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
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.
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 time
Various packages are available to automatically create publication-ready tables from raw Stata output
  • Do not enter the stars and other stuff on your own...
  • ...Use estout or a similar set of commands (my favourite one, which requires some learning but is very flexible, is outreg2)

Always use a do file (strictly necessary) and logs (highly advisable) to keep track of your regressions and results
  • One do file to create the database through merging (for merging datasets, use joinby)
  • One do file for the regressions
  • You can use do files of do files
  • You will need to come back to your specifications as you revise material for publication
Research design and methodology

• Firm-level research in trade is still in its infancy, but....

• Diminishing returns are setting in for work that just looks at export or import premia
  • The basic results are now well-established
  • Some value in replicating them for different countries
  • Publication possibilities are limited to national or regional journals

• The trick to making a good publication is in finding an interesting research question that can be answered well with firm-level data, but not so well with other types of data
Research design and methodology (ct’d)

• Potentially under-researched areas with firm-level data include:
  • Services
  • Behind the border barriers
  • Corruption and governance
  • Regulatory barriers to trade
  • Importance of networks and connectivity in international trade
  • Trade facilitation
  • Links between trade and innovation
  • …

• This does not mean that there is no research, simply that there is some space for more good contributions
Research design and methodology (ct’d)

• Start with a good question

• Identify the relevant data and conduct an exploratory analysis using descriptive statistics, graphs, and simple regressions

• If the data seem to be telling an interesting story, push the analysis further in terms of technique, but...

• ...Use the simplest technique that is consistent with your research design and data: there is no advantage in using complicated techniques if the data do not call for them

• And finally, remember that robustness checks are necessary to support your core results