

Comparing Multilevel Models and Fixed Effects Regression

Andy Grogan-Kaylor

8 Apr 2021 14:21:19

Background

This example draws from the Stata documentation for the `xtreg` command.

Multilevel models for longitudinal data, and fixed effects regression provide two alternative methods for analyzing longitudinal data.

Briefly...

- Multilevel models use both within person and between person variation, and provide statistical control for observed variables that are included in the model.
- Fixed effect regressions use only within person variation. As a consequence, fixed effects regression is unable to provide parameter estimates for time invariant variables, even when they are included in the statistical model. Fixed effects regressions provide statistical controls for all time invariant variables, whether observed or unobserved.

Get The Data (use)

```
. clear all

. use https://www.stata-press.com/data/r16/nlswork, clear
(National Longitudinal Survey. Young Women 14-26 years of age in 1968)
```

Describe the Key Variables (describe)

```
. describe ln_w grade age race union south
```

variable name	storage type	display format	value label	variable label
ln_wage	float	%9.0g		ln(wage/GNP deflator)
grade	byte	%8.0g		current grade completed
age	byte	%8.0g		age in current year
race	byte	%8.0g	racelbl	race
union	byte	%8.0g		1 if union
south	byte	%8.0g		1 if south

Multilevel Model (mixed y x || id:)

The model uses within and between person variation. Estimates are provided for all variables. The model only controls for variables that are included in the model.

```
. mixed ln_w grade age i.race union south || idcode:
Performing EM optimization:
Performing gradient-based optimization:
Iteration 0:   log likelihood =  -5486.826
Iteration 1:   log likelihood =  -5486.826
Computing standard errors:
Mixed-effects ML regression              Number of obs      =    19,224
Group variable: idcode                   Number of groups   =     4,148
                                         Obs per group:
                                         min =             1
                                         avg =             4.6
                                         max =             12
                                         Wald chi2(6)       =    3471.83
                                         Prob > chi2        =     0.0000

Log likelihood =  -5486.826
```

ln_wage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
grade	.0781541	.0021992	35.54	0.000	.0738438	.0824644
age	.0137491	.0003907	35.19	0.000	.0129833	.0145149
race						
black	-.0405347	.0126091	-3.21	0.001	-.0652482	-.0158212
other	.0404357	.0508123	0.80	0.426	-.0591545	.140026
union	.1243977	.0065614	18.96	0.000	.1115375	.1372579
south	-.1019453	.0090188	-11.30	0.000	-.1196219	-.0842687
_cons	.3110752	.0314868	9.88	0.000	.2493622	.3727882

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
idcode: Identity				
var(_cons)	.0998265	.0027427	.0945931	.1053494
var(Residual)	.0691308	.0007996	.0675813	.0707159

```
LR test vs. linear model: chibar2(01) = 8473.10      Prob >= chibar2 = 0.0000

. est store MLM
```

Fixed Effects Regression (xtreg y x, i(id) fe)

The model uses only within person variation. Estimates are only provided for within person change over time. The model controls for all time invariant variables whether observed or unobserved.

```
. xtreg ln_w grade age i.race union south, i(idcode) fe
note: grade omitted because of collinearity
note: 2.race omitted because of collinearity
note: 3.race omitted because of collinearity
Fixed-effects (within) regression              Number of obs      =    19,224
Group variable: idcode                   Number of groups   =     4,148
R-sq:                                     Obs per group:
      within = 0.0983                               min =             1
      between = 0.0712                               avg =             4.6
```

```

overall = 0.0847
corr(u_i, Xb) = 0.0599
max = 12
F(3,15073) = 547.57
Prob > F = 0.0000

```

ln_wage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
grade	0	(omitted)				
age	.0153807	.0004154	37.03	0.000	.0145665	.0161949
race						
black	0	(omitted)				
other	0	(omitted)				
union	.1034851	.0070913	14.59	0.000	.0895853	.1173849
south	-.0759973	.0135167	-5.62	0.000	-.1024917	-.0495029
_cons	1.279453	.0143464	89.18	0.000	1.251332	1.307573
sigma_u	.41784013					
sigma_e	.2618843					
rho	.71796552	(fraction of variance due to u_i)				

```

F test that all u_i=0: F(4147, 15073) = 9.60
Prob > F = 0.0000

```

```

. est store FE

```

Compare The Two Sets of Estimates (estimates table)

```

. est table MLM FE, star equations(1)

```

Variable	MLM	FE
#1		
grade	.07815409***	(omitted)
age	.01374911***	.01538067***
race		
black	-.04053471**	(omitted)
other	.04043574	(omitted)
union	.12439767***	.10348514***
south	-.10194526***	-.07599732***
_cons	.31107518***	1.2794525***
lns1_1_1		
_cons	-1.1521609***	
lnsig_e		
_cons	-1.3358773***	

```

legend: * p<0.05; ** p<0.01; *** p<0.001

```