



University of
Nottingham
UK | CHINA | MALAYSIA

Autumn 2018

Econometric Theory I

Computer Lab Class II

Juergen Amann

juergen.amann@nottingham.ac.uk

Wednesday 12:00 - 13:00, C42 SCGB

Last week

- Using `earnings.xls` we analysed the effect of education and work experience on earnings.

Last week

- Using `earnings.xls` we analysed the effect of education and work experience on earnings.
- We estimated

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

where

- earnings_i : earnings of individual i in USD per hour
- educ_i : education of individual i as in years
- workexp_i : work experience of individual i in years
- u_i : idiosyncratic error of i .

Last week

- Using earnings.xls we analysed the effect of education and work experience on earnings.
- We estimated

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

where

- earnings_i : earnings of individual i in USD per hour
 - educ_i : education of individual i as in years
 - workexp_i : work experience of individual i in years
 - u_i : idiosyncratic error of i .
- We found educ and workexp to be highly significant and positive.

This week

Two exercises:

- **Exercise 1:** Examine the evidence of a gender pay gap.

This week

Two exercises:

- **Exercise 1:** Examine the evidence of a gender pay gap.
 - Dummy variables.

This week

Two exercises:

- **Exercise 1:** Examine the evidence of a gender pay gap.
 - Dummy variables.
 - What does it mean when we '*control for (a) variable(s)*'?

This week

Two exercises:

- **Exercise 1:** Examine the evidence of a gender pay gap.
 - Dummy variables.
 - What does it mean when we '*control for (a) variable(s)*'?
- **Exercise 2:** Examine the determinants of educational attainment.

This week

Two exercises:

- **Exercise 1:** Examine the evidence of a gender pay gap.
 - Dummy variables.
 - What does it mean when we '*control for (a) variable(s)*'?
- **Exercise 2:** Examine the determinants of educational attainment.
 - Joint parameter hypothesis testing of coefficients; e.g.
 $H_0 : \beta_2 = \beta_3$.

This week

Two exercises:

- **Exercise 1:** Examine the evidence of a gender pay gap.
 - Dummy variables.
 - What does it mean when we '*control for (a) variable(s)*'?
- **Exercise 2:** Examine the determinants of educational attainment.
 - Joint parameter hypothesis testing of coefficients; e.g.
 $H_0 : \beta_2 = \beta_3$.
 - Restricted Least Squares.

Exercise 1

- We examine evidence of a gender pay gap.

Exercise 1

- We examine evidence of a gender pay gap.
- Our data set `earnings2.xls` contains the same variables as `earnings.xls` plus two **dummy variables**:

Exercise 1

- We examine evidence of a gender pay gap.
- Our data set `earnings2.xls` contains the same variables as `earnings.xls` plus two **dummy variables**:
 - $\text{female}_i = \begin{cases} 1 & \text{if individual } i \text{ is female} \\ 0 & \text{otherwise} \end{cases}$

Exercise 1

- We examine evidence of a gender pay gap.
- Our data set `earnings2.xls` contains the same variables as `earnings.xls` plus two **dummy variables**:
 - $\text{female}_i = \begin{cases} 1 & \text{if individual } i \text{ is female} \\ 0 & \text{otherwise} \end{cases}$
 - $\text{male}_i = \begin{cases} 1 & \text{if individual } i \text{ is male} \\ 0 & \text{otherwise} \end{cases}$

Exercise 1

- We examine evidence of a gender pay gap.
- Our data set `earnings2.xls` contains the same variables as `earnings.xls` plus two **dummy variables**:
 - $\text{female}_i = \begin{cases} 1 & \text{if individual } i \text{ is female} \\ 0 & \text{otherwise} \end{cases}$
 - $\text{male}_i = \begin{cases} 1 & \text{if individual } i \text{ is male} \\ 0 & \text{otherwise} \end{cases}$
- We want to know if there is a significant difference in earnings between men and women **controlling for education and work experience**.

Exercise 1

- Launch Stata.
- Download and import data `earnings2.xls` into Stata.

Exercise 1

- Launch Stata.
- Download and import data `earnings2.xls` into Stata.
- First, estimate last week's regression again:

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

Exercise 1

- Launch Stata.
- Download and import data earnings2.xls into Stata.
- First, estimate last week's regression again:

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

Type: `regress EARNINGS EDUC WORKEXP`

Exercise 1

- Launch Stata.
- Download and import data `earnings2.xls` into Stata.
- First, estimate last week's regression again:

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

Type: `regress EARNINGS EDUC WORKEXP`

- How can we use the dummy variables to examine whether there is evidence of a gender pay gap?

Exercise 1

- Launch Stata.
- Download and import data `earnings2.xls` into Stata.
- First, estimate last week's regression again:

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

Type: `regress EARNINGS EDUC WORKEXP`

- How can we use the dummy variables to examine whether there is evidence of a gender pay gap?

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + \beta_4 \text{female}_i + u_i$$

Exercise 1

- Launch Stata.
- Download and import data earnings2.xls into Stata.
- First, estimate last week's regression again:

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + u_i$$

Type: `regress EARNINGS EDUC WORKEXP`

- How can we use the dummy variables to examine whether there is evidence of a gender pay gap?

$$\text{earnings}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{workexp}_i + \beta_4 \text{female}_i + u_i$$

Type: `regress EARNINGS EDUC WORKEXP FEMALE`

Exercise 1

Stata Code

```
. regress EARNINGS EDUC WORKEXP
```

Exercise 1

Stata Code

```
. regress EARNINGS EDUC WORKEXP
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.678125	.2336497	11.46	0.000	2.219146	3.137105
WORKEXP	.5624327	.1285136	4.38	0.000	.3099816	.8148837
_cons	-26.48501	4.27251	-6.20	0.000	-34.87789	-18.09213

Exercise 1

Stata Code

```
. regress EARNINGS EDUC WORKEXP
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.678125	.2336497	11.46	0.000	2.219146	3.137105
WORKEXP	.5624327	.1285136	4.38	0.000	.3099816	.8148837
_cons	-26.48501	4.27251	-6.20	0.000	-34.87789	-18.09213

```
. regress EARNINGS EDUC WORKEXP FEMALE
```


Exercise 1

Stata Code

```
. regress EARNINGS EDUC WORKEXP
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.678125	.2336497	11.46	0.000	2.219146	3.137105
WORKEXP	.5624327	.1285136	4.38	0.000	.3099816	.8148837
_cons	-26.48501	4.27251	-6.20	0.000	-34.87789	-18.09213

```
. regress EARNINGS EDUC WORKEXP FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.591137	.2285497	11.34	0.000	2.142174	3.0401
WORKEXP	.4056773	.1288199	3.15	0.002	.1526236	.658731
FEMALE	-5.90905	1.113972	-5.30	0.000	-8.097337	-3.720764
_cons	-19.69195	4.36076	-4.52	0.000	-28.25822	-11.12567

Exercise 1

- In the previous regression FEMALE captures the difference between male and female earnings.

Exercise 1

- In the previous regression FEMALE captures the difference between male and female earnings.
- We find a **negative and significant** effect for this parameter:

Exercise 1

- In the previous regression FEMALE captures the difference between male and female earnings.
- We find a **negative and significant** effect for this parameter:
 - Controlling for education and work experience, women earn significantly **less** than men.

Exercise 1

- In the previous regression FEMALE captures the difference between male and female earnings.
- We find a **negative and significant** effect for this parameter:
 - Controlling for education and work experience, women earn significantly **less** than men.
 - More specifically:

$$\widehat{earnings}_i = \begin{cases} (-19.69 - 5.91) + 2.59 * educ + 0.41 * workexp & \text{if female}_i = 1 \\ -19.69 & + 2.59 * educ + 0.41 * workexp & \text{if female}_i = 0 \end{cases}$$

Exercise 1

- In the previous regression FEMALE captures the difference between male and female earnings.
- We find a **negative and significant** effect for this parameter:
 - Controlling for education and work experience, women earn significantly **less** than men.
 - More specifically:

$$\widehat{earnings}_i = \begin{cases} (-19.69 - 5.91) + 2.59 * educ + 0.41 * workexp & \text{if female}_i = 1 \\ -19.69 & + 2.59 * educ + 0.41 * workexp & \text{if female}_i = 0 \end{cases}$$

- Structural stability in cross-section setting:

Exercise 1

- In the previous regression FEMALE captures the difference between male and female earnings.
- We find a **negative and significant** effect for this parameter:
 - Controlling for education and work experience, women earn significantly **less** than men.
 - More specifically:

$$\widehat{earnings}_i = \begin{cases} (-19.69 - 5.91) + 2.59 * educ + 0.41 * workexp & \text{if female}_i = 1 \\ -19.69 + 2.59 * educ + 0.41 * workexp & \text{if female}_i = 0 \end{cases}$$

- Structural stability in cross-section setting:
 - A different relationship exists for different cross-sectional groups, i.e. there is a significant difference in earnings between men and women.

How important are controls in this example?

```
. regress EARNINGS EDUC WORKEXP FEMALE
```


How important are controls in this example?

```
. regress EARNINGS EDUC WORKEXP FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.591137	.2285497	11.34	0.000	2.142174	3.0401
WORKEXP	.4056773	.1288199	3.15	0.002	.1526236	.658731
FEMALE	-5.90905	1.113972	-5.30	0.000	-8.097337	-3.720764
_cons	-19.69195	4.36076	-4.52	0.000	-28.25822	-11.12567

How important are controls in this example?

```
. regress EARNINGS EDUC WORKEXP FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.591137	.2285497	11.34	0.000	2.142174	3.0401
WORKEXP	.4056773	.1288199	3.15	0.002	.1526236	.658731
FEMALE	-5.90905	1.113972	-5.30	0.000	-8.097337	-3.720764
_cons	-19.69195	4.36076	-4.52	0.000	-28.25822	-11.12567

```
. regress EARNINGS FEMALE
```

How important are controls in this example?

```
. regress EARNINGS EDUC WORKEXP FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.591137	.2285497	11.34	0.000	2.142174	3.0401
WORKEXP	.4056773	.1288199	3.15	0.002	.1526236	.658731
FEMALE	-5.90905	1.113972	-5.30	0.000	-8.097337	-3.720764
_cons	-19.69195	4.36076	-4.52	0.000	-28.25822	-11.12567

```
. regress EARNINGS FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
FEMALE	-6.956519	1.205095	-5.77	0.000	-9.323786	-4.589251
_cons	23.11448	.8521306	27.13	0.000	21.44057	24.78839

How important are controls in this example?

```
. regress EARNINGS EDUC WORKEXP FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.591137	.2285497	11.34	0.000	2.142174	3.0401
WORKEXP	.4056773	.1288199	3.15	0.002	.1526236	.658731
FEMALE	-5.90905	1.113972	-5.30	0.000	-8.097337	-3.720764
_cons	-19.69195	4.36076	-4.52	0.000	-28.25822	-11.12567

```
. regress EARNINGS FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
FEMALE	-6.956519	1.205095	-5.77	0.000	-9.323786	-4.589251
_cons	23.11448	.8521306	27.13	0.000	21.44057	24.78839

- The gender pay gap is bigger if we do **not** control for educ and workexp.

How important are controls in this example?

```
. regress EARNINGS EDUC WORKEXP FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EDUC	2.591137	.2285497	11.34	0.000	2.142174	3.0401
WORKEXP	.4056773	.1288199	3.15	0.002	.1526236	.658731
FEMALE	-5.90905	1.113972	-5.30	0.000	-8.097337	-3.720764
_cons	-19.69195	4.36076	-4.52	0.000	-28.25822	-11.12567

```
. regress EARNINGS FEMALE
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
FEMALE	-6.956519	1.205095	-5.77	0.000	-9.323786	-4.589251
_cons	23.11448	.8521306	27.13	0.000	21.44057	24.78839

- The gender pay gap is bigger if we do **not** control for educ and workexp.
- This is what is meant when you hear '*The gender pay gap shrinks when we take into account ...*' in our case education and work experience.
- Also, there's more to the question than it seems! [▶ Let's take a look!](#)

Exercise 2

- Let's examine the determinants of educational attainment.
- Download `education.xls` and import it into Stata.

Exercise 2

- Let's examine the determinants of educational attainment.
- Download `education.xls` and import it into Stata.
- We estimate

$$\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$$

where

- educ_i : education of individual i in years
- aptitude_i : test score of individual i attained on aptitude test
- mothereduc_i : years i 's mother spent in full-time education
- fathereduc_i : years i 's father spent in full-time education
- u_i : idiosyncratic error of i .

Exercise 2

Check for significance and interpret results!

Exercise 2

Check for significance and interpret results!

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

Exercise 2

Check for significance and interpret results!

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

Exercise 2

Check for significance and interpret results!

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

- All coefficients are positive.
- Coefficient for mothereduc is insignificant.

Exercise 2

- Use an F-test to test the hypothesis that a father's and mother's educational attainments have the *same* impact on individual i 's educational attainment.

Exercise 2

- Use an F-test to test the hypothesis that a father's and mother's educational attainments have the *same* impact on individual i 's educational attainment.
- In the context of our model:

Exercise 2

- Use an F-test to test the hypothesis that a father's and mother's educational attainments have the *same* impact on individual i 's educational attainment.
- In the context of our model:
 - $\beta_f = \beta_m \Leftrightarrow \beta_f - \beta_m = 0$

Exercise 2

- Use an F-test to test the hypothesis that a father's and mother's educational attainments have the *same* impact on individual i 's educational attainment.
- In the context of our model:
 - $\beta_f = \beta_m \Leftrightarrow \beta_f - \beta_m = 0$
 - $H_0 : \beta_f - \beta_m = 0, H_a : \beta_f - \beta_m \neq 0$

Exercise 2

- Use an F-test to test the hypothesis that a father's and mother's educational attainments have the *same* impact on individual i 's educational attainment.
- In the context of our model:
 - $\beta_f = \beta_m \Leftrightarrow \beta_f - \beta_m = 0$
 - $H_0 : \beta_f - \beta_m = 0, H_a : \beta_f - \beta_m \neq 0$
- $F_{1,T-k-1} = \left(\frac{(\hat{\beta}_f - \hat{\beta}_m) - (\beta_f - \beta_m)}{\sqrt{se_{\hat{\beta}_f}^2 + se_{\hat{\beta}_m}^2 + 2 \times se_{\hat{\beta}_f} se_{\hat{\beta}_m}}} \right)^2, H_0 : \left(\frac{(\hat{\beta}_f - \hat{\beta}_m)}{\sqrt{se_{\hat{\beta}_f}^2 + se_{\hat{\beta}_m}^2 + 2 \times se_{\hat{\beta}_f} se_{\hat{\beta}_m}}} \right)^2$

Exercise 2

- Use an F-test to test the hypothesis that a father's and mother's educational attainments have the *same* impact on individual i 's educational attainment.
- In the context of our model:
 - $\beta_f = \beta_m \Leftrightarrow \beta_f - \beta_m = 0$
 - $H_0 : \beta_f - \beta_m = 0, H_a : \beta_f - \beta_m \neq 0$
- $F_{1, T-k-1} = \left(\frac{(\hat{\beta}_f - \hat{\beta}_m) - (\beta_f - \beta_m)}{\sqrt{se_{\hat{\beta}_f}^2 + se_{\hat{\beta}_m}^2 + 2 \times se_{\hat{\beta}_f} se_{\hat{\beta}_m}}} \right)^2, H_0 : \left(\frac{(\hat{\beta}_f - \hat{\beta}_m)}{\sqrt{se_{\hat{\beta}_f}^2 + se_{\hat{\beta}_m}^2 + 2 \times se_{\hat{\beta}_f} se_{\hat{\beta}_m}}} \right)^2$
 - $se^2 = \hat{\sigma}^2$, the sample variance.
 - $\hat{\beta}_j$, the estimated coefficient.
 - As before: Observed Value – Value Predicted under H_0 (here equal to 0) divided by the estimated standard error of the estimator.

Exercise 2

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

Exercise 2

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

Exercise 2

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

```
. test MOTHEREDUC = FATHEREDUC
```

Exercise 2

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

```
. test MOTHEREDUC = FATHEREDUC
```

```
( 1) MOTHEREDUC - FATHEREDUC = 0
```

```
F( 1, 536) = 0.90  
Prob > F = 0.3440
```

- We **fail** to reject H_0 as F-statistic is smaller than the critical value.
- Important: This is **not** a test of joint significance of both coefficients!
- You can also get the above F-statistic 'by hand'. [▶ See how it's done!](#)

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.
- More intuitive and faster if we note:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.
- More intuitive and faster if we note:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$
 - If $\beta_3 = \beta_4$ then our model becomes:

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.
- More intuitive and faster if we note:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$
 - If $\beta_3 = \beta_4$ then our model becomes:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \underbrace{(\text{mothereduc}_i + \text{fathereduc}_i)}_{\text{parentseduc}_i} + u_i$

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.
- More intuitive and faster if we note:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$
 - If $\beta_3 = \beta_4$ then our model becomes:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \underbrace{(\text{mothereduc}_i + \text{fathereduc}_i)}_{\text{parentseduc}_i} + u_i$

```
. generate PARENTSEDUC = MOTHEREDUC + FATHEREDUC
```

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.
- More intuitive and faster if we note:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$
 - If $\beta_3 = \beta_4$ then our model becomes:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \underbrace{(\text{mothereduc}_i + \text{fathereduc}_i)}_{\text{parentseduc}_i} + u_i$

```
. generate PARENTSEDUC = MOTHEREDUC + FATHEREDUC  
  
. regress EDUC APTITUDE PARENTSEDUC
```

Exercise 2

- Re-estimate the regression imposing the constraint that $\beta_f = \beta_m \Leftrightarrow \beta_3 = \beta_4$.
- We just saw how super tedious this is using Stata's menu.
- More intuitive and faster if we note:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \text{mothereduc}_i + \beta_4 \text{fathereduc}_i + u_i$
 - If $\beta_3 = \beta_4$ then our model becomes:
 - $\text{educ}_i = \beta_1 + \beta_2 \text{aptitude}_i + \beta_3 \underbrace{(\text{mothereduc}_i + \text{fathereduc}_i)}_{\text{parentseduc}_i} + u_i$

```
. generate PARENTSEDUC = MOTHEREDUC + FATHEREDUC  
  
. regress EDUC APTITUDE PARENTSEDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1253106	.0098434	12.73	0.000	.1059743	.1446469
PARENTSEDUC	.0828368	.0164247	5.04	0.000	.0505722	.1151014
_cons	5.29617	.4817972	10.99	0.000	4.349731	6.242608

Thank you and see you in
two weeks!

Juergen Amann

juergen.amann@nottingham.ac.uk

Wednesday 12:00 - 13:00, C42 SCGB

Exercise 1: Gender differences [◀ Go back](#)

In our data set men have (on average) higher education and more work experience:

```
. tabstat EARNINGS EDUC WORKEXP, stat(mean sd) long by(FEMALE)
```

Exercise 1: Gender differences [◀ Go back](#)

In our data set men have (on average) higher education and more work experience:

```
. tabstat EARNINGS EDUC WORKEXP, stat(mean sd) long by(FEMALE)
```

FEMALE	stats	EARNINGS	EDUC	WORKEXP
0	mean	23.11448	13.72222	17.87201
	sd	16.05073	2.575381	3.993107
1	mean	16.15796	13.62222	15.9287
	sd	11.59666	2.297135	4.641399
Total	mean	19.63622	13.67222	16.90036
	sd	14.41566	2.438476	4.433377

- There is more variation in earnings and workexp for men (standard deviation).

Exercise 1: Gender differences [◀ Go back](#)

In our data set men have (on average) higher education and more work experience:

```
. tabstat EARNINGS EDUC WORKEXP, stat(mean sd) long by(FEMALE)
```

FEMALE	stats	EARNINGS	EDUC	WORKEXP
0	mean	23.11448	13.72222	17.87201
	sd	16.05073	2.575381	3.993107
1	mean	16.15796	13.62222	15.9287
	sd	11.59666	2.297135	4.641399
Total	mean	19.63622	13.67222	16.90036
	sd	14.41566	2.438476	4.433377

- There is more variation in earnings and workexp for men (standard deviation).
- Do the blue numbers look familiar? Check results when running

```
. regress EARNINGS FEMALE
```

Exercise 2: F-statistic 'by hand'

[◀ Go back](#)

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

Exercise 2: F-statistic 'by hand'

[◀ Go back](#)

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

Exercise 2: F-statistic 'by hand'

[◀ Go back](#)

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

```
. vce
```

Exercise 2: F-statistic 'by hand' [◀ Go back](#)

```
. regress EDUC APTITUDE MOTHEREDUC FATHEREDUC
```

EDUC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
APTITUDE	.1257087	.0098533	12.76	0.000	.1063528	.1450646
MOTHEREDUC	.0492425	.0390901	1.26	0.208	-.027546	.1260309
FATHEREDUC	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

```
. vce
```

e(V)	APTITUDE	MOTHEREDUC	FATHEREDUC	_cons
APTITUDE	.00009709			
MOTHEREDUC	-.00008909	.00152803		
FATHEREDUC	-.00006315	-.00066072	.00095804	
_cons	-.00320754	-.00529709	-.00044575	.23835441

Remember under H_0 : $\left[(\hat{\beta}_f - \hat{\beta}_m) \times \left(\sqrt{se_{\beta_f}^2 + se_{\beta_m}^2 + 2 \times se_{\beta_f} se_{\beta_m}} \right)^{-1} \right]^2$

```
. display ((.0492425 - .1076825) / (sqrt(.00152803 + .00095804 + 2 * .00066072)))^2
.89697298 <- this is the F-statistic F( 1, 536) = 0.90!
```