Econ 413R: Computational Economics

Spring Term 2013

Bootstrapping

1 Introduction to Bootstrapping

When conducting statistical inference we rarely know the exact finite sample distributions of interest so we use approximate distributions and hope they are helpful. We often use asymptotic theory (e.g., a Central Limit Theorem) to obtain approximations. An alternative to asymptotic theory is the use of so-called bootstrap methods. Rather than relying on a known asymptotic distribution, the bootstrap relies on the known empirical distribution as an approximation. Bootstrap methods are generally rather straightforward and often provide better approximations than asymptotic methods. With the increasing power of computers, bootstrap methods have become increasingly popular.

The bootstrap approach began with a seminal paper by Bradley Efron (Efron, 1979). An enormous literature has followed that original contribution. A good introductory reference is Efron and Tibshirani (1993)[1]

[1]In addition I have relied on Killian (2008).

1.1 A General Statistical Problem

Consider a random sample, $\mathbf{x} = \{x_1, x_2, \ldots, x_t\}$, from an unknown distribution, $F : F \to \mathbf{x} = \{x_1, x_2, \ldots, x_t\}$. We want to estimate a parameter of interest, $\theta = t(F), [\text{e.g.,} \mu = \int_{-\infty}^{\infty}(xdF(x)]$ on the basis of x: $\theta = s(\mathbf{x})$ e.g.,$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i]$. We would like to know the properties of the random variable $\hat{\theta}$. For example, what is the variance of $\hat{\theta}$?

Ideally, we would like to draw a number of additional random samples from F, compute $\hat{\theta}$ in each case and use these estimates to investigate its distribution. Unfortunately, it is generally impossible to do so.

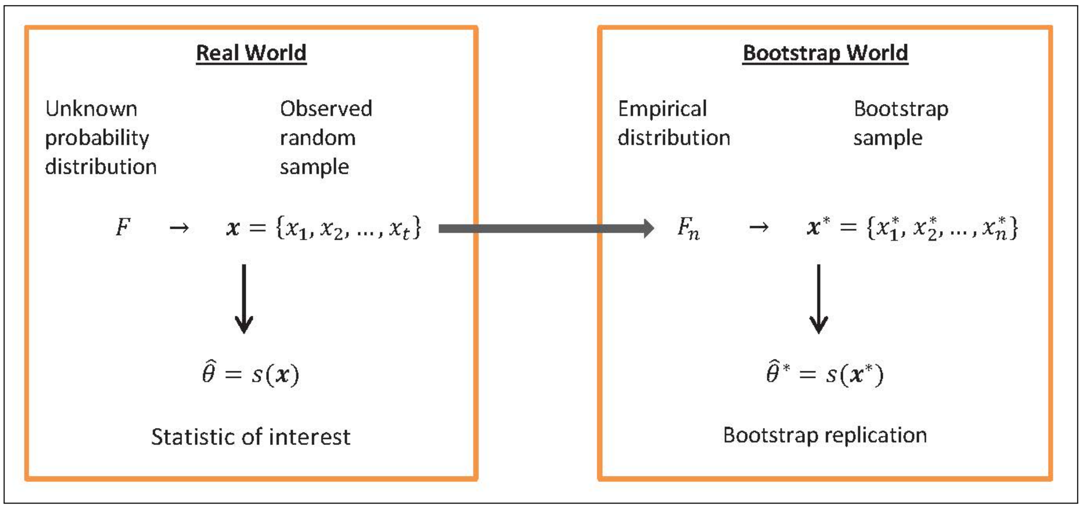
1.2 The Bootstrap Analogy

The bootstrap is based on an analogy in which the observed sample data take on the role of the population data. See figure 1 (adapted from Efron and Tibshirani) in which the random sample given above comes from the "Real World" but we do our analysis in the "Bootstrap World."

From the original sample $\mathbf{x} = \{x_1, x_2, \ldots, x_t\}$ we construct a *known* empirical distribution, F_n, and draw a random sample (with replacement): $\mathbf{x}^* = x_1^*, x_2^*, \ldots, x_n^* : F_n \to \mathbf{x}^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$.

This is a **bootstrap sample**. Using this sample we compute $\hat{\theta}^* = s(\mathbf{x}^*)$, a **bootstrap replication**. Note that since the empirical distribution is known we can draw as many bootstrap samples as we desire so we obtain B bootstrap samples: $\mathbf{x}^{*b}, b = 1, 2, \ldots, B$. From these we compute B bootstrap replications: $\hat{\theta}b^* = s(\mathbf{x}^{*b}), b = 1, 2, \ldots, B$. We use the distribution of $\hat{\theta}^*$ to draw inferences about the distribution of $\hat{\theta}$.

Figure 1: This figure illustrates the bootstrap analogy between the "Real World" and the "Bootstrap World." In the Real World, we have a single random sample from an *unknown* population distribution from which we want to calculate a statistic of interest. In the Bootstrap World, we have a *known* empirical distribution from which we obtain as many bootstrap samples and bootstrap replications as desired.



As we can see, the critical step in this process is obtaining the empirical distribution, F_n, from the observed sample, as an estimate of the unknown population distribution, F. Everything else proceeds by analogy: F_n yields $\mathbf{x}^*$ by random sampling just as F yields $\mathbf{x}$ by random sampling, $\hat{\theta}^*$ is obtained from $\mathbf{x}^*$ in exactly the same way as $\hat{\theta}$ is obtained from $\mathbf{x}$.

An obvious way to construct the empirical distribution from x is to assign probability $1/n$ to each of the observed values of $x_i, i = 1, 2, \ldots, n$. Thus, $F_n(x) = \#(x_i \leq x)/n$, a step function. It can be shown that F_n is a consistent estimate of F; i.e., as $n \to \infty$, $F_n(x)$ gets arbitrarily close to $F(x)$.

2 An Illustrative Example

Suppose we have a random sample of size n drawn from an unknown distribution, F. $F \to x = \{x_1, x_2, \ldots, x_t\}$ and we are interested in properties, say the variance, of the sample mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Unfortunately, we cannot obtain additional samples from F so how can we estimate the variance of this statistic?

2.1 Traditional Approach

The traditional approach is to determine the variance of the sample mean from $VAR(\bar{x}) = \frac{1}{n^2}\sum_{i=1}^{n} VAR(x_i) = \frac{\sigma^2}{n}$. But, since we do not know the value of $\sigma^2 = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n}$, we replace it with the estimator $\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{n-1}$ resulting in $VAR(\bar{x}) = \frac{\sigma^2}{n}$. I am using uppercase *VAR* to denote the unknown true population variance and lowercase *var* to denote estimators of *VAR*. The accuracy of $var(\bar{x})$ as an estimate of $VAR(\bar{x})$ depends on how accurate $\hat{\sigma}^2$ is as an estimate of σ^2. In many cases, $\hat{\sigma}^2$ may be unreliable.

2.2 Bootstrap Approach

Though we cannot resample from the true population distribution, the bootstrap approach takes advantage of the fact that we can resample from the known empirical distribution constructed from the original sample and described above: $F_n(x) = \sharp(x_i \leq x)/n$. The bootstrap estimate of the variance of the sample mean is obtained by the following **simulation algorithm**:

(1) Using a random number generator, create B bootstrap samples by drawing samples of size n with replacement from the original sample:

$x^{*b} = \{x_1^{*b}, x_2^{*b}, \ldots, x_n^{*b}\}, b = 1, 2, \ldots, B$

(2) For each bootstrap sample, compute bootstrap replications for $\bar{x}^{*b}$:

$\bar{x}^{*b}\ \frac{1}{n}\sum_{i=1}^{n} x_i^{*b}, b = 1, 2, \ldots, B$

(3) From the B bootstrap replications in (2), calculate the simulated bootstrap variance estimate:

$var_B^*(\bar{x}) = \frac{1}{B-1}\sum_{b=1}^{B}\{\bar{x}^{*b} - [\sum_{b=1}^{B}(\bar{x}^{*b}/B]\}^2$

Note the following:

i. As B approaches ∞, $VAR_B^*(\bar{x})$ approaches $VAR^*(\bar{x})$. $B = 200$ is enough to get a good estimate of the mean or variance. For bootstrap confidence intervals, at least 1000 replications should be used.

ii. A 95% confidence interval for $\hat{\theta}$ can be obtained by taking the 2.5 and 97.5 percentiles of the distribution of the bootstrap replications $\hat{\theta}_b^*, b = 1, 2, \ldots, B$. E.g., compute 1000 bootstrap replications of $\bar{x}^{*b}$ and, ordering them from smallest to largest, take the 25th and 975th elements as confidence interval endpoints.

iii. Bootstrap methods are generally quite useful but that ultimately depends on the implicit assumption that the resampling properties of $(\hat{\theta}^* - \hat{\theta})$ are similar to the sampling properties of $(\hat{\theta} - \theta)$. Note that the population parameter θ is unknown in the Real World but the pseudo-population parameter $\hat{\theta}$ is known in the Bootstrap World.

3 Bootstrap Resampling in Linear Regression Models with Nonstochastic Regressors

Consider a univariate regression model represented by:

$$y = X\beta + u \tag{3.1}$$

where y is a $T \times 1$ vector of observations on a dependent variable, X is a $T \times R$ matrix of observations on R nonstochastic regressors including a constant (vector of ones), β is an $R \times 1$vector of regression coefficients, and u is a $T \times 1$ 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 $\hat{u} = y - X\hat{\beta}$. We can obtain bootstrap estimates of the regression coefficients as follows:[2]

i. For replication b, draw a random sample of size T from the OLS residuals, $\hat{u}$, to obtain the $T \times 1$ vector u_b^* which we add to $X\hat{\beta}$ to get y_b^*:

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

   (I have changed notation slightly.)

ii. Obtain bootstrap estimates as

   $\hat{\beta}_b^* = (X'X)^{-1}X'y_b^* = \hat{\beta} + (X'X)^{-1}X'u_b^*$

   Note that we don't need to explicitly compute y_b^* if we are only interested in $\hat{\beta}_b^*$.

[2]We will look at obtaining bootstrap estimates of the error variance below since there is an additional consideration we must take into account.

iii. Repeat steps (i) and (ii) for $b = 1, 2, \ldots, B$ and use the resulting bootstrap estimates $\beta_1^*, \beta_2^*, \ldots, \beta_B^*$ to construct the empirical distribution of the estimator $\hat{\beta}^*$.

As we will see below, more care must be taken when we have lagged endogenous variables on the right-hand-side of a regression equation in order to preserve the dependence that exists in the data. There is much more that could be said about obtaining bootstrap samples in other cases but we will spend the rest of the notes looking only at the case of bootstrapping in structural vector autoregression (SVAR) models.

4 Bootstrapping in SVAR Models

In a recent paper, Phillips and Spencer (2011) examined the widespread use of bootstrapping to obtain confidence intervals (CIs) for impulse response functions (IRFs) in SVAR models. They demonstrated that the methods commonly used in the literature result in systematic bias in bootstrapped confidence intervals. Furthermore, they showed how to reduce the bias through a straightforward scale adjustment. The problem arises because of a bias in the usual bootstrap estimate of the VAR error covariance matrix which is used in computing the bootstrap IRF CIs. We will follow the paper closely in looking at the source of this bias which will suggest the appropriate correction.[3]

[3] Much of what follows is taken directly from Phillips and Spencer (2011)

4.1 A Source of Bias

4.1.1 Standard Regression Models

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 (4.1)

$$y = X\beta + u \tag{4.1}$$

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 $\hat{\sigma}^2$ 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., $E\{\frac{u'u}{T}\} = \sigma^2$. On the other hand, $E\{\frac{\hat{u}'\hat{u}}{T}\} = \frac{T-R}{T}E\{\frac{u'u}{T}\}$, which reflects that, on average, the squared residuals are $\frac{T-R}{T}$ times as large as the squared errors[4]. Thus, to obtain an unbiased estimate, we must rescale each residual by $\sqrt{\frac{T-R}{T}}$ and then compute the average squared *rescaled* residual giving the usual unbiased estimate for σ^2, $\hat{\sigma}^2 = \frac{1}{T-R}\hat{u}'\hat{u}$.

Now consider obtaining a bootstrap variance estimate in this simple case. The bootstrap methodology relies on an analogy between the unknown population probability distribution of the "real world" and the known empirical

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

distribution in the "bootstrap world."[5] The bootstrap analyst hopes to learn about the population distribution of $\hat{\sigma}^2 - \sigma$ by examining the distribution of $\tilde{\sigma}^2 - \hat{\sigma}$ where $\hat{\sigma}^2$ is the "pseudo-population" variance of the empirical distribution (not a random variable in the bootstrap world) and $\tilde{\sigma}^2$ is a candidate bootstrap variance estimate (which, of course, is a random variable in the bootstrap world).[6] Typically, this is done by drawing many samples from the pseudo-population given by the original sample. Because we can resample as many times as we want, we can estimate the mean of $\tilde{\sigma}^2 - \hat{\sigma}$ and, thus, the bias of $\tilde{\sigma}^2$ using Monte Carlo experiments.

Pursuing this bootstrap analogy and recalling the insight discussed above, we might expect an analogous degrees of freedom adjustment to be helpful for bootstrap variance estimates. This has been confirmed for the simple regression model by Freedman and Peters (1984, p. 99) and Peters and Freedman (1984, p. 408).

Suppose we obtain *bootstrap* estimates of the error variance as follows. For bootstrap replications $b = 1, \ldots, B$, generate

$$y_b^* = X\hat{\beta} + u_b^* \tag{4.2}$$

where the elements of u_b^* are drawn with replacement from the OLS residuals, $\hat{u}$. Then, apply OLS to equation (4.2) to get bootstrap estimates of $\hat{\beta}$ (not β), which we denote $\tilde{\beta}_b$ and bootstrap residuals, $\tilde{u}_b$. In the bootstrap, the variance estimate, $\tilde{\sigma}_b^2$, is an estimate of $\hat{\sigma}^2$ (not σ^2), the "population" error variance in the pseudo-population given by the error variance of the original

[5] See (Efron and Tibshirani, 1993, especially Chapter 8), for discussion of this analogy.
[6] So, in the bootstrap world, $\tilde{\sigma}^2$ is an estimate of $\hat{\sigma}^2$.

OLS residuals, $\hat{u}$. The usual bootstrap variance estimate is given by $\tilde{\sigma}^2_{b,1} = \frac{1}{T-R}\tilde{u}_b'\tilde{u}_b$.

In this case, we can get some analytical insight for the properties of $\tilde{\sigma}^2_{b,1}$ by conditioning on the unknown population distribution. Proceeding as above, we note that though $E\{\frac{\hat{u}'\hat{u}}{T-R}\} = \sigma^2$, $E\{\frac{\tilde{u}_b'\tilde{u}_b}{T-R}\} = \frac{T-R}{T}E\{\frac{u_b^{*\prime}u_b^*}{T-R}\} = \frac{T-R}{T}\sigma^2$ since the elements of u_b^* are drawn randomly from $\hat{u}$. This reflects that, on average, the squared bootstrap residuals are $\frac{T-R}{T}$ times as large as the squared OLS residuals which are the pseudo-population errors. Consequently, we suggest that a better bootstrap estimate might be given by $\tilde{\sigma}^2_{b,2} = \frac{1}{(T-R)^2}\tilde{u}_b'\tilde{u}_b = \frac{1}{T-R}\tilde{\sigma}^2_{b,1}$. This is the same rescaling suggested by Freedman and Peters (1984) and Peters and Freedman (1984).

If this analogy holds exactly, we would expect the size of the (proportional) bias for the natural estimator to be $-R/T$[7]. While this vanishes asymptotically, it can be important in small samples when R is large relative to T. To illustrate, we conduct a Monte Carlo experiment in which we simulate obtaining bootstrap estimates of the error variance in a univariate regression model like (4.1). We estimate models with nine regressors including a constant term, $R = 9$, for three sample sizes: $T = 30, 50, 100$[8]. Consequently, the expected bias for $\tilde{\sigma}^2_{b,1}$ is -30%, -18% and -9% respectively.

[7]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., $E\{\breve{\sigma}^2\} = \frac{T-R}{T}\sigma^2$. Thus, the proportional bias is $-R/T$. Now, when we bootstrap and obtain the MLE of $\breve{\sigma}^2_b$, σ^2, $\tilde{\breve{\sigma}}^2_b = \frac{1}{T}\tilde{u}'\tilde{u}$, the bias is magnified since we have a biased estimate of a biased estimate. $\tilde{\breve{\sigma}}^2_b = \left(\frac{T-R}{T}\right)^2\tilde{\sigma}^2_{b,2}$, so $E\{\tilde{\breve{\sigma}}^2_b\} = \left(\frac{T-R}{T}\right)^2\sigma^2$ and the expected proportional bias is $\left(\frac{T-R}{T}\right)^2 - 1 = \frac{R^2-2TR}{T^2}$ which is negative and larger (in absolute value) than $-R/T$.

[8]The values of the regressors are 1.1, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3 with the first element being the constant term; $\sigma^2 = 0.81$.

Table 1: Bootstrap error variance estimates in standard univariate linear regression model with $R = 9$; number of Monte Carlo trials = 1000, number of bootstrap draws = 200. True value of variance = 0.81.

Estimator	**Sample Size**	**Mean Estimate**	**Proportional Difference** $(-R/T)$	**Estimated Bias**[a]
$\hat{\sigma}^2$	30	0.8051		
$\tilde{\sigma}^2_{b,1}$	30	0.5633	-30.0%	-30.03%
$\tilde{\sigma}^2_{b,2}$	30	0.8047	0	-0.05%
$\hat{\sigma}^2$	50	0.8053		
$\tilde{\sigma}^2_{b,1}$	50	0.6603	-18.0%	-18.01%
$\tilde{\sigma}^2_{b,2}$	50	0.8053	0	0.00%
$\hat{\sigma}^2$	100	0.8054		
$\tilde{\sigma}^2_{b,1}$	100	0.7327	-9.0%	-9.03%
$\tilde{\sigma}^2_{b,2}$	100	0.8052	0	-0.02%

[a] The estimated bootstrap bias is the average difference between the relevant bootstrap error variance estimate and $\hat{\sigma}^2$, the pseudo-population variance across the 1000 Monte Carlo trials.

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 (4.1) by OLS, and compute the usual population-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 $\tilde{\sigma}^2_{b,2}$,[9] we take each of the 1000 Monte Carlo samples and obtain 200 bootstrap estimates in each case. The average values are reported in Table 1 for our three sample sizes. We call this the estimated bootstrap bias.

The results in Table 1 confirm our expectation very nicely. The "natural" bootstrap estimator, $\tilde{\sigma}^2_{b,1}$, has bias approximately equal to $-R/T$ while $\tilde{\sigma}^2_{b,2}$

[9] Note that Table 1 reports the bias relative to, $\hat{\sigma}^2$, the pseudo-population variance.

is approximately unbiased.

This bias in the standard bootstrap "error" variance carries over exactly to the case of a multivariate seemingly unrelated regression model with non-stochastic regressors. To confirm the theory, we have conducted simple Monte Carlo experiments similar to those undertaken for the univariate regression model discussed above. To save space, we do not report the results here but simply indicate that the conclusions are the same.[10]

4.1.2 Autoregressive Models

Consider a univariate AR(p) with a constant term, ν, so that $R = p + 1$:

$$y_t = \nu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + u_t; t = 1 - p, \ldots, 0, 1, \ldots, T \tag{4.3}$$

where u_t is white noise with variance σ^2 and T is the number of usable observations. Because the regressors are stochastic, the finite sample theory of the previous section does not apply. However, following Stine (1987, p.1074) and Berkowitz and Kilian (2000, p. 5), we might speculate (correctly) that similar bias problems exist for bootstrap estimators of the error variance in this case.

Since analytical results are not available, we examine the finite-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 (4.3) in which $p = 8$ so $R = 9$.[11] For each of three sample sizes, $T = 30, 50, 100$, we

[10]The results are reported in an unpublished appendix available on request.

[11]The model coefficients are 0.008, 0.25, 0.11, -0.03, -0.004, -0.12, 0.03, -0.02, -0.08 with the first element being the constant term; $\sigma^2 = 0.81$.

Table 2: Bootstrap error variance estimates in an AR(8) model with a constant term ($R = 9$); number of Monte Carlo trials = 1000, number of bootstrap draws = 200. True value of variance =0.81.

Estimator	**Sample Size**	**Mean Estimate**	**Proportional Difference** ($-R/T$)	**Estimated Bias**[a]
$\hat{\sigma}^2$	30	0.8476		
$\tilde{\sigma}^2_{b,1}$	30	0.6178	-30.0%	-27.11%
$\tilde{\sigma}^2_{b,2}$	30	0.8826	0	4.13%
$\hat{\sigma}^2$	50	0.8129		
$\tilde{\sigma}^2_{b,1}$	50	0.6766	-18.0%	-16.77%
$\tilde{\sigma}^2_{b,2}$	50	0.8252	0	1.51%
$\hat{\sigma}^2$	100	0.8120		
$\tilde{\sigma}^2_{b,1}$	100	0.7418	-9.0%	-8.65%
$\tilde{\sigma}^2_{b,2}$	100	0.8152	0	0.39%

[a]The estimated bootstrap bias is the average difference between the relevant bootstrap error variance estimate and $\hat{\sigma}^2$, the pseudo-population variance across the 1000 Monte Carlo trials.

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 (4.3) 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 (relative to $\hat{\sigma}^2$) of the two bootstrap error variance estimates, $\tilde{\sigma}^2_{b,1}$ and $\tilde{\sigma}^2_{b,2}$, we obtain 200 bootstrap estimates for each of the Monte Carlo samples[12]. The average values are reported in Table 2 for each of our three sample sizes.

The results are quite informative. The heuristically expected 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

[12]For each bootstrap iteration, we obtain the initial p observations $\{y_{-p-1}, \ldots, y_0\}$ by drawing (with replacement) from the original generated sample $\{y_t\}_{-p+1}^{T}$.

error variance given by $\tilde{\sigma}^2_{b,1}$is biased and thus likely to result in significant distortion when the number of slope coefficients is large relative to the sample size.

Proceeding by analogy, we expect these bias results to carry over to the case of a (non-structural) 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 $\hat{\Sigma} = \frac{1}{T-R}\hat{U}'\hat{U}$ where $\hat{U}$is the $T \times K$ matrix of OLS residuals and $R = Kp+1$. The "natural" but perhaps biased bootstrap estimator of $\hat{\Sigma}$ is $\tilde{\Sigma}_{b.1} = \frac{1}{T-R}\tilde{U}_b'\tilde{U}_b$ where $\tilde{U}_b$ is the matrix of bootstrap residuals from the b^{th} bootstrap iteration. The degrees of freedom adjusted (DFA) bootstrap estimator of Σ is $\tilde{\Sigma}_{b.2} = \frac{T}{(T-R)^2}\tilde{U}_b'\tilde{U}_b$.

We have investigated the bootstrap error variance bias for a two-equation VAR(8) model with a constant term using Monte Carlo methods similar to those described above and find the bias to be quite close to the bias expected from the above heuristic analysis above. To conserve space, we do not report the results here since they are quite similar to those reported for the AR(8) model above.[13] In particular, the bias for $\tilde{\Sigma}_{b.1}$ is approximately $-\frac{Kp+1}{T}$ where K is the number of equations (variables) in the VAR(p). For a two-equation VAR(8) model, this implies an approximate bias of -17% for each element of Σ when $T = 100$.[14]

[13]The results are reported in an unpublished appendix available on request.

[14]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 magnified to -31%. See footnote 11.

4.2 Bootstrapping IRFs for SVARs

The downward bias of the standard bootstrap estimator of the VAR error covariance matrix is of particular concern when we are interested in drawing inferences about IRFs from a SVAR model since the IRFs are nonlinear functions of both VAR slope parameters and the elements of the error covariance matrix.[15]. In this section we show how bias in the bootstrap estimate of the VAR error covariance matrix affects bootstrap IRFs and, thus, bootstrap CIs.

To estimate a SVAR model of the behavior of a $K \times 1$ vector of variables, $\{y_t\}$, and the corresponding IRFs, we begin by specifying a finite-order *reduced form* VAR model which can always be estimated:

$$B(L)y_t = u_t \tag{4.4}$$

where $B(L)$ is a matrix of p-order polynomials in the lag operator, L, u_t is a vector of K reduced form errors, and $E\{U_t U_t'\} = \Sigma$. In general, consistent OLS estimates of $B(L)$ and Σ can be obtained: $\hat{B}(L)$ and $\hat{\Sigma} = \frac{1}{T-R}\hat{U}'\hat{U}$, where $\hat{U}$, is the $T \times K$ matrix of OLS residuals, and $R = Kp + 1$ since we assume a constant term.

The reduced form moving average representation is obtained by inverting (4.4):

$$y_t = B(L)^{-1}u_t = C(L)u_t \tag{4.5}$$

[15]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).

where $C_0 = I$, the identity matrix. The shocks in this representation are one-step-ahead forecast errors and do not, in general, correspond to structural economic shocks for which we want to obtain impulse response functions.

We assume that there exists a structural moving average representation of the model from which the IRFs can be obtained:

$$y_t = A(L)\varepsilon_t \tag{4.6}$$

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 restrictions. The elements of the matrix polynomial $A(L)$ give the impulse response functions: $a_{i,j,\ell}$ $i, j = 1, \ldots, K, \ell = 0, 1, \ldots$ indicates the response of variable i in ℓ periods to a one unit (standard deviation) movement in the j^{th} structural shock today.

Though the IRFs are frequently the objects of interest in macroeconomic analysis, they cannot generally be estimated directly from time series data since the SVAR model (4.6) is not identified without further restrictions. However, equating terms in (4.4) and (4.6) allows us to conclude the following:

$$u_t = A_0\varepsilon_t \tag{4.7}$$

$$A_\ell = C_\ell A_0, \ell = 1, 2, \ldots \tag{4.8}$$

Thus, it is clear that knowledge of the K^2 elements of A_0 is sufficient to obtain the IRF.

From (4.7) and the assumption that $E\{\varepsilon_t\varepsilon_t'\} = I_K$, we infer the key relationship between the covariance matrices of the structural and reduced form errors:

$$\Sigma = A_0 A_0' \tag{4.9}$$

Symmetry of Σ provides $\frac{K(K+1)}{2}$ restrictions on A_0. With $\frac{K(K-1)}{2}$ additional restrictions, A_0 can be identified and IRFs computed. Equations (4.8) and (4.9) assure us that the estimated IRFs depend on estimates of $B(L)$ and Σ. I.e., $\hat{a}_{i,j,\ell} = g(\hat{\beta}, \hat{\sigma}), i, j = 1, \ldots, K, \ell = 1, 2, \ldots$ where $\hat{\beta} = vec(\hat{B})$, a $K(Kp + 1) \times 1$vector, and $\hat{\sigma} = vech(\hat{\Sigma})$, a $\frac{K(K+1)}{2} \times 1$ vector and the form of the nonlinear function g depends on the identification strategy. Consequently, 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 $\tilde{\beta}$ and $\tilde{\sigma}$: $\tilde{a}_{i,j,\ell} = g(\tilde{\beta}, \tilde{\sigma}), i, j = 1, \ldots, K, \ell = 1, 2, \ldots$.

We can see from this that there are several potential sources of bias for the bootstrapped IRFs and, thus, bootstrap confidence intervals for the original IRFs. The source we focus on here arises when the *bootstrap* estimate of $sigma$, $\tilde{\sigma}$, is biased for $\hat{\sigma}$, the elements of the pseudo-population covariance matrix. How much difference does the appropriate degrees or freedom adjusted bootstrap estimation of the error covariance matrix make for bootstrap estimates of the IRF? We provide some intuitive analytics to address this question.

From equation (4.8) we infer that the bootstrap estimates of the IRF are given by (4.10)

$$\tilde{A}_\ell = \tilde{C}_\ell \tilde{A}_0, \ell = 1, 2, \ldots \tag{4.10}$$

where $\tilde{A}_\ell$, $\tilde{C}_\ell$, and $\tilde{A}_0$ are bootstrap estimates. Any bias in $\tilde{A}_0$ will, thus, likely carry over to all the $\tilde{A}_\ell$. We see from equation (4.9) that $\tilde{A}_0$ depends only on the bootstrap estimate of Σ, $\tilde{\Sigma}$, which, based on results reported above, we expect to be systematically biased. Consequently, it will be instructive to consider how potential bias in the bootstrapped estimates of the VAR covariance matrix can affect the bootstrapped IRF.

The original sample estimate of the VAR error covariance matrix is $\hat{\Sigma} = \frac{1}{T-R}\hat{U}'\hat{U}$. As in the previous section, we consider two alternative bootstrap estimates of $\hat{\Sigma}$ based on the pseudo-population. The standard bootstrap estimate is given by $\tilde{\Sigma}_{b,1} = \frac{1}{T-R}\tilde{U}_b'\tilde{U}_b$ where $\tilde{U}_b$ is the $T \times K$ matrix of bootstrap residuals from the b^{th} bootstrap iteration. The DF-adjusted bootstrap estimate is given by $\tilde{\Sigma}_{b,2} = \frac{T}{(T-R)^2}\tilde{U}_b'\tilde{U}_b$. So,

$$\tilde{\Sigma}_{b,1} = (1+b)\tilde{\Sigma}_{b,2} \tag{4.11}$$

where $b = -R/T$ is the proportional difference between the standard and DF-adjusted bootstrap estimates of $\hat{\Sigma}$. Based on the results of the previous section (including the Monte Carlo evidence alluded to), we might expect b to approximate the bias in $\tilde{\Sigma}_{b,1}$.

We can use equation (4.11) to derive the implied proportional difference between the corresponding bootstrapped IRFs. Equation (4.9) implies that $\tilde{\Sigma}_{b,1} = \tilde{A}_{0,1}\tilde{A}_{0,1}'$ and $\tilde{\Sigma}_{b,2} = \tilde{A}_{0,2}\tilde{A}_{0,2}'$. So, from (4.11) we have $\tilde{A}_{0,1}\tilde{A}_{0,1}' = (1+b)\tilde{A}_{0,2}\tilde{A}_{0,2}'$ which, in turn, implies that

$$\tilde{A}_{0,1} = \sqrt{1+b}\,\tilde{A}_{0,2} = (1+a)\tilde{A}_{0,2} \tag{4.12}$$

Equating $\sqrt{1+b}$ and $1+a$ in (4.12) implies that b and a are related by

$$a = \sqrt{1+b} - 1 \tag{4.13}$$

Since $-1 < b = -R/T < 0$, we see that $b < a < 0$. As T increases without bound, b and a approach zero so the standard and DF-adjusted IRFs are asymptotically equivalent.

Now, consider how this proportional difference in the bootstrap estimate of $\hat{A}_0$ affects the alternative bootstrap IRFs. First, it is important to recognize that $\hat{C}_\ell$ does not depend on which bootstrap estimate of $\hat{\Sigma}$ we choose. Thus, as implied by equation (4.8), the two IRFs are given by

$$\tilde{A}_{\ell,j} = \tilde{C}_\ell \tilde{A}_{0,j}, j = 1, 2, \ell = 1, 2, \ldots \tag{4.14}$$

In particular,

$$\tilde{A}_{\ell,2} = \tilde{C}_\ell \tilde{A}_{0,2} = \tilde{C}_\ell \frac{1}{1+a} \tilde{A}_{0,1} = \frac{1}{1+a} \tilde{A}_{\ell,1} \tag{4.15}$$

where the second equality follows from (4.11) and the final equality follows from (4.10). From equation (4.15), it follows that the proportional difference between the two IRFs is the same for all values of ℓ:

$$\frac{\tilde{A}_{\ell,1} - \tilde{A}_{\ell,2}}{\tilde{A}_{\ell,2}} = a, \;\; \ell = 1, 2, \ldots \tag{4.16}$$

Thus, the bootstrap IRF proportional difference is constant and equal to a 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 elements of the standard bootstrap covariance matrix estimate is 17% less than the DF-adjusted estimate and the corresponding IRFs differ by 9% of the DF-adjusted IRF.[16]

4.3 An Example

As indicated earlier,[17] the procedures that generate this bias seem to be quite common in the empirical SVAR literature. In this section we illustrate its effect in practice by replicating the biased results obtained in a single influential paper by Christiano, Eichenbaum, and Evans (1999) (CEE). We then compute the corresponding DF-adjusted IRF and associated bootstrap confidence intervals to draw our comparison. Finally, we examine coverage accuracy by comparing the coverage rates for standard bootstrap CIs with those for the DF-adjusted CI.

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.[18] 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

[16] These are the implied values of b and a in percentage terms. As indicated above, $b = -R/T$, and a can be computed from (4.13)

[17] See footnote 7 above

[18] The seven variables they include, in recursive order, are the log of real GDP, the log of the implicit GDP deflator, a series of smoothed changes in a commodity price index, the federal funds rate, the log of nonborrowed reserves plus extended credit, the log of total reserves, and the log of M1.

period.[19] 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.[20] As seen in Figure 2 here and Figure 2 of Christiano, Eichenbaum, and Evans (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.

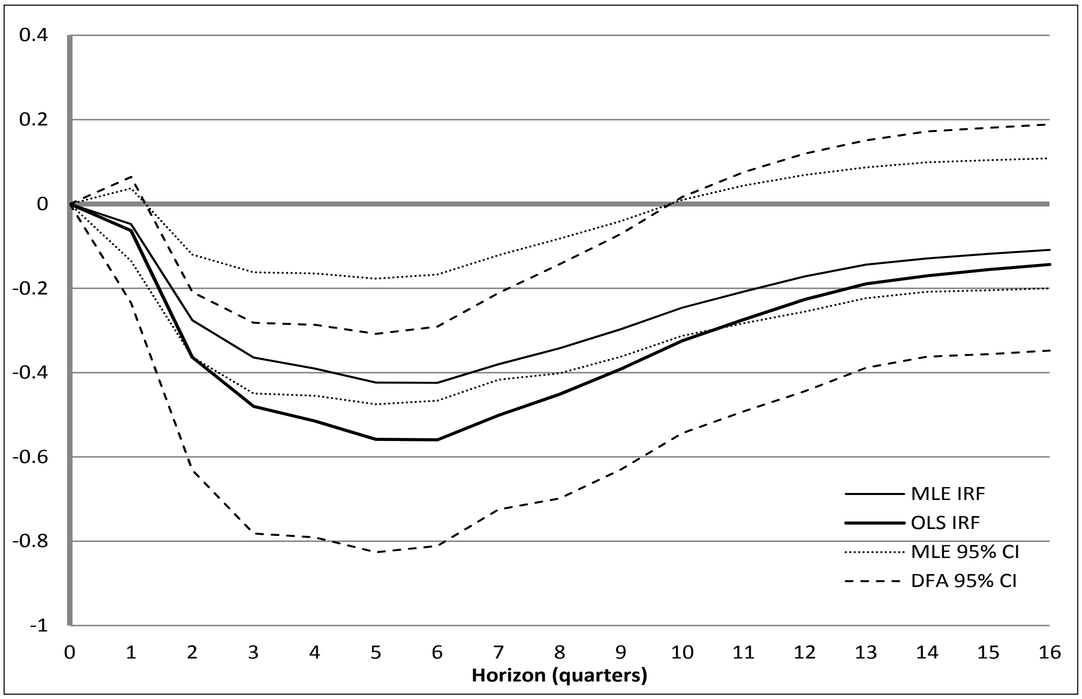
To illustrate the effect of bias due to MLE and the further bias due to the CEE bootstrap IRFs, we estimate the CEE model once again but this time including the degrees of freedom correction we suggest in this paper. These results for the first-stage IRF and the bootstrap confidence intervals are also reported in Figure 2. We first notice 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.[21] We also

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

[20]This is because, as equation (4.11) shows, the proportional difference between the bootstrap estimates of the A_0 matrix is a multiplicative scalar that affects all elements of the matrix the same way. Equation (4.13) shows that this same proportional difference will carry over to every IRF. The complete results are reported in an unpublished appendix available on request

[21]Indeed, we will always draw the same conclusion about statistical significance when our interest is in whether or not the IRF is significantly different from zero. This is a consequence of the fact, illustrated in the previous section equation (4.15), that the DF-adjusted bootstrap IRF is proportional to the standard IRF at all horizons with the constant of proportionality positive but less than one. Accordingly, both confidence interval bounds will cross the horizontal axis (zero line) at exactly the same horizons. This implies that the

Figure 2: 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 OLS IRF; CEE use MLE. The dotted lines give the MLE bootstrap 95% confidence intervals and the dashed lines give the DF-adjusted 95% confidence intervals.



notice that using the OLS rather than the MLE estimate of original error covariance matrix causes the corresponding OLS IRF to lie entirely below the MLE IRF obtained by CEE.

In addition, we see that the confidence intervals also shift significantly when we adjust the degrees of freedom in the bootstrap estimates of the

range over which the IRF is significantly greater or less than zero will be the same whether or not a degrees of freedom adjustment is applied. Adjusting the degrees of freedom can lead to a reversal of conclusion, however, if the null hypothesis takes on a value other than zero.

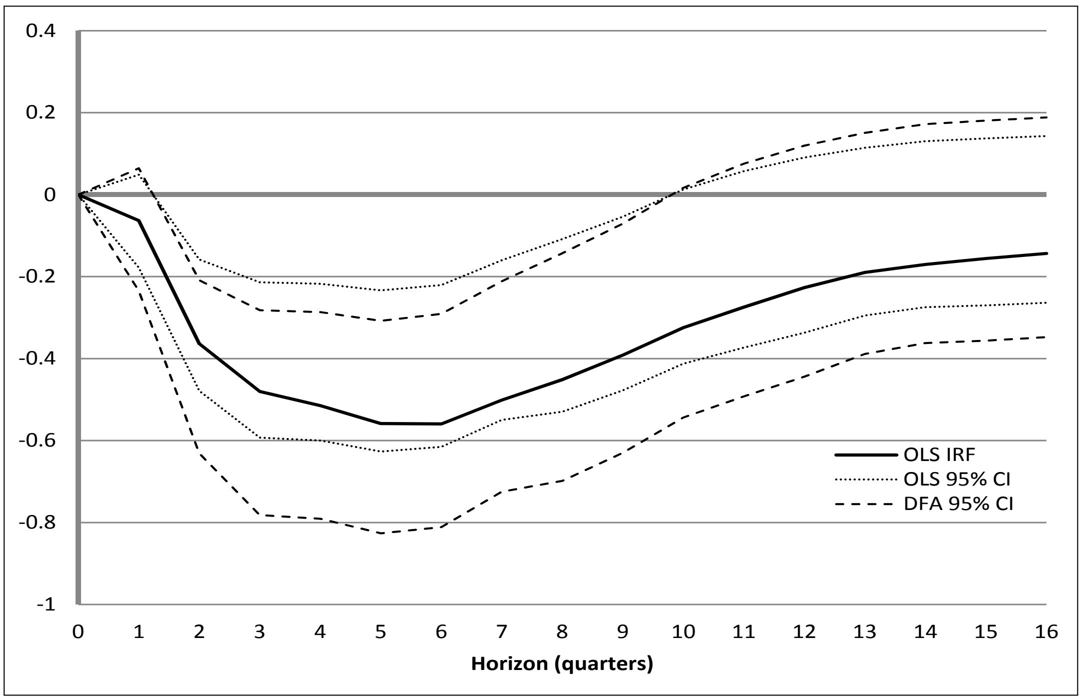
error covariance matrix. We note three consequences. First, we see that for much of the time horizon, the DF-adjusted OLS IRF actually lies below the CEE 95% confidence intervals[22] Second, we see that adjusting the degrees of freedom has greatly reduced the asymmetry in the confidence intervals. Third, we notice that between 2 and 11 quarters, the upper 95% confidence bounds are farther away from zero after degrees of freedom adjustment. This provides stronger evidence supporting 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 3. In the typical approach incorporating the natural OLS degrees of freedom correction, the original IRF is already DF-adjusted so we only have a single IRF estimate. However, the typical procedure does result in biased bootstrap confidence intervals. As in Figure 2, we again see that the typical biased procedure results in quite asymmetric confidence intervals which are, in part, a consequence of the bias; the DF-adjusted confidence intervals exhibit much less asymmetry. Also, as noted in the discussion of Figure 2, over a range of intermediate horizons, the upper bound of the DF-adjusted 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 adjusting the degrees of freedom in both the original IRF and especially in the bootstrap confidence interval estimates

[22]This leads us to conjecture that the often puzzling asymmetry in IRF CIs found in the literature is largely due to the bias documented in this paper.

Figure 3: 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 bootstrap 95% confidence intervals not adjusted for degrees of freedom and the dashed lines give the DF-adjusted 95% confidence intervals.



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

Of course, for the degrees of freedom adjustment we recommend to be of practical value, we must have confidence that it will result in greater coverage accuracy for the resulting CIs. Accordingly, we conclude this section by reporting the results of a series of Monte Carlo experiments that investigate the coverage rates of alternative bootstrap CIs. To avoid the potential arbitrariness of an ad hoc data generating process (DGP), we treat the benchmark

CEE model as our initial DGP from which we obtain the "true" IRF.[23] Using that model and assuming jointly normal errors with the CEE estimated covariance matrix, we generate 1000 Monte Carlo trials of the same length as the CEE sample. Once again, to keep the analysis focused, we look only at the IRF representing the effect of a negative monetary policy shock on output.[24] For each Monte Carlo trial, we then take 200 bootstrap replications and construct three sets of 95% bootstrap IRF confidence intervals: MLE (following CEE), standard OLS, and DF-adjusted. We then report the coverage rates[25] for each of these respective confidence intervals across the 1000 trials.

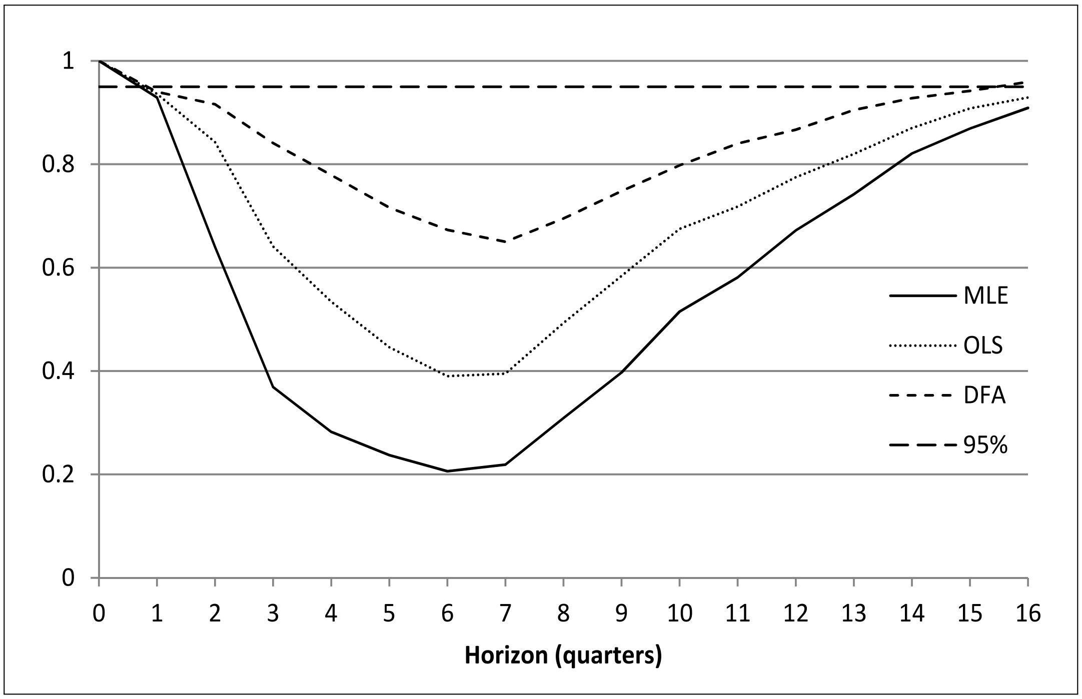
Figure 4 reports the results for the benchmark CEE model DGP along with a reference line at 0.95 reflecting the 95% nominal value of the confidence intervals. We label the methods: MLE, OLS, and DFA. We notice that none of the methods yields coverage rates that are consistently near the ideal value of 0.95 in this baseline case but the DFA method we recommend is uniformly superior to the traditionally-used alternatives. Coverage rates for the DFA method fall to about 0.65 but are generally above 0.75. The MLE has particularly poor coverage rates for intermediate horizons, falling as low as 0.2 while coverage rates fall in between for the OLS method. These poor coverage rates may reflect the fact that the bias we account for here is not

[23] Kilian and Chang (2000) argue that the results of studies that focus on simple ad hoc (e.g., bivariate) VAR models may not generalize to higher dimensional models that are typical of actual applied work. In their study investigating coverage rates, they use three leading models in the literature, including the CEE model, as data generating processes.

[24] For comparison, see the upper left graph in Figure 3 of Kilian and Chang (2000).

[25] By coverage rate we mean the fraction of Monte Carlo trials for which the respective confidence interval includes the true IRF. We evaluate the coverage rate at each point of the IRF horizon.

Figure 4: A comparison of coverage rates for three alternative 95% bootstrap confidence intervals: MLE, OLS, and DFA. Applied to the impulse response function showing the effect of monetary policy on output in the baseline CEE model as originally parameterized.



the only bias affecting the results. Furthermore, it would not be appropriate to generalize that coverage rates for all IRFs are likely to be as poor as those reported in Figure 4. The evidence reported by Kilian and Chang (2000) for the CEE model suggests that the IRF representing the effects of a monetary policy shock on output exhibited lower coverage rates than other IRFs. We have confirmed this for our case.[26] It turns out that the IRF we are most interested in has the poorest coverage rates.

[26]In the appendix available from the authors, we report coverage rate results for all the IRFs corresponding to the effects of a monetary policy (federal funds rate) shock. For the other IRFs as well, the DFA method yields higher coverage rates.

To investigate the robustness of the finding that the DFA method gives greater coverage accuracy, we consider alternative variations on the benchmark CEE model. We first consider alternative parameterizations of the error covariance matrix and then examine variation in key slope coefficients. We also consider the effect of increasing the sample size and the consequences for DGPs with non-normal errors.

The first three alternative DGPs retain the slope parameter of the baseline CEE model but change the values of the elements of the error covariance matrix. The first of these doubles all those values, the second halves them, and the third sets all the off-diagonal elements to zero. To conserve space, and since the coverage rates were only marginally affected in each of these three cases, the results are not reported here but are available from the authors in an unpublished appendix. As in the baseline model, coverage rates were generally poor but the DFA method improved considerably on the OLS method and especially the MLE method.

Because of the impracticality of varying the very large number of slope parameters for the seven-equation, four-lag CEE model in a broadly systematic way, we focus only on the parameters that are most likely to affect the IRF of principal interest. In particular, we vary the first order autoregressive and cross-autoregressive coefficients relating to output and the federal funds rate. We first try a parameterization which is the same as the baseline model except that it halves the benchmark values of these four first order parameters. We also examine two further parameterizations which halve only the first order autoregressive slope coefficients for output and the federal funds rate respectively. The result for the third of these is reported in Figure 5.[27] We see that halving the size of the coefficient on the first lagged value of the federal funds rate results in considerable improvement in coverage rates for all three methods with the performance of the recommended DFA method being quite good. For all horizons, the coverage rate is above 0.85 and often above 0.9. The coverage rates for the other parameterizations also improve relative to the baseline model with the halving of all four of the relevant first order slope coefficients producing slightly better coverage rates and the model that halves only the first order coefficients producing slightly better coverage rates and the model that halves only the first order autoregressive coefficient on output doing a little worse than the results reported in Figure 5.

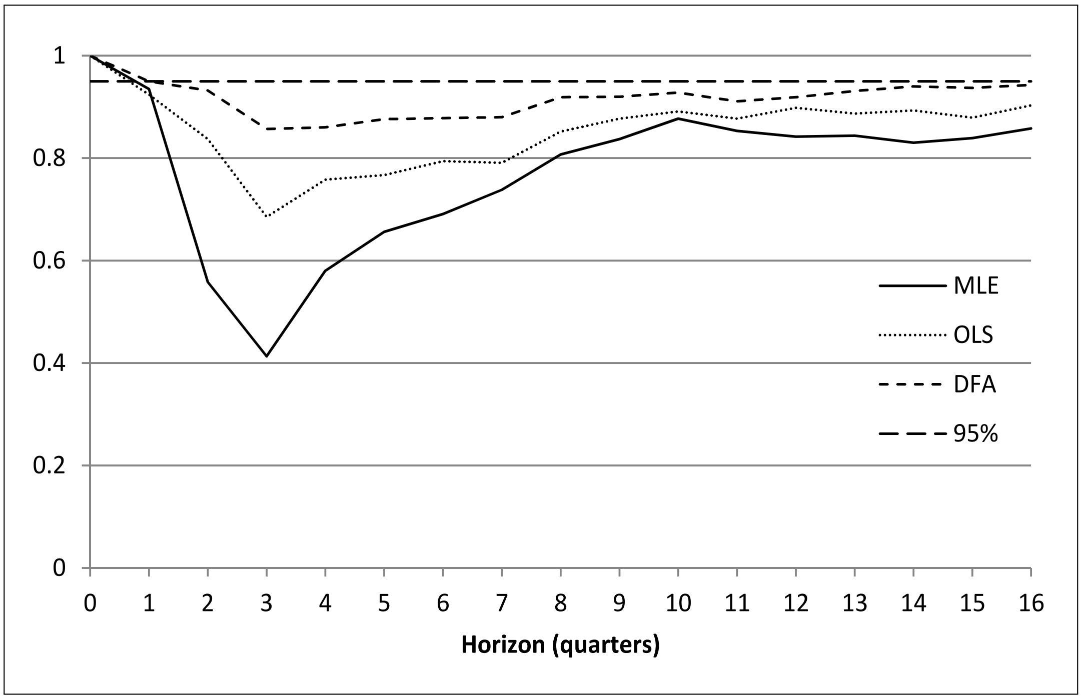
Since the bias we are investigating shrinks as sample size increases, we examine the effect on coverage rates of increasing sample size in the baseline CEE model. We first double the usable CEE sample size (T) from 116 to 232 and report the results in Figure 6. Not surprisingly, when compared with Figure 4, we see that the coverage rates improve considerably, with rates for the DFA method exceeding 0.8 for most of the horizon. The DFA method also continues to be greatly superior to the OLS and MLE methods for estimating the bootstrap error covariance matrix. It is interesting to note that this larger sample size is typical of most recent empirical work using post-war U.S. quarterly data. When we quadruple the sample size (T=464), the coverage rate for the DFA method generally exceeds 0.90.[28]

As a final consideration, we consider DGPs with non-normal errors. In

[27] The others are included in the unpublished appendix available on request.

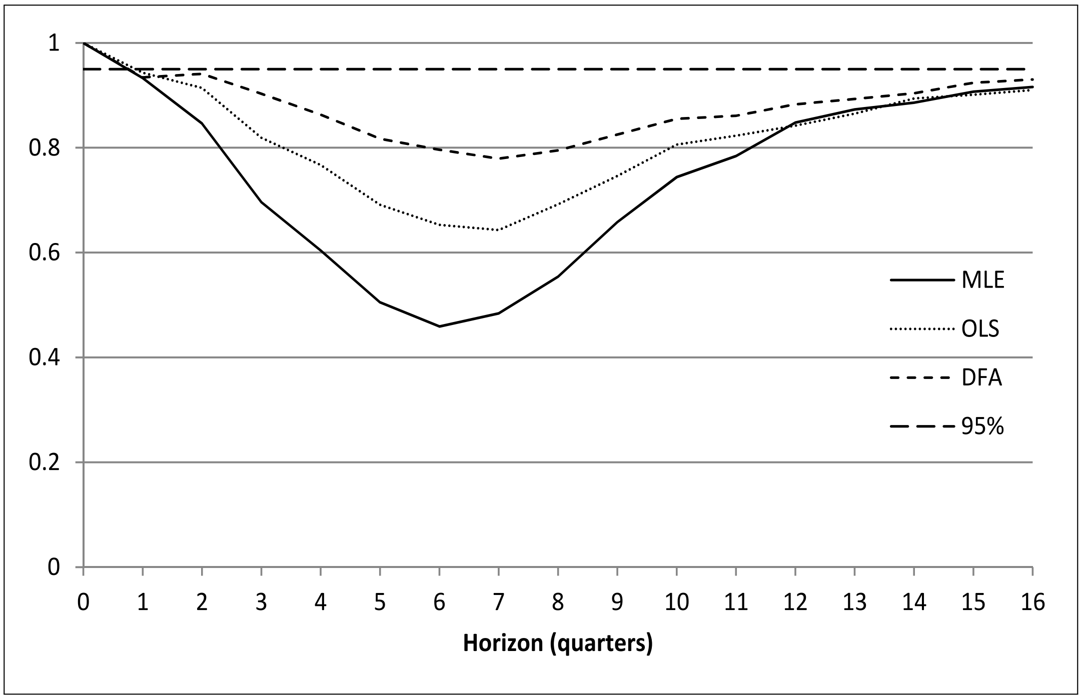
[28] Results are included in the unpublished appendix.

Figure 5: A comparison of coverage rates for three alternative 95% bootstrap confidence intervals: MLE, OLS, and DFA. Applied to the impulse response function showing the effect of monetary policy on output in the CEE model setting the first order autoregressive slope coefficient for the federal funds rate at half its size in the baseline model.



particular, we are interested in what happens if the errors come from a distribution with fatter tails than the normal. We investigate errors generated by two t-distributions, one with ten degrees of freedom and one with five degrees of freedom. The results for the latter are given in Figure 7. We see that coverage rates improve, relative to Figure 4, for all three methods with the DFA method maintaining its superiority. It exhibits coverage rates generally quite close to the nominal level of 0.95. The results for the t-distribution with ten degrees of freedom (reported in the unpublished appendix) also represent an

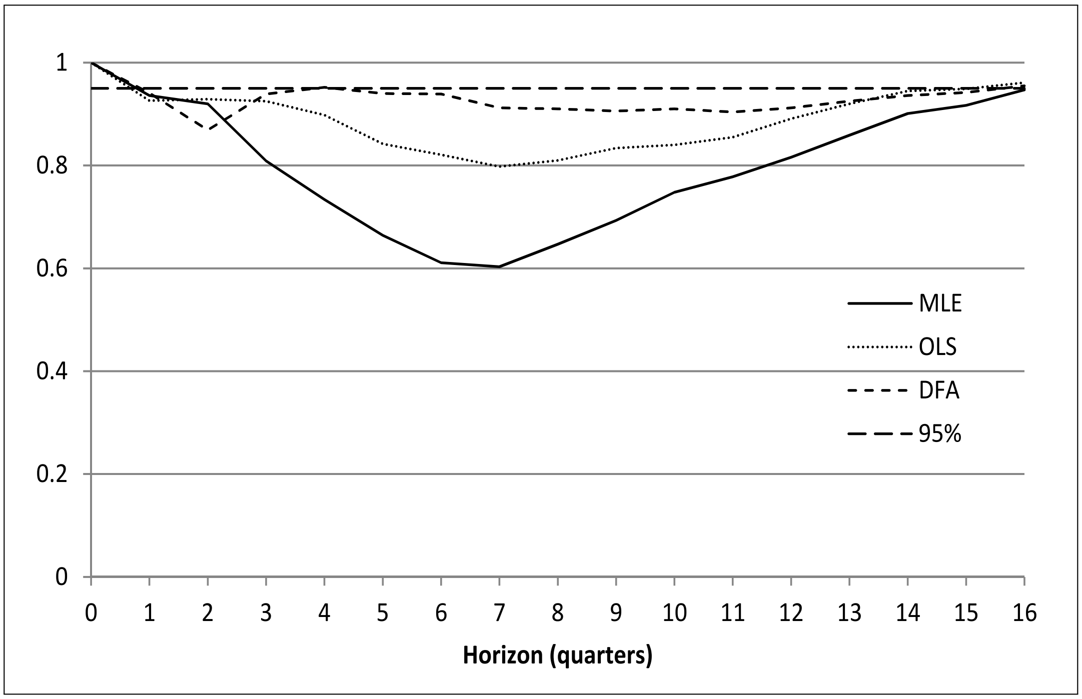
Figure 6: A comparison of coverage rates for three alternative 95% bootstrap confidence intervals: MLE, OLS, and DFA. Applied to the impulse response function showing the effect of monetary policy on output in the CEE model with the sample size doubled (T=232).



improvement over the baseline case but not as much as we see in Figure 7.

In summary, we conclude that the straightforward DFA method for obtaining bootstrap estimates results in considerably improved coverage accuracy. Since the bias we discuss here reflects only one of several potential sources of bias, we are not surprised to see that even the DFA method often results in poor coverage accuracy. However, the evidence reported here suggests that, as sample size approaches that of most modern macroeconomic research, and/or if the distribution generating the errors has fatter tails than a normal distribution, coverage rates for the DFA method may not only be

Figure 7: A comparison of coverage rates for three alternative 95% bootstrap confidence intervals: MLE, OLS, and DFA. Applied to the impulse response function showing the effect of monetary policy on output in the CEE model with errors generated from a t-distribution with 5 degrees of freedom.



improved but become reasonable. Furthermore, these potential gains are available with a simple degrees of freedom adjustment.

4.4 Conclusion

We have discussed a commonly occurring source of bias in bootstrap estimates of confidence intervals for IRFs in SVARs arising from the downward bias in the traditional bootstrap estimate of the VAR covariance matrix. Since the bootstrap IRFs depend on these biased estimates, they are systematically distorted along with the implied bootstrap IRF percentile confidence intervals. This distortion is potentially large but, fortunately, can be readily ameliorated by an additional degrees of freedom adjustment when estimating the VAR covariance matrix. Furthermore, the results of a series of Monte Carlo experiments suggest that we can expect the degrees of freedom adjusted confidence intervals to exhibit improved coverage accuracy relative to traditionally-used confidence intervals.

Exercises

Homework 1

Using the bootstrap to obtain confidence bands for IRFs

Obtain the 95% confidence bands for the four IRFs in Problem 2 from the Structural VAR chapter. Use the degrees of freedom adjustment suggested in Phillips and Spencer (2011).

References

Berkowitz, J., and L. Kilian (2000): "Recent developments in bootstrapping time series.," *Econometric Reviews*, 19, 1–48.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): *The Handbook of Macroeconomics, vol. 1A*chap. Monetary policy shocks: What have we learned and to what end?, pp. 65–148. North Holland, Amsterdam.

Davidson, R., and J. G. MacKinnon (1993): *Estimation and inference in econometrics.* Oxford University Press, New York.

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

Efron, B., and R. J. Tibshirani (1993): *An introduction to the bootstrap.* Chapman & Hall, New York.

Freedman, D. A., and S. C. Peters (1984): "Bootstrapping a regression equation: Some empirical results," *Journal of the American Statistical Association*, 79, 97–206.

Inoue, A., and L. Kilian (2002): "Bootstrapping smooth functions of slope parameters and innovation variances in VAR(∞) models," *International Economic Review*, 43, 309–331.

Kilian, L., and P. Chang (2000): "How accurate are confidence intervals for impulse responses in large VAR models?," *Economics Letters*, 69, 299–307.

Killian, L. (2008): "Notes for Economics 676: Macroeconometrics, University of Michigan," .

Peters, S. C., and D. A. Freedman (1984): "Some notes on the bootstrap in regression problems," *Journal of Business and Economics Statistics*, 2, 406–409.

Phillips, K. L., and D. E. Spencer (2011): "Bootstrapping Structural VARs: Avoiding a Potential Bias in Confidence Intervals for Impulse Response Functions," *Journal of Macroeconomics*, 33(4), 582–594.

Stine, R. A. (1987): "Estimating properties of autoregressive forecasts," *Journal of the American Statistical Association*, 82, 1072–1078.