Time Series Cross-Section

Jan Rovny

Sciences Po, Paris

Time Series Cross-Section

- parallel observation of different units (e.g. countries) over time
- time-ordered observations across different units
- expect theoretical *unit homogeneity* – comparable causal processes across units

Time Series Cross-Section

particular data structure of N units and T time periods. Total observations $= N * T$

Unit	Time
1	1
1	2
1	3
1	...
1	T
2	1
2	2
2	3
2	...
2	T
...	...
N	1
N	2
N	3
N	...
N	T

Problems in TSCS

- unit effects: each unit has some idiosyncrasies that are not modeled $\rightarrow$ omitted variable bias
- autocorrelation: temporal error aggregation processes we know from Time Series 1 (ARIMA) $\rightarrow$ biased s.e.
- heteroscedasticity: given the differences in units, likely unequal variances between them: $\sigma_i^2 \neq \sigma_j^2$ for units i and j (e.g. spending variation across U.S. and Eritrea) $\rightarrow$ biased s.e.
- errors may be correlated (not-independent) across:
    - spatially proximate units (they are similar or interacting)
    - units at the same time (common exogenous event at time t)

  $\rightarrow$ units not independent of one anther $\rightarrow$ biased s.e.

What is the most serious problem?

Unit Effects

- historical, contextual idiosyncrasies of units
- unit effects suggest that 'all else is not equal' or *ceteris non paribus*

Test whether unit variable explains variation in the DV:

`anova<-aov(DV~unit, data=D)`

`summary(anova)` if we observe low p-value (< 0.05), it show significant unit effects

TSCS set up

TSCS equation:

$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + a_i + u_{it}$

x_{it} are predictors that vary over time and across units
z_i are predictors that vary across units only. They are time-invariant!

The overall residual, normally ϵ_{it}, is split into two parts:

- a_i is unobserved variance of the DV that varies across units, but not over time = unit effects
- u_{it} is error that varies over time and across units

Solution 1) Fixed effects – time-demeaned

Deal with unit effects is by subtracting the mean values of the variables across countries over time. Start with basic TSCS:
$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + a_i + u_{it}$ (1)

Taking the mean of this equation over time produces this:
$\bar{y}_{it} = \beta_1 \bar{x}_{it} + \beta_2 z_i + a_i + \bar{u}_{it}$ (2)
(the over-time mean of z_i is z_i, and of a_i is a_i – they are time-invariant)

Subtract equation (2) from equation (1), we get the *time-demeaned equation*:
$y_{it} - \bar{y}_i = \beta_1(x_{it} - \bar{x}_i) + u_{it} - \bar{u}_i$ (3)

Equation (3) is the *fixed effects estimator*, also known as *within estimator*. It estimates over-time effects *within* units.

Solution 1) Fixed effects – dummy variables

Equivalent alternative:
$y_{it} = \beta_0 + \beta_1 x_{it} + \gamma C_i + u_{it}$ (4)

C_i are N dummy variables = 1 for unit i, otherwise 0.

We are modeling the unit effects as 'country dummies.' This is the *dummy variable estimator*.

It produces the same results as the fixed effects estimator.

Fixed effects: problems

- inefficiency: too many means or dummies estimated
- brutality: takes "general ignorance" (ϵ_{it}) and turns it into "specific ignorance" (a_i). We atheoreticly remove 'context'.
- what happened to z_i? Time-invariant variables cannot be estimated!
- fixed effects cannot estimate level effects, explaining the actual levels reached by the dependent variable

Solution 2) Random effects

Random effects, like in multi-level models, assume time observations as nested within units.
Unit effects a_i seen as ignorance which is illogical to model.

Random effects thus proceed from:
$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \epsilon_{it}$ (5)
$\epsilon_{it} = a_i + u_{it}$, thus including unit effects.
This is problematic...

Solution 2) Random effects

Random effects equation:
$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \epsilon_{it}$ (5)

moving unit effects into the residual introduces two problems:

- 1) autocorrelation: a_i is in the error at each time point t.
- 2) if a_i is correlated with predictors $x_{it} \rightarrow$ omitted variable bias.

Fix: transform the error to account for serial autocorrelation.

- 1) use the knowledge of the size of error variance caused by unit effects a_i.
- 2) we wave our hands, and ... assume that $Cov(a_i, x_{it}) = 0$. This is a very, very, very strong assumption...

The random effects estimator

Random effects equation:
$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \epsilon_{it}$ (5)

Next, define a transformation coefficient λ:
$\lambda = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_a^2}}$

The transformation of equation (5) is then:
$y_{it} - \lambda \bar{y}_i = \beta_0(1-\lambda) + \beta_1(x_{it} - \lambda \bar{x}_i) + \beta_2(z_i - \lambda z_i) + (\epsilon_{it} - \lambda \bar{\epsilon}_i)$ (6)

The random effects estimator

- the *quasi-demeaned* equation (6) is similar to the *demeaned* equation (3), but the means are multiplied by λ
- the transformation subtracts a fraction of the time average,
- the fraction depends on σ_u, σ_a and T.
- this means that we are using our knowledge of the *within* error variance σ_u^2 (estimated with fixed effects) and the total variance σ^2 (estimated with OLS).
- time-invariant predictors z_i are now possible.
- when $\lambda = 1$, the RE model is identical to an FE model.
- when $\lambda = 0$ the RE model is identical to OLS estimation.
- the bigger the variance of the unobserved effect, the closer RE is to FE, the smaller the variance of the unobserved effect, the closer RE is to OLS.

Panel Corrected Standard Errors (PCSE)

- Beck and Katz (1995): TSCS with OLS, but correct s.e. to make them robust to contemporaneous error correlation and autocorrelation
- PCSE have the same coefficients as OLS, but different s.e.
- the problem is that they deal with error issues only, esp. contemporaneous error correlation
- we can do this by modeling specific time shocks (e.g. dummy for Covid year)
- does not resolve unit effects!

Lagged Dependent Variable (LDV)

- To fix big problems, Beck and Katz propose the use of lagged dependent variable (LDV) in OLS model
- $y_{it} = \beta_0 + \alpha y_{t-1} + \beta_1 x_{it} + \beta_2 z_i + \epsilon_{it}$ (7)
- Here y_{t-1} is the lagged dependent variable
- LDV models history of y, doing two things in one:
- 1) models temporal dynamics (autoregression)
- 2) as it models history in y, it captures unit effects (past levels of y)
- But the control for unit effects is as strong as the dynamics in y. If α is very small (say < 0.5), then control for unit effects is also small.
- But if dynamics are strong and α is large, then LDV explains a lot of the variance, and our x variables have little leverage...

Integrated time series

- Need to consider temporal dynamics across each panel
- Key concern is trending, as all trends predict all trends
- Check data for stationarity with Dickey Fuller test

Estimating integrated time series

- best solution is differencing
- it models not the levels, but the differences between values in the series
- need to consider level of integration I(1), then use first difference

difference model looks like this:

$$\Delta y = \beta_0 + \beta_1 \Delta x1_{it} + \beta_2 \Delta x2_{it} + ... + \epsilon_{it}$$

Where Δy is the first difference in y: $\Delta y = y_t - y_{t-1}$ and so on.