Recent Advances in the Field of Trade Theory and Policy Analysis Using Micro-Level Data

July 2012
Bangkok, Thailand

Cosimo Beverelli
(World Trade Organization)
Instrumental variables models in Stata

- Cross section
- Panel data
- References
Cross-section

• Stata has a built in command for instrumental variables regression, *ivregress*
  
  • *ivregress 2sls depvar indepvar1 indepvar2 (endogvar1 endogvar2 = iv1 iv2 iv3...) ..., first options*

• There is a user-developed extension with a number of desirable features, *ivreg2*
  
  • *ivreg2 depvar indepvar1 indepvar2 (endogvar1 endogvar2 = iv1 iv2 iv3...) ..., first options*

• These two commands produce the same output, although the input format is a little different

• *ivreg2* automatically provides additional diagnostic statistics, and is usually preferable
Cross-section (ct’d)

• Specifying the *gmm* option with either command results in more efficient estimates for overidentified models, but no difference for just-identified models

• Always use the *first* and *ffirst* options to check the first stage regression results, and ensure that instruments are not weak

• Always try to overidentify your model (i.e., include at least one more instrument than potentially endogenous variable) and test the exclusion restriction using the Hansen/Sargan test

• Always perform an endogeneity test of endogenous regressors (*endog(endogvar)* option)
Panel data

- Stata has built in panel data models using instrumental variables, as well as user-built extensions:
  - `xtivreg, re and fe`
  - `xtivreg2, re and fe`

- The general format, including the use of parentheses for the instruments, is as for `ivregress` and `ivreg2`

- Results using the two commands are identical
- `xtivreg2` contains additional diagnostic statistics (similar to `ivreg2`), and is generally preferable
References

- Baum et al. (2003)
- Baum et al. (2007)
- Nichols (2007)