right)$, and then rescale by $\left(\frac{T}{T-R}\right)$.

Either procedure results in the unbiased estimator $\tilde{\sigma}_{b(2)}^2$ given by (4)[5].

Even though the estimator $\tilde{\sigma}_{b(1)}^2$ may seem a natural choice, it will result in a biased

[5] The virtue of rescaling bootstrap residuals has been noted in the literature before. See, e.g., Freedman and Peters (1984, p. 99), Peters and Freedman (1984, p. 408), Stine (1987, p. 1074) and Berkowitz and Kilian (2000, p. 5). In spite of this recognition, it is generally ignored in practice by empirical macroeconomists.

estimate of $\hat{\sigma}^2$ while $\tilde{\sigma}^2_{b(2)}$ gives an unbiased estimate. The size of the (proportional) bias for the natural estimator is $-R/T$[6]. This vanishes asymptotically but can be important in small samples where *R* is large relative to *T*. To illustrate this, we conduct simple Monte Carlo experiments in which we simulate obtaining bootstrap estimates of the error variance in a simple univariate regression model like (1). We choose to estimate models with nine regressors including a constant term, $R = 9$, for three sample sizes: $T = 30,\ 50,\ 100$. Consequently, the true bias for $\tilde{\sigma}^2_{b(1)}$ is -30%, -18% and -9% respectively. For each sample size, we draw 1000 samples of size *T* from a normal distribution with mean zero and variance 0.81. For each of these Monte Carlo draws we generate observations for *y*, estimate (1) by OLS and compute the usual unbiased estimate of the error variance, $\hat{\sigma}^2$. The average estimate is given in Table 1. To examine the bias of the two bootstrap error variance estimates, $\tilde{\sigma}^2_{b(1)}$ and $\tilde{\sigma}^2_{b(2)}$, we take each of the 1000 Monte Carlo samples and obtain 1000 bootstrap estimates in each case. The average values are reported in Table 1 for our three sample sizes.

The results reported in Table 1 confirm the theory very nicely. The "natural" bootstrap estimator, $\tilde{\sigma}^2_{b(1)}$, has bias approximately equal to $-R/T$ while the other estimators are unbiased.

These theoretical results can be readily extended to a multivariate seemingly unrelated regression model of dimension *K* as well. In anticipation of extension to VAR models, we assume that the regressors are the same in each equation. The model is given by

(5) $y_i = X\beta_i + u_i,\ i = 1,\ldots,K$

where y_i and u_i are $T \times 1$ vectors, X is a $T \times R$ matrix, and the β_i are $R \times 1$ vectors. Define

---

[6] It should be noted that bias arising from maximum likelihood estimation (MLE) of the error variance will be even larger. As is well known, the MLE of σ^2, $\breve{\sigma}^2 = \frac{1}{T}(\hat{u}'\hat{u})$, is biased; i.e., $E\left(\breve{\sigma}^2\right) = \left(\frac{T-R}{T}\right)\sigma^2$. Thus, the proportional bias is $-R/T$. Now, when we bootstrap and obtain the MLE of $\breve{\sigma}^2$, $\tilde{\breve{\sigma}}^2_b = \frac{1}{T}\left(\tilde{u}_b{}'\tilde{u}_b\right)$, the bias is magnified since we have a biased estimate of a biased estimate. $\tilde{\breve{\sigma}}^2_b = \frac{(T-R)^2}{T^2}\tilde{\sigma}^2_{b(2)}$, so $E\left(\tilde{\breve{\sigma}}^2_b\right) = \frac{(T-R)^2}{T^2}\sigma^2$ and the proportional bias is $\left(\frac{(T-R)^2}{T^2}\right) - 1 = \frac{R^2 - 2TR}{T^2}$ which is negative and larger (in absolute value) than $-R/T$.

the $TK \times 1$ vector $u = \left(u_1' \ldots u_K'\right)'$ and the $T \times K$ matrix $U = \left[u_1 \ldots u_K\right]$. We assume that $E\left(uu'\right) = \Sigma \otimes I_T$ where Σ is the contemporaneous covariance matrix. The OLS estimate of Σ is $\hat{\Sigma} = \frac{1}{T-R}\hat{U}'\hat{U}$ where $\hat{U}$ is the matrix of OLS residuals. The alternative bootstrap estimators are $\tilde{\tilde{\Sigma}}_{b(1)} = \frac{1}{T-R}\tilde{U}_b'\tilde{U}_b$ (biased) and $\tilde{\tilde{\Sigma}}_{b(2)} = \frac{T}{\left(T-R\right)^2}\tilde{U}_b'\tilde{U}_b = \frac{T}{\left(T-R\right)}\tilde{\tilde{\Sigma}}_{b(1)}$ (unbiased) where $\tilde{U}_b$ is the matrix of residuals from the bth bootstrap iteration. Thus the elements of the "natural" bootstrap estimate of Σ, $\tilde{\tilde{\Sigma}}_{b(1)}$, are all biased by a factor of $-R/T$. To confirm the theory, we have conducted simple Monte Carlo experiments similar to those undertaken for the univariate regression model and discussed above. To save space, we do not report the results here but simply indicate that the conclusions are the same[7].

B. The AR(p) Model

Consider a univariate autoregressive model of order *p* [AR(p)] with a constant term so that $R = p+1$:

(6) $$y_t = \nu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + u_t; \quad T = -p+1, \ldots, 0, 1, \ldots, T$$

where u_t is white noise with variance σ^2. Because the regressors are stochastic, the finite sample theory for the standard regression model above does not apply. However, we might speculate (correctly) that similar bias problems exist for bootstrap estimators of the error variance in this case.

Since exact analytical results are not available, we examine the small-sample bias issue for the AR(*p*) model using a Monte Carlo exercise similar to the one described above. We generate data for, and estimate, a model like (6) in which $p = 8$ so $R = 9$. For each of three sample sizes, $T = 30,\ 50,\ 100$, we draw 1000 samples for u_t of size *T+p* from a normal distribution with mean zero and variance 0.81. For each of these Monte Carlo draws we generate observations for *y*, estimate (1) by OLS, and compute the usual estimate of the error variance, $\hat{\sigma}^2$. The average estimate is given in Table 2. To examine the bias of the two

[7] The results are available on request.

bootstrap error variance estimates, $\tilde{\sigma}^2_{b(1)}$ and $\tilde{\sigma}^2_{b(2)}$, we obtain 1000 bootstrap estimates for each of the Monte Carlo samples[8]. The average values are reported in Table 2 for each of our three sample sizes.

The results are quite informative. The theoretical bias for the corresponding standard linear regression is a rather good guide for the bias in the AR(*p*) model. We confirm that the bootstrap estimator of the error variance given by $\tilde{\sigma}^2_{b(1)}$ is biased and thus likely to lead to incorrect inference when the number of slope coefficients is large relative to the sample size.

We expect these bias results to carry over to the case of a VAR(*p*) with *K* variables. In that case, our interest is the $K \times K$ error (innovation) covariance matrix Σ. Assuming a constant term, the usual degrees-of-freedom-corrected OLS estimator for Σ is $\hat{\Sigma} = \left(\frac{1}{T-Kp-1}\right)\hat{U}'\hat{U}$ where $\hat{U}$ is the $T \times K$ matrix of OLS residuals. The "natural" *bootstrap* estimator of Σ is $\tilde{\Sigma}_{b(1)} = \left(\frac{1}{T-Kp-1}\right)\tilde{U}_b{}'\tilde{U}_b$ where $\tilde{U}_b$ is the $T \times K$ matrix of bootstrap residuals from the b^{th} bootstrap iteration. The alternative bootstrap estimator of Σ is $\tilde{\Sigma}_{b(2)} = \left(\frac{T}{(T-Kp-1)^2}\right)\tilde{U}_b{}'\tilde{U}_b$. $\tilde{\Sigma}_{b(2)}$ corresponds to the unbiased bootstrap error covariance matrix in the standard multiple regression model, described by (5) above, with $R = Kp+1$.

We have investigated the bootstrap error variance bias for a two-equation VAR(8) model with a constant term using Monte Carlo methods and find the bias to be quite close to the theoretical bias from the corresponding multivariate regression model. To conserve space, we do not report the results here since they are quite similar to those reported for the AR(8) model above[9]. In particular, the bias for $\tilde{\Sigma}_{b(1)}$ is approximately $-\left(\frac{Kp+1}{T}\right)$ where *K* is the number of equations (variables) in the VAR(*p*). For a two-equation VAR(8) model, this implies

---

[8] For each bootstrap iteration, we obtain the initial *p* observations $\{y_{-p+1},..., y_o\}$ by drawing (with replacement) from the original generated sample $\{y_t\}^T_{-p+1}$.

[9] Results are available on request.

an approximate bias of -17% for each element of Σ when $T = 100$ [10].

III. Bootstrapping IRFs for SVARs

The previous section demonstrates that the natural bootstrap estimator of the VAR error covariance matrix, $\tilde{\Sigma}_{b(1)}$, is systematically biased and indicates an appropriate bias correction which yields the estimator $\tilde{\Sigma}_{b(2)}$. Of course, this is of no concern if we are interested only in the properties of statistics that do not depend on the error covariance estimates. If, for example, we are interested only in conducting hypothesis tests for the slope coefficients in a VAR, we could base appropriate tests on the bootstrap percentile confidence intervals for the slope coefficients. These do not require estimates of the error covariance matrix and, thus, would not suffer from the bias discussed here. However, in SVAR models, we are often interested in drawing inferences about IRFs which are highly nonlinear functions of both VAR slope parameters *and* the elements of the error covariance matrix[11]. Consequently, the bias we have discussed is potentially important for constructing accurate bootstrap confidence intervals for IRFs in practice. Indeed, in this section we show exactly how any bias in the bootstrap estimate of Σ creates a specific bias in the bootstrap IRFs and thus the bootstrap confidence intervals.

To illustrate this, consider a SVAR model explaining the behavior of a $K \times 1$ vector of variables, y_t. The IRFs are obtained from the moving average representation of the model:

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

where ε_t is a vector of *K* structural shocks and we make the standard assumption that $E(\varepsilon_t \varepsilon_t') = I_K$. This assumption provides a normalization as well as a set of identifying

[10] Note that if, for this VAR model, we had computed MLE rather than OLS estimates of Σ in both the initial and bootstrap stages, the approximate bias for the elements of the bootstrap estimate of Σ would have been -31%. See footnote 5.

[11] Other objects of frequent interest that are also nonlinear functions of VAR slope parameters and elements of the error covariance matrix are forecast error variance decompositions and measures of predictability. Thus, related bootstrap confidence or prediction intervals would also suffer from the bias we discuss here. See Inoue and Kilian (2002).

restrictions. The elements of the matrix polynomial $A(L)$ give the impulse response functions: $a_{ij,l}$ $(i, j = 1,...K;\ l = 1,...)$ indicates the response of variable *i* to a one unit (standard deviation) movement in the j^{th} structural shock *l* periods later. These IRFs are frequently the objects of interest in macroeconomic analysis. However, they cannot generally be estimated directly from time series data since the SVAR model (7) is not identified without further restrictions.

To obtain an estimate of the SVAR and thus the IRFs, we begin by specifying a finite-order reduced form VAR model which can always be estimated:

(8) $B(L)y_t = u_t$

where $B(L)$ is a matrix of polynomials of order *p* and $E(u_t u_t') = \Sigma$. In general, estimates of the parameters in $B(L)$ and Σ can be obtained by OLS. The reduced form moving average representation can be obtained by inverting (8):

(9) $y_t = B(L)^{-1} u_t = C(L) u_t$

Equating terms in (7) and (9) allows us to conclude the following:

(10) $u_t = A_0 \varepsilon_t$

(11) $A_l = C_l A_0 \quad l = 1,...$

It is clear from these relationships that knowledge of the K^2 elements of A_0 is sufficient to obtain the IRF.

From (10) we infer the key relationship between the covariance matrices of the structural and reduced form errors:

(12) $\Sigma = A_0 A_0'$

Due to the symmetry of Σ, this equation provides $\left(\frac{K(K+1)}{2}\right)$ restrictions. With $\left(\frac{K(K-1)}{2}\right)$ additional restrictions, A_0 can be identified and the impulse response functions computed. Note that (11) and (12) assure us that the estimated IRFs depend on the estimates of both $B(L)$ and Σ. I.e., $\hat{a}_{ij,l} = g(\hat{\beta}, \hat{\sigma})$ $(i, j = 1,...K;\ l = 1,...)$ where $\hat{\beta} = vec(\hat{B})$, a $K(Kp+1) \times 1$ vector, and $\hat{\sigma} = vech(\hat{\Sigma})$, a $\left(\frac{K(K+1)}{2}\right) \times 1$ vector. As a consequence, the properties of the IRFs depend on the properties of $\hat{\beta}$ and $\hat{\sigma}$. Similarly, the properties of the bootstrap IRFs depend in the same way on the properties of the bootstrap estimates of β and σ: $\tilde{a}_{ij,l} = g(\tilde{\beta}, \tilde{\sigma})$ $(i, j = 1,\ldots K;\ l = 1,\ldots)$.

We can see from this that there are five potential sources of bias for the bootstrapped IRFs and, thus, bootstrap confidence intervals for the original IRFs.

- Since the IFR is a nonlinear function of β and σ, the bootstrap IRF may be biased even if the original OLS and bootstrap estimates of β and σ are unbiased. Furthermore, this nonlinearity can potentially exacerbate or attenuate the consequences of biased estimates of β and σ.
- The original OLS estimate of β is biased. This is the problem reviewed by Kilian (1998). It is potentially especially perverse because the IRF bias is magnified for the bootstrap.
- The original OLS estimate of σ is biased. If this bias is large, it would have the same magnification effects. Kilian (1998, p. 219, fn 3) argues that any variance estimate bias due to coefficient bias is of second order [$O(T^{-2})$] and ignores it.
- The *bootstrap* estimate of β is biased for $\hat{\beta}$ beyond the original bias in the original OLS estimate of β. This is unlikely to be incrementally important.
- The *bootstrap* estimate of σ is biased for $\hat{\sigma}$ beyond any original bias in the original OLS estimates of β and σ. This is the problem we focus on in this paper. It is important because the bias comes from inappropriate degrees of freedom adjustment. It can be large and, furthermore, is easily corrected.

The obvious question is: How much difference does the appropriate bootstrap estimation of the error covariance matrix make for bootstrap estimates of the IRF? We can obtain an exact analytical answer to this question.

From equation (11) we infer that the bootstrap estimates of the IRF are given by

(13) $$\tilde{A}_l = \tilde{C}_l \tilde{A}_0 \quad l = 1,\ldots$$

where $\tilde{A}_l$, $\tilde{C}_l$, and $\tilde{A}_0$ are bootstrap estimates. If $\tilde{C}_l$ and $\tilde{A}_0$ are biased we expect $\tilde{A}_l$ to be affected. Note that $\tilde{C}_l$ depends only on the bootstrap estimates of B_l, $\tilde{B}_l$; equation (9). On the other hand, we see from equation (12) that $\tilde{A}_0$ depends only on the bootstrap estimate of Σ, $\tilde{\Sigma}$. We first derive the "bias"[12] for $\tilde{A}_0$ that arises from a biased estimate of Σ and then derive the resulting bias for the IRF.

Let $\hat{\Sigma} = \hat{A}_0 \hat{A}_0'$ be an unbiased estimate of Σ so that $\hat{A}_0$ is the corresponding "unbiased" estimate of A_0. Then $\tilde{\Sigma} = \tilde{A}_o \tilde{A}_o' = (1+b)\hat{\Sigma}$ where the scalar *b* reflects the proportional bias in $\tilde{\Sigma}$, a potentially biased bootstrap estimate. Thus,

(14) $$\tilde{A}_0 = (1+b)^{1/2} \hat{A}_0 = (1+a)\hat{A}_0$$

where *a* is the "bias" in $\tilde{A}_0$. Equating $(1+b)^{1/2}$ and $(1+a)$ in (14) allows us to deduce the following relationship between the bias of $\tilde{\Sigma}$, *b*, and the "bias" of $\tilde{A}_0$, *a*:

(15) $$a = (1+b)^{1/2} - 1$$

When $-1 < b \leq 0$, as it is in our case, we see than $a \geq b$ and thus the "bias" for $\tilde{A}_0$ is closer to zero than the bias for $\tilde{\Sigma}$.

Now, consider how this "bias" in the bootstrap estimate $\tilde{A}_0$ affects the bootstrap IRF given by equation (13). To isolate the effect of a bias in $\tilde{\Sigma}$, we assume that C_l is known (or at least $\tilde{C}_l$ is unbiased). Suppose $\hat{A}_0$ is an unbiased estimate of A_0. We can rewrite equation (14) as

(16) $$\hat{A}_0 = \left(\frac{1}{1+a}\right)\tilde{A}_0$$

where *a* is the scalar proportional bias for $\tilde{A}_0$. The "unbiased" IRF is then given by

(17) $$\hat{A}_l = C_l \hat{A}_0 = \left(\frac{1}{1+a}\right) C_l \tilde{A}_0, \ l = 1, \ldots$$

from which we can infer the proportional "bias" for the terms in the IRF

[12] We put the term "bias" in quotes here because $\tilde{A}_0$ and $\tilde{\Sigma}$ are related by a quadratic equation. Consequently, we cannot infer the true bias of $\tilde{A}_0$ from the bias of $\tilde{\Sigma}$. The bias we derive is therefore only approximate. We will use this notational device for the next few paragraphs.

(18) $$\frac{\tilde{A}_l - \hat{A}_l}{\hat{A}_l} = a,\ l = 1,\ldots.$$

Thus, the bootstrap IRF "bias" is constant and equal to the bias for $\tilde{A}_0$ for the entire IRF horizon. So, for example, if we have a SVAR model with *K=2*, *p=8*, *T=100* and a constant term, the "bias" for $\tilde{\Sigma}_{b(1)}$ is -17% and the "bias" for the IRF is -9%.

IV. An Example

The bias discussed here is pervasive in the empirical SVAR literature. To illustrate the effect of this bias in practice, we have chosen to replicate the biased results obtained in a single influential paper by Chistiano, Eichenbaum, and Evans, CEE, (1999). We then compute the corresponding bias-corrected IRF and associated bootstrap confidence intervals to draw our comparison.

In their paper, CEE examine the effects of monetary policy shocks on several economic variables of interest using models imposing a recursive structure to identify the relevant shocks. Their first benchmark model includes a constant term and four lags (*p=4*) of seven variables (*K=7*) with the federal funds rate as the chosen monetary policy instrument. They estimate their models using quarterly data over the period 1965:3-1995:2. Given the loss of observations due to the four lags in the VAR, *T*=116 in our notation. We replicate their results by estimating their model over the same sample period.[13] For illustrative purposes, we report only the IRF indicating the effects of a negative monetary policy shock on output. While this is an IRF of particular interest, the same bias will be present in all the other 48 IRFs as well. As seen in Figure 1 here and Figure 2 of CEE (1999, p. 86), given a positive federal funds rate shock, "after a delay of 2 quarters, there is a sustained decline in real GDP " (p. 87). We note that CEE use MLE to estimate the VAR error covariance estimate so the estimated IRF will be biased. Furthermore, we see that the bootstrap confidence intervals reflect considerable asymmetry, which we shall see momentarily, is partially due to bias in the confidence intervals arising from biased bootstrap IRF estimates.

[13] Indeed, we have estimated the CEE model using their data and their RATS program which Larry Christiano has generously made available on his website.

To illustrate the effect of bias due to MLE and the further bias due to the CEE bootstrap IRFs, we estimate the CEE model once again but this time including the appropriate degrees of freedom corrections. These results for the first-stage IRF and the bootstrap confidence intervals are also reported in Figure 1. The first thing we notice is that the fundamental conclusion regarding the IRF is unchanged: a contractionary federal funds rate shock will, after a lag, have a sustained negative effect on real GDP.[14] We also notice that, due to the degrees of freedom correction in the original error covariance matrix estimate, the bias-corrected IRF lies entirely below the CEE IRF.

In addition, we see that the confidence intervals also shift significantly when we correct the bias in the bootstrap estimates of the error covariance matrix. We note three consequences. First, we see that for much of the time horizon, the bias-corrected IRF actually lies below the biased 95% confidence intervals. Second, we see that correcting for our bias has greatly reduced the asymmetry in the confidence intervals. Third, we notice that between 2 and 11 quarters, the upper 95% confidence bounds are farther away from zero after bias correction. This strengthens the conclusion that a contractionary monetary policy has a significant negative effect on output over that horizon.

Since part of the distortion in the CEE results is a consequence of their choice to use MLE estimates of the error covariance matrix, we also illustrate how much distortion remains when we use OLS estimates. The results are reported in Figure 2. In the typical approach which incorporates the natural OLS degrees of freedom correction, the IRF is already bias-corrected so we only have a single IRF estimate. However, the typical procedure does result in biased bootstrap confidence intervals. As in Figure 1, we again see that the typical biased procedure results in quite asymmetric confidence intervals which are, in part, a consequence of the bias; the bias-corrected confidence intervals exhibit much less asymmetry. Also, as noted in the discussion of Figure 1, over an intermediate horizon, the upper bound of the bias-corrected confidence intervals lie below their biased counterparts giving us greater confidence in our conclusion that a monetary contraction has a significant negative effect on output.

These examples illustrate that correcting for bias in both the original IRF and especially

[14] Indeed, because the bias we are reporting is *proportional* to the bootstrap IRFs, we will always draw the same conclusion when out interes is in whether or not the IRF is significantly different from zero.

in the bootstrap confidence intervals can remove distortions that change the quantitative (if not qualitative) conclusions when SVAR models are used.

V. Conclusion

This paper has discussed a commonly occurring bias in bootstrap estimates of confidence intervals for IRFs in SVARs. The source of that bias is the systematic bias in the traditional bootstrap estimate of the VAR covariance matrix. Since the bootstrap IRFs depend on these biased estimates, they are systematically biased as well. Consequently the implied bootstrap IRF percentile confidence intervals inherit the same bias. This bias in potentially large but, fortunately, is easily corrected by accounting for the fact that the natural bootstrap estimate of the VAR covariance matrix must include an additional degrees of freedom adjustment.

References

Berkowitz, J., Kilian, L., 2000. "Recent Developments in Bootstrapping Time Series." *Econometric Reviews* 19, 1-48.

Blanchard, O., Quah, D., 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* 79, 655-673.

Christiano, L. J., Eichenbaum, M., Evans, C. L., 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" In: Taylor, J.B., Woodford, M., (Eds.), *Handbook of Macroeconomics*. Elsevier, Amsterdam, 65-148.

Christiano, L. J., Eichebaum, M., Vigfusson, R., 2006. "Assessing Structural VARs." In: Acemoglu, D., Rogoff, K., Woodford, M., (Eds.), *NBER Macroeconomics Annual*. MIT Press, Cambridge, Massachusetts, 1-72.

Davidson, R., MacKinnon, J. G., 1993. *Estimation and Inference in Econometrics*. Oxford University Press, New York.

Freedman, D. A., Peters, S. C., 1984. "Bootstrapping a Regression Equation: Some Empirical Results." *Journal of the American Statistical Association* 79, 97-206.

Inoue, A., Kilian, L., 2002. "Bootstrapping Smooth Functions of Slope Parameters and Innovation Variances in VAR(∞) Models." *International Economic Review* 43, 309-331.

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

Lütkepohl, H., 1990. "Asymptotic Distribution of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models." ." *Review of Economics and Statistics* 72, 116-125.

Mittnik, S., Zadrosny, P. A., 1993. "Asymptotic Distribution of Impulse Responses, Step Responses, and Variance Decompositions of Estimated Linear Dynamic Models." *Econometrica* 61, 857-870.

Peters, S. C., Freedman, D. A., 1984. "Some Notes on the Bootstrap in Regression Problems." *Journal of Business and Economics Statistics* 2, 406-409.

Runkle, D. E., 1987. "Vector Autoregression and Reality." *Journal of Business and Economics Statistics* 5, 437-442.

Sims, C. A., Zha, T., 1999. "Error Bands for Impulse Responses." *Econometrica* 67, 1113-1155.

Stine, R. A., 1987. "Estimating Properties of Autoregressive Forecasts." *Journal of the American Statistical Association* 82, 1072-1078.

Table 1: Bias of error variance estimate in standard univariate linear regression model with $R = 9$; number of Monte Carlo replications = 1,000, number of bootstrap draws = 1,000. True value of variance =0.81.

Variance estimate	Sample Size	Theoretical bias (%)	Mean estimate	Bias (%)
$\hat{\sigma}^2$	30	0	0.8207	1.33%
$\tilde{\sigma}^2_{b(1)}$	30	-30.0%	0.5745	-29.07%
$\tilde{\sigma}^2_{b(2)}$	30	0	0.8208	1.33%
$\hat{\sigma}^2$	50	0	0.8196	1.19%
$\tilde{\sigma}^2_{b(1)}$	50	-18.0 %	0.6720	-17.03%
$\tilde{\sigma}^2_{b(2)}$	50	0	0.8195	1.18%
$\hat{\sigma}^2$	100	0	0.8096	-0.05%
$\tilde{\sigma}^2_{b(1)}$	100	-9.0%	0.7686	-9.0%
$\tilde{\sigma}^2_{b(2)}$	100	0	0.8097	-0.04%

Table 2: Bias of error variance estimate in an AR(8) model with a constant term ($R = 9$); number of Monte Carlo replications = 1,000, number of bootstrap draws = 1,000. True value of variance =0.81.

Variance estimate	Sample Size	"Theoretical" bias (%)[15]	Mean estimate	Bias (%)
$\hat{\sigma}^2$	30	0	0.8323	2.753%
$\tilde{\sigma}^2_{b(1)}$	30	-30.0%	0.5989	-26.06%
$\tilde{\sigma}^2_{b(2)}$	30	0	0.8556	5.63%
$\hat{\sigma}^2$	50	0	0.8316	2.67%
$\tilde{\sigma}^2_{b(1)}$	50	-18.0 %	0.6891	-14.93%
$\tilde{\sigma}^2_{b(2)}$	50	0	0.8404	3.75%
$\hat{\sigma}^2$	100	0	0.8145	0.55%
$\tilde{\sigma}^2_{b(1)}$	100	-9.0%	0.7431	-8.26%
$\tilde{\sigma}^2_{b(2)}$	100	0	0.8166	0.81%

[15] This is the theoretical bias for the corresponding ($R = 9$) standard regression model.

Figure 1: Impulse response functions showing the effect of a contractionary monetary policy on real GDP with 95% confidence intervals. The solid line gives the original MLE IRF and the long-dashed bold line gives the bias-corrected OLS IRF; CEE use MLE. The dotted lines give the MLE bootstrap 95% confidence intervals and the dashed lines give the bias-corrected 95% confidence intervals.

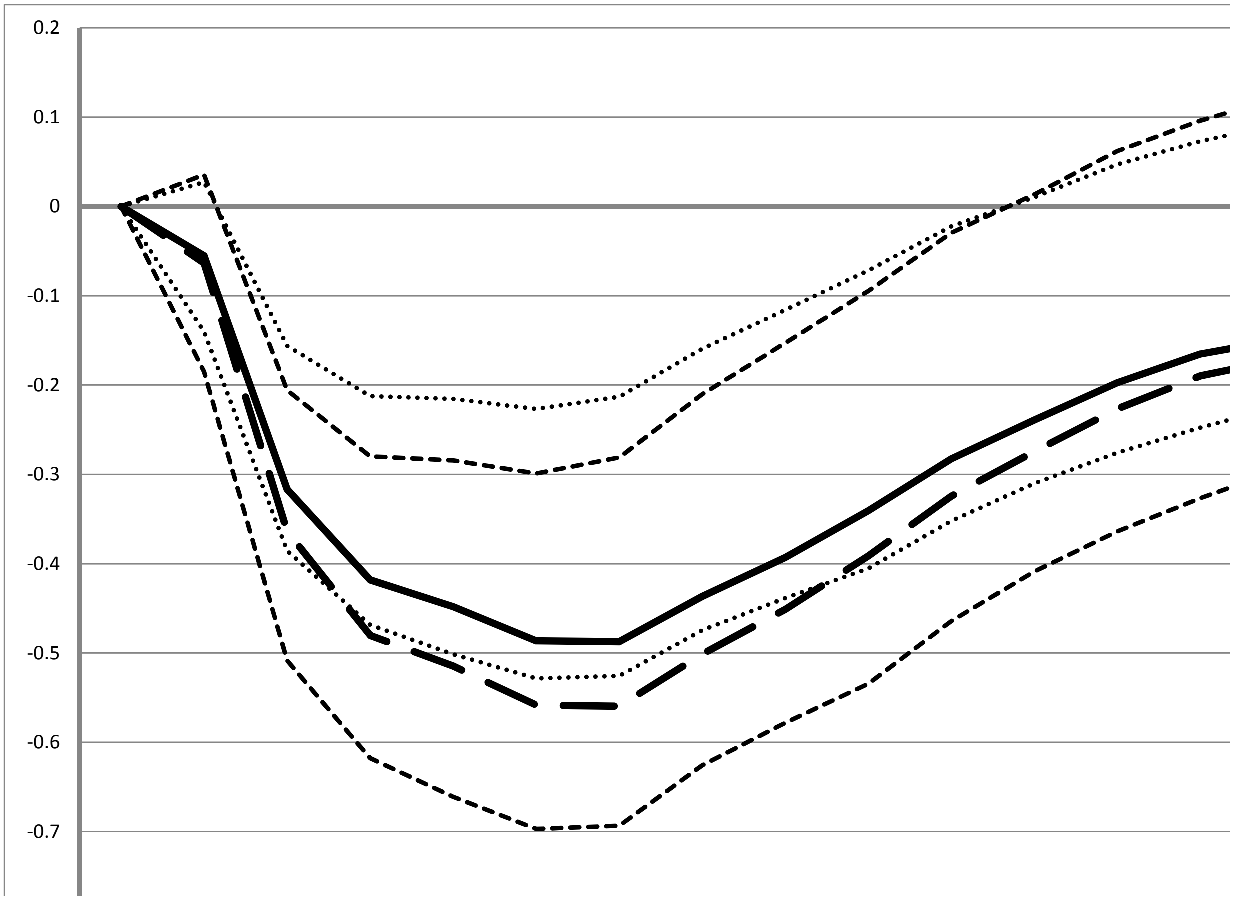

Figure 2: Impulse response function showing the effect of a contractionary monetary policy on real GDP with 95% confidence intervals. The solid line gives the original OLS IRF. The dotted lines give the typical but biased bootstrap 95% confidence intervals and the dashed lines give the bias corrected 95% confidence intervals.