The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$ 98	1-year subscription	$138
2-year subscription	$165	2-year subscription	$245
3-year subscription	$225	3-year subscription	$345
1-year student subscription	$ 75	1-year student subscription	$ 99
1-year university library subscription	$125	1-year university library subscription	$165
2-year university library subscription	$215	2-year university library subscription	$295
3-year university library subscription	$315	3-year university library subscription	$435
1-year institutional subscription	$245	1-year institutional subscription	$285
2-year institutional subscription	$445	2-year institutional subscription	$525
3-year institutional subscription	$645	3-year institutional subscription	$765
Electronic only		**Electronic only**	
1-year subscription	$ 75	1-year subscription	$ 75
2-year subscription	$125	2-year subscription	$125
3-year subscription	$165	3-year subscription	$165
1-year student subscription	$ 45	1-year student subscription	$ 45

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The Stata Journal (2013)
13, Number 3, pp. 528–546

Valid tests when instrumental variables do not perfectly satisfy the exclusion restriction

Andrés Riquelme
North Carolina State University
Department of Economics
Raleigh, NC
jariquel@ncsu.edu

Daniel Berkowitz
University of Pittsburgh
Department of Economics
Pittsburgh, PA
dmberk@pitt.edu

Mehmet Caner
North Carolina State University
Department of Economics
Raleigh, NC
mcaner@ncsu.edu

Abstract. There is a growing consensus that it is difficult to pick instruments that perfectly satisfy the exclusion restriction. Drawing on results from Berkowitz, Caner, and Fang (2012, *Journal of Econometrics* 166: 255–266), we provide in this article a nontechnical summary of how valid inferences can be made when instrumental variables come close to satisfying the exclusion restriction. Although the Anderson–Rubin (1949, *Annals of Mathematical Statistics* 20: 46–63) test statistic is robust to weak identification, it assumes that the instruments are perfectly orthogonal to the structural error term and is therefore oversized under mild violations of the orthogonality condition. The fractionally resampled Anderson–Rubin (FAR) test is a modification of the Anderson–Rubin test that accounts for violations of the orthogonality condition. We show that in small samples, the size of the resampling block of the FAR test can be modified to obtain valid critical values and analyze its size and power properties. We focus on power and not on size-adjusted power because the FAR test uses only one critical value in its application. We also describe user-written commands to implement the Anderson–Rubin and FAR tests in Stata.

Keywords: st0307, far, fractionally resampled Anderson–Rubin test, exclusion restriction, instrumental variables, near exogeneity

1 Introduction

Instrumental-variable methods are used in economics to study major questions, including the impact of institutions on economic performance and the returns to schooling. Valid instruments must be relevant and exogenous. In the case of relevance, substantial progress has been made in understanding the asymptotic properties of weak instruments. Stock and Wright (2000) show how the Anderson–Rubin (1949) test (the AR test) can be used to draw valid inferences when the instruments are weak.

In the case of exogeneity, however, researchers are becoming more concerned about the difficulty of picking instruments that perfectly satisfy the exclusion restriction. For example, in an influential study of the impact of institutions on long-term growth, Acemoglu, Johnson, and Robinson (2001) use early settler mortality data from as far back as the fifteenth and sixteenth centuries as an instrument for contemporary institutions.[1] Glaeser et al. (2004) argue that the early settlers brought their attitudes about education to their colonies, affecting the long-term growth through their influence on human capital accumulation. Similarly, draft lotteries (Angrist 1990) and whether a man grew up in the vicinity of a four-year college (Card 1995) are influential instruments for estimating the returns to schooling. In each case, however, there are good reasons to believe that the exclusion restriction is not necessarily perfect (see Wooldridge [2010, 95–96]).

In this article, we provide a nontechnical summary of the new test statistic derived in Berkowitz, Caner, and Fang (2012) for instruments that come "close" to satisfying the exclusion restriction but do not satisfy it perfectly. In our analysis, we use the AR test because it is robust to weak identification. However, because the AR test uses the overly strong assumption that an instrument is perfectly exogenous, it can have bad small-sample properties (Caner 2010; Guggenberger 2012). The fractionally resampled AR (FAR) test modifies the AR test on the basis of results from Wu (1990, sec. 2), accounting for the extent to which an instrument violates the orthogonality condition and is not oversized in large samples.

The rest of the article is organized as follows: Section 2 describes the AR test in a setup that allows for instruments that do not perfectly satisfy the orthogonality condition. Section 3 summarizes the FAR test and shows how the block size for the FAR test can be adjusted to improve the test size and power. Section 4 describes the syntax and output of our user-written Stata command and details the different available options through an example from Acemoglu, Johnson, and Robinson (2001, 2011). Section 5 presents the results of size and power simulations under different levels of violation of the orthogonality condition. Section 6 concludes.

2 Inferences when instruments are not perfectly exogenous

Consider the following setup:

$$y = WB + Y\theta_0 + u \tag{1}$$

$$Y = W\Gamma + Z\Pi + V \tag{2}$$

In this system of equations, y is an $n \times 1$ vector of outcomes, n is the sample size, Y is an $n \times m$ matrix of endogenous variables, and Z is an $n \times k$ matrix of instruments.

1. We ignore the controversy about the construction of the early settler mortality variable. For this debate, see Acemoglu, Johnson, and Robinson (2008) and Albouy (2008).

For example, y can be long-term gross national product (GNP) per capita, n can be the number of countries that are former colonies, and Y can be a set of contemporary institutions. In Acemoglu and Johnson (2005), $m = 2$ and includes property rights and contract enforcement. For simplicity, and without loss of generality, we consider the case where Y is an $n \times 1$ vector of property rights institutions.

There are a host of exogenous covariates in W, which is an $n \times l$ matrix. For example, if $l = 3$, then W could include GNP, human capital, and temperature in 1960. The coefficients obtained for θ_0 and Π in (1) and (2) will remain the same after projecting out W from the system. By using the projection matrix $P = W(W'W)^{-1}W'$, we define

$$\begin{aligned} y_W &= y - Py \\ Z_W &= Z - PZ \\ Y_W &= Y - PY \end{aligned}$$

And the system of equations in (1) and (2) can be written as

$$y_W = Y_W \theta_0 + u_W \tag{3}$$

$$Y_W = Z_W \Pi + V_W \tag{4}$$

Thus the vector W of covariates can be ignored.

In Acemoglu, Johnson, and Robinson (2001), the parameter of interest θ_0 in (3) is the impact of institutions on long-term growth. Because long-term GNP per capita also influences institutions and because there are potentially omitted variables in the residual u_W that influence both institutions and GNP per capita, the variable Y is endogenous. Technically, this means that $\mathrm{cov}(Y_W, u_W) \neq 0$. To correct for the endogeneity of institutions, one uses an instrument or a set of instruments, Z_W, as an exogenous source of variation for institutions. The instruments satisfy the condition

$$E(Z_{Wi} V'_{Wi}) = 0, \qquad i = 1, \ldots, n \tag{5}$$

There is much literature for drawing inferences when instruments are weak but still sufficiently relevant (Stock and Wright 2000), and there are now commands for implementing valid tests in Stata (see Moreira and Poi [2003]). Here we consider tests for instruments that are not perfectly exogenous, in which case the standard t statistic and the AR test for testing $H_0 : \theta = \theta_0$ have massive size distortions (Berkowitz, Caner, and Fang 2008) because they assume orthogonality as in (5). More realistically, a set of instruments may exhibit near exogeneity as follows:

$$E(Z_{Wi} u_{Wi}) = \frac{C}{\sqrt{n}} \tag{6}$$

Equation (6) allows for a slight covariance between the instruments and the error term. C is a $k \times 1$ vector (one component for each instrument), and each element of C, denoted by C_j $(j = 1, \ldots, k)$, is a constant. The sign of each C_j depends on the sign of the covariance between the jth instrument and the error term. For example, when $k = 2$, then we can have $C = (-1, 2)'$. For simplicity and without loss of generality, we assume that the upper and lower bounds of the set containing the C_j values are the same for all the instruments. Further technical details are described in section 2 of Berkowitz, Caner, and Fang (2012).

To test the null hypothesis $H_0 : \theta = \theta_0$, the AR test is preferred for several reasons. First, it can be used when the instruments are weak. Moreover, Guggenberger (2012) shows that the AR test is the best choice for limiting size distortion when the exclusion restriction is slightly violated. Caner (2010) also shows that the AR test is slightly oversized in a framework of many instruments.

Let the $n \times 1$ vector of residuals of the structural equation under the null be denoted $u_W(\theta_0)$:

$$u_W(\theta_0) = y_W - Y_W \theta_0$$

Then the AR test for testing $H_0 : \theta = \theta_0$ assumes that $C = 0$ (that is, the instruments perfectly satisfy the exclusion restriction). The test statistic is given by

$$\text{AR}(\theta_0) = n \times \overline{S}'_n(\theta_0) \widehat{\Omega}^{-1} \overline{S}_n(\theta_0) \tag{7}$$

where $\widehat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} Z_{Wi} Z'_{Wi} \{u_W(\theta_0)\}^2$ and $\overline{S}_n(\theta_0) = [\{Z'_W u_W(\theta_0)\}/n]$ can be interpreted as the $k \times 1$ vector of estimated covariances between the instruments and the residuals in the structural equation under the null hypothesis $H_0 : \theta = \theta_0$.

The limiting distribution of the AR test is central chi-squared with k degrees of freedom. Berkowitz, Caner, and Fang (2008, 2012) show that the AR test overrejects the null when the orthogonality condition is not perfectly satisfied. Moreover, in small samples, the test can be oversized even when the correlation between the instruments and structural error is close to 0. This size distortion gets worse as the correlation between an instrument and the structural error terms gets stronger. This problem arises because the AR test assumes that $C = 0$ in (6). In the next section, we explain how the FAR test accounts for $C \neq 0$ and thus allows the researcher to draw valid but conservative inferences.

3 The FAR test

The FAR test uses Wu's (1990) jackknife histogram estimator to recover the limits of the population mean of θ by taking a subset of size b from the n observations in the full sample. There are $\binom{n}{b}$ blocks of size b with equal probability of being selected, and these are drawn via simple random sampling without replacement. To test the null hypothesis $H_0 : \theta = \theta_0$, we need to estimate $Z'_W u_W(\theta_0)$. Following Berkowitz, Caner, and Fang

(2012), we use the subscript * to label the resampled estimates. Using this notation, we can write the FAR test as

$$\text{FAR}(\theta_0) = \frac{b\overline{S}_b'(\theta_0)\widehat{\Omega}^{-1}\overline{S}_b(\theta_0)}{(1-f)}$$

where $\overline{S}_b(\theta_0) = \sum_{i=1}^{b} Z_i u_i / b$ and f is the fraction of the sample that generates the block of size b.[2] Note that $\widehat{\Omega}$ is obtained from the full sample and replaced by $(1-f)\widehat{\Omega}$ in each iteration. From theorem 1 in Berkowitz, Caner, and Fang (2012, 258), under suitable assumptions, the statistic $J_b(t) = P_*\{\text{FAR}(\theta_0) \leq t\}$, where P_* stands for the resampled probability, converges to $\phi_{mf}(t)$, the cumulative distribution of

$$\left(1 + \frac{\sqrt{f}}{\sqrt{1-f}}\right)^2 \chi^2_{k,nc} \tag{8}$$

where $\chi^2_{k,nc}$ is the noncentral χ^2 with k degrees of freedom and noncentrality parameter $nc = \{1/(1+2\sqrt{f}\sqrt{1-f})\}\{(C\Omega^{-1}C)/2\}$. If half of the sample is resampled, then $f = 1/2$ and the limit in (8) becomes

$$4\chi^2_k + 4C'\Omega^{-1}L + C'\Omega^{-1}C \tag{9}$$

where $L \equiv N(0,1)$, whereas the AR test limit is

$$\chi^2_k + 2C'\Omega^{-1}L + C'\Omega^{-1}C$$

Equation (9) is used for testing $H_0 : \theta = \theta_0$ when $C \neq 0$ and corrects for the size distortions obtained in the standard AR test. This version of the FAR test is very conservative, especially in small samples. To correct for this, theorem 1 in Berkowitz, Caner, and Fang (2012, 258) shows that the resampled fraction f can be modified,

$$f_n = 1/2 - \kappa_n \tag{10}$$

where $\kappa_n > 0$ is a data-driven deterministic sequence converging to 0. In practice, $\kappa_n = \kappa/\sqrt{n}$ is used. For example, if $n = 100$ and $\kappa = 2.5$, then $\kappa_n = 2.5/\sqrt{100} = 0.25$ and $f_n = 0.25$, so each resampling consists of 25 observations. $\kappa_n = 2.5/\sqrt{n}$ provides good power in our simulations, and $\kappa_n = 3/\sqrt{n}$ is recommended when the researcher is confident that the instrument comes close to perfectly satisfying the exclusion restriction.

Our user-written `far` command takes advantage of the flexibility and fast execution of the Mata language to perform the resampling process and estimate the FAR test in an efficient way. The command is introduced in the next section.

2. For practical purposes, `b = ceil(f*n)` should be implemented (see [D] **functions**).

4 The far command

4.1 Syntax

`far` *depvar* [*varlist1*] `(`*varlist2* = *varlist_iv*`)` [*if*] [*in*] [, `reps(`#`)` `kappa(`#`)` `theta(`*numlist1*`)` `ci` `level(`#`)` `grid(`*numlist2*`)`]

4.2 Description

The `far` command performs the FAR test (Berkowitz, Caner, and Fang 2012) for the joint significance of the endogenous regressors in an instrumental-variables regression of *depvar* using the optional controls in *varlist1*, the endogenous regressors in *varlist2*, and the instrumental variables in *varlist_iv*.

4.3 Options

`reps(`#`)` specifies the number of repetitions of the resampling procedure. A large number of repetitions is necessary for the results in section 3 to be valid. The default is `reps(10000)`, and it gives fast and reliable estimates in small samples ($n < 100$). If the number of repetitions is not large enough, the FAR test p-values may vary.

`kappa(`#`)` specifies the value of the κ constant. Note that $\kappa_n = \kappa/\sqrt{n}$ in (10). Any positive real number may be used. The default is `kappa(3)` (see section 5 for justification of the selected default value).

`theta(`*numlist1*`)` allows for a user-defined hypothesis test. *numlist1* is a list of values for the endogenous parameters to be tested (one for each endogenous variable). If `theta()` is not specified, the `far` command will perform a significance test (all the values in *numlist1* will be set as 0). By implementing this option, the user can invert the FAR test to find confidence intervals for θ_0.

`ci` enables the user to test for a grid of different values of θ_0 and search for the $(1-\alpha)\%$ confidence interval for the true scalar θ. The significance level and the grid can be customized by using the options `level(`#`)` and `grid(`*numlist2*`)`. This option is available when there is only one endogenous variable.

`level(`#`)` is the significance level for the test in the grid search. The default is `level(95)`.

`grid(`*numlist2*`)` specifies the grid for the values of θ_0 to be tested. *numlist2* consists of three elements: the minimum level, the maximum level, and the increments of the grid. The default is `grid(-30, 30, 0.01)`.

4.4 Stored results

`far` stores the following in `r()`:

Scalars			
`r(n)`	number of observations	`r(kappa)`	the constant κ
`r(ar)`	full-sample AR statistic	`r(k)`	number of instruments
`r(arp)`	full-sample p-value	`r(l)`	number of controls
`r(farp)`	FAR p-value	`r(m)`	number of endogenous variables
`r(reps)`	resampling repetitions		
Macros			
`r(cmdline)`	command as typed	`r(endogenous)`	list of endogenous variables
`r(depvar)`	name of dependent variable	`r(instruments)`	list of instruments
`r(title)`	title in estimation output	`r(grid)`	grid values
`r(exogenous)`	list of controls		
Matrices			
`r(theta)`	endogenous parameters tested	`r(ci)`	fractionally resampled p-values for the parameters in the grid

4.5 Example

Acemoglu, Johnson, and Robinson (2001) use two-stage least-squares methods to estimate the effect of institutions on long-term economic growth. Their baseline dataset consists of 64 countries that are former European colonies. They use the log of per-capita gross domestic product with purchasing-power-corrected prices (`logpgp95`) as the measure of long-term growth, an index of protection against expropriation from 1985 to 1995 (`avexpr`) as the measure of institutions, and the log early settler mortality of colonizers (`logem4`) as the instrument for institutions. The fundamental identifying assumption then is that early settler mortality influences long-term growth exclusively through the quality of contemporary institutions [see (5)].

One important control that Acemoglu, Johnson, and Robinson include in their robustness checks is the incidence of malaria in 1994 (`malfal94`). There are two missing values for this variable, which reduces the sample size to 62. This control is critical for their exclusion restriction because it offsets the potential impact of early settler mortality through the contemporary disease environment. However, even after controlling for the contemporary disease environment, there are still reasons to argue that the exclusion restriction that Acemoglu, Johnson, and Robinson use is not perfect (see, for example, Glaeser et al. [2004]). Thus we relax the strict exclusion restriction in (5) and allow for the early settler mortality instrument to exhibit near exogeneity as in (6). We compare the AR and FAR tests to examine how the potential correlation between the instrument and structural error will affect inference.[3]

3. Acemoglu, Johnson, and Robinson (2011) point out that the inclusion of the variable `malfal94` is "highly problematic" because the current prevalence of malaria is endogenous. In our example, we include `malfal94` to show how our `far` command can easily incorporate control variables in the first and second stages.

In the next two command lines, we load the local data file `fardata.dta` and call the `far` command for the specified instrumental-variable regression:[4]

```
. use fardata
. far logpgp95 malfal94 (avexp = logem4)
Fractionally resampled Anderson and Rubin test.
```

	Full sample statistic	Full sample p-value	FAR p-value	reps	N
AR-test	5.5421	0.0186	0.1462	10000	62

The output displays the full-sample AR statistic, the full-sample and the fractionally resampled p-values, the number of resampling repetitions, and the number of observations. In this case, under the full-sample AR test, the hypothesis $H_0 : \theta_0 = 0$ is rejected at 5% (with a p-value of 1.86%), but the FAR test does not reject it. This is consistent with the result in (9), which shows that the FAR test is more conservative.

To show other available options for the `far` command, we perform the same hypothesis test, but this time, we increase the number of repetitions to 100,000 and set $\kappa = 2$. Note that the null is not rejected at 15% under the FAR test after decreasing κ:

```
. far logpgp95 malfal94 (avexp = logem4), reps(100000) kappa(2)
Fractionally resampled Anderson and Rubin test.
```

	Full sample statistic	Full sample p-value	FAR p-value	reps	N
AR-test	5.5421	0.0186	0.1505	100000	62

To test if the θ_0 parameter is equal to, say, 3, we use

```
. far logpgp95 malfal94 (avexp = logem4), theta(3)
Fractionally resampled Anderson and Rubin test.
```

	Full sample statistic	Full sample p-value	FAR p-value	reps	N
AR-test	2.5611	0.1095	0.3315	10000	62

Note that the p-value of the FAR test increases when testing $H_0 : \theta_0 = 3$. We have rejected that θ_0 is equal to 0 and 3 already. To look for the θ_0 values for which the null hypothesis is not rejected at some fixed α significance level, we can perform a grid search.

4. To obtain the results presented in this article, we set the initial value of the random-number seed to 1111 at the beginning of the Stata session (see [R] **set**).

We implement the grid search by using the `ci` option. To test the null under the default grid,[5] we simply use

```
. far logpgp malfal94 (avexp = logem4), ci
  (output omitted)
```

We are not presenting the default grid here because of its extension.[6] The user can list the grid stored in the `r(ci)` matrix to inspect it. It is enough to say that all the FAR p-values are greater than 0.05; thus the 95% confidence interval for θ_0 obtained from this search is $[-\infty, +\infty]$. A portion of the default grid can be displayed using the following lines:

```
. far logpgp95 malfal94 (avexp = logem4), ci grid(-1,1,0.1)
Fractionally resampled Anderson and Rubin test.
```

	Full sample statistic	Full sample p-value	FAR p-value	reps	N
AR-test	5.5421	0.0186	0.1495	10000	62

```
. matrix list r(ci)
r(ci)[21,3]
     theta  FAR-p   test
 r1     -1  .2191      1
 r2    -.9  .2166      1
 r3    -.8  .2095      1
 r4    -.7   .205      1
 r5    -.6  .1967      1
 r6    -.5  .1921      1
 r7    -.4  .1791      1
 r8    -.3  .1782      1
 r9    -.2  .1706      1
r10    -.1  .1621      1
r11      0  .1536      1
r12     .1   .146      1
r13     .2  .1437      1
r14     .3  .1617      1
r15     .4  .2446      1
r16     .5  .3633      1
r17     .6  .6744      1
r18     .7  .9634      1
r19     .8  .7547      1
r20     .9  .6336      1
r21      1  .5617      1
```

5. This is equivalent to executing the following command:
`far logpgp malfal94 (avexp = logem4), ci grid(-30, 30, 0.01) level(95).`
6. The default grid has 6,001 consecutive hypothesis tests.

The first column of the `r(ci)` matrix contains the grid of θ_0 values defined by the `grid(`*numlist2*`)` option. The second column corresponds to the FAR test p-values at each of the different θ_0 values. The third column contains a dummy variable that takes the value of 1 if the corresponding θ_0 is included in the confidence interval defined by the `level` option (this occurs if the p-value in column 2 is greater than the critical α level). The default confidence level corresponds to an α level of 5%; therefore, the elements in the third column will be 1 if the corresponding FAR p-value is greater than 0.05.

In the next example, we derive a bounded confidence interval. In light of the debate between Albouy (2008) and Acemoglu, Johnson, and Robinson (2001), Acemoglu, Johnson, and Robinson (2011) recommend capping the settler mortality at 250 per 1,000 per annum. We can generate a transformed variable, estimate the AR and FAR tests, and perform the grid search in two command lines. We increased the number of resampling repetitions to 100,000 to improve the precision of the estimated interval, and we set $\kappa = 3.1$ to show the full usage of the grid search:

```
. generate malaria250 = min(malfal94, 0.250) if malfal != .
(2 missing values generated)
. far logpgp95 malaria250 (avexp = logem4), kappa(3.1) reps(100000)
Fractionally resampled Anderson and Rubin test.
```

	Full sample statistic	Full sample p-value	FAR p-value	reps	N
AR-test	9.2185	0.0024	0.0180	100000	62

After we cap mortality at 250, the FAR p-value is 1.8%, and we reject $H_0 : \theta_0 = 0$. The grid search gives a 95% confidence interval for θ_0 of $[0.34, 4.39]$. Acemoglu, Johnson, and Robinson (2011) obtained the confidence interval $[0.27, 0.95]$ by using the AR test and including other covariates. Ours is more conservative, but it does not suffer the small-sample problems discussed in section 2.

The lower limit of the confidence interval can be obtained using the following command line:

```
. far logpgp95 malaria250 (avexp = logem4), kappa(3.1) reps(10000) ci
> grid(.3,.5,.01)
Fractionally resampled Anderson and Rubin test.
```

	Full sample statistic	Full sample p-value	FAR p-value	reps	N
AR-test	9.2185	0.0024	0.0173	10000	62

```
. matrix list r(ci)
r(ci)[21,3]
     theta  FAR-p   test
 r1     .3  .0402      0
 r2    .31  .0397      0
 r3    .32  .0426      0
 r4    .33  .0464      0
 r5    .34   .053      1
 r6    .35   .054      1
 r7    .36  .0595      1
 r8    .37  .0694      1
 r9    .38    .07      1
r10    .39  .0892      1
r11     .4  .0939      1
r12    .41  .0974      1
r13    .42  .1123      1
r14    .43  .1257      1
r15    .44  .1459      1
r16    .45  .1592      1
r17    .46  .1761      1
r18    .47  .2027      1
r19    .48  .2243      1
r20    .49  .2423      1
r21     .5    .27      1
```

By inspecting the grid, we can see that the lower limit of the interval is 0.34 by using the indicators in the third column.

To make the dummy in the third column take the value of 1 on the basis of the FAR p-values rounded to two decimal places, the user must set the confidence level to 95.5. In this example, the rounded lower limit is 0.33.

Similarly, the upper limit can be obtained by

```
. far logpgp malaria250 (avexp=logem4), reps(100000) ci kappa(3.1)
> grid(4.3,4.5,0.01)
  (output omitted)
```

This last result is for illustrative purposes; it needs to be carefully considered. With a sample of 62 observations, selecting $\kappa = 3.1$ corresponds to a resampled fraction $f = 0.11$, which implies a block size of $b = 7$. This fraction is too small. As the block size diminishes, the resampling technique turns into a subsampling procedure. Berkowitz, Caner, and Fang (2012) (section 4) show that as $f \rightarrow 0$, the AR test is always oversized. We choose $\kappa = 3.1$ only because it generates a bounded interval, although in our simulations, we find that the best combinations of size and power are obtained by selecting subsample sizes between 20% and 25% of the total observations. κ values that generate $f < 0.2$ generate unreliable and unstable results, but this topic should be further examined. In our example, the best choice is $\kappa < 2$. The statistical implications of the estimates obtained by this smaller κ value and the best choice of the block size are beyond the scope of this article.

To empirically obtain valid confidence intervals, we suggest exploring the default grid under κ values that correspond to f above 0.2 to check the overall sequence of the test results and then fine-tune the grid intervals. By using this heuristic approach, we needed three trials to find the presented bounded interval for this dataset. In general, the confidence set can be bounded, disjointed, or even infinite if the model is misspecified, which implies that the grid search might become excessively time consuming. We believe that the option of user-defined grids gives the researcher enough flexibility for finding a solution that is not too time consuming and not too computationally intensive.

5 Simulations

To choose the default value of the constant κ in the `far` command, we simulate the system of equations in (3) and (4) under different scenarios in which the exclusion condition is violated and explore the FAR test size and power properties. We choose scenarios similar to those in Berkowitz, Caner, and Fang (2012) but with smaller correlations between the structural error and the instruments. For empirical purposes, we assume that the researcher chooses imperfect instruments that come close to satisfying the exclusion restriction, so the covariance between the instruments and the residuals in the structural equation is very small but nonzero. The data for z_i, u_i, and v_i are generated from a joint normal distribution $N(0, \Sigma)$, where

$$\Sigma = \begin{bmatrix} 1 & \sigma_{zu} & 0 \\ \sigma_{zu} & 1 & 0.9 \\ 0 & 0.9 & 1 \end{bmatrix} \tag{11}$$

and $\sigma_{zu} = \text{cov}(Z_{Wi} u_{Wi})$.

In (11), we set up $\sigma_z^2 = \sigma_u^2 = \sigma_v^2 = 1$, $\sigma_{zv} = 0$, and $\sigma_{uv} = 0.9$. Note that the upper-left 2×2 submatrix corresponds to our simulated version of the Ω matrix in (7). We also set up σ_{zu} in three different ways:

In the first setup, we have σ_{zu} local to 0 as in (6),

$$\sigma_{zu} = \frac{h}{\sqrt{n}}$$

and we choose h equal to 0.5 and 1 for the simulations. The larger h becomes, the worse is the selected instrument.

The second setup corresponds to σ_{zu} constant,

$$\sigma_{zu} = D$$

and we choose D equal to 0.1 and 0.25 for the simulations.

In the third setup, we have σ_{zu} consistent with the bounds of the compact set containing C,

$$\sigma_{zu} = \frac{an^{1/3}}{n^{1/2}}$$

and a is equal to 0.25 and 0.5 in the simulations.

To explore the size properties of the FAR test, we simulate one endogenous variable ($m = 1$), one instrument ($k = 1$), and two controls ($l = 2$), one of them being a constant: $B = (1, 2)'$. To model strong identification, we set $\Pi = 2$ in (4). We get results (not reported) similar to the weak identification case. The sample size n is equal to 100 and 200, and κ is equal to 1.5, 2, 2.5, ..., 6, so κ_n is equal to $1.5/\sqrt{n}$, $2/\sqrt{n}$, $2.5/\sqrt{n}$, ..., $6/\sqrt{n}$ in (10). We give the data a heteroskedastic structure by using the following error form:

$$u_{*Wi} = abs(Z_{Wi})u_{Wi}$$

Each scenario was simulated 1,000 times with 1,000 resampling iterations. The results for the setups 1, 2, and 3 for the size of the test are presented in table 1. We found the same patterns as those found by Berkowitz, Caner, and Fang (2012). Given that our correlations are smaller, the test is undersized when κ_n is equal to $1.5/\sqrt{n}$ and $2/\sqrt{n}$, but the undersize is corrected when κ_n is equal to $2.5/\sqrt{n}$ and $3/\sqrt{n}$, especially in setups 2 and 3 when the sample size is 100. Note that when $n = 200$, the FAR test is undersized in all the setups because of its conservative nature.

Table 1. Size of the FAR test at $\theta_0 = 0$

	Size at 10%													
	$n = 100$							$n = 200$						
$\kappa =$	1.5	2.0	2.5	3.0	3.5	4.0	4.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0
Setup 1:	$\sigma_{zu} = h/\sqrt{n}$													
$h = 0.5$	0.0	0.1	0.3	2.0	3.3	8.0	16.2	0.0	0.7	1.7	2.3	3.1	6.1	7.7
$h = 1.0$	0.0	0.2	1.3	2.0	6.7	12.3	21.4	0.3	0.7	2.3	4.8	6.3	9.4	12.3
Setup 2:	$\sigma_{zu} = D$													
$D = 0.1$	0.0	0.0	0.6	2.9	6.4	11.5	20.9	0.4	1.3	4.2	6.7	8.4	14.9	19.8
$D = 0.25$	0.1	1.6	5.6	12.4	24.4	36.4	51.8	7.0	14.9	29.2	36.4	50.9	59.4	64.9
Setup 3:	$\sigma_{zu} = an^{1/3}/n^{1/2}$													
$a = 0.25$	0.0	0.0	0.8	3.6	7.4	14.0	23.4	0.4	1.4	4.5	7.3	8.5	15.5	20.8
$a = 0.5$	0.1	1.6	5.2	10.4	21.9	32.1	49.2	3.1	8.7	18.7	25.0	36.0	45.1	52.2
	Size at 5%													
Setup 1:	$\sigma_{zu} = h/\sqrt{n}$													
$h = 0.5$	0.0	0.0	0.0	0.5	0.5	1.8	4.6	0.0	0.0	0.2	0.7	0.2	1.8	1.7
$h = 1.0$	0.0	0.0	0.0	0.4	1.2	2.8	5.8	0.0	0.0	0.1	0.6	0.8	2.2	4.4
Setup 2:	$\sigma_{zu} = D$													
$D = 0.1$	0.0	0.0	0.0	0.6	1.0	2.7	6.7	0.0	0.2	0.6	1.6	2.0	4.2	6.0
$D = 0.25$	0.0	0.1	0.4	2.5	6.2	12.8	27.5	0.1	3.0	7.2	11.4	20.0	30.1	37.1
Setup 3:	$\sigma_{zu} = an^{1/3}/n^{1/2}$													
$a = 0.25$	0.0	0.0	0.0	0.6	1.3	3.6	7.9	0.0	0.2	0.7	1.6	2.0	4.4	6.2
$a = 0.5$	0.0	0.1	0.3	1.9	5.5	10.5	23.5	0.0	1.2	4.0	7.9	12.1	18.2	25.0

The correction factor in (10) is calculated for the different values of κ. $\Pi = 2$ in (4). Each result corresponds to 1,000 heteroskedastic simulations and 1,000 resampling iterations.

To explore the power properties of the FAR test, we simulate scenarios with θ_0 equal to -2, -1.5, -1, -0.5, 0.5, 1, 1.5, and 2 and tested for $\theta_0 = 0$. The results are presented in the tables 2, 3, and 4. We focus on power and not on size-adjusted power because the FAR test uses only one critical value in its application. The simulation exercise shows the test has low power when θ_0 is equal to -0.5 and 0.5 and κ_n is equal to $1.5/\sqrt{n}$ and $2/\sqrt{n}$. The power improves when κ_n is equal to $2.5/\sqrt{n}$ and $3/\sqrt{n}$. Considering these results, we decided to set $\kappa = 3$ as the default value in the `far` command. This κ value is the one that gave us the best size and power combinations and corresponds to a resampling fraction $f = 0.2$ when $n = 100$. The researcher can easily adjust κ to obtain resampling fractions above the 20% of the total sample. Lower f values generate unreliable and unstable results, as discussed in section 4. Further discussion and other setups for the covariance matrix can be found in Berkowitz, Caner, and Fang (2012).

Table 2. Power of the FAR test at $\theta_0 = 0$, covariance setup 1

		θ							
		-2	-1.5	-1	-0.5	0.5	1	1.5	2
κ	h	$n = 100$							
1.5	0.5	96.8	95.4	84.7	14.4	16.8	72.0	90.1	92.2
	1.0	97.3	93.8	84.5	16.4	9.0	72.8	90.7	93.5
2.0	0.5	99.7	99.1	98.0	51.8	54.5	95.1	98.7	99.6
	1.0	99.4	99.3	98.3	59.1	46.1	94.3	98.9	99.5
2.5	0.5	100.0	99.7	99.2	78.4	80.3	98.9	99.9	99.8
	1.0	99.9	100.0	99.4	82.1	74.2	98.7	99.7	99.9
3.0	0.5	100.0	99.9	99.8	88.5	89.8	99.5	99.7	100.0
	1.0	100.0	100.0	99.5	92.6	85.7	99.5	99.9	100.0
3.5	0.5	100.0	100.0	99.9	92.4	94.2	99.6	99.9	100.0
	1.0	100.0	100.0	99.8	93.6	92.1	99.8	100.0	100.0
4.0	0.5	100.0	100.0	100.0	96.2	97.1	99.8	100.0	100.0
	1.0	100.0	99.9	99.9	97.2	96.6	99.8	99.9	100.0
4.5	0.5	100.0	100.0	100.0	98.0	98.7	100.0	99.8	99.9
	1.0	100.0	100.0	99.9	98.8	98.0	99.8	99.9	100.0
		$n = 200$							
1.5	0.5	99.9	98.0	91.5	8.9	5.6	77.9	94.6	98.3
	1.0	99.0	98.7	91.9	5.1	8.9	80.7	95.5	98.3
2.0	0.5	100.0	99.9	99.8	59.6	52.7	99.1	99.7	99.9
	1.0	100.0	100.0	99.7	50.0	62.9	99.2	99.9	100.0
2.5	0.5	100.0	100.0	99.9	87.0	86.3	99.7	100.0	100.0
	1.0	100.0	100.0	99.9	83.7	90.0	99.8	100.0	100.0
3.0	0.5	100.0	100.0	100.0	96.7	97.2	100.0	100.0	100.0
	1.0	100.0	100.0	100.0	95.2	97.8	100.0	100.0	100.0
3.5	0.5	100.0	100.0	100.0	98.9	99.2	100.0	100.0	100.0
	1.0	100.0	100.0	100.0	98.2	99.2	100.0	100.0	100.0
4.0	0.5	100.0	100.0	100.0	99.0	99.7	100.0	100.0	100.0
	1.0	100.0	100.0	100.0	98.8	99.3	100.0	100.0	100.0
4.5	0.5	100.0	100.0	100.0	100.0	99.7	100.0	100.0	100.0
	1.0	100.0	100.0	99.9	99.6	99.9	100.0	100.0	100.0
5.0	0.5	100.0	100.0	100.0	99.9	99.7	100.0	100.0	100.0
	1.0	100.0	100.0	100.0	99.8	99.7	100.0	100.0	100.0
5.5	0.5	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0
	1.0	100.0	100.0	100.0	99.6	100.0	100.0	100.0	100.0
6.0	0.5	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0
	1.0	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0

Setup 1 corresponds to $\mathrm{cov}(Z_{Wi}u_{Wi}) = h/\sqrt{n}$. The corresponding correction factor for the three setups is calculated as in (10) for the different values of κ. Each result corresponds to 1,000 heteroskedastic simulations and 1,000 resampling iterations.

Table 3. Power of the FAR test at $\theta_0 = 0$, covariance setup 2

		θ							
		-2	-1.5	-1	-0.5	0.5	1	1.5	2
κ	D	$n = 100$							
1.5	0.1	97.2	95.5	84.0	10.9	20.3	73.1	89.9	92.2
	0.25	98.5	96.2	82.3	2.3	28.7	77.4	91.9	92.6
2.0	0.1	99.8	99.5	97.8	43.7	62.0	95.2	98.7	99.5
	0.25	99.8	99.4	97.1	21.7	72.6	95.9	98.8	99.1
2.5	0.1	100.0	99.7	99.1	70.8	84.7	99.0	99.9	99.9
	0.25	99.9	100.0	99.3	46.3	93.7	99.3	99.8	99.8
3.0	0.1	100.0	99.9	99.8	83.9	91.6	99.7	99.8	99.9
	0.25	100.0	99.9	99.6	63.9	95.7	99.2	99.9	100.0
3.5	0.1	100.0	100.0	99.9	89.2	95.8	99.7	99.9	100.0
	0.25	100.0	100.0	99.8	72.7	97.4	99.8	99.9	100.0
4.0	0.1	100.0	100.0	100.0	94.6	98.0	99.8	100.0	100.0
	0.25	100.0	100.0	99.9	82.6	99.4	99.9	99.9	99.9
4.5	0.1	100.0	100.0	100.0	96.7	99.1	100.0	99.9	99.9
	0.25	100.0	100.0	100.0	88.1	99.6	99.9	100.0	100.0
		$n = 200$							
1.5	0.1	99.9	98.1	91.5	5.2	10.7	79.5	94.7	98.5
	0.25	99.4	98.9	90.2	0.4	24.8	84.8	95.4	98.4
2.0	0.1	100.0	99.9	99.8	45.7	65.9	99.2	99.6	99.9
	0.25	100.0	100.0	99.6	13.9	85.1	99.6	99.9	100.0
2.5	0.1	100.0	100.0	99.9	77.4	91.8	99.8	100.0	100.0
	0.25	100.0	100.0	99.9	45.0	96.3	99.8	100.0	100.0
3.0	0.1	100.0	100.0	100.0	92.4	98.6	100.0	100.0	100.0
	0.25	100.0	100.0	100.0	65.8	99.5	100.0	100.0	100.0
3.5	0.1	100.0	100.0	100.0	97.0	99.5	100.0	100.0	100.0
	0.25	100.0	100.0	100.0	82.2	99.9	100.0	100.0	100.0
4.0	0.1	100.0	100.0	100.0	98.0	100.0	100.0	100.0	100.0
	0.25	100.0	100.0	100.0	88.6	99.9	100.0	100.0	100.0
4.5	0.1	100.0	100.0	100.0	99.6	99.8	100.0	100.0	100.0
	0.25	100.0	100.0	99.9	94.1	99.9	100.0	100.0	100.0
5.0	0.1	100.0	100.0	100.0	99.8	99.9	100.0	100.0	100.0
	0.25	100.0	100.0	100.0	95.8	99.9	100.0	100.0	100.0
5.5	0.1	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0
	0.25	100.0	100.0	100.0	96.3	100.0	100.0	100.0	100.0
6.0	0.1	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0
	0.25	100.0	100.0	100.0	97.8	99.9	100.0	100.0	100.0

Setup 2 corresponds to $\mathrm{cov}(Z_{Wi}u_{Wi}) = D$. The corresponding correction factor for the three setups is calculated as in (10) for the different values of κ. Each result corresponds to 1,000 heteroskedastic simulations and 1,000 resampling iterations.

Table 4. Power of the FAR test at $\theta_0 = 0$, covariance setup 3

		θ							
		-2	-1.5	-1	-0.5	0.5	1	1.5	2
κ	a	$n = 100$							
1.5	.25	97.3	95.3	83.9	9.8	21.3	73.3	89.6	92.0
	.5	98.4	96.1	82.3	2.7	27.1	77.5	91.8	92.8
2.0	.25	99.9	99.5	98.0	41.8	64.1	95.4	98.7	99.5
	.5	99.7	99.4	97.2	24.2	71.5	95.9	98.7	99.2
2.5	.25	100.0	99.7	99.2	67.8	85.6	99.0	99.9	99.9
	.5	99.9	100.0	99.3	49.4	92.8	99.3	99.8	99.8
3.0	.25	100.0	99.9	99.7	81.1	92.3	99.7	99.8	99.9
	.5	100.0	99.9	99.6	66.3	95.6	99.2	99.9	100.0
3.5	.25	100.0	100.0	99.9	87.8	96.0	99.7	99.9	100.0
	.5	100.0	100.0	99.8	75.5	97.1	99.8	99.9	100.0
4.0	.25	100.0	100.0	100.0	94.0	98.4	99.8	100.0	100.0
	.5	100.0	100.0	99.9	84.9	99.2	99.9	99.9	100.0
4.5	.25	100.0	100.0	100.0	96.0	99.3	100.0	99.9	99.9
	.5	100.0	100.0	100.0	89.3	99.6	99.9	100.0	100.0
		$n = 200$							
1.5	.25	99.9	98.2	91.4	5.2	11.1	79.5	94.7	98.5
	.5	99.3	98.9	91.0	0.6	20.5	83.5	95.6	98.5
2.0	.25	100.0	99.9	99.8	44.9	66.2	99.2	99.6	99.9
	.5	100.0	100.0	99.6	20.0	81.5	99.3	99.9	100.0
2.5	.25	100.0	100.0	99.9	76.9	91.8	99.8	100.0	100.0
	.5	100.0	100.0	99.9	55.3	95.8	99.8	100.0	100.0
3.0	.25	100.0	100.0	100.0	92.3	98.6	100.0	100.0	100.0
	.5	100.0	100.0	100.0	76.3	99.4	100.0	100.0	100.0
3.5	.25	100.0	100.0	100.0	96.8	99.5	100.0	100.0	100.0
	.5	100.0	100.0	100.0	88.6	99.9	100.0	100.0	100.0
4.0	.25	100.0	100.0	100.0	98.0	100.0	100.0	100.0	100.0
	.5	100.0	100.0	100.0	92.5	99.8	100.0	100.0	100.0
4.5	.25	100.0	100.0	100.0	99.6	99.8	100.0	100.0	100.0
	.5	100.0	100.0	99.9	97.3	99.8	100.0	100.0	100.0
5.0	.25	100.0	100.0	100.0	99.7	99.9	100.0	100.0	100.0
	.5	100.0	100.0	100.0	98.1	99.9	100.0	100.0	100.0
5.5	.25	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0
	.5	100.0	100.0	100.0	98.2	100.0	100.0	100.0	100.0
6.0	.25	100.0	100.0	100.0	99.8	100.0	100.0	100.0	100.0
	.5	100.0	100.0	100.0	99.0	99.9	100.0	100.0	100.0

Setup 3 corresponds to $\text{cov}(Z_{Wi}u_{Wi}) = an^{1/3}/n^{1/2}$. The corresponding correction factor for the three setups is calculated as in (10) for the different values of κ. Each result corresponds to 1,000 heteroskedastic simulations and 1,000 resampling iterations.

6 Conclusion

We have shown how the FAR test can be used to draw valid inferences when the instruments do not perfectly satisfy the exclusion condition. Our simulations for $n = 100$ exhibit good size and power combinations when we select approximately 20%–25% of the total sample for the resampling block sizes. This corresponds to $\kappa_n = 3/\sqrt{n}$ in (6). κ values that generate smaller block sizes are not recommended. By taking advantage of the speed of the Mata language, the `far` test can be easily performed in Stata, allowing researchers to overcome the small-sample problems of the AR test in a fast and user-friendly manner.

7 Acknowledgment

We are grateful to our anonymous reviewer for detailed and constructive comments.

8 References

Acemoglu, D., and S. Johnson. 2005. Unbundling institutions. *Journal of Political Economy* 113: 949–995.

Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91: 1369–1401.

———. 2008. Reply to the revised (2008) version of David Albouy's "The colonial origins of comparative development: An investigation of the settler mortality data". http://baselinescenario.files.wordpress.com/2009/09/reply-to-albouy-oct-2008-final-sept-2009.pdf.

———. 2011. Hither thou shalt come, but no further: Reply to "The colonial origins of comparative development: An empirical investigation: Comment". Working Paper 16966, National Bureau of Economic Research.

Albouy, D. Y. 2008. The colonial origins of comparative development: An investigation of the settler mortality data. Working Paper 14130, National Bureau of Economic Research.

Anderson, T. W., and H. Rubin. 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 20: 46–63.

Angrist, J. D. 1990. Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. *American Economic Review* 80: 313–336.

Berkowitz, D., M. Caner, and Y. Fang. 2008. Are "nearly exogenous instruments" reliable? *Economics Letters* 101: 20–23.

———. 2012. The validity of instruments revisited. *Journal of Econometrics* 166: 255–266.

Caner, M. 2010. Near exogeneity and weak identification in generalized empirical likelihood estimators: Many moment asymptotics. Working paper, Department of Economics, North Carolina State University.

Card, D. E. 1995. Using geographic variation in college proximity to estimate the return to schooling. In *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, ed. L. N. Christofides, E. K. Grant, and R. Swindinsky, 201–222. Toronto, Canada: University of Toronto Press.

Glaeser, E. L., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2004. Do institutions cause growth? *Journal of Economic Growth* 9: 271–303.

Guggenberger, P. 2012. On the asymptotic size distortion of tests when instruments locally violate the exogeneity assumption. *Econometric Theory* 28: 387–421.

Moreira, M. J., and B. P. Poi. 2003. Implementing tests with correct size in the simultaneous equations model. *Stata Journal* 3: 57–70.

Stock, J. H., and J. H. Wright. 2000. GMM with weak identification. *Econometrica* 68: 1055–1096.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Wu, C. F. J. 1990. On the asymptotic properties of the jackknife histogram. *Annals of Statistics* 18: 1438–1452.

About the authors

Andrés Riquelme is a doctoral student in economics at North Carolina State University. His research interests include econometrics, labor economics, and nonmarket valuation.

Daniel Berkowitz is a professor of economics at the University of Pittsburgh. His research interests include comparative institutions, applied microeconomics, law and economics, and postsocialist transition.

Mehmet Caner is a professor of econometrics at North Carolina State University. His research interests include econometrics and international finance.