

Stata Commands for pooled time-series and cross-sections
PLSC 597D: Pooled Time-Series Cross-Sectional Analysis.

Suzanna De Boef
Spring 2004

1 Basics

All the Stata commands for pooled time-series cross-sections is listed under the “xt” commands. Before you use these, however, you need to tell Stata that you have this two dimensional data structure. To do so we specify the variables associated with time and with unit in each do-file that you write to run xt commands:

```
. iis county  
. tis year
```

TSCS data must be in what Stata calls “long” format. Long format looks like this:

state	year	dep	indep1	indep2	indep3
1	1900	13.6	75	1	100
1	1901	13.9	82	1	110
⋮	⋮	⋮	⋮	⋮	⋮
2	1900	12.2	75	0	90
2	1901	12.1	82	0	95
⋮	⋮	⋮	⋮	⋮	⋮

“Wide” format looks like this:

state	year	dep	indep1	indep2	indep3
1	1900	13.6	75	1	100
2	1900	12.2	75	0	90
⋮	⋮	⋮	⋮	⋮	⋮
1	1901	13.9	82	1	110
2	1902	12.1	82	0	95
⋮	⋮	⋮	⋮	⋮	⋮

To go from wide to long (or long to wide) you use the “reshape” command. See the Stata help file for details on this command.

2 Key -xt- commands

- xtsum: gives you the same information as the summarize command — gives means and standard deviations and does so both within units and between them.

- xttab: similar to tabulate, but again breaks down both within units and between them.
- xtreg: To fit fixed or random effects models, options include
 - fe estimates a fixed effects model,
 - re estimates a gls random effects models (default),
- xtgls: options include
 - panel() indicates the type of panel (unit) variability; options include:
 - * iid: independent, common-variance panels, yielding a single estimate of σ^2 ,
 - * heteroscedastic: uncorrelated units, each with its own variance estimate, yielding N distinct estimates of σ_i^2 ,
 - * correlated: heteroscedastic, spatially correlated panels, yielding $\frac{N(N-1)}{2} + N$ parameters ($\frac{N(N-1)}{2}$ covariances σ_{it}, σ_{jt} and N variances estimates σ_i^2),
 - corr() indicates the within-unit temporal correlation:
 - * independent gives estimates no temporal correlation,
 - * ar1 gives estimates with common AR(1) errors, estimates a single value of $\hat{\rho}$.
 - * psar1 gives panel-specific AR(1) estimates, yielding N separate estimates of $\hat{\rho}_i$.
 - A summary of the various xtgls options and the number of parameters each estimates like:

	No AR(1)	Common $\hat{\rho}$	Separate $\hat{\rho}_i$
$\sigma_i^2 = \sigma^2, \text{cov}(\sigma_{it}, \sigma_{jt})=0$	$k + 1$	$k + 2$	$N + k + 1$
$\sigma_i^2 \neq \sigma^2, \text{cov}(\sigma_{it}, \sigma_{jt})=0$	$N + k$	$N + k + 1$	$2N + k$
$\sigma_i^2 \neq \sigma^2, \text{cov}(\sigma_{it}, \sigma_{jt}) \neq 0$	$\frac{N(N-1)}{2} + N + k$	$\frac{N(N-1)}{2} + N + k + 1$	$\frac{N(N-1)}{2} + 2N + k$

- xtpcse: estimates models with Beck and Katz’s (1995) “panel corrected standard errors”:
 - corr is the same as for xtgls. xtpcse will automatically do a Prais Winston regression to deal with temporal correlation if either of the two AR options are specified,
 - hetonly corresponds to the panel(heteroscedastic) option in xtgls,
 - independent corresponds to the panel(independent) option in xtgls,
 - the default is heteroscedastic, spatially correlated panels.
- xtregar: estimates fixed and random effects models with AR(1) errors. This is similar to the xtreg command, in that one specifies fe or re for fixed or random effects models, respectively. The model also gives estimates of fixed or random effects and common $\hat{\rho}$.
- Useful tests:
 - xttest2: used after xtgls or xtreg,fe: implements a Breusch-Pagan test for cross-unit correlation. Uses estimated residuals from the most recent model to test whether the cross-unit correlations are identically zero. The test is distributed χ^2 with $\frac{N(N-1)}{2}$ degrees of freedom.
 - xttest3: modified Wald statistics that tests the hypothesis that $\sigma_i^2 = \sigma^2$ for all i (homoscedasticity). The statistic is distributed χ^2 with N degrees of freedom.