MQE: Economic Inference from Data: Module 2: Fixed Effects

Claire Duquennois

6/9/2020

Module 2: Fixed Effects

- ▶ Data Structures
- ▶ Fixed Effects

-A simulation

-Fixed effects as demeaned data

-Thinking about variation

-Example: Crime and Unemployment

Controling for unobservables

We saw with AGG(2006) that even with many covariates, unobservables are a problem.

Certain types of data allow us to control for more of these unobservables by using fixed effects.

Example:

$$Income_i = \beta_0 + \beta_1 Schooling_i + \epsilon$$

β_1 cannot be interpreted as causal: big OVB problems, even with lots of control variables. Unlikely to have good measures of 'ability', 'enthusiasm', 'grit'...

What if I can control for unchanging individual characteristics?

Data Structures: Cross-Section

Individual	Income	Schooling	Female
1	22000	12	1
2	57000	16	1
. . .	. . .	. . .	. . .
N	15000	12	0

Each individual is observed once.

Data Structures: Panel Data

Individual	Income	Schooling	Female	Year
1	22000	12	1	2001
1	23000	12	1	2002
2	57000	16	1	2001
2	63000	17	1	2002
. . .	. . .	. . .	. . .	. . .
N	15000	12	0	2001
N	13000	12	0	2002

Each individual is observed multiple times.

Data Structures: Panel Data Subscripts

Unique observations must be identified by both the individual and time dimensons. . . notice the new subscripts:

$$Income_{it} = \beta_0 + \beta_1 Schooling_{it} + \epsilon.$$

Data Structures: Panel Data

Panel Data can be

-**balanced**: same number of observations for each unit

-**unbalanced**: some units are observed more often then others
(probably good to look into why)

Review: Indicator (Dummy) Variables

If I have multiple Female observation and multiple non-female observations I can control for the effect of being female on wages:

$$Income_{it} = \beta_0 + \beta_1 Schooling_{it} + \beta_2 Female_i + \epsilon.$$

Fixed Effects as Individual Indicator Variables

Indiv	Income	School	Female	Year	Indiv1	Indiv2	. . .	IndivN
1	22000	12	1	2007	1	0	0	0
1	23000	12	1	2008	1	0	0	0
2	57000	16	1	2007	0	1	0	0
2	63000	17	1	2008	0	1	0	0
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
N	15000	12	0	2007	0	0	0	1
N	13000	12	0	2008	0	0	0	1

Fixed Effects as Individual Indicator Variables

I can estimate:

$Inc_{it} = \beta_0 + \beta_1 School_{it} + \beta_2 Fem_i + \beta_{a1} Ind1_i + \beta_{a2} Ind2_i + ... + \beta_{aN-1} Ind(N-1)_i + \epsilon.$

What do the β_{ak} coefficients tell me?

Also:

-Why do the *IndN* indicators only have an *i* subscript?

-What is the implied assumption if *Fem* only has an *i* subscript?

-Why are there only (N-1) individual dummies?

Fixed Effects as Individual Indicator Variables

What will these individual controls control for?

-β_{a1} will control for the effect of being individual 1 on income that is not explained by that person's gender or schooling.

-Any **time invariant** characteristic that affects individual 1's income, such as ability, grit, enthusiasm. . . will be controlled for by adding this individual dummy variable.

-These controls are known as individual **fixed effects**.

For notational convenience:

$$Income_{it} = \beta_0 + \beta_1 Schooling_{it} + \beta_2 Female_i + \gamma_i + \epsilon.$$

Fixed Effects

With my panel data, what else can I control for?

$$Income_{it} = \beta_0 + \beta_1 Schooling_{it} + \beta_2 Female_i + \gamma_i + \tau_t + \epsilon.$$

-What is τ_t?

-What is this estimation equivalent to?