Applying Panel Data Methods to Other Data Structures

Professor Jan Minx

Hertie School of Governance

Stats II, Lecture ?, ? 2015

Motivation

- Panel data methods can be used with data structures that do not involve time
- Hierarchical data structures contain clusters of observation which share commom characteristics
- When these characteristics are unobservable and correlated with other explanatory variables, pooled OLS will give us estimates that are biased and inefficient

Motivation

- Consider a geographical dataset that observes variables for small areas (in this case MSOAs, or Middle Layer Super Output Areas)
- Each small area belongs to a local authority
- If local authority attributes that we cannot observe affect our other variables, we will get biased and inefficient estimates using OLS

Pooled OLS

reg energy_consumption income_est
outtex, file(ols.tex) labels level detail legend key(stab) replace
est store ols

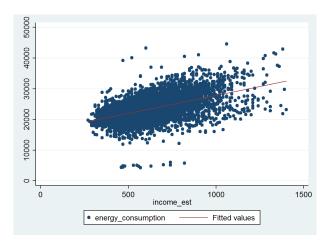
Table: Estimation results: regress

Variable	Coefficient	(Std. Err.)	
income_est	11.681**	(0.222)	
Intercept	16142.067**	(139.382)	
N	7133		

N	7	133	
R^2	0	.28	
F _(1,7131)	2766.551		
Significance levels :	† : 10%	*:5%	** : 1%

Pooled OLS

```
graph twoway (scatter energy_consumption income_est) ///
(lfit energy_consumption income_est)
```



Fixed Effects

xtset LA_CODE MSOA_CODE
xtreg energy_consumption income_est, fe
outtex, file(fe.tex) labels level detail legend key(stab) replace
est store fe

Table: Estimation results: xtreg

Variable	Coefficient	(Std. Err.)	
income_est	20.111**	(0.237)	
Intercept	11040.143**	(145.751)	
N	7133		
D ²	0.516		

Fixed Effects

```
xi: regress energy_consumption income_est i.LA_CODE
predict energy_consumption_fitted
separate energy_consumption_by(LA_CODE)
separate energy_consumption_fitted, by(LA_CODE)
graph twoway (scatter energy_consumption1-energy_consumption80 income_est) /// (line energy_consumption_fit
income_est), legend(off)
```

