

# Panel Data

Kaden Grace   Lucas Gurgel   John Paul Lynn

PS 4 Presentation, May 9, 2023



THE UNIVERSITY OF  
TENNESSEE  
KNOXVILLE

# Q1 - Summary of Data

Variable	mean	sd	min	max
Years of Experience	19.85378	10.96637	1	51
Weeks Worked.	46.81152	5.129098	5	52
Blue-Collar Occupation	.5111645	.4999354	0	1
Working in Industry	.3954382	.4890033	0	1
South Resident	.2902761	.4539442	0	1
Standt. Metrop. Stat Area	.6537815	.475821	0	1
Marital Status	.8144058	.3888256	0	1
Gender (1=female)	.112605	.3161473	0	1
Wage Set by Union	.3639856	.4812023	0	1
Years of Education	12.84538	2.787995	4	17
Black	.0722689	.2589637	0	1
log Wage	6.676346	.4615122	4.60517	8.537
id	298	171.7821	1	595
t	4	2.00024	1	7
t= 1 to 7 (dummies)			0	1
exp2	514.405	496.9962	1	2601
Observations	4165			

## Q2 - List of All Variables

. codebook, compact

Variable	Obs	Unique	Mean	Min	Max	Label
exp	4165	51	19.85378	1	51	Years of Experience
wks	4165	41	46.81152	5	52	Weeks Worked.
occ	4165	2	.5111645	0	1	Blue-Collar Occupation
ind	4165	2	.3954382	0	1	Working in Industry
south	4165	2	.2902761	0	1	South Resident
smsa	4165	2	.6537815	0	1	Standt. Metrop. Stat Area
ms	4165	2	.8144058	0	1	Marital Status
fem	4165	2	.112605	0	1	Gender (1=female)
union	4165	2	.3639856	0	1	Wage Set by Union
ed	4165	14	12.84538	4	17	Years of Education
blk	4165	2	.0722689	0	1	Black
lwage	4165	1017	6.676346	4.60517	8.537	log Wage
id	4165	595	298	1	595	
t	4165	7	4	1	7	
tdum1	4165	2	.1428571	0	1	t== 1.0000
tdum2	4165	2	.1428571	0	1	t== 2.0000
tdum3	4165	2	.1428571	0	1	t== 3.0000
tdum4	4165	2	.1428571	0	1	t== 4.0000
tdum5	4165	2	.1428571	0	1	t== 5.0000
tdum6	4165	2	.1428571	0	1	t== 6.0000
tdum7	4165	2	.1428571	0	1	t== 7.0000
exp2	4165	51	514.405	1	2601	

## Q3 - Individual Identifier and Time Identifier

We set our data set as panel data through the command below ("id" stands for the individual and "t" stands for time):

```
. xtset id t
```

```
Panel variable: id (strongly balanced)
```

```
Time variable: t, 1 to 7
```

```
Delta: 1 unit
```



## Q4 - Panel Data Description

We may have a description of our panel data through the command `xtdescribe`:

```
. xtdescribe
```

```
id: 1, 2, ..., 595          n =      595
t: 1, 2, ..., 7             T =       7
Delta(t) = 1 unit
Span(t) = 7 periods
(id*t uniquely identifies each observation)
```

```
Distribution of T_i:  min      5%      25%      50%      75%      95%      max
                    7         7         7         7         7         7         7
```

Freq.	Percent	Cum.	Pattern
595	100.00	100.00	1111111
595	100.00		XXXXXXX

## Q4 - Panel Data Description

Main takeaways:

- (i) id is the identification of individuals.
- (ii) t is the time period that in this database goes from 1 to 7.
- (iii) Balanced means that for each individual we have the same number of observation.

## Q5 - Regressions RE and FE

$$\log Wage_{it} = \alpha_i + \beta_1 * exp_{it} + \beta_2 * wks_{it} + \beta_3 * exp_{it}^2 + \mu_{it}$$

	model_RE	model_FE
Years of Experience	0.0919*** (0.00)	0.1138*** (0.00)
Weeks Worked.	0.0009 (0.00)	0.0008 (0.00)
exp2	-0.0008*** (0.00)	-0.0004*** (0.00)
Constant	5.1980*** (0.05)	4.5964*** (0.04)

## Q5 - Hausman Test

```
. hausman model_FE model_RE, sigmamore
```

	— Coefficients —		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
	(b) model_FE	(B) model_RE		
exp	.1137879	.0919066	.0218813	.0011118
wks	.0008359	.0008628	-.0000269	.0000936
exp2	-.0004244	-.0007519	.0003275	.0000248

b = consistent under  $H_0$  and  $H_a$ ; obtained from xtreg

B = inconsistent under  $H_a$ , efficient under  $H_0$ ; obtained from xtreg

Test:  $H_0$ : difference in coefficients not systematic

```
chi2(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
        =      1457.90
Prob>chi2 =      0.0000
```

Given our  $Prob > chi2 = 0.000$ , we reject  $H_a$ . This tells us that the Random effects will be inconsistent, thus Fixed effects is more appropriate in this case.



## Q5 - Hausman Test

	$H_0$ Not Rejected	$H_0$ Rejected
RE	Consistent, Efficient	Inconsistent
FE	Consistent, Inefficient	Consistent



## Q6 - Fonseca and Van Doornik, 2022

- (i) Outcome variables: wage and allocation of high and low skill workers.
- (ii)  $Reform_t$  corresponds to  $d2_t$  and  $HighEnforcement_c$  corresponds to  $prog_i$ .
- (iii) Regression equation:

$$g(Y_{icst}) = \beta_1 Reform_t * HighEnforcement_c + \beta_2 X_{it} + \kappa_i + \theta_{st} + \epsilon_{icst}$$

(iv) Three main results:

- ➊ Increased credit access causes firms to hire relatively more skilled labor (4 percent increase in skilled labor share),
- ➋ Increased credit access leads to a higher wage for skilled labor (3.8 percent increase in skill premium), and
- ➌ The causal mechanism is complementarity between capital and skilled labor.