

Week 6: Weak IV

JJ Chen

February 20, 2015

Plan

- ▶ Today we'll discuss
 1. 2SLS finite sample bias
 2. MC simulation of finite sample bias
 3. Stata implementations of specification tests
- ▶ Questions?

Recap: What're Weak IVs?

- ▶ Assume we have valid instruments, so 2SLS estimators are consistent
 - ▶ Stata command:
 - ▶ `glo ivmodel y (s = z1 z2) x1 x2`
 - ▶ `ivregress 2sls $ivmodel, r`
- ▶ But our concern is with whether the instruments are weak
 - ▶ That is, whether the true 1st coefficient is close to zero
 - ▶ Or whether the (conditional) joint correlation of the endog var with instruments are small
- ▶ Recall that there are two simple measures of joint correlation
 1. R^2 from regressing the endog var on Z s and exo ctrl
 2. F-stat from the same regression

What Happens When Instruments are Weak?

- ▶ Trade-off of Use all the possible instruments (satisfied ER and RC)
 - ▶ Get a more *efficient* estimator
 - ▶ But lead to larger *finite sample bias* (Hahn and Hausman (2003))
- ▶ Usually we discuss finite sample bias in over-id cases
 - ▶ Over-id: $\# \text{ instruments } (J) > \# \text{ endog var } (K)$
 - ▶ Just-id: $\# \text{ instruments } (J) = \# \text{ endog var } (K)$
 - ▶ In Just-id (one endog) models, the first moment of the IV estimator doesn't exist
 - ▶ The M th moment of a 2SLS estimator only exist if $M < J - K + 1$ (Kinal (1980))

The 2SLS Estimator

- ▶ Consider a simple setting: one endog $x_{N \times 1}$ and a few instruments $Z_{N \times Q}$:
 - ▶ $y = \rho x + \eta$
 - ▶ $x = Z\pi + \xi$
- ▶ We have the followings:
 - ▶ η_i is correlated with ξ_i
 - ▶ By construction, ξ_i is orthogonal to the instruments z_i'
 - ▶ By assumption (ER), η_i is orthogonal to the instruments z_i'
- ▶ The 2SLS estimator is defined as
 - ▶ $\hat{\rho}^{2SLS} = (x' P_Z x)^{-1} x' P_Z y = \rho + (x' P_Z x)^{-1} x' P_Z \eta$
 - ▶ $P_Z = Z(Z'Z)^{-1}Z'$ is the projection matrix

The Bias of 2SLS

- ▶ Substitute the 1st for x in $x'P_Z\eta$:
 - ▶ $\hat{\rho}^{2SLS} - \rho = (x'P_Zx)^{-1}\pi'Z\eta + (x'P_Zx)^{-1}\xi'P_Z\eta$
- ▶ Take expectation and Bekker approximation:
 - ▶ $E(\hat{\rho}^{2SLS} - \rho) \approx (E(x'P_Zx))^{-1}E(\pi'Z\eta) + (E(x'P_Zx))^{-1}E(\xi'P_Z\eta) = (E(x'P_Zx))^{-1}E(\xi'P_Z\eta)$
 - ▶ The first term of RHS can be canceled since z_i is orthogonal to η_i : $E(\pi'Z\eta) = 0$
- ▶ Substitute the 1st for x again in $x'P_Zx$ gives
 - ▶ $E(\hat{\rho}^{2SLS} - \rho) \approx [E(\pi'Z'Z\pi) + E(\xi'P_Z\xi)]^{-1}E(\xi'P_Z\eta)$
 - ▶ Thus we see 2SLS is biased because η_i is correlated with ξ_i : $E(\xi'P_Z\eta) \neq 0$

Tricks for Getting $E(\xi' P_Z \eta) \neq 0$

- ▶ Notice that $\xi_i' P_Z \eta$ is just a scalar, so that we can apply the trace trick (cyclic permutations and $\text{tr}(P_Z) = \text{rk}(P_Z) = Q$) we learned last semester:

$$\begin{aligned}
 E(\xi' P_Z \eta) &= E(\text{tr}(\xi' P_Z \eta)) \\
 &= E(\text{tr}(P_Z \eta \xi')) \\
 &= \text{tr}(P_Z E(\eta \xi')) \\
 &= \text{tr}(P_Z \sigma_{\eta\xi}^2 I) \\
 &= \sigma_{\eta\xi}^2 Q.
 \end{aligned}$$

- ▶ Similarly, $E(\xi' P_Z \xi) = \sigma_{\xi^2}^2 Q$.

Simplify the Bias Term

- ▶ We can further simplify the term and get a more intuitive formula

$$\begin{aligned}
 E\left(\hat{\rho}^{2SLS} - \rho\right) &\approx [E(\pi'Z'Z\pi) + E(\xi'P_Z\xi)]^{-1} E(\xi'P_Z\eta) \\
 &= \sigma_{\eta\xi}^2 Q [E(\pi'Z'Z\pi) + \sigma_{\xi^2}^2 Q]^{-1} \\
 &= \frac{\sigma_{\eta\xi}}{\sigma_{\xi}^2} \left[\frac{E(\pi'Z'Z\pi)/Q}{\sigma_{\xi}^2} + 1 \right] \\
 &= \frac{\sigma_{\eta\xi}}{\sigma_{\xi}^2} \frac{1}{F + 1}
 \end{aligned}$$

- ▶ Notice that when the 1st F-stat is small, the bias term gets larger; and if the instruments are too weak, the 2SLS bias term gets closer to OLS bias

Rule of Thumb and LIML

- ▶ Rule of thumb: $F < 10$ indicates weak instruments (Staiger and Stock (1997))
- ▶ Alternative estimators that are asym equivalent to 2SLS, but have better finite sample properties: the LIML estimator
 - ▶ LIML has a smaller bias than 2SLS (and GMM)
 - ▶ Assume joint normality of errors in the structural and 1st eqn
 - ▶ Put different weights on instruments
 - ▶ The 1st moment is still not exist for just-id models
 - ▶ Stata command:
 - ▶ `glo ivmodel y (s = z1 z2) x1 x2`
 - ▶ `ivregress liml $ivmodel, r`

Set up

- ▶ Following our textbook, we'll simulate some data and compare performance of the OLS estimator, IV estimator (just-id), 2SLS estimator (over-id), and LIML (over-id)
- ▶ Suppose we have the following DGP
 - ▶ $y_i = \beta s_i + \eta_i$
 - ▶ $s_i = \sum_{j=0}^{20} \pi_j z_{ij} + \xi_i$
 - ▶ $\beta = 1, \pi_0 = 0.1, \pi_1 = \dots = \pi_{20} = 0$
 - ▶ assume η_i and ξ_i are jointly normal distributed

To Get the Idea, Here is One Random Sample

```

N <- 10000 # set obs 10000
ability <- 0.8944272 * rnorm(N)
eta <- ability + 0.4472136 * rnorm(N)
xi <- ability + 0.4472136 * rnorm(N)
z0 <- rnorm(N)
z1 <- rnorm(N); z2 <- rnorm(N); z3 <- rnorm(N)
s <- 0.1 * z0 + xi # 1st eqn
y <- 1 * s + eta # structural eqn

```

- ▶ s is endogenous. Why? What's the OV?
- ▶ z0 is a valid instrument. Why?
- ▶ z1, z2, z3 seems like “fake” instruments. Why?

R Code: Regressions

```
ols <- lm(y ~ s)
iv <- ivreg(y ~ s | z0)
kable(texreg(list(ols, iv)))
```

	Model 1	Model 2
(Intercept)	−0.00 (0.01)	−0.00 (0.01)
s	1.79*** (0.01)	1.09*** (0.09)
R ²	0.90	0.76
Adj. R ²	0.90	0.76
Num. obs.	10000	10000

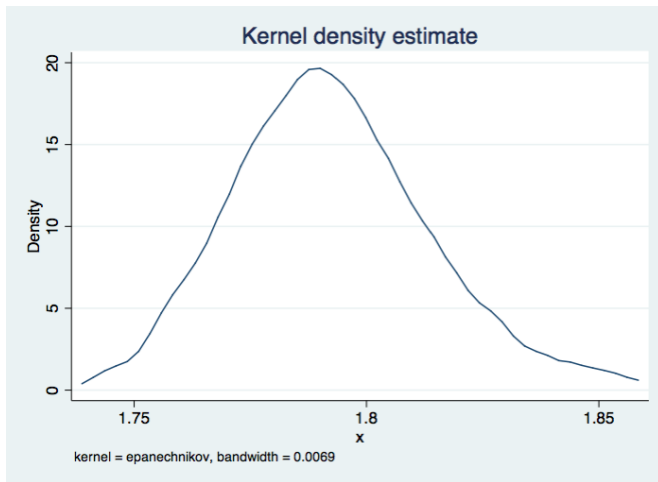
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 1: Statistical models

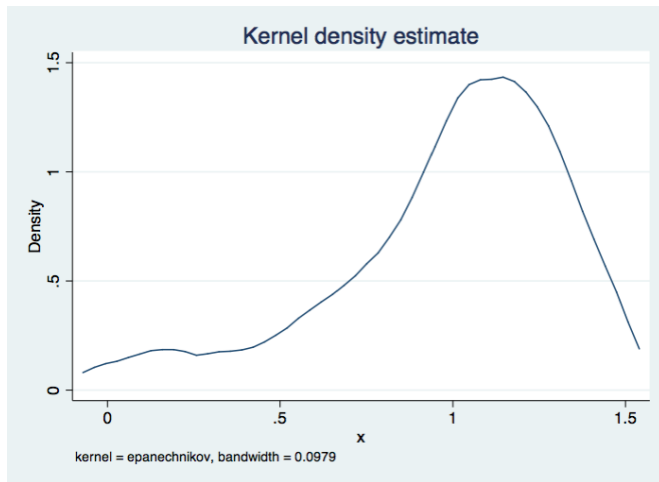
100 Sample

- ▶ Now suppose we draw another 99 sample and estimate
 1. OLS regression
 2. 2SLS with one valid instrument (just-id)
 3. 2SLS with two instruments (only one is valid)
 4. LIML with two instruments (only one is valid)
 5. 2SLS with twenty instruments (only one is valid)
 6. LIML with twenty instruments (only one is valid)
 7. 2SLS with twenty instruments (all invalid)
 8. LIML with twenty instruments (all invalid)

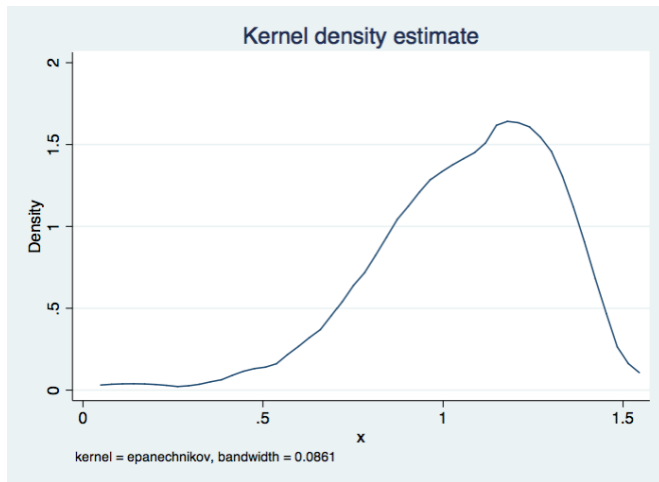
OLS is biased



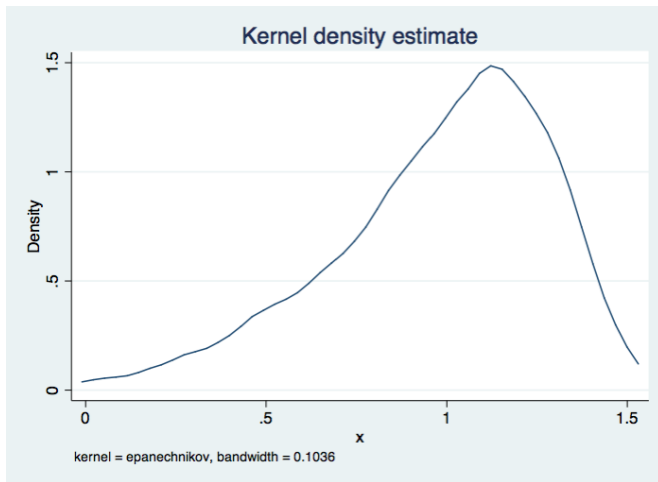
IV is consistent (with one valid instrument)



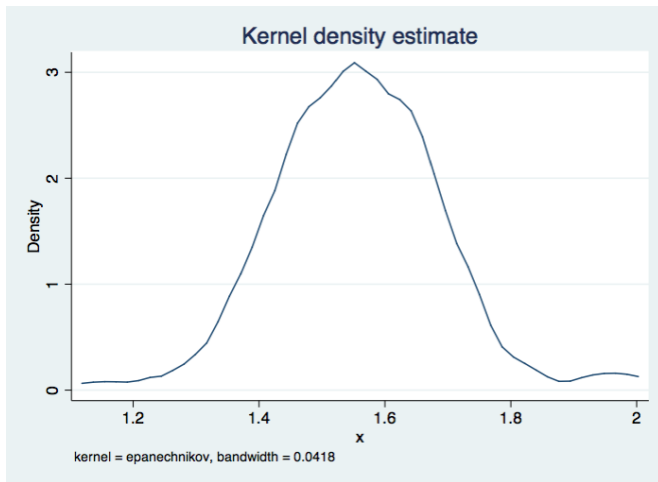
2SLS (two instruments, one valid)



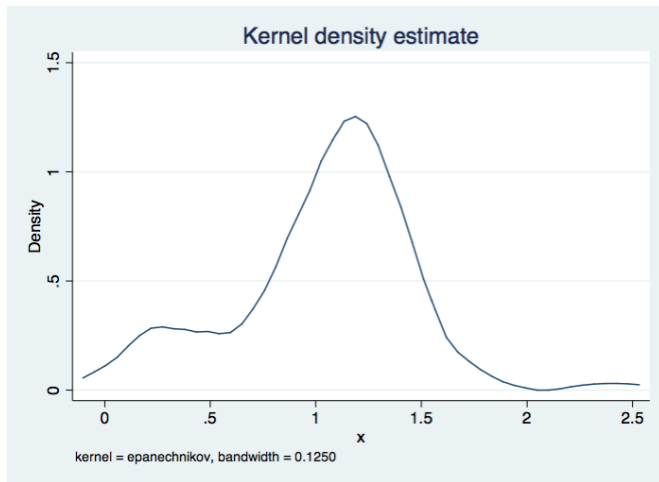
LIML (two instruments, one valid)



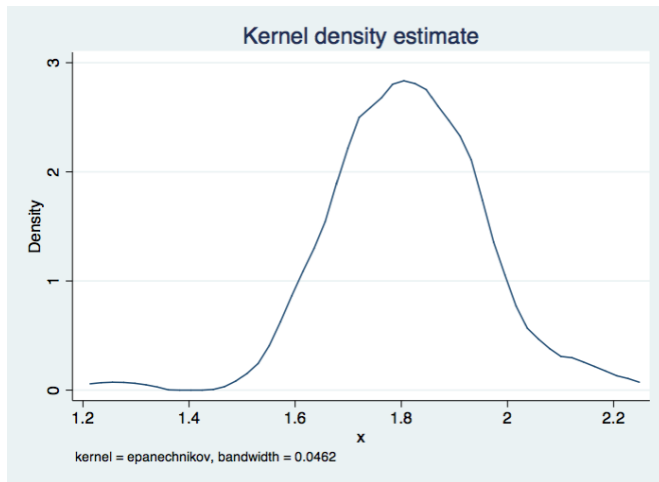
2SLS (twenty instruments, one valid)



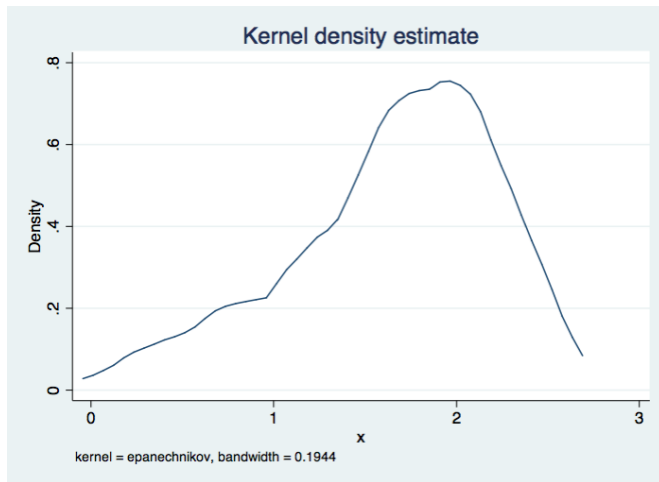
LIML (twenty instruments, one valid)



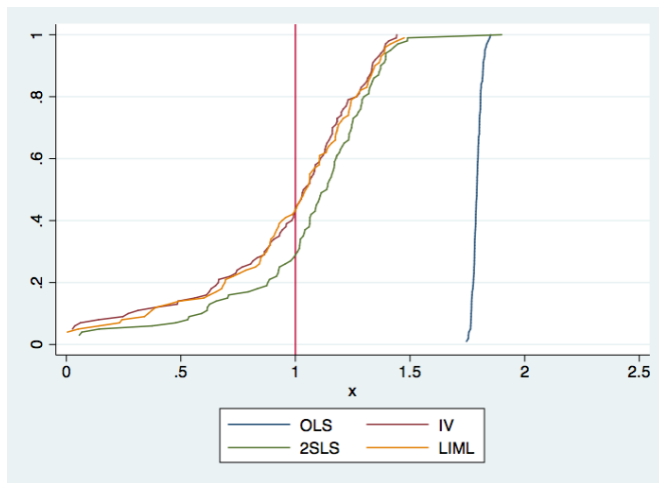
2SLS (twenty instruments, all invalid)



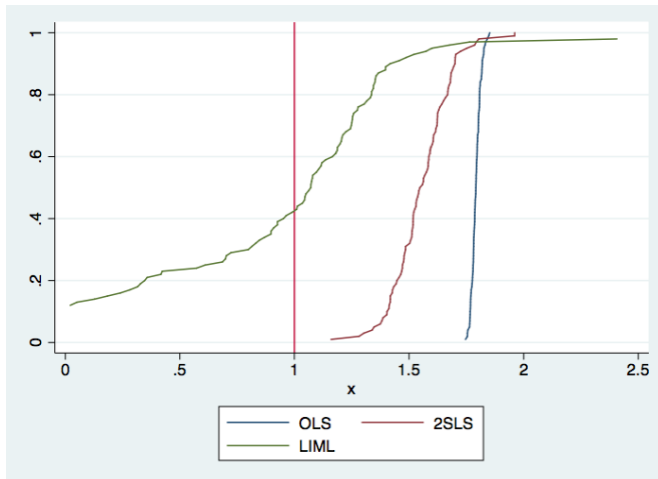
LIML (twenty instruments, all invalid)



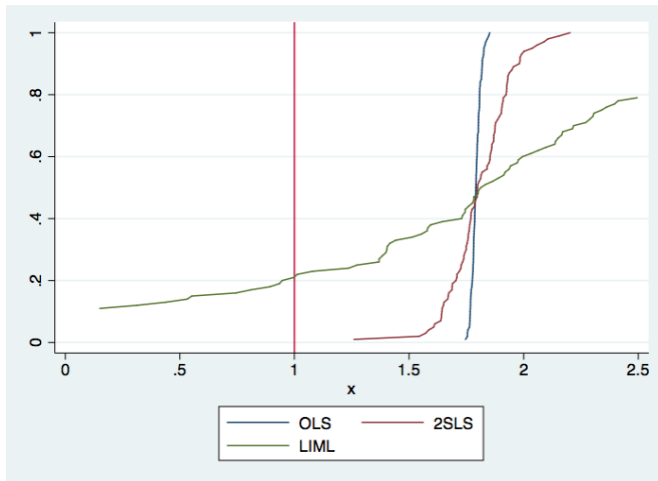
First 4 Cases in one graph



2SLS and LIML with 20 instruments (One Valid Instruments)



2SLS and LIML with 20 instruments (all invalid)



DWH test of endogeneity

- ▶ Intuition: If there's little difference between OLS and IV, then there is no need to instrument.
- ▶ Null: Regressors of interest are exogenous (In class, the null is stated as “ $\hat{\rho}_I$ and $\hat{\rho}_O$ are both consistent, but $\hat{\rho}_O$ is more efficient”.)

Tests of over-id restrictions

- ▶ Or Hansen's test, Sargan's test, Hansen-Sargen's test
- ▶ Intuition: If multiple instrument, we can construct separate IV estimates of the same parameter. If these estimates are “too different” from one another, then we can reject some part of the model.
- ▶ Null: All instruments are valid; formally, $H_0 : E(Z_i \eta_i) = 0$.
Rejecting the null suggests that at least one of the instrument is not satisfied the ER.

Control Function and OV Test

- ▶ Intuition: If there are no omitted variable, predicted 1st error ($\hat{x}i_i$) should have a zero coefficient in the regression of Y_i on S_i and $\hat{x}i_i$
- ▶ Null: Regressor of interest are endogenous (see handout)

Stata Commands

- ▶ Stata offers three post-estimation commands
 - ▶ `glo ivmodel y (s = z1 z2) x1 x2`
 - ▶ `ivregress 2sls $ivmodel, r`
 - ▶ `estat firststage`
 - ▶ `estat endog s`
 - ▶ `estat overid`
- ▶ Example: AK91, see handout

Discussion

- ▶ Here is a midterm question from last year:
- ▶ Suppose you are interested in estimating the regression $Y_i = \alpha + \beta X_i + e_i$, but you are worried that X_i is endogenous. You have two instruments for X_i : Z_{1i} and Z_{0i} .
 - ▶ You are confident that the ER holds for Z_{1i} but are worried that it's a weak instrument.
 - ▶ You are less confident the exclusion restriction holds for Z_{2i}
- ▶ Questions:
 1. How would you estimate β ? (possible ivmodels?)
 2. Explain any potential advantages or limitations of your choice.

References I

Hahn, Jinyong, and Jerry Hausman. 2003. “Weak Instruments: Diagnosis and Cures in Empirical Econometrics.” *American Economic Review*. JSTOR, 118–25.

Kinal, Terrence W. 1980. “The Existence of Moments of K-Class Estimators.” *Econometrica: Journal of the Econometric Society*. JSTOR, 241–49.

Staiger, Douglas, and James H Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica: Journal of the Econometric Society*. JSTOR, 557–86.