

LAB0 - Introduction, and Exploring OLS Estimates in TSCS

Making “do files” (scripts; batch files), good practice

Analysis in Stata done in any of three (or more!) ways:

1. Menu
2. Command Line
3. “do files”

Effective practice suggestions:

- Use menus or command line to test commands, copy into do file
- Make comments with `*`, `//`, `/* */`
- Make the “script” so that it does all of the data analysis in a single pass
- Make log files of interactive sessions, cut & paste to do file:

Typical Start to a DO FILE

```
clear
capture log close
log using "[...path\filename]", replace text
use "[...path...]/garmit_esspanel1.dta"
```

But we’re going to work interactively here to start...

—Exceptionally useful starting point for exploring Stata’s explicitly “xt” (panel/tscs) designed features:

. help xt

—Also useful are (in recent versions, see also **help fvvarlist**):

. help xi

—Data from Garrett & Mitchell: “Globalization, government spending and taxation in the OECD,” *European Journal of Political Research* 39(2) (2001): 145–177. Explore variable list.

. use "...[PATH]...\garmit_esspanel1.dta", clear

—Can also use pull-down menu, or can also download from web like so:

. use "http://www-personal.umich.edu/~franzese/garmit_esspanel1.dta"

—Good idea to get in habit of logging your work every session. Can always delete records later if you wish, but you cannot go back to create a log as easily from within a session and not at all once you’ve closed one.

. log using "...[PATH]...\TAMU.TSCSwrkshp.Lab0.log", append

—This dataset has already been set-up in lots of ways, but let’s review that process a bit:

. help xtset

. xtset

panel variable: cc (strongly balanced)

time variable: year, 1961 to 1994

delta: 1 unit

The 18 CC’s & Country’s are:

- | | | |
|--------------|-----------------|--------------------|
| 1. Australia | 7. France | 13. New Zealand |
| 2. Austria | 8. Germany | 14. Norway |
| 3. Belgium | 9. Ireland | 15. Sweden |
| 4. Canada | 10. Italy | 16. Switzerland |
| 5. Denmark | 11. Japan | 17. United Kingdom |
| 6. Finland | 12. Netherlands | 18. United States |

—All right, let's estimate a model already!

```
. reg spend unem growthpc depratio left cdem trade lowage fdi
```

Source	SS	df	MS	Number of obs	=	529
Model	33411.5828	8	4176.44785	F(8, 520)	=	91.12
Residual	23834.1148	520	45.8348362	Prob > F	=	0.0000
				R-squared	=	0.5837
				Adj R-squared	=	0.5772
Total	57245.6976	528	108.419882	Root MSE	=	6.7701

spend	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
unem	.9320547	.0928529	10.04	0.000	.7496417	1.114468
growthpc	-.9012817	.1303971	-6.91	0.000	-1.157451	-.6451119
depratio	-.4446718	.1235697	-3.60	0.000	-.6874289	-.2019147
left	.058016	.0087215	6.65	0.000	.0408823	.0751497
cdem	.0287597	.0131838	2.18	0.030	.0028596	.0546597
trade	.1293114	.014866	8.70	0.000	.1001066	.1585163
lowage	-.157155	.0478007	-3.29	0.001	-.2510612	-.0632488
fdi	.0927398	.2159476	0.43	0.668	-.3314971	.5169768
_cons	48.81442	4.943615	9.87	0.000	39.1025	58.52633

—One low-tech, but sometimes quite useful post-estimation strategy is simple graphical exploration of your estimates, your estimated residuals in particular. For Stata's many built-in post-estimation tools for graphical & statistical exploration, see menu item: Statistics/Linear models and related/Regression diagnostics... Also see:

```
. help regress postestimation
```

And:

```
. help regress postestimation ts
```

```
. help xtreg postestimation ts
```

You may also wish to create graphics related to your estimated residuals yourself. Useful for this are to create a variable like the following:

```
. predict e_FullPooledOLS, residuals      (83 missing values generated)
```

A simple line-plot of the residuals by country is now available thus:

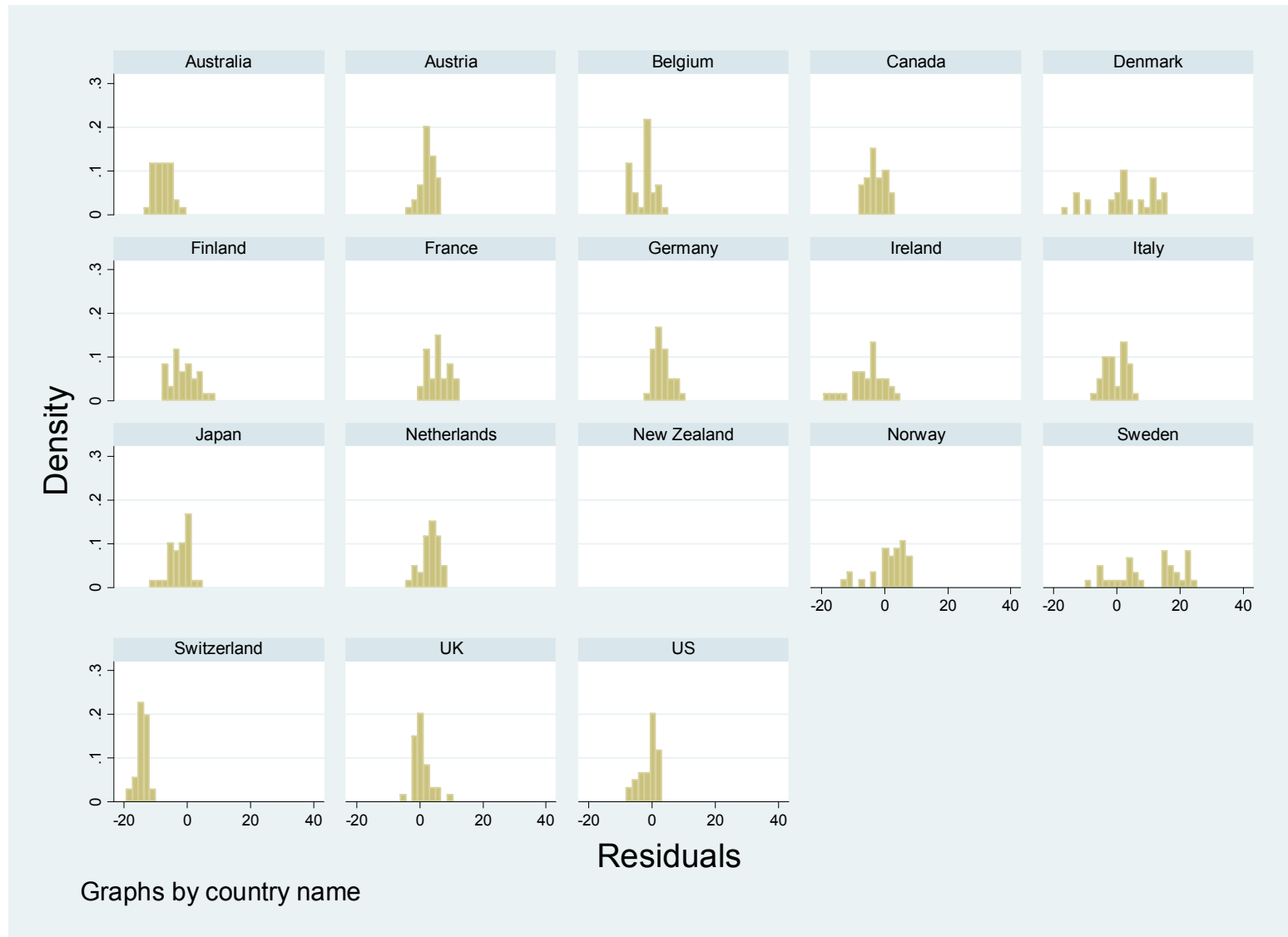
```
. xtline e_FullPooledOLS, i(country) t(year)
```



HOMEWORK: What patterns should you see in the means, trends, & spreads of these residuals within & across units & over time if the C(N)LRM applied? What's wrong with these pictures compared to those expectations?

Another useful plot for exploring residuals are histograms. **HOMEWORK: Same questions here:**

```
. histogram e_FullPooledOLS, by(cc)
```



HOMEWORK: A similarly useful plot would be a box-and-whiskers by country. Generate.

Another useful variable to generate for post-estimation graphical and other exploration is this:

```
. gen e2_FullPooledOLS=e_FullPooledOLS^2      (83 missing values generated)
```

HOMEWORK: *What is this an estimate of?* [I didn't say anything about it's properties. Generally, not great, but does have one underemphasized, but tremendously useful property: *existence!*]

HOMEWORK: How might you use e and/or e^2 to test for
...insufficiently modeled (orthogonal) country-specific factors? (i.e., the sufficiency of a common intercept?)
...heteroskedasticity that would bias OLS estimates of coefficient-estimate variance-covariance matrices?
...heteroskedasticity by country? ...by period? ...autoregressive conditional heteroscedasticity (ARCH)?
...(time-)serial correlation?

Look at these:

```
. help estat hetttest  
. help estat imtest  
. help estat szroeter
```

Here's another that sounds magically wonderful:

```
. help estat ovtest
```

HOMEWORK: Claims to be a test for omitted variables. What is it? What does it actually test?

More Explorations of Typical Heterogeneity in TSCS Data

```
. use "...[PATH]...\debt_699.dta", clear
. use "http://www-personal.umich.edu/~franzese/debt_699.dta"
```

Generalized (Normal) Linear-Regression Model:

```
. reg debtx growth ue enop cog if debtx~=.
. predict LSresids , res
. reg LSresids us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al nz
```

(Lots of) Unit Heterogeneity; Try Fixed Effects:

```
. reg debtx us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al nz
growth ue enop cog
. predict LSDVresids, res
. gen LSDVresids2=LSresids^2
. reg LSDVresids2 us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al
nz
```

Still lot of panel heteroskedasticity; return primary model to memory:

```
. reg debtx us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al nz
growth ue enop cog
```

Stata's Battery of Postestimation Commands:

```
. help regress postestimation
```

dfbeta	DFBETA influence statistics	acprplot	augmented component-plus-residual plot
estat hetttest	tests for heteroskedasticity		
estat imtest	information matrix test	avplot	added-variable plot
estat ovtest	Ramsey reg specification-error test for omitted vars	avplots	all added-variable plots in one image
estat szroeter	Szroeter's rank test for heteroskedasticity	cprplot lvr2plot	component-plus-residual plot leverage-versus-squared-residual plot
estat vif	variance inflation factors for the independent variables	rvfplot rvpplot	residual-versus-fitted plot residual-versus-predictor plot

Cook-Weisberg:

```
. estat hetttest
```

Augmented Cook-Weisberg:

```
. estat hetttest, rhs
```

White's General test:

```
. help whitetst [install if needed]
```

```
. whitetst
```

Or alternative route to White's test:

```
. help estat imtest
```

```
. estat imtest , white
```

Current form of our old friend -- , robust: *(What is this, actually?)*

```
. reg debtx us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al nz  
growth ue enop cog, vce(robust)
```

Panel WLS (Feasible PWLS):

```
xtset ctry year
```

```
xtgls debtx us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al nz  
growth ue enop cog, p(h)
```

Exploiting Stata's xi for TSCS Post-Estimation Exploration

Compare the following:

```
. reg debtx us ja ge fr it uk ca au be de fi gr ir ne no po sp sw sz al nz  
growth ue enop cog  
. xi: reg debtx i.ctry growth ue enop cog
```

Now try this:

```
. xi: reg debtx i.ctry*growth i.ctry*ue i.ctry*enop i.ctry*cog
```

It works, but is very messy...

You could then do, e.g. [WARNING: Messy, time-consuming...]:

```
. avplots or . avplot growth or . acprplot growth , low
```