On Bootstrap Validity for the Test of Overidentifying Restrictions with Many Instruments and Heteroskedasticity

Wang, Wenjie

21 December 2020

On Bootstrap Validity for the Test of Overidentifying Restrictions with Many Instruments and Heteroskedasticity

Wenjie Wang *

December 21, 2020

Abstract

This note studies the asymptotic validity of bootstrapping the test of over-identifying restrictions under many/many weak instruments and heteroskedasticity. We show that the wild bootstrap consistently estimates the null limiting distributions of a jackknife overidentification statistic under this asymptotic framework. In particular, such bootstrap validity holds even when the bootstrap procedure fails to mimic well the distribution of the jackknife instrumental variable estimator, an important component of the statistic of interest. Monte Carlo simulations show that the wild bootstrap provides a more reliable method than that based on asymptotic critical values to approximate the null distributions of interests under many/many weak instruments and heteroskedasticity.

JEL classification: C12, C15, C26.

Keywords: Bootstrap, Overidentification Tests, Many Instruments, Weak Instruments, Heteroskedasticity.

*Division of Economics, School of Social Sciences, Nanyang Technological University, HSS-04-65, 14 Nanyang Drive, 637332, Singapore. E-mail address: wang.wj@ntu.edu.sg.

1 Introduction

Empirical applications of instrumental variables (IV) estimation often produce imprecise results. It is now well understood that standard first-order asymptotic theory breaks down when the instruments are weakly correlated with the endogenous regressors, and commonly used IV estimators can lose consistency; cf., Dufour (1997) and Staiger and Stock (1997) among others. However, as has been demonstrated by Chao and Swanson (2005), having many instruments in such a weakly identified situation can help to improve estimation accuracy. Indeed, using a large number of instruments can enhance the growth of the so-called concentration parameter even if each individual instrument is only weakly correlated with the endogenous explanatory variables. For empirical example, Hansen, Hausman, and Newey (2008) reveal in an application from Angrist and Krueger (1991) that using 180 instruments, rather than 3, substantially improves estimator accuracy.

For implementing inferences in the context of many/many weak instruments, Hansen et al. (2008) provide Corrected Standard Errors (CSE) under conditional homoskedasticity, and Chao, Swanson, Hausman, Newey, and Woutersen (2012) and Hausman, Newey, Woutersen, Chao, and Swanson (2012) further extend the CSE to the general heteroskedastic case. For specification tests, Anatolyev and Gospodinov (2011) propose modifications of the J test of overidentifying restrictions so that the test can be robust to many instruments under conditional homoskedasticity. Moreover, Chao, Hausman, Newey, Swanson, and Woutersen (2014) give a jackknife version of the overidentification test, which is asymptotically valid even under heteroskedasticity. For a comprehensive review of the related literature, see Anatolyev (2019) and the references therein. In addition, the literature on bootstrap for the IV model includes Davidson and MacKinnon (2008, 2010, 2014), Moreira, Porter, and Suarez (2009), Wang and Kaffo (2016), Kaffo and Wang (2017), Wang and Tchatoka (2018), Finlay and Magnusson (2019), among others.

In this note, we study the wild bootstrap as an alternative inference method for the overidentification test and show its asymptotic validity under many/many weak instruments and heteroskedasticity. In particular, we show that the validity of the wild bootstrap procedure holds even in the case when the bootstrap fails to mimic well the distribution of the heteroskedastic-robust version of the Fuller (1977) estimator proposed by Hausman et al. (2012), an important component of the jackknife test statistic. Simulations show that the bootstrap test has finite-sample size properties superior to the test based on asymptotic critical values provided in Chao et al. (2014), which can be rather conservative in many settings.

2 Model, assumptions and test statistics

Following Chao et al. (2012), Hausman et al. (2012) and Chao et al. (2014), we consider a standard linear IV regression model given by

$$\begin{aligned} y &= X\delta + \epsilon, && (1) \\ X &= \Gamma + U, && (2) \end{aligned}$$

where y and X are, respectively, an $n \times 1$ vector and an $n \times G$ matrix of observations on the endogenous variables, Γ is the $n \times G$ reduced form matrix, and ϵ and V are, respectively, an $n \times 1$ vector and an $n \times G$ matrix of random disturbances. The estimation of δ will be based on an $n \times K$ matrix, Z, of instrumental variable observations with $rank(Z) = K$, and we treat Z as deterministic. Also denote $P = Z(Z'Z)^{-1}Z'$ and $M = I_n - P$, where I_n is an identity matrix with dimension n. Throughout this paper, we consider the case where G, the dimension of δ, is small relative to n, but we let $K \to \infty$ as $n \to \infty$ to model the effect of having many/many weak instruments.

For the IV estimator of δ, we consider the heteroskedasticity-robust version of the Fuller (1977) estimator proposed by Hausman et al. (2012), referred to as HFUL. To define HFUL, let P_{ij} denote the ij-th element of P, and $\bar{X} = [y, X]$. Let $\tilde{\alpha}$ be the smallest eigenvalue of $\left(\bar{X}'\bar{X}\right)^{-1}\left(\bar{X}'P\bar{X} - \sum_{i=1}^{n} P_{ii}\bar{X}_i\bar{X}_i'\right)$. Let $\hat{\alpha} = [\tilde{\alpha} - (1-\tilde{\alpha})/T]/[1 - (1-\tilde{\alpha})/T]$. HFUL is given by

$$\hat{\delta} = \left(X'PX - \sum_{i=1}^{n} P_{ii}X_iX_i' - \hat{\alpha}X'X\right)^{-1}\left(X'Py - \sum_{i=1}^{n} P_{ii}X_iy_i - \hat{\alpha}X'y\right). \quad (3)$$

To define the jackknife overidentification statistic of Chao et al. (2014), let $\hat{\epsilon}_i = y_i - X_i'\hat{\delta}$, $\hat{\epsilon} = (\hat{\epsilon}_1, ..., \hat{\epsilon}_n)'$, $\hat{\epsilon}(2) = (\hat{\epsilon}_1^2, ..., \hat{\epsilon}_n^2)'$, and $P(2)$ be the n-dimensional square matrix with ij-th component equal to P_{ij}^2. The jackknife test statistic is

$$\begin{aligned} \hat{T} &= \frac{\hat{\epsilon}'P\hat{\epsilon} - \sum_{i=1}^{n} P_{ii}\hat{\epsilon}_i^2}{\sqrt{\hat{V}}} + K = \frac{\sum_{i\neq j}\hat{\epsilon}_iP_{ij}\hat{\epsilon}_j}{\sqrt{\hat{V}}} + K, \\ \hat{V} &= \frac{\hat{\epsilon}(2)'P(2)\hat{\epsilon}(2) - \sum_{i=1}^{n} P_{ii}^2\hat{\epsilon}_i^4}{K} = \frac{\sum_{i\neq j}\hat{\epsilon}_i^2P_{ij}^2\hat{\epsilon}_j^2}{K}, \quad (4) \end{aligned}$$

where $\sum_{i\neq j}$ denotes the double sum over all i not equal to j. Chao et al. (2014) show that the jackknife overidentification test with critical region

$$\hat{T} \geq q_{K-G}(1-\alpha), \quad (5)$$

where $q_r(\tau)$ denotes the τ-th quantile of the chi-squared distribution with r degrees of freedom, has asymptotic rejection probability equal to α under many/many weak instruments and heteroskedasticity. For the bootstrap test, we also consider an unstudentized version of the test statistic, which can be defined as

$$\hat{T}_u = \sum_{i \neq j} \hat{\epsilon}_i P_{ij} \hat{\epsilon}_j. \tag{6}$$

The model and data are assumed to satisfy the following regularity conditions. Let Z_i', ϵ_i, U_i' and Γ_i' denote the i-th row of Z, ϵ, U, and Γ, respectively.

Assumption 1

(i) Z includes among its columns a vector of ones, $rank(Z) = K$, and there is a constant C such that $P_{ii} \leq C < 1$ (i=1, ..., n), $K \to \infty$.

(ii) $\Gamma_i = S_i z_i/\sqrt{n}$ where $S_n = \tilde{S} diag(\mu_{1n}, ..., \mu_{Gn})$ and $\tilde{S}$ is nonsingular. Also, for each j either $\mu_{jn} = \sqrt{n}$ or $\mu_{jn}/\sqrt{n} \to 0$, $\mu_n = \min_{1 \leq j \leq G} \mu_{jn} \to \infty$, and $\sqrt{K}/\mu_n^2 \to 0$. Also, there is $C > 0$ such that $\left\| \sum_{i=1}^n z_i z_i'/n \right\| \leq C$ and $\lambda_{min} \left(\sum_{i=1}^n z_i z_i'/n \right) \geq 1/C$, for n sufficiently large.

(iii) There is a constant C such that $(\epsilon_1, U_1), ..., (\epsilon_n, U_n)$ are independent, with $E[\epsilon_i] = 0$, $E[U_i] = 0$, $E[\epsilon_i^2] < C$, $E[||U_i||^2] \leq C$, $Var((\epsilon_i, U_i')') = diag(\tilde{\Omega}_i, 0)$, and $\lambda_{min} \left(\sum_{i=1}^n \tilde{\Omega}_i/n \right) \geq 1/C$.

(iv) There is π_{K_n} such that $\sum_{i=1}^n \left\| z_i - \pi_{K_n} Z_i \right\|^2 /n \to 0$.

(v) There is a constant, $C > 0$, such that with probability one, $\sum_{i=1}^n ||z_i||^4/n^2 \to 0$, $E[\epsilon_i^4] \leq C$ and $E[||U_i||^4] \leq C$.

(vi) $\mu_n S_n^{-1} \to S_0$ and either (I) $K/\mu_n^2 \to \alpha$ for finite α or (II) $K/\mu_n^2 \to \infty$. Also, each of the following exists:

$$\begin{aligned} H_P &= \lim_{n \to \infty} \sum_{i=1}^n (1 - P_{ii}) z_i z_i'/n, \\ \Sigma_P &= \lim_{n \to \infty} \sum_{i=1}^n (1 - P_{ii})^2 z_i z_i' \sigma_i^2/n, \\ \Psi &= \lim_{n \to \infty} \sum_{i \neq j} P_{ij}^2 \left(\sigma_i^2 E[\tilde{U}_j \tilde{U}_j'] + E[\tilde{U}_i \epsilon_i] E[\epsilon_j \tilde{U}_j'] \right) /K, \end{aligned}$$

where $\sigma_i^2 = E[\epsilon_i^2]$, $\gamma_n = \sum_{i=1}^n E[U_i \epsilon_i] / \sum_{i=1}^n \sigma_i^2$, and $\tilde{U} = U - \epsilon \gamma_n'$ having i-th row $\tilde{U}_i'$.

Chao et al. (2014) show that the chi-square approximation in (5) is asymptotically correct under Assumption 1 as $\hat{T} \geq q_{K-G}(1-\alpha)$ if and only if

$$\frac{\sum_{i \neq j} \hat{\epsilon}_i P_{ij} \hat{\epsilon}_j}{\sqrt{2K \hat{V}_n}} \geq \frac{q_{K-G}(1-\alpha) - K}{\sqrt{2K}} = \sqrt{\frac{K-G}{K}} \left(\frac{q_{K-G}(1-\alpha) - (K-G)}{\sqrt{2(K-G)}} \right) - \frac{G}{\sqrt{2K}} \to q_{K-G}(1-\alpha). \tag{7}$$

3 Wild bootstrap for jackknife overidentificaiton test

In this section, we study the wild bootstrap as an alternative to the asymptotic critical values for the jackknife overidentification test, and we show the validity of the bootstrap test under many/many weak instruments and heteroskedasticity. Specifically, the bootstrap procedure is as follows:

Step 1: The bootstrap error term ϵ_i^* is obtained by

$$\epsilon_i^* = \hat{\epsilon}_i w_i^*,\ i = 1, ..., n, \tag{8}$$

where w_i^* is a Rademacher random variable with $P(w_i^* = 1) = P(w_i^* = -1) = 1/2$.

Step 2: The bootstrap analogue of y_i is obtained by

$$y_i^* = X_i'\hat{\delta} + \epsilon_i^*,\ i = 1, ..., n. \tag{9}$$

Step 3: Obtain $\hat{\epsilon}^* = y^* - X\hat{\delta}^*$, where $\hat{\delta}^*$, the bootstrap analogue of HFUL, is computed using (y^*, X, Z). Then, we construct the bootstrap analogue of the jackknife overidentification test statistic

$$\begin{aligned} \hat{T}^* &= \frac{\sum_{i \neq j} \hat{\epsilon}_i^* P_{ij} \hat{\epsilon}_j^*}{\sqrt{\hat{V}^*}} + K, \\ \hat{V}^* &= \frac{\sum_{i \neq j} \hat{\epsilon}_i^{*2} P_{ij}^2 \hat{\epsilon}_j^{*2}}{K}, \end{aligned} \tag{10}$$

and its unstudentized version

$$\hat{T}_u^* = \sum_{i \neq j} \hat{\epsilon}_i^* P_{ij} \hat{\epsilon}_j^*. \tag{11}$$

Step 4: Repeat Steps 1-4 B times, and compute the bootstrap P values $\hat{p}_T^* = B^{-1} \sum_{b=1}^{B} I\{\hat{T}_b^* \geq \hat{T}\}$ and $\hat{p}_{Tu}^* = B^{-1} \sum_{b=1}^{B} I\{\hat{T}_{u,b}^* \geq \hat{T}_u\}$. We reject the null hypothesis of no misspecification if the bootstrap P value is smaller than α.

The following result states the asymptotic validity for these bootstrap tests.

Theorem 3.1 *Suppose that Assumption 1 holds. Then,*

$$\sup_{x \in R} \left| P^* \left(\hat{T}^* \leq x \right) - P \left(\hat{T} \leq x \right) \right| \to^p 0,$$

and

$$\sup_{x \in R} \left| P^* \left(\hat{T}_u^* \leq x \right) - P \left(\hat{T}_u \leq x \right) \right| \to^p 0,$$

where P^* denotes the probability measure induced by the wild bootstrap procedure in (8)-(11).

We note that terms involving $\hat{\delta}^*$ appear in the formula of $\hat{T}^*$ and $\hat{T}_u^*$. Recently, Wang and Kaffo (2016) show that under many/many weak instruments and conditional homoskedasticity, a double-equation residual-based bootstrap procedure that re-samples both structural form equation (1) and first-stage equation (2) fails to consistently estimate the limiting distribution of the limited information maximum likelihood estimator (LIML) and Fuller (1977) estimator. Similarly, the single-equation wild bootstrap procedure in (8)-(11) cannot consistently estimate the limiting distribution of HFUL under the current many/many weak instrument sequence and heteroskedasticity. However, such bootstrap inconsistency related to $\hat{\delta}^*$ does not affect the validity of bootstrapping the jackknife overidentification tests. The reason is that under the current asymptotic framework, the terms involving $\hat{\delta}^*$ are all of relatively small order in the bootstrap world, and do not affect the first-order (conditional) limiting distributions of the test statistics.

We also emphasize that the many/many weak instrument asymptotics considered here, in which K and the overall instrument strength (e.g., measured by the concentration parameter) is required to tend to infinity, is very different from the weak instrument asymptotics, in which K is assumed to be fixed and the instruments are weak in the sense of Staiger and Stock (1997). It is well known that under this weak instrument asymptotics, the IV estimators, including HFUL, are inconsistent and have highly nonstandard limiting distributions. The overidentification tests considered here, no matter based on asymptotic or bootstrap critical values, will also be invalid under this framework. Indeed, Wang and Tchatoka (2018) and Wang (2020) show the inconsistency of residual-based and nonparametric bootstrap procedures for a subvector version of the Anderson-Rubin test, which has a formula closely related to the overidentification test, under the weak instrument asymptotics.

4 Monte Carlo simulations

We conduct some simulations using the model in (1)-(2). Following Davidson and MacKinnon (2010, 2014), we use the following data generating process:

$$\begin{aligned} y_i &= \beta X_i + n^{1/2}|w_i|\varepsilon_i, \\ X_i &= a w_i + U_i,\ U_i = \rho n^{1/2}|w_i|\epsilon_i + rV_i, \end{aligned}$$

where $r = (1-\rho^2)^{1/2}$, and ϵ_i and V_i are random variables with mean 0 and variance 1. The endogeneity parameter is set at $\rho \in \{0.1, 0.5, 0.9\}$. The instrument matrix Z is distributed $N(0, I_K)$, and w is an n-vector with $\|w\|^2 = 1$ and $w \in \mathcal{S}(Z)$, the subspace spanned by the columns of the instruments Z. As pointed out in their papers,the only property of Z that matters here is the subspace spanned by the columns of Z and in their setting, all the explanatory power comes from the vector w and the other columns of Z are simply noise. We consider $n = 100$ or 500, $K \in \{5, 15, 25, 35, 45\}$, and the overall instrument strength is captured by the parameter a, which is set at $a \in \{8, 12, 16\}$. In addition, $\delta = 1$, $B = 199$, $\alpha = 5\%$, and the number of Monte Carlo replications equals 5000.

The results for $n = 100$ and $n = 500$ are shown in Figures 1 and 2, respectively. Note that the overidentification tests based on the asymptotic critical values tend to under-reject, especially when the number of K is small (e.g., when $K = 5$). The unstudentized bootstrap tests improve upon the asymptotic tests with regard to conservativeness but can over-reject when $n = 100$ and the degree of endogeneity is small (e.g., when $\rho = 0.1$). The studentized bootstrap tests have the best overall performance with null rejection frequencies very close to the nominal level.

Figure 1: Null empirical rejection frequencies, $n = 100$

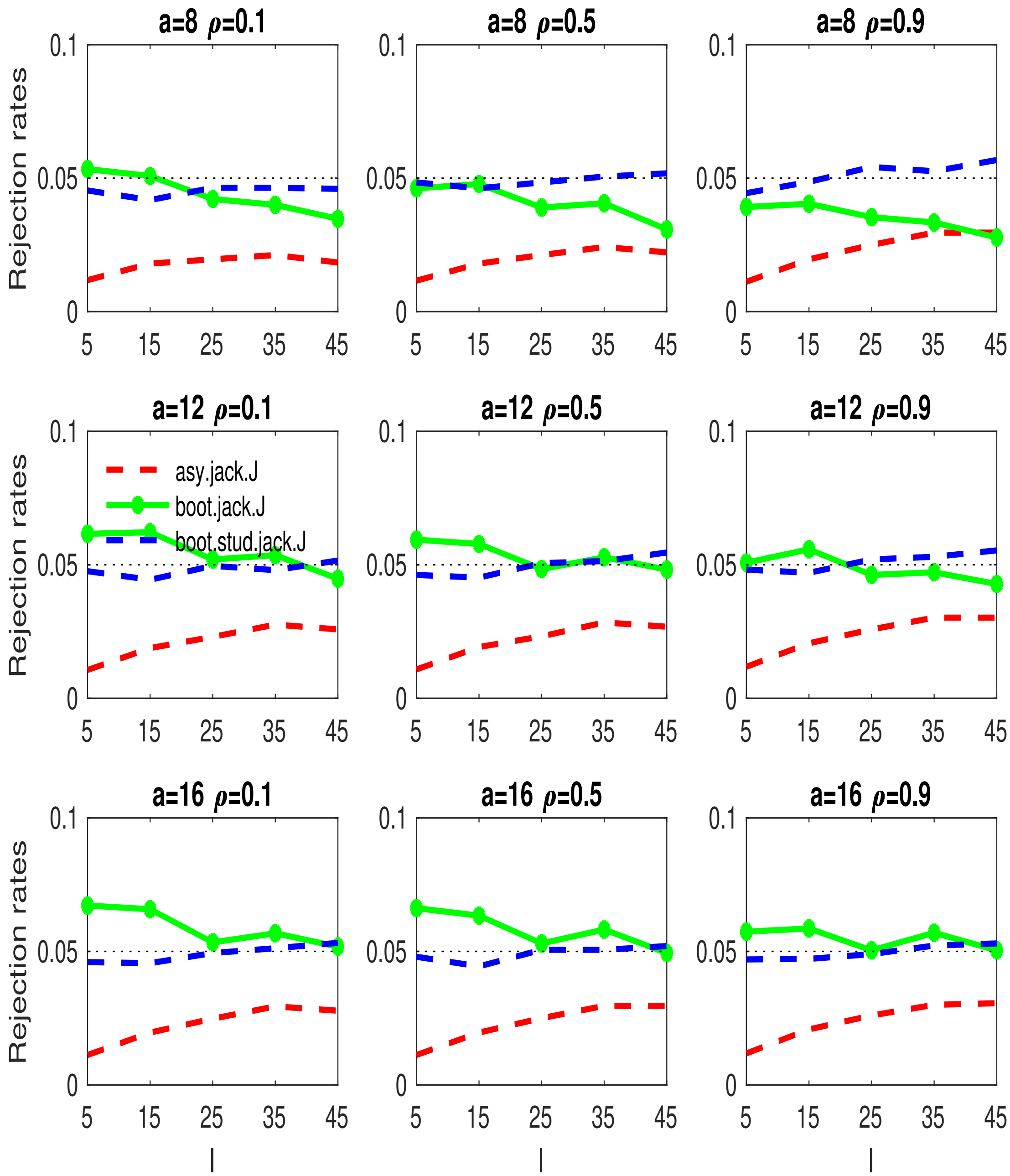
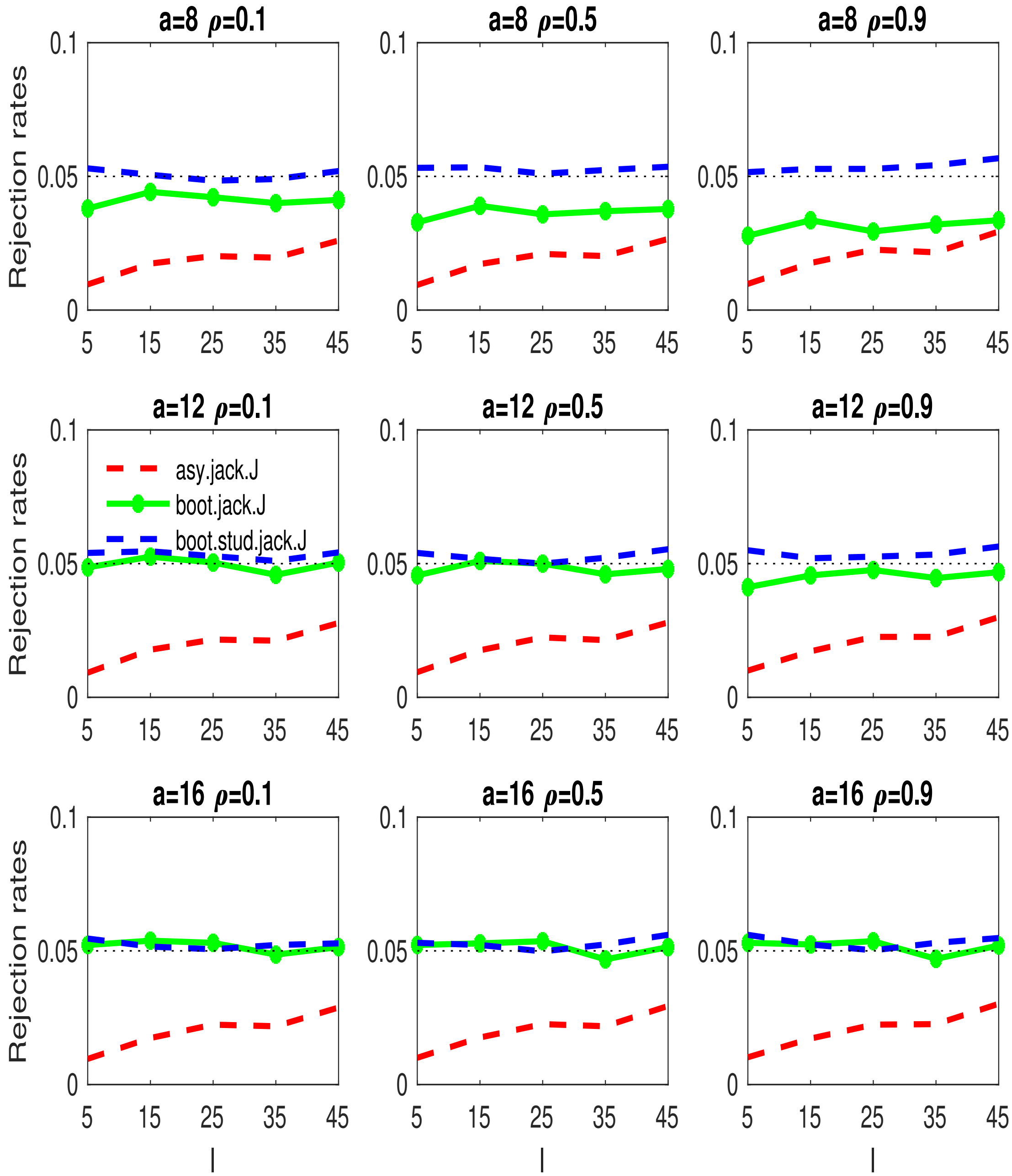


Figure 2: Null empirical rejection frequencies, $n = 500$



References

Anatolyev, S. (2019): "Many instruments and/or regressors: A friendly guide," *Journal of Economic Surveys*, 33(2), 689–726.

Anatolyev, S., and N. Gospodinov (2011): "Specification Testing In Models With Many Instruments," *Econometric Theory*, 27(02), 427–441.

Angrist, J. D., and A. B. Krueger (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 106(4), 979–1014.

Chang, Y., and J. Y. Park (2003): "A sieve bootstrap for the test of a unit root," *Journal of Time Series Analysis*, 24(4), 379–400.

Chao, J. C., J. A. Hausman, W. K. Newey, N. R. Swanson, and T. Woutersen (2014): "Testing overidentifying restrictions with many instruments and heteroskedasticity," *Journal of Econometrics*, 178, 15–21.

Chao, J. C., and N. R. Swanson (2005): "Consistent Estimation with a Large Number of Weak Instruments," *Econometrica*, 73(5), 1673–1692.

Chao, J. C., N. R. Swanson, J. A. Hausman, W. K. Newey, and T. Woutersen (2012): "Asymptotic Distribution Of Jive In A Heteroskedastic Iv Regression With Many Instruments," *Econometric Theory*, 28(01), 42–86.

Davidson, R., and J. G. MacKinnon (2008): "Bootstrap inference in a linear equation estimated by instrumental variables," *Econometrics Journal*, 11(3), 443–477.

——— (2010): "Wild Bootstrap Tests for IV Regression," *Journal of Business & Economic Statistics*, 28(1), 128–144.

——— (2014): "Bootstrap confidence sets with weak instruments," *Econometric Reviews*, 33(5-6), 651–675.

Dufour, J.-M. (1997): "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models," *Econometrica*, 65(6), 1365–1388.

Finlay, K., and L. M. Magnusson (2019): "Two applications of wild bootstrap methods to improve inference in cluster-IV models," *Journal of Applied Econometrics*, 34(6), 911–933.

Fuller, W. A. (1977): "Some properties of a modification of the limited information estimator," *Econometrica*, 45(4), 939–53.

Hansen, C., J. Hausman, and W. Newey (2008): "Estimation With Many Instrumental Variables," *Journal of Business & Economic Statistics*, 26, 398–422.

Hausman, J. A., W. K. Newey, T. Woutersen, J. C. Chao, and N. R. Swanson (2012): "Instrumental variable estimation with heteroskedasticity and many instruments," *Quantitative Economics*, 3(2), 211–255.

Kaffo, M., and W. Wang (2017): "On bootstrap validity for specification testing with many weak instruments," *Economics Letters*, 157, 107–111.

Moreira, M. J., J. R. Porter, and G. A. Suarez (2009): "Bootstrap validity for the score test when instruments may be weak," *Journal of Econometrics*, 149(1), 52–64.

Staiger, D., and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65(3), 557–586.

Wang, W. (2020): "On the inconsistency of nonparametric bootstraps for the subvector Anderson–Rubin test," *Economics Letters*, p. 109157.

Wang, W., and M. Kaffo (2016): "Bootstrap inference for instrumental variable models with many weak instruments," *Journal of Econometrics*, 192(1), 231–268.

Wang, W., and F. D. Tchatoka (2018): "On bootstrap inconsistency and Bonferroni-based size-correction for the subset Anderson–Rubin test under conditional homoskedasticity," *Journal of Econometrics*, 207(1), 188–211.

A Mathematical Proofs

The following notations are used for the bootstrap asymptotics (see Chang and Park (2003) for similar notation and for several useful bootstrap asymptotic properties): for any bootstrap statistic T^* we write $T^* \rightarrow^{P^*} 0$ in probability if for any $\delta > 0$, $\epsilon > 0$, $lim_{n\rightarrow\infty} P\left[P^*\left(|T^*| > \delta\right) > \epsilon\right] = 0$, i.e., $P^*\left(|T^*| > \delta\right) = o_P(1)$. Also, we write $T^* = O_{P^*}\left(n^{\varphi}\right)$ in probability if and only if for any $\delta > 0$ there exists a $M_\delta < \infty$ such that $lim_{n\rightarrow\infty} P\left[P^*\left(|n^{-\varphi}T^*| > M_\delta\right) > \delta\right] = 0$, i.e., for any $\delta > 0$ there exists a $M_\delta < \infty$ such that $P^*\left(|n^{-\varphi}T^*| > M_\delta\right) = o_P(1)$. Finally, we write $T^* \rightarrow^{d^*} T$ in probability if, conditional on the sample, T^* weakly converges to T under P^*, for all samples contained in a set with probability converging to one. Specifically, we write $T^* \rightarrow^{d^*} T$ in probability if and only if $E^*\left(f(T^*)\right) \rightarrow E\left(f(T)\right)$ in probability for any bounded and uniformly continuous function f. To be concise, we sometimes use the short version $T^* \rightarrow^{P^*} 0$ to say that $T^* \rightarrow^{P^*} 0$ in probability, and use $T^* = O_{P^*}\left(n^{\varphi}\right)$ for $T^* = O_{P^*}\left(n^{\varphi}\right)$ in probability.

Proof of Theorem 3.1. We focus on the proof for the studentized version of the bootstrap test. The proof of the unstudentized version is very similar, thus omitted. First, note that

$$
\begin{aligned}
&\frac{\sum_{i\neq j} \hat{\epsilon}_i^* P_{ij} \hat{\epsilon}_j^*}{\sqrt{K}} \\
= \; & \sum_{i\neq j} \left(\epsilon_i^* - X_i'(\hat{\delta}^* - \hat{\delta})\right)' P_{ij} \left(\epsilon_i^* - X_i'(\hat{\delta}^* - \hat{\delta})\right) / \sqrt{K} \\
= \; & \frac{\sum_{i\neq j} \epsilon_i^* P_{ij} \epsilon_j^*}{\sqrt{K}} + \left(S_n^{-1} \sum_{i\neq j} X_i P_{ij} X_j' S_n^{-1'}\right) S_n'(\hat{\delta}^* - \hat{\delta})/\sqrt{K} + 2(\hat{\delta}^* - \hat{\delta})' S_n \left(S_n^{-1} \sum_{i\neq j} X_i P_{ij} \epsilon_j^*\right) / \sqrt{K}.
\end{aligned}
\tag{A.12}
$$

Second, by using similar arguments as those in Wang and Kaffo (2016), we can show that for both Case (I) ($K/\mu_n^2 \rightarrow \alpha < \infty$) and Case (II) ($K/\mu_n^2 \rightarrow \infty$), we have $S_n'(\hat{\delta}^* - \hat{\delta}) = O_{P^*}(1)$. In particular, let $\tilde{\alpha}^*(\delta) = \sum_{i\neq j} \epsilon_i^*(\delta) P_{ij} \epsilon_j^*(\delta) / \epsilon^*(\delta)'\epsilon^*(\delta)$, where $\epsilon_i^*(\delta) = y_i^* - X_i'\delta$, and

$$
\begin{aligned}
\hat{D}^*(\delta) &= -\left[\frac{\epsilon^*(\delta)'\epsilon^*(\delta)}{2}\right] \frac{\partial}{\partial \delta} \left[\frac{\sum_{i\neq j} \epsilon_i^*(\delta) P_{ij} \epsilon_j^*(\delta)}{\epsilon^*(\delta)'\epsilon^*(\delta)}\right] \\
&= \sum_{i\neq j} X_i P_{ij} \epsilon_j^*(\delta) - \epsilon^*(\delta)'\epsilon^*(\delta) \tilde{\alpha}^*(\delta) \tilde{\gamma}^*(\delta),
\end{aligned}
\tag{A.13}
$$

where $\tilde{\gamma}^*(\delta) = X'\epsilon^*(\delta)/\epsilon^*(\delta)'\epsilon^*(\delta)$. Note that $S_n'(\hat{\delta}^* - \hat{\delta}) = \left(S_n^{-1}(\partial \hat{D}^*(\bar{\delta}^*)/\partial\delta) S_n^{-1'}\right)^{-1} S_n^{-1} \hat{D}^*(\hat{\delta})$, where $\bar{\delta}^*$ lies on the line joining $\hat{\delta}^*$ and $\hat{\delta}$. Also note that by Markov inequality and the current wild bootstrap procedure,

$$
\begin{aligned}
{\epsilon^*}'\epsilon^*/n &= \sum_{i=1}^n E^*[\epsilon_i^{*2}]/n + O_{P^*}\left(1/\sqrt{n}\right) = \sum_{i=1}^n \hat{\epsilon}_i^2/n + O_{P^*}\left(1/\sqrt{n}\right) = \sum_{i=1}^n \sigma_i^2/n + O_{P^*}\left(1/\sqrt{n}\right), \\
X'\epsilon^*/n &= \sum_{i=1}^n E^*[X_i \epsilon_i^*]/n + O_{P^*}(1/\sqrt{n}) = O_{P^*}\left(1/\sqrt{n}\right),
\end{aligned}
\tag{A.14}
$$

as $E^*[\epsilon_i^{*2}] = \hat{\epsilon}_i^2$ and $E^*[X_i\epsilon_i^*] = X_i E^*[\epsilon_i^*] = 0$. Also, let $\tilde{\alpha}^*$ and $\tilde{\gamma}^*$ denote $\tilde{\alpha}^*(\hat{\delta})$ and $\tilde{\gamma}^*(\hat{\delta})$, respectively. Then, we obtain that $\tilde{\alpha}^* = o_{P^*}(\mu_n^2/n)$ under the same arguments as in the proof for Lemma A5 of Hausman et al. (2012), and we have

$$
\begin{aligned}
S_n^{-1}\hat{D}^*(\hat{\delta}) &= S_n^{-1}\left(X'P\epsilon^* - \sum_{i=1}^n P_{ii}X_i\epsilon_i^* + \epsilon^{*'}\epsilon^*\tilde{\alpha}^*\tilde{\gamma}^*\right) \\
&= S_n^{-1}\left(X'P\epsilon^* - \sum_{i=1}^n P_{ii}X_i\epsilon_i^*\right) + o_{P^*}(1) \\
&= S_n^{-1}\sum_{i=1}^n \left(\hat{X}_i - P_{ii}X_i\right)\epsilon_i^* + o_{P^*}(1) \to^{d^*} N\left(0, H_P\right), \qquad (A.15)
\end{aligned}
$$

by $S_n^{-1}\left(\sum_{i=1}^n(\hat{X}_i - P_{ii}X_i)(\hat{X}_i - P_{ii}X_i)'\hat{\epsilon}_i^2\right)S_n^{-1'} \to^P H_P$, where $\hat{X}_i = \sum_{j=1}^n P_{ij}X_j$, and by following the arguments in the proof of Lemma A8, Theorem 2 and Theorem 3 of Hausman et al. (2012). Notice that the asymptotic variance-covariance matrix of $\hat{\delta}^*$ is different from that of $\hat{\delta}$, which has an extra term Ψ. There is no bootstrap analogue for Ψ because under the current wild bootstrap procedure, ϵ_i^* is independent from the first-stage equation in the bootstrap world. Similarly, by following the arguments in the proof of Lemma A7 of Hausman et al. (2012), we obtain

$$
-S_n^{-1}(\partial\hat{D}^*(\bar{\delta}^*)/\partial\delta)S_n^{-1'} = H_n + o_{P^*}(1). \qquad (A.16)
$$

Then, we have for both Case (I) and Case (II),

$$
\frac{\sum_{i\neq j}\hat{\epsilon}_i^* P_{ij}\hat{\epsilon}_j^*}{\sqrt{K}} = \frac{\sum_{i\neq j}\epsilon_i^* P_{ij}\epsilon_j^*}{\sqrt{K}} + o_{P^*}(1), \qquad (A.17)
$$

by using $S_n'(\hat{\delta}^* - \hat{\delta}) = O_{P^*}(1)$, $S_n^{-1}\sum_{i\neq j}X_iP_{ij}X_j'S_n^{-1'} = O_P(1) = O_{P^*}(1)$, and $S_n^{-1}\sum_{i\neq j}X_iP_{ij}\epsilon_j^* = O_{P^*}(1)$.

Next, note that $\sigma_i^{*2} \equiv E^*[\epsilon_i^{*2}] = \hat{\epsilon}_i^2 \geq C$ with probability approaching one (w.p.a.1) by Assumption 1(iii), and $P_{ii} \leq C < 1$ by Assumption 1(i) so that

$$
V_n^* = \frac{\sum_{i\neq j}\sigma_i^{*2}P_{ij}\sigma_j^{*2}}{K} > C\left(\frac{\sum_{i=1}^n\sum_{j=1}^n P_{ij}^2}{K} - \frac{\sum_{i=1}^n P_{ii}^2}{K}\right) = C\frac{\sum_{i=1}^n P_{ii}(1-P_{ii})}{K} > C(1-C) > 0. \qquad (A.18)
$$

In addition, we note that $E^*[\epsilon_i^{*4}] = \hat{\epsilon}_i^4$ is bounded in probability by Assumption 1(v), and $E^*\left[\sum_{i\neq j}(\epsilon_i^*P_{ij}\epsilon_j^*)^2\right] = KV_n^*$. It follows by Lemma A2 of Chao et al. (2012) that

$$
\frac{\sum_{i\neq j}\epsilon_i^*P_{ij}\epsilon_j^*}{\sqrt{KV_n^*}} \to^{d^*} N(0,1) \text{ in probability.} \qquad (A.19)
$$

Furthermore, by $\hat{\delta}^* \to^{P^*} \hat{\delta}$, we obtain that w.p.a.1, $\left\|\hat{\delta}^* - \hat{\delta}\right\|^2 \leq \left\|\hat{\delta}^* - \hat{\delta}\right\|$ and

$$
|\hat{\epsilon}_i^{*2} - \epsilon_i^{*2}| \leq 2||X_i||\left\|\hat{\delta}^* - \hat{\delta}\right\| + ||X_i||^2\left\|\hat{\delta}^* - \hat{\delta}\right\|^2 \leq d_i\left\|\hat{\delta}^* - \hat{\delta}\right\|, \qquad (A.20)
$$

for $d_i = 3(1 + ||X_i||^2)$. Also by $\sum_{i=1}^n \sum_{j=1}^n P_{ij}^2 = \sum_{i=1}^n P_{ii} = K$, we have $E\left[\sum_{i\neq j} P_{ij}^2 d_i d_j\right]/K \leq C\sum_{i\neq j} P_{ij}^2/K \leq C$, so that by Markov inequality $\sum_{i\neq j} P_{ij}^2 d_i d_j/K = O_P(1)$, which implies that $\sum_{i\neq j} P_{ij}^2 d_i d_j/K = O_{P^*}(1)$. Similarly, since $\hat{\epsilon}_i^2$ is bounded in probability, we have w.p.a.1,

$$E^*\left[\sum_{i\neq j} P_{ij}^2 \epsilon_i^{*2} d_j\right]/K \leq C\sum_{i\neq j} P_{ij}^2/K \leq C. \tag{A.21}$$

Therefore, for $\hat{V}_n^* = \sum_{i\neq j} P_{ij}^2 \hat{\epsilon}_i^{*2}\hat{\epsilon}_j^{*2}/K$ and $\tilde{V}_n^* = \sum_{i\neq j} P_{ij}^2 \epsilon_i^{*2}\epsilon_j^{*2}/K$ we have

$$\begin{aligned} \left|\hat{V}_n^* - \tilde{V}_n^*\right| &\leq \sum_{i\neq j} P_{ij}^2 \left|\hat{\epsilon}_i^{*2}\hat{\epsilon}_j^{*2} - \epsilon_i^{*2}\epsilon_j^{*2}\right|/K \\ &\leq \left\|\hat{\delta}^* - \hat{\delta}\right\|^2 \sum_{i\neq j} P_{ij}^2 d_i d_j/K + 2\left\|\hat{\delta}^* - \hat{\delta}\right\| \sum_{i\neq j} P_{ij}^2 \epsilon_i^{*2} d_j/K \to^{P^*} 0. \end{aligned} \tag{A.22}$$

In addition, note that by the choice of w_i^* for the wild bootstrap procedure, $V_n^* = \sum_{i\neq j} P_{ij}^2 \hat{\epsilon}_i^2 \hat{\epsilon}_j^2/K = \sum_{i\neq j} P_{ij}^2 \epsilon_i^{*2}\epsilon_j^{*2}/K = \tilde{V}_n^*$. Then by the triangular inequality we have $\hat{V}_n^* - V_n^* \to^{P^*} 0$.

Therefore by the Slutzky Theorem and (A.19),

$$\frac{\sum_{i\neq j} \hat{\epsilon}_i^* P_{ij} \hat{\epsilon}_j^*}{\sqrt{K\hat{V}_n^*}} = \frac{\sum_{i\neq j} \epsilon_i^* P_{ij} \epsilon_j^*}{\sqrt{K\hat{V}_n^*}} + \frac{o_{P^*}(1)}{\sqrt{\hat{V}_n^*}} = \sqrt{\frac{V_n^*}{\hat{V}_n^*}} \frac{\sum_{i\neq j} \epsilon_i^* P_{ij} \epsilon_j^*}{\sqrt{KV_n^*}} + o_{P^*}(1) \to^{d^*} N(0,1), \tag{A.23}$$

in probability. The result of bootstrap validity follows by Polya's Theorem. ■