





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

July 2012 Bangkok, Thailand

Cosimo Beverelli (World Trade Organization)

- Enterprise Surveys database
- Stata T&T (tips and traps)
- Conclusion: research design and methodology

Enterprise Surveys database

- The Enterprise Surveys dataset offers an attractive alternative to traditional firm-level data sources
 - Cross-country (135 countries, 130,000 firms)
 - Standardized methodology
 - Inclusion of data relevant to policy
 - Covers producers of goods and services
- These data always need to be used carefully, however:
 - Accounting problems and under-reporting of sales for tax reasons
 - Possible non-comparability of business climate data across countries
 - Over-sampling of large firms and exporters
 - Often difficult to obtain robust measures of productivity

Enterprise Surveys database (ct'd)

- The Enterprise Surveys data have been used in a variety of contexts in published research:
 - Identification of firm-level premia for exporting and importing
 - Examination of the determinants of export performance (propensity and intensity)
 - Analysis of the links between the business climate or trade facilitation and trade performance
 - Identification of the determinants of trade-related corruption
- Remember that you CANNOT use the example data from this course to do real research – they are fictional and have been altered from the original source
 - Contact the Enterprise Surveys team directly to have access to the original (real) data
 - The data are freely accessible to researchers upon agreeing to terms regarding confidentiality of individual survey responses

Stata T&T (tips and traps)

- Stata is ideally suited to working with large firm-level datasets like the Enterprise Surveys data
 - Ideally, use Stata MP to work with large datasets
- Start with descriptive statistics and graphical methods:
 - summarize (sum)
 - tabulate (tab)
 - correlate (corr)
 - *histogram* (*hist*) and *kdensity*
 - *twoway* [+ scatter, lfit, kdensity, etc.]
- Always try to tell your story with simple statistics or, even better, a graph or two before moving to the econometrics
- Make use of the *if* command to exploit interesting splits in the data

Stata T&T (tips and traps) (ct'd)

- When working with Enterprise Surveys data, you will mostly be using panel data techniques to control for unobserved heterogeneity:
 - Countries or regions
 - Years
 - Industries
 - Combinations of the above
- Make sure your results are robust to different panel data assumptions, and try to push the unobserved heterogeneity as far as it can go
- For example, results with fixed effects by country-industry-year are stronger than those with fixed effects by country, fixed effects by industry, and fixed effects by year

Stata T&T (tips and traps) (ct'd)

- When working with panel data, you will be using *xt* commands:
 - xtset
 - xtreg
 - xtlogit
 - *xtivreg* and *xtivreg2*
 - etc.
- The literature mostly uses fixed effects, and this should be the starting point for your research
- Random effects is a more restricted model, and so should only be used if absolutely necessary and if the data support it
- The Hausman test can be used as a guide, but it is often unreliable in practice...
- ...The test of overidentifying restrictions should be preferred

Stata T&T (tips and traps) (ct'd)

- When working with fixed effects, use the *xt* commands whenever possible for one dimension:
 - Takes care of clustering
 - Makes computation faster
- For multiple dimensions of fixed effects, you will need to enter some dimensions manually
 - quietly tab, gen(newvar_)
 - Use wildcards (*) in the regression command
 - Or use the *i*. command
 - Use *xtset* to have the *xt* command take care of the dimension with the largest number of fixed effects, thereby maximizing the reduction in computation



https://www.yunbaogao.cn/report/index/report?reportId=5_7494

