3: Robust Inference with Weak Instruments

July 22, 2018

The Story So Far...

- Conventional t-test-based confidence intervals can under-cover true parameter value when instruments are weak
- Effective First-stage F-statistic provides a guide to bias
    - But screening applications on F-statistics can induce size distortions
- This section: identification-robust confidence sets
    - Ensure correct coverage regardless of instrument strength
    - No need to screen on first stage
        - Avoids pretesting bias
        - Avoids throwing away applications with valid instruments just because weak
        - Confidence sets can be informative even with weak instruments

Reminder: Normal Model

- To discuss these issues, continue to consider the normal model

$$\begin{pmatrix} \hat{\delta} \\ \hat{\pi} \end{pmatrix} \sim N\left(\begin{pmatrix} \delta \\ \pi \end{pmatrix}, \Sigma \right)$$

  where
  - $\hat{\delta}$ is the reduced-form OLS coefficient
  - $\hat{\pi}$ is the first-stage OLS coefficient
  - Σ is known
- IV model implies $\delta = \pi\beta$

Negative Result

- Initial Question: can we obtain correct coverage by adjusting our standard errors?
  - Confidence interval $\left[\hat{\beta} \pm b\left(\hat{\delta}, \hat{\pi}\right)\right]$ for some $b(\cdot, \cdot)$
  - Answer: no (unless $b\left(\hat{\delta}, \hat{\pi}\right)$ can be infinite)
- Gleser and Hwang (1989) and Dufour (1997) show that for any robust confidence set CS with coverage $1 - \alpha$,

$$Pr_{\beta,\pi}\left\{\beta \in CS\right\} \geq 1 - \alpha \text{ for all } \beta,\ \pi,$$

  we must have

$$Pr_{\beta,\pi}\left\{CS \text{ has infinite length}\right\} > 0 \text{ for all } \beta,\ \pi$$

  - Inutition: in case with $\pi = 0$, must cover every value β with probability $1 - \alpha$
- Adjusting our (finite) standard errors isn't enough: need alternative approach

Test Inversion

- Leading alternative: test inversion
- Idea: Define a family of tests $\phi(\cdot)$ where
  - $\phi(\beta_0)$ test for $H_0 : \beta = \beta_0$
  - $\phi(\beta_0) = 1$ if reject H_0, 0 otherwise
- Suppose $\phi(\beta_0)$ has size α for all β_0, i.e.

$$E_{\beta_0,\pi}[\phi(\beta_0)] \leq \alpha \text{ for all } \beta_0, \pi$$

- If we form CS by collecting the non-rejected values

$$CS = \{\beta : \phi(\beta) = 0\}$$

  then CS has coverage $1 - \alpha$
  - Called test inversion
- Hence, to form an identification-robust confidence set, we only need to form identification-robust tests of $H_0 : \beta = \beta_0$

Restriction Implied By IV Model

- To implement test inversion, need to find a test
- To construct robust test, use restrictions that hold regardless of instrument strength
    - IV model implies that $\delta - \pi\beta = 0$
- Under $H_0 : \beta = \beta_0$,

$$\hat{\delta} - \hat{\pi}\beta_0 \sim N\left(0, \Omega\left(\beta_0\right)\right)$$

for

$$\Omega(\beta_0) = \Sigma_{\delta\delta} - \beta_0(\Sigma_{\delta\pi} + \Sigma_{\pi\delta}) + \beta_0^2\Sigma_{\pi\pi}$$

- Holds regardless of instrument strength

AR Statistic

- Building on this observation, can introduce AR statistic

$$AR(\beta_0) = \left(\hat{\delta} - \hat{\pi}\beta_0\right)' \Omega(\beta_0)^{-1} \left(\hat{\delta} - \hat{\pi}\beta_0\right)$$

  - Originally introduced by Anderson and Rubin (1949) for homoskedastic normal case
  - Here, generalization to non-homoskedastic case
- Under $H_0 : \beta = \beta_0$, $AR(\beta_0) \sim \chi^2_k$ for all π
  - Recall $k = \dim(Z_i)$
- AR test $\phi_{AR}(\beta_0) = 1\left\{AR(\beta_0) > \chi^2_{k,1-\alpha}\right\}$
  - $\chi^2_{k,1-\alpha}$ the $1-\alpha$ quantile of a χ^2_k distribution
- AR Confidence set $CS_{AR} = \left\{\beta : AR(\beta) \le \chi^2_{k,1-\alpha}\right\}$
- AR test and CS fully robust to weak instruments

The Form of AR Confidence Sets

- CS_{AR} can behave in counterintuitive ways
- In just-identified setting ($k = \dim(X) = 1$) can take form of:
  - bounded interval: $CS_{AR} = [a, b]$
  - real line: $CS_{AR} = (-\infty, \infty)$
  - real line excluding bounded interval: $CS_{AR} = (-\infty, a] \cup [b, \infty)$
- In over-identified settings can also be empty, $CS_{AR} = \emptyset$
  - In overidentifed non-homoskedastic settings, can take additional forms

The Form of AR Confidence Sets

- Infinite confidence sets strange-looking...
  - but have natural explanation
- Unboundedness consistent with Gleser and Hwang (1989), Dufour (1997)
- Moreover, can show that

$$\lim_{\beta_0 \to \pm\infty} AR(\beta_0) = \hat{\pi}' \Sigma_{\pi\pi}^{-1} \hat{\pi} = k \cdot F^R$$

  - Implies that CS_{AR} unbounded if and only if first-stage F-test cannot reject $\pi = 0$ at level α
- Unbounded AR confidence sets arise only when cannot reject that model totally unidentified

The Form of AR Confidence Sets

- Empty confidence sets more awkward
- Arise from fact that AR tests $H_0 : \delta = \pi\beta_0$. Can be decomposed into
  - Parameteric restriction $\beta = \beta_0$
  - Overidentifying restrictions $\delta \propto \pi$ if $k > 1$
    - If $k = 1$, no overidentifying restrictions to test
- Empty AR confidence sets can be interpreted as a rejection of the overidentifying restrictions
  - Unfortunate feature: how to interpret a small CS?
  - Confidence sets non-empty with probability one in just-identified case

Optimality of AR in Just-Identified Models

- In just-identified case with single endogenous regressor, AR is optimal
  - 101 out of 230 specifications in our AER sample are just-identified with a single endogenous regressor
- Moreira (2009) shows that AR test uniformly most powerful unbiased
- AR equivalent to two-sided t-test when instruments are strong
- In just-identified settings, strong case for using AR CS
  - Optimal among CS robust to weak instruments
  - No loss of power relative to t-test if instruments strong

AR Tests in Applications

- To examine practical impact of using CS_{AR}, return to our AER sample
- Limit attention to just-identified specifications with single endogenous regressor where can estimate variance-covariance matrix of $\left(\hat{\delta}, \hat{\pi}\right)$
    - Yields 36 specifications
- Comparing 95% t and AR confidence sets, find infinite AR CS in two cases. In remaining cases:
    - AR confidence sets 56.5% longer on average in all specifications
    - 20.3% longer on average in specifications that report F>10
    - 0.04% longer on average in specifications that report F>50

AR Tests in Overidentified Models

- Strong argument for using AR in just-identified settings
- AR tests and CS perform worse in over-identified settings
  - As already noted, CS_{AR} may be empty
- Also inefficient under strong instruments
  - Tests violations of $H_0 : \beta = \beta_0$, and of overidentifying restrictions
  - If only care about parametric restrictions, "wastes" degrees of freedom

Improving Efficiency in Over-identified Settings

- To obtain efficiency under strong instruments, need alternative tests
- For example, tests based on t-statistic

$$t\left(\beta_0\right) = \frac{\left|\hat{\beta} - \beta_0\right|}{\hat{\sigma}_{\hat{\beta}}}$$

- Problem: distribution of $t\left(\beta_0\right)$ under $H_0 : \beta = \beta_0$ depends on π
  - Already know this: distribution of t-statistic depends on instrument strength
  - Since π unknown, not clear what critical values to use with $t\left(\beta_0\right)$

Conditional Critical Values

- Moreira (2003): conditional critical values
  - Originally for homoskedastic case. Here discuss generalization
- Idea: Find a sufficient statistic $D(\beta_0)$ for π under $H_0 : \beta = \beta_0$
  - Means conditional distribution of $t(\beta_0)|D(\beta_0)$ doesn't depend on π under H_0
  - Once condition on $D(\beta_0)$, can compute data-dependent critical values $c_\alpha(D(\beta_0))$
- Question: how to find $D(\beta_0)$

Conditional Critical Values

- Idea for sufficient statistic: separate parts of $\left(\hat{\delta}, \hat{\pi}\right)$ that do/don't depend on π
- Define
$$g\left(\beta\right) = \hat{\delta} - \hat{\pi}\beta$$
- Let
$$D(\beta) = \hat{\pi} - \left(\Sigma_{\pi\delta} - \Sigma_{\pi\pi}\beta\right)\Omega(\beta)^{-1}g(\beta),$$
denote $\hat{\pi}$ orthogonalized with respect to $g\left(\beta\right)$
- Under $H_0 : \beta = \beta_0$
$$\begin{pmatrix} g(\beta_0) \\ D(\beta_0) \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \pi \end{pmatrix}, \begin{pmatrix} \Omega\left(\beta_0\right) & 0 \\ 0 & \Psi(\beta_0) \end{pmatrix}\right)$$
- Conditional distribution of $g(\beta_0)$ given $D\left(\beta_0\right)$ doesn't depend on π under $H_0 : \beta = \beta_0$
$$g\left(\beta_0\right)|D\left(\beta_0\right) \sim N\left(0, \Omega\left(\beta_0\right)\right)$$
  - $\left(g(\beta_0), D\left(\beta_0\right)\right)$ one-to-one transformation of $\left(\hat{\delta}, \hat{\pi}\right)$
  - $\Rightarrow D\left(\beta_0\right)$ is sufficient statistic for π

Conditional Critical Values

- To construct conditional distribution of $t(\beta_0)|D(\beta_0)$:
  1. Fix $D(\beta_0)$ at observed value
  2. Repeatedly draw $g^*(\beta_0) \sim N(0, \Omega(\beta_0))$
  3. Construct $\left(\hat{\delta}^*, \hat{\pi}^*\right)$ from $(g^*(\beta_0), D(\beta_0))$
  4. Calculate $t^*(\beta_0)$ based on $\left(\hat{\delta}^*, \hat{\pi}^*\right)$
- Conditional critical value $c_\alpha(D(\beta_0))$: $1-\alpha$ quantile of $t^*(\beta_0)$

Conditional Critical Values

- Conditional t-test that rejects when $t(\beta_0)$ exceeds $c_\alpha(D(\beta_0))$

$$\phi(\beta_0) = 1\{t(\beta_0) > c_\alpha(D(\beta_0))\}$$

  is fully robust to weak instruments,

$$E_{\beta_0,\pi}[\phi(\beta_0)] = \alpha \text{ for all } \pi$$

- Conditioning not specific to $t(\beta_0)$, works for any statistic $s(\beta_0)$
  - In each case construct data-dependent critical value $c_\alpha(D(\beta_0))$
  - Yields tests that control size
- Question what statistic $s(\beta_0)$ to use

Alternative Test Statistics

Many possible choices of statistic $s(\beta_0)$

- t-statistic

$$t(\beta_0) = \frac{\left|\hat{\beta} - \beta_0\right|}{\hat{\sigma}_{\hat{\beta}}}$$

- Score statistic (Kleibergen 2002, 2005)

$$K(\beta_0) = g(\beta_0)' \Omega(\beta_0)^{-1} D(\beta_0) \times$$

$$\left(D(\beta_0)' \Omega(\beta_0)^{-1} D(\beta_0)\right)^{-1} D(\beta_0)' \Omega(\beta_0)^{-1} g(\beta_0)$$

- AR statistic

$$AR(\beta_0) = g(\beta_0)' \Omega(\beta_0)^{-1} g(\beta_0)$$

- LR statistic

$$LR(\beta_0) = 2\left(\max_{\beta,\pi} \ell(\beta,\pi) - \max_{\pi} \ell(\beta_0,\pi)\right)$$

Properties

- Different test statistics imply different $c_\alpha(D(\beta_0))$
  - For t and LR, need data-dependent critical values
  - Conditional distributions $AR(\beta_0)|D(\beta_0)$ and $K(\beta_0)|D(\beta_0)$ don't depend on $D(\beta_0)$
    - Can use χ^2_k and χ^2_1 critical values, respectively
- Conditional t, K, and LR tests efficient under strong instruments
  - AR test inefficient in overidentified models
- Under weak instruments, also yield different power properties
  - Conditional two-sided t-test poor power
  - K test sometimes poor power
  - Conditional LR (CLR) test performs well (in homoskedastic case)
- See D. Andrews Moreira and Stock (2006), (2007) Plots

Near-Optimality of CLR Test

- D. Andrews, Moreira, and Stock (2006) show that CLR test near-optimal
  - In homoskedastic case with single endogenous regressor
- Power close to upper bound for a natural class of tests over a wide range of parameter values
- Consensus in literature that CLR is a good test for homoskedastic settings
  - ... but homoskedasticity assumption unappealing

Tests for Non-Homoskedastic Models

- Variety of CLR extensions for non-homoskedastic case
  - D. Andrews Moreira and Stock (2004), Kleibergen (2005), D. Andrews and Guggenberger (2015), I. Andrews (2016), I. Andrews and Mikusheva (2016)
  - All efficient with strong instruments, but only simulation evidence on power with weak instruments
- Alternative: tests proposed that maximize weighted average power
  - Maximize integral of power function with respect to some weights
  - Moreira and Moreira (2015), Montiel Olea (2017), Moreira and Ridder (2018)
  - Question: what are "right" weights?
- Many options, but so far no consensus on what tests should be used in over-identified and non-homoskedastic models
  - In just-identified setting, use AR
  - In over-identified settings, use something that's efficient under strong instruments

Two-Step Confidence Sets

- Robust confidence sets not widely used in practice
  - When reported, usually only after authors find evidence of weak instruments
  - In AER sample, reported in 2 papers. Minimal first-stage F of 2.3 and 6.3, respectively
- If only report robust confidence set when F small, can view as constructing confidence set in two steps
  - If $F \geq 10$, report t-statistic confidence set
  - If $F < 10$, report robust CS
- Screening applications on first-stage F can generate very bad behavior. Does two-step CS do the same?
  - Positive results for F^N in homoskedastic case based on Stock and Yogo (2005)
  - Negative result for F^N in non-homoskedastic case based on Montiel Olea and Pflueger (2013), for F^R with conventional critical values based on I. Andrews (2018)
    - Negative results based on extreme forms of non-homoskedasticity: open question how bad in practice

Implementation

- Implementations of some weak-IV tests are in Stata package weakiv, available on SSC
  - Finlay, Magnusson, and Schaffer
  - Versions of CLR, AR, K, and other tests applicable to non-homoskedastic models
  - Can be used with fixed effects, clustered standard errors, etc.
  - Stata Journal article on previous version of package: Finlay and Magnusson (2009)

Summary

- A number of tests and confidence sets are available that are fully robust to weak instruments
  - Avoid pretesting bias, discarding applications
  - Many efficient under strong instruments
- In just-identified models, strong case for using AR CS
  - Covers many applications
- In over-identified models, less clear
  - CLR if assume homoskedastic
  - No consensus for non-homoskedastic case
  - ...other than using something efficient under strong instruments

References

- Anderson T, Rubin H. 1949. Estimators for the parameters of a single equation in a complete set of stochastic equations. *Annals of Mathematical Statistics* 21:570–582
- Andrews D, Guggenberger P. 2015. Identification- and singularity-robust inference for moment condition models. Unpublished Manuscript
- Andrews D, Moreira M, Stock J. 2004. Optimal invariant similar tests of instrumental variables regression. Unpublished Manuscript
- Andrews D, Moreira M, Stock J. 2006. Optimal two-sided invariant similar tests of instrumental variables regression. *Econometrica* 74:715–752
- Andrews D, Moreira M, Stock J. 2007. Performance of conditional wald tests in iv regression with weak instruments . *Econometrica* 74:715–752
- Andrews I. 2016. Conditional linear combination tests for weakly identified models. *Econometrica* 84:2155–2182

References

- Andrews I, Mikusheva A. 2016. Conditional inference with a functional nuisance parameter. *Econometrica* 84:1571–1612.
- Dufour J. 1997. Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica* 65:1365–1387
- Finlay K, Magnusson L. 2009. Implementing weak-instrument robust tests for a general class of instrumental-variables models. *Stata Journal* 9, 398-421
- Geser L, Hwang J. 1987. The nonexistence of 100(1-α)% confidence sets of finite expected diameter in errors-in-variables and related models. *Journal of the American Statistical Association* 15:1341–1362
- Kleibergen F. 2002. Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* 70:1781–1803
- Kleibergen F. 2005. Testing parameters in gmm without assuming they are identified. *Econometrica* 73:1103–1123

References

- Montiel Olea J. 2017. Admissible, similar tests: A characterization. Unpublished Manuscript
- Montiel Olea J, Pflueger C. 2013. A robust test for weak instruments. *Journal of Business and Economic Statistics* 31:358–369
- Moreira H, Moreira M. 2015. Optimal two-sided tests for instrumental variables regression with heteroskedastic and autocorrelated errors. Unpublished Manuscript
- Moreira M. 2003. A conditional likelihood ratio test for structural models. *Econometrica* 71:1027–1048
- Moreira M. 2009. Tests with correct size when instruments can be arbitrarily weak. *Journal of Econometrics* 152:131–140
- Moreira M, Ridder G. 2017. Optimal invariant tests in an instrumental variables regression with heteroskedastic and autocorrelated errors. Unpublished Manuscript

Power Comparisons

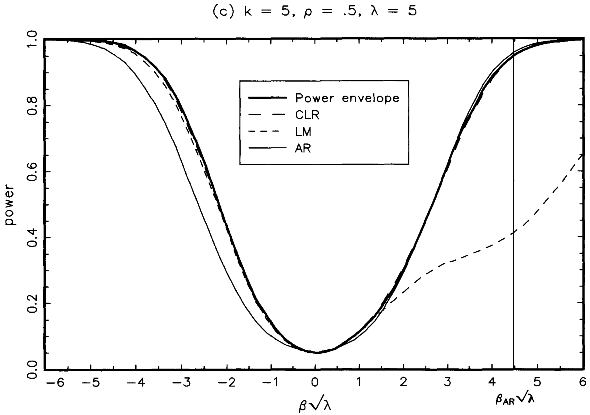


Figure: Power of AR, K, and C LR tests in homoskedastic case (from D. Andrews, Moreira, and Stock (2006))

Power Comparisons

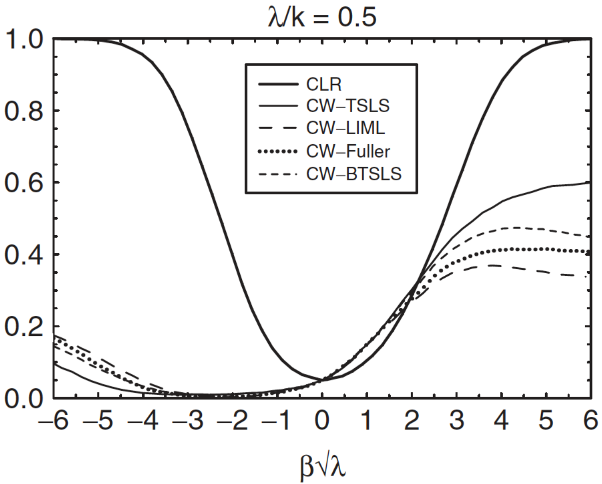


Figure: Power of Conditional t-and LR-tests in homokedastic (from D. Andrews, Moreira, and Stock (2007)) ▸ Return

4: Open Issues and Recent Research

July 22, 2018

Outline

Two goals for this section

1. Examine practical importance of issues covered so far
   - Simulations calibrated to specifications published in AER
2. Discuss other open questions and recent research on weak instruments

AER Simulation Specifications

- To assess practical importance of weak instrument issues, calibrate simulations to AER data
- Specifications from AER articles (excluding Papers and Proceedings) from 2014-2018 that:
  1. Published in main text
  2. Allow us to estimate variance matrix Σ of $\left(\hat{\delta}, \hat{\pi}\right)$
     - Mostly papers with replication data
     - In one other case, back out Σ from published results
- Yields 124 specifications from 8 papers
  - All specifications have a single endogenous variable

Simulation Design

- To focus attention on weak instrument issues, simulations use normal model

$$\begin{pmatrix} \hat{\delta} \\ \hat{\pi} \end{pmatrix} \sim N\left(\begin{pmatrix} \delta \\ \pi \end{pmatrix}, \Sigma \right)$$

  with Σ known, $\delta = \pi\beta$
- Simulations fix β, π, and Σ at estimated values
- Abstracts away from:
  - Non-normality of $\hat{\delta}$, $\hat{\pi}$
  - Estimation error in Σ
  - Will return to these later
- Any disortions must arise from weak instruments

Distribution of t-Statistics

- Theoretical results show t-tests can perform poorly when instruments weak
  - Distribution of t-statistics may not be centered at zero
  - Rejection probability of 5% t-tests may be much larger
- In each of our 124 AER specifications simulate
  - Median t-statistic

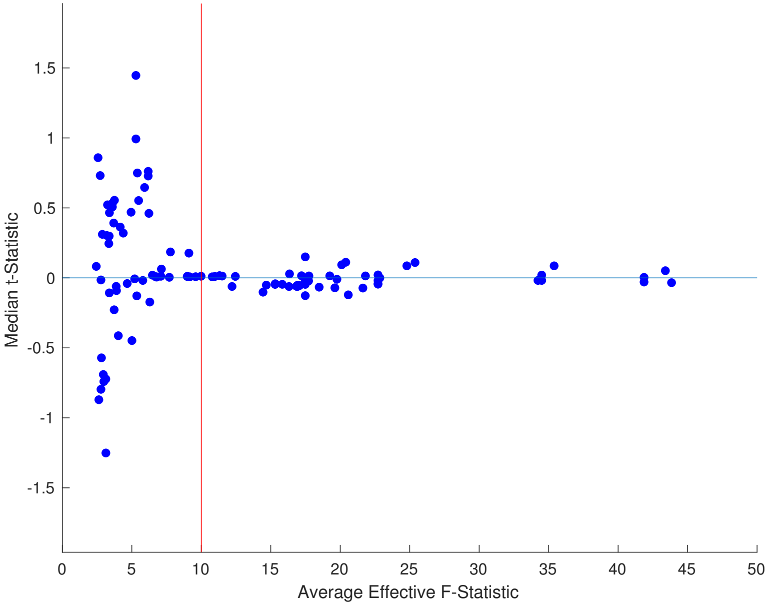
$$Med\left(\frac{\hat{\beta}-\beta_0}{\hat{\sigma}_{\hat{\beta}}}\right)$$

  - Size of 5% t-tests

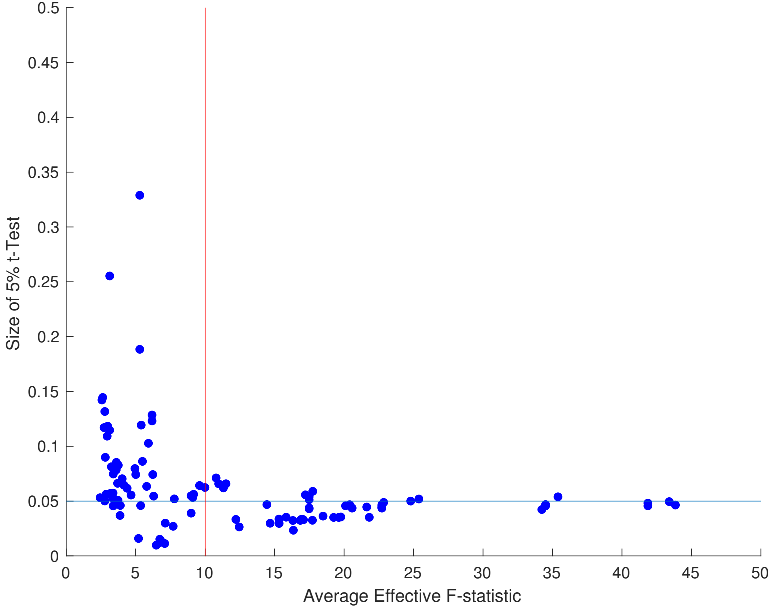
$$Pr\left\{\frac{\left|\hat{\beta}-\beta_0\right|}{\hat{\sigma}_{\hat{\beta}}} > 1.96\right\}$$

- Plot against average effective first-stage F-stat
  - Little action for $E\left[F^E\right] > 50$. Limit plot to $E\left[F^E\right] \leq 50$
  - Includes 106 of 124 specifications

Distribution of t-Statistics



Distribution of t-Statistics



Distribution of t-Statistics

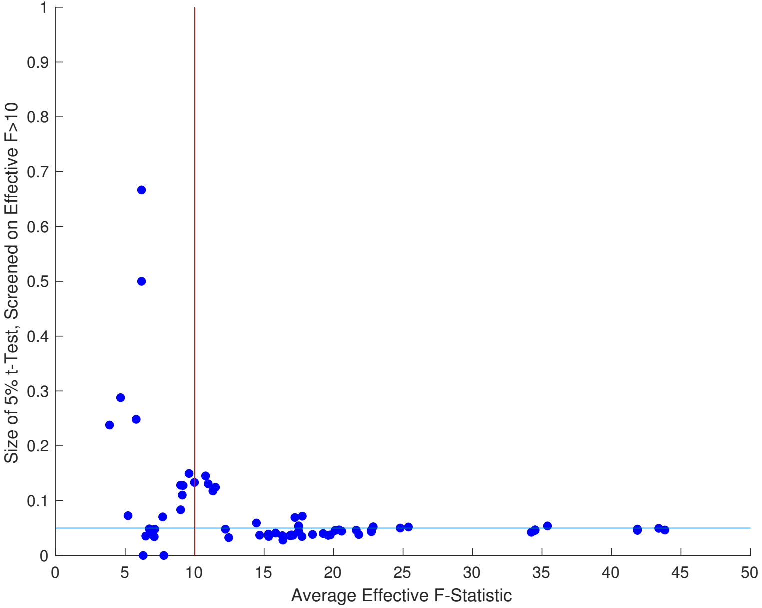
- Weak instrument issues apparent in some specifications
  - Median t-statistic far from zero
  - Size of nominal 5% t-test much larger than 5%
- Problems limited to specifications with a small average effective F-stat
  - No large distortions in specifications with $E\left[F^E\right] > 10$
  - Population rule of thumb seems to work pretty well
    - Not a theorem!

Weak instrument issues appear relevant for some recently-published specifications, but only in cases with $E\left[F^E\right]$ small

Screening on F-Statistics

- Given that average effective F-statistics seem to capture weak instrument issues, tempting to screen applications on F
  - e.g. only pursue applications with $F^E > 10$
- Distribution of F-statistics in AER sample suggests may be common
- As already discussed, can introduce size distortions
- Examine effect in our AER specifications
  - For each specification, calculate size of 5% t-test conditional on $F^E > 10$

Screening on F-Statistics



Screening on F-Statistics

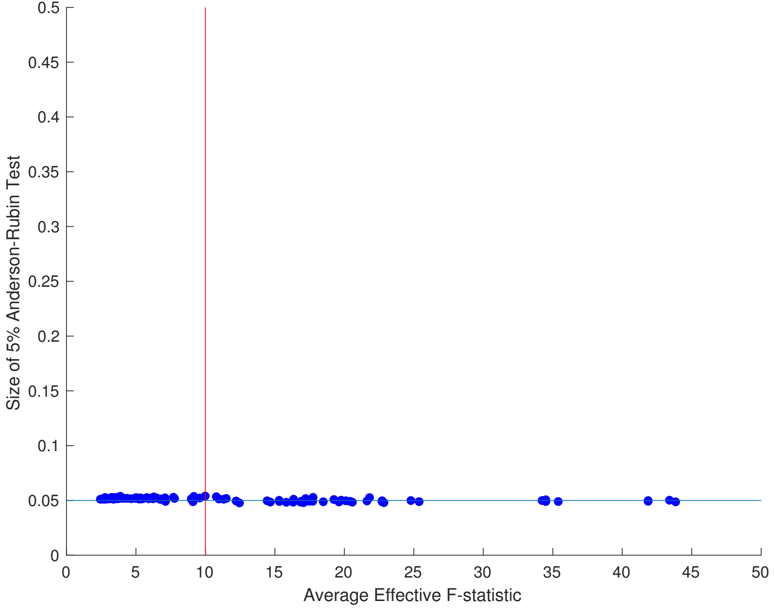
- Screening leads to much larger size for some specifications
  - Not specific to F^E, same issues appear for F^N, F^R
  - Not specific to threshold of 10: if move threshold, get distortions in neighborhood of new cutoff

Screening on F-statistics can make published results less reliable

Robust Confidence Sets

- Rather than screening applications on F^E, can compute robust confidence sets
  - Guaranteed to have correct coverage regardless of instrument strength
- For illustration, here consider Anderson-Rubin (AR)
  - Plot size of AR tests in AER specifications

Robust Confidence Sets



Robust Confidence Sets

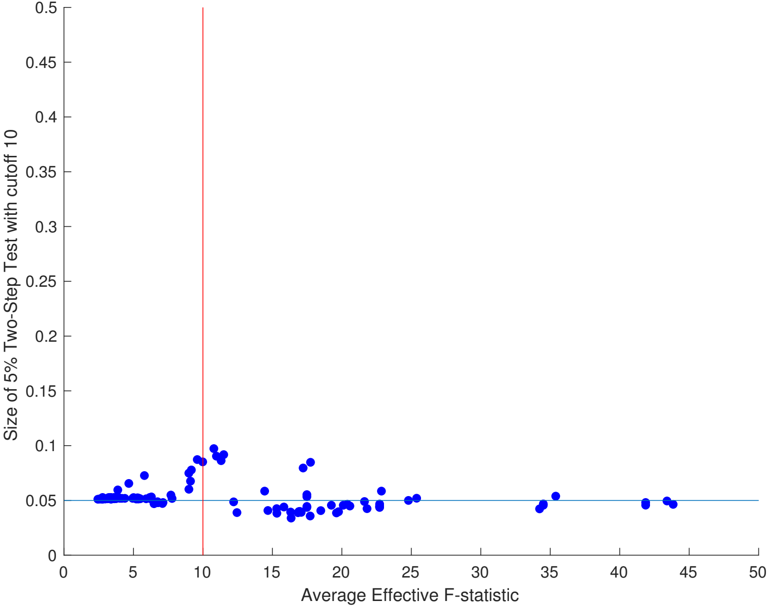
- AR size is flat at 5% regardless of instrument strength
  - AR also efficient in just-identified case
- For over-identified models, variety of robust tests and confidence sets available
  - Many ensure efficiency in strongly-identified case
  - All again ensure correct size regardless of instrument strength

Robust confidence sets eliminate size distortions from weak instruments

Two-Step Confidence Sets

- Robust confidence sets currently little-used in practice
  - When used, often because weak identification is suspected
- When used in this way, can be viewed as alternative to screening on F-statistic. For example
  - If $F^E \geq 10$, report t-statistic
  - If $F^E < 10$, report AR
- Alternatively, could use Montiel Olea and Pflueger (2013) critical values
- May introduce size distortions, but not clear how large
  - Examine performance in our AER specifications

Two-Step Confidence Sets



Two-Step Confidence Sets

- Observe some size distortions for specifications with $E\left[F^E\right] \approx 10$
  - True size never above 10% in these simulations
  - Results similar if instead use Montiel Olea and Pflueger (2013) critical values
- Also not a theorem!

Deciding to use a robust confidence set based on F^E isn't theoretically guaranteed to work, but cost appears small in our AER specifications

Summary of Simulation Results

1. Weak instruments appear to be a problem in some published specifications
2. Bad behavior largely limited to specifications with $E\left[F^E\right] < 10$
3. Screening on F^E can amplify problems
4. Robust confidence sets eliminate size distortions
5. Choosing whether to report robust CS based on F^E introduces some distortions, but small

Questions from Simulations: Performance of F-Statistics

Simulation results suggest some questions

1. Theoretical justification for F^E in Montiel Olea and Pflueger (2013) only concerns bias. Appears to also diagnose size problems reasonably well. Can this be formalized?
2. Two-step confidence sets based on F^E work reasonably well in simulations. Can this be formalized?

Questions from Simulations: Normal Approximation

- Our simulations take $\left(\hat{\delta}, \hat{\pi}\right)$ to be normally distributed with known variance
  - Focuses attention solely on distortions from weak instruments
- Results from Young (2018) suggest may be problematic
  - Based on sample of papers from AEA journals (larger than, but overlaps with, our AER sample)
- Young (2018) finds that
  1. A small number of observations have a large influence on estimates and p-values
  2. Variance estimates $\hat{\Sigma}$ often extremely noisy in simulation
  3. As a result of noisy estimates $\hat{\Sigma}$, AR tests can have large size distortions in over-identified settings
- Further exploration of interaction between weak instruments and issues discussed by Young (2018) of considerable interest

Other Research: Subvector Inference

- Some applications have more than one endogenous regressor
  - 19 out of 230 specifications in AER sample
- Most tests previously discussed extend to tests of $\dim(X) \times 1$ vector β in settings with multiple endogenous variables
  - Imply joint confidence sets for full vector
- Joint confidence sets rarely reported in strong-instrument settings. Instead, usually report e.g. estimates and standard errors for each element of β separately
  - Write $\beta = (\beta_1, \beta_2)$, and want confidence set for β_1 alone
- If assume instruments strong for β_2, simple solution
  - "Strong for β_2" meaning strong if treat β_1 as known
  - Plug appropriate estimate $\hat{\beta}_2(\beta_1)$ into robust test statistics (see e.g. Stock and Wright (2000))
- If instruments weak for β_2, hard problem

Other Research: Subvector Inference

- One option projection method
  - Form joint confidence set for (β_1, β_2), and collect implied set of values for β_1
  - Can have very low power
- Several recent papers seeking to improve power of projection method
  - Smaller critical values for AR statistic in homoskedastic case: Guggenberger et al. (2012)
  - Modified projection approach to improve power in well-identified case: Chaudhuri and Zivot (2011), D. Andrews (2017)
- Active area of research

Subvector Inference: Implementation

- Stata package weakiv can
  - Compute joint confidence sets for β
  - Plug in estimates for strongly identified β_2
  - Implement projection method
- Package twostepweakiv (also on SSC/Github) implements refined projection based on Chaudhuri and Zivot (2011)
  - Stata Journal article: Sun (Forthcoming)
  - Nearly-efficient inference on β_1 under strong identification
  - Also implements two-step CS with guaranteed coverage

Other Research: Nonlinear Models

- All the results discussed for IV apply directly to linear GMM
  - GMM moments linear in parameters
- Many (though not all) generalize to nonlinear GMM
  - No known analog of first-stage F-statistic
    - Alternative for approach detecting weak identification: I. Andrews (2018)
  - Many procedures for robust inference, e.g. Stock and Wright (2000), Kleibergen (2005), D. Andrews and Guggenberger (2015), I. Andrews and Mikusheva (2016)

The End

Thank you!

References

- Andrews D. 2017. Identification-robust subvector inference. Unpublished Manuscript
- Andrews D, Guggenberger P. 2015. Identification- and singularity-robust inference for moment condition models. Unpublished Manuscript
- Andrews I. 2018. Valid two-step identification-robust confidence sets for gmm. *Review of Economics and Statistics* 100:337–348
- Andrews I, Mikusheva A. 2016. Conditional inference with a functional nuisance parameter. *Econometrica* 84:1571–1612.
- Chaudhuri S, Zivot E. 2011. A new method of projection-based inference in gmm with weakly identified nuisance parameters. *Journal of Econometrics* 164:239–251
- Guggenberger P, Kleibergen F, Mavroeidis S, Chen L. 2012. On the asymptotic sizes of subset anderson–rubin and lagrange multiplier tests in linear instrumental variables regression. *Econometrica* 80:2649–2666

References

- Kleibergen F. 2005. Testing parameters in gmm without assuming they are identified. *Econometrica* 73:1103–1123
- Montiel Olea J, Pflueger C. 2013. A robust test for weak instruments. *Journal of Business and Economic Statistics* 31:358–369
- Sun L. Forthcoming. Implementing valid two-step identification-robust confidence sets for linear instrumental-variables models. *Stata Journal*
- Stock J, Wright J. 2000. Gmm with weak identification. *Econometrica* 68:1055–1096
- Young A. 2018. Consistency without inference: Instrumental variables in practical application. Unpublished Manuscript