ECON 1190: Applied Econometrics 2: Module 2: Fixed Effects

Claire Duquennois

Module 2: Fixed Effects

- Data Structures
- Fixed Effects
- A simulation
- Fixed effects as demeaned data
- Thinking about variation
- Example: Crime and Unemployment

Controlling 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 dimensions... 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.

If I have multiple Married observation and multiple non-married observations I can control for the effect of being married on wages.

$$Income_{it} = \beta_0 + \beta_1 Schooling_{it} + \beta_2 Female_i + \beta_3 Married_{it} + \epsilon.$$

Fixed Effects as Individual Indicator Variables

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

Fixed Effects as Individual Indicator Variables

I can estimate:

$$Inc_{it} = \beta_0 + \beta_1 Sch_{it} + \beta_2 Fem_i + \beta_3 Mar_{it} + \beta_{a1} Ind1_i + \beta_{a2} Ind2_i + ... + \beta_{aN-1} Ind(N-1)_i + \epsilon.$$

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?
- Will I be able to estimate the β_2 on Fem_i?
- Will I be able to estimate the β_3 on Mar_{it}?

Fixed Effects as Individual Indicator Variables

What will these individual controls control for?

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 marital status 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:

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

Fixed Effects

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

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

- What is λ_t?
- What is this estimation equivalent to?

Fixed Effects and variation:

If I estimate

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

Where (who?) is my identifying variation coming from for estimating β_1?

Does this matter? How does it change our interpretation of the estimate?

A Simulation:

You are a principle of a small school composed of four classrooms. You have just implemented a new option available to teachers for students to spend some small group reading time with a para-educator. You would like to know how this reading time is affecting reading scores.

You have data for ten students in each class that tells you:

- the class the student is in
- whether they participated in small group reading
- their reading score.

Generating Simulated Data

I will work with a simulated dataset to show how the use of fixed effects can help us recover the true treatment effect.

I start by loading the dplyr package and "setting the seed":

```
#install.packages("dplyr")
#install.packages("lfe")
#install.packages("stargazer")

library(dplyr)
library(lfe)
library(stargazer)
library(fastDummies)

set.seed(1999)
```

A Simulation:

I generate a vector of class identifiers and a random error term.

```
class<-c(1,2,3,4)
class<-rep(class, each=10)
scores<-as.data.frame(class)
scores$error<-rnorm(40, mean=0, sd=5)

#note: if you are not working in markdown you would just write head(scores)
knitr::kable(head(scores))
```

class	error
1	3.6633624
1	-0.1891486
1	6.0150457
1	7.3490101
1	0.6684515
1	2.5991362

A Simulation:

I simulate some selection into treatment. The probability of getting treated is

- 0.8 for students in classrooms 3 and 4
- 0.2 in classrooms 1 and 2.

```
scores$treat1<-rbinom(40,1,0.2)
scores$treat2<-rbinom(40,1,0.8)
scores$treat[scores$class%in%c(1,2)]<-scores$treat1[scores$class%in%c(1,2)]
scores$treat[scores$class%in%c(3,4)]<-scores$treat2[scores$class%in%c(3,4)]

knitr::kable(head(scores))
```

class	error	treat1	treat2	treat
1	3.6633624	0	1	0
1	-0.1891486	0	1	0
1	6.0150457	0	1	0
1	7.3490101	1	0	1
1	0.6684515	0	1	0
1	2.5991362	0	1	0

A Simulation:

I drop unneeded variables and generate a dummy variable for each classroom

```
scores<-scores%>%select(class,error,treat)
scores <- dummy_cols(scores, select_columns = "class")

knitr::kable(head(scores))
```

class	error	treat	class_1	class_2	class_3	class_4
1	3.6633624	0	1	0	0	0
1	-0.1891486	0	1	0	0	0
1	6.0150457	0	1	0	0	0
1	7.3490101	1	1	0	0	0
1	0.6684515	0	1	0	0	0
1	2.5991362	0	1	0	0	0

A Simulation:

Finally! I simulate the DGP (Data Generating Process):

- The true treatment effect = 15
- students in classrooms 1 and 2 have higher reading scores
- students in classrooms 3 and 4 have lower reading scores.

```
scores$classFE<-NA
scores$classFE[scores$class_1==1]<-0
scores$classFE[scores$class_2==1]<-10
scores$classFE[scores$class_3==1]<-(-30)
scores$classFE[scores$class_4==1]<-(-35)


scores$score<-80+15*scores$treat+scores$classFE+scores$error

knitr::kable(head(scores))
```

class	error	treat	class_1	class_2	class_3	class_4	classFE	score
1	3.6633624	0	1	0	0	0	0	83.66336
1	-0.1891486	0	1	0	0	0	0	79.81085
1	6.0150457	0	1	0	0	0	0	86.01505
1	7.3490101	1	1	0	0	0	0	102.34901
1	0.6684515	0	1	0	0	0	0	80.66845
1	2.5991362	0	1	0	0	0	0	82.59914

A Simulation:

I estimate three specifications. The first:

$$Score_{ci} = \beta_0 + \beta_1 Treat_{ci} + \epsilon$$

```
nofe<-felm(score~treat,scores)
```

A Simulation:

The second:

$$Score_{ci} = \beta_0 + \beta_1 Treat_{ci} + \beta_2 Class2_c + \beta_3 Class3_c + \beta_4 Class4_c + \epsilon$$

```
dummies<-felm(score~treat+class_2+class_3+class_4, scores)
```

A Simulation:

The third:

$$Score_{ci} = \beta_0 + \beta_1 Treat_{ci} + \kappa_c + \epsilon$$

where κ_c is a classroom fixed effect.

```
fe<-felm(score~treat|class,scores)
```

A Simulation:

```
stargazer(nofe, dummies, fe,header=FALSE, type='latex')
```

Table 8

	Dependent variable:		
	score		
	(1)	(2)	(3)
treat	-1.097	15.881^{***}	15.881^{***}
	(5.949)	(1.914)	(1.914)
class_2		9.338^{***}	
		(2.345)	
class_3		-32.126^{***}	
		(2.437)	
class_4		-35.624^{***}	
		(2.673)	
Constant	75.284^{***}	80.974^{***}	
	(4.310)	(1.680)	
Observations	40	40	40
R^2	0.001	0.930	0.930
Adjusted R^2	-0.025	0.922	0.922
Residual Std. Error	18.789 (df = 38)	5.172 (df = 35)	5.172 (df = 35)
Note:			*p<0.1; **p<0.05; ***p<0.01

A Simulation:

Recall: $\beta_1 = 15$ (the true treatment effect) $\Rightarrow \hat{\beta}_1^{nofe}$ is very biased!

Why?

$\Rightarrow$ Top Hat

A Simulation:

Recall: $\beta_1 = 15$ (the true treatment effect) $\Rightarrow \hat{\beta}_1^{nofe}$ is very biased!

Why?

- The classes are an important omitted variable: $cor(Score, Class3/4) < 0$ and $cor(Treat, Class3/4) > 0$ creating substantial downward bias.

We can correct for this in two (equivalent) ways:

- adding the dummy variables for the class to the regression,
- adding a class fixed effect.

Either approach returns an identical unbiased estimate such that $E[\hat{\beta}_1] = \beta_1$.

Fixed Effects as Demeaned Data:

Fixed effect estimates are also known as the **within estimator**, because it identifies β using within-unit variation.

$\Rightarrow$ we only using the variation that exists **within the classroom** to estimate the treatment effect.

This is the equivalent of "correcting" our data by demeaning each observation using it's classroom mean, so that the corrected data represents deviations from the classroom mean.

Fixed Effects as Demeaned Data:

Our fixed effect estimation is

$$y_{ci} = \beta_1 x_{ci} + \kappa_c + \epsilon_{ci}$$

For each class, the average across the students is

$$\bar{y}_i = \beta_1 \bar{x}_i + \kappa_c + \bar{\epsilon}_i$$

Subtracting this from the fixed effect model gives

$$y_{ic} - \bar{y}_i = \beta_1 (x_{ic} - \bar{x}_i) + (\epsilon_{ic} - \bar{\epsilon}_i)$$

Fixed Effects as Demeaned Data:

1. Calculate the mean score, and the mean treatment, in each classroom

```
#getting the mean score in each classroom
cl_mean<-scores %>%
    group_by(class) %>%
    summarize(Classmean = mean(score, na.rm=TRUE), treatmean=mean(treat, na.rm=TRUE))

knitr::kable(head(cl_mean))
```

class	Classmean	treatmean
1	84.14982	0.2
2	96.66454	0.4
3	58.37619	0.6
4	59.64236	0.9

Fixed Effects as Demeaned Data:

2. Merging the means into full data

```
scores<-left_join(scores, cl_mean, by = "class")

knitr::kable(head(scores))
```

class	error	treat	class_1	class_2	class_3	class_4	classFE	score	Classmean	treatmean
1	3.6633624	0	1	0	0	0	0	83.66336	84.14982	0.2
1	-0.1891486	0	1	0	0	0	0	79.81085	84.14982	0.2
1	6.0150457	0	1	0	0	0	0	86.01505	84.14982	0.2
1	7.3490101	1	1	0	0	0	0	102.34901	84.14982	0.2
1	0.6684515	0	1	0	0	0	0	80.66845	84.14982	0.2
1	2.5991362	0	1	0	0	0	0	82.59914	84.14982	0.2

Fixed Effects as Demeaned Data:

3. Calculating the demeaned score

```
scores$demeansc<-scores$score-scores$Classmean
scores$demeantrt<-scores$treat-scores$treatmean

knitr::kable(head(scores[,c("class","treat","score","Classmean","treatmean","demeansc","demeantrt")]))
```

class	treat	score	Classmean	treatmean	demeansc	demeantrt
1	0	83.66336	84.14982	0.2	-0.4864594	-0.2
1	0	79.81085	84.14982	0.2	-4.3389704	-0.2
1	0	86.01505	84.14982	0.2	1.8652238	-0.2
1	1	102.34901	84.14982	0.2	18.1991883	0.8
1	0	80.66845	84.14982	0.2	-3.4813704	-0.2
1	0	82.59914	84.14982	0.2	-1.5506856	-0.2

Fixed Effects as Demeaned Data:

4. Running the basic regression on the demeaned scores

```
regdemean<-felm(demeansc~demeantrt, scores)
stargazer( fe, regdemean,header=FALSE, type='latex', omit.stat=c("all" ))
```

Table 12

	Dependent variable:	
	score	demeansc
	(1)	(2)
treat	15.881***	
	(1.914)	
demeantrt		15.881***
		(1.837)
Constant		0.000
		(0.785)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Careful: the standard errors on the demeaned regression are incorrect because the cases are not independent of each other.

Variation

What would happen if none of the students in classes 1 and 2 went to the small reading group and all of the students in class 3 and 4 did?

Variation

Creating a new treatment variable to reflect this:

```
scores$treat2[scores$class%in%c(1,2)]<-0
scores$treat2[scores$class%in%c(3,4)]<-1

scores$score2<-80+15*scores$treat2+scores$classFE+scores$error

nofe2<-felm(score2~treat2,scores)
dummies2<-felm(score2~treat2+class_2+class_3+class_4, scores)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
#fe2<-felm(score2~treat2|class,scores)
```

Variation

```
stargazer(nofe2, dummies2, header=FALSE,  type='latex')
```

Table 13

	Dependent variable:	
	score2	
	(1)	(2)
treat2	-23.148^{***}	-20.007^{***}
	(1.951)	(2.288)
class_2		9.515^{***}
		(2.288)
class_3		3.234
		(2.288)
class_4		
Constant	85.907^{***}	81.150^{***}
	(1.380)	(1.618)
Observations	40	40
R^2	0.787	0.862
Adjusted R^2	0.782	0.850
Residual Std. Error	6.171 (df = 38)	5.115 (df = 36)
Note:		*p<0.1; **p<0.05; ***p<0.01

Example: Crime and Unemployment

You are interested in the relationship between unemployment and crime.

You have data on the crime and unemployment rates for 46 cities for 1982 and 1987.

I start by using the data from the 1987 cross section and run the following simple regression of the crime rate on unemployment,

$$crimerate_i = \beta_0 + \beta_1 unemployment_i + \epsilon$$

Example: Crime and Unemployment

```
#install.packages("wooldridge")
library(wooldridge)
```

```
## Warning: package 'wooldridge' was built under R version 4.1.3
```

```
#note: this dataset comes from the wooldridge textbook. Conveniently there is an R package that
#includes all the wooldridge datasets.

crime<-data('crime2')
crime<-crime2
```

Example: Crime and Unemployment

```
regcrime<-felm(crmrte~unem, crime[crime$year=="87",])
summary(regcrime)
```

```
##
## Call:
##    felm(formula = crmrte ~ unem, data = crime[crime$year == "87",      ])
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -57.55 -27.01 -10.56  18.01  79.75
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  128.378     20.757   6.185  1.8e-07 ***
## unem          -4.161      3.416  -1.218     0.23
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 34.6 on 44 degrees of freedom
## Multiple R-squared(full model): 0.03262   Adjusted R-squared: 0.01063
## Multiple R-squared(proj model): 0.03262   Adjusted R-squared: 0.01063
## F-statistic(full model):1.483 on 1 and 44 DF, p-value: 0.2297
## F-statistic(proj model): 1.483 on 1 and 44 DF, p-value: 0.2297
```

Example: Crime and Unemployment

Weird.

The culprit? Probably omitted variables.

Reflex: lets add controls for more observable city characteristics: the area of the city, if the city is in the west, police officers per square mile, expenditure on law enforcement, per capita income...

$$crmrte_i = \beta_0 + \beta_1 unemp_i + \beta_2 area_i + \beta_3 west_i + \beta_4 offarea_i + \beta_5 lawexp_i + \beta_6 pcinc_i + \epsilon$$

Example: Crime and Unemployment

```
regcrime2<-felm(crmrte~unem+area+west+offarea+lawexpc+pcinc, crime[crime$year=="87",])
summary(regcrime2)
```

```
##
## Call:
##    felm(formula = crmrte ~ unem + area + west + offarea + lawexpc +      pcinc, data = crime[crime$year
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -50.847 -21.511  -6.829  18.940  75.114
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 140.06017   51.16000   2.738  0.00927 **
## unem         -6.70024    3.71634  -1.803  0.07913 .
## area          0.05867    0.04757   1.233  0.22491
## west        -21.96336   12.27535  -1.789  0.08135 .
## offarea      -0.11442    0.66876  -0.171  0.86504
## lawexpc       0.02137    0.01859   1.149  0.25736
## pcinc        -0.00185    0.00352  -0.526  0.60215
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 34.27 on 39 degrees of freedom
## Multiple R-squared(full model): 0.1587   Adjusted R-squared: 0.02932
## Multiple R-squared(proj model): 0.1587   Adjusted R-squared: 0.02932
## F-statistic(full model):1.227 on 6 and 39 DF, p-value: 0.3138
## F-statistic(proj model): 1.227 on 6 and 39 DF, p-value: 0.3138
```

Example: Crime and Unemployment

Even weirder.

So many potential omitted variables...

What if we can **capture all unobserved, time invariant factors** about a city that might affect crime rates?

Use data for 1987 and 1982, and add city **fixed effects,**(α_i).

$$\begin{aligned} crmrte_{it} = &\beta_0 + \beta_1 unemp_{it} \\ &+ \beta_2 area_i + \beta_3 west_i + \beta_4 offarea_{it} + \beta_5 lawexp_{it} + \beta_6 pcinc_{it} \\ &+ \alpha_i + \epsilon \end{aligned}$$

Example: Crime and Unemployment

```r
#note: the data does not have a unique city identifier. I am assuming the area of the city is
#1)time-invariant and
#2) uniquely identifies the 46 cities.
#The line of code below generates a unique identifier

crime <- transform(crime,city=as.numeric(factor(area)))

#I check that my assumptions were correct by seeing if I have 2 observations for 46 cities.
table(crime$city)
```

```
##
##  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
##  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2
## 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46
##  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2
```

Example: Crime and Unemployment

```
regcrime3<-felm(crmrte~unem+area+west+offarea+lawexpc+pcinc|city, crime)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
summary(regcrime3)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
##
## Call:
##    felm(formula = crmrte ~ unem + area + west + offarea + lawexpc +      pcinc | city, data = crime)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -27.36  -6.85   0.00   6.85  27.36
##
## Coefficients:
##          Estimate Std. Error t value Pr(>|t|)
## unem     1.491297   0.795972   1.874   0.0680 .
## area          NaN         NA     NaN      NaN
## west          NaN         NA     NaN      NaN
## offarea  1.348882   1.805672   0.747   0.4592
## lawexpc -0.005076   0.013915  -0.365   0.7171
## pcinc    0.003821   0.001644   2.324   0.0251 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.66 on 42 degrees of freedom
## Multiple R-squared(full model): 0.8887   Adjusted R-squared: 0.7588
## Multiple R-squared(proj model):  0.18   Adjusted R-squared: -0.7767
## F-statistic(full model):6.843 on 49 and 42 DF, p-value: 1.888e-09
## F-statistic(proj model): 1.537 on 6 and 42 DF, p-value: 0.19
```

Example: Crime and Unemployment

Interpret the coefficient on unemployment.

Why were we not able to estimate a coefficient for $area_i$ and $west_i$?

$\Rightarrow$ Top Hat

Example: Crime and Unemployment

Example: Crime and Unemployment

Suppose I am concerned about how national factors could be affecting all cities simultaneously.

Add a year fixed effect, (λ_t).

$$\begin{aligned}
crmrte_{it} =& \beta_0 + \beta_1 unemp_{it} \\
& + \beta_2 area_i + \beta_3 west_i + \beta_4 offarea_{it} + \beta_5 lawexp_{it} + \beta_6 pcinc_{it} \\
& + \alpha_i + \lambda_t + \epsilon
\end{aligned}$$

Example: Crime and Unemployment

```r
regcrime4<-felm(crmrte~unem+area+west+offarea+lawexpc+pcinc|city+year, crime)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```r
summary(regcrime4)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
##
## Call:
##    felm(formula = crmrte ~ unem + area + west + offarea + lawexpc +      pcinc | city + year, data = cr
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -23.641  -7.441   0.000   7.441  23.641
##
## Coefficients:
##          Estimate Std. Error t value Pr(>|t|)
## unem     2.931904   1.133562   2.586   0.0133 *
## area          NaN         NA     NaN      NaN
## west          NaN         NA     NaN      NaN
## offarea  1.838022   1.785312   1.030   0.3093
## lawexpc -0.006982   0.013632  -0.512   0.6113
## pcinc   -0.005697   0.005683  -1.002   0.3220
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.31 on 41 degrees of freedom
## Multiple R-squared(full model): 0.8964   Adjusted R-squared:  0.77
## Multiple R-squared(proj model): 0.1709   Adjusted R-squared: -0.8402
## F-statistic(full model):7.094 on 50 and 41 DF, p-value: 1.405e-09
## F-statistic(proj model): 1.408 on 6 and 41 DF, p-value: 0.2347
```

Example: Crime and Unemployment

Is this estimate causal?

$\Rightarrow$ Top Hat

Example: Crime and Unemployment

Is this estimate causal?

What we have controlled for?

- The city fixed effects: controls for any time invariant factors that always affect the crime rates in a city in a similar way. Such as...
- The time fixed effects: controls for any patterns that are common to all cities in a given year. Such as...
- In addition to this we are also controlling for some observable time varying variables: officers in an area, law enforcement expenditures per capita and income per capita.

So are these estimates causal?

What kind of omitted variables should we still be concerned about?
$\Rightarrow$ Top Hat

Example: Crime and Unemployment

Any variable that changes within a city across years and that is correlated with both unemployment and crime rates could still be biasing our results.

This could be things like school funding, decriminalization of marijuana, housing costs... to name just a few.

Example: Crime and Unemployment Variation

Suppose I get ambitious and want to control for all these factors as well.

I decide I am going to generate a city-by-year fixed effect, (γ_{it}), to control for these time variant omitted variables. I estimate,

$$\begin{aligned} crmrte_{it} = &\beta_0 + \beta_1 unemp_{it} \\ &+ \beta_2 area_i + \beta_3 west_i + \beta_4 offarea_{it} + \beta_5 lawexp_{it} + \beta_6 pcinc_{it} \\ &+ \gamma_{it} + \epsilon \end{aligned}$$

Example: Crime and Unemployment Variation

```
crime$city_year<-paste(crime$city, crime$year, sep="_")

#Note: the following regression will not run!
#regcrime5<-felm(crmrte~unem+area+west+offarea+lawexpc+pcinc|city_year, crime)
#summary(regcrime5)
```

Why?!?

What kind of data would I need to be able to estimate this?