

# PS2

Mattar Klein (328780168)

*I am trying out using markstat package for better reproducability, All code and answers will be in the same document, I will include the .stmd file and the boot file I use for running it (you will need to change the directories)*

*Also, all files can be found on my github at:*

*<https://github.com/MattarKlein/Econometrics-B.git>*

## Import Data

```
. clear all
. use "star_problem_set.dta"
```

## Question 1

### (a) Run Regression Basic Regression

```
. quietly gen sc = .
. quietly gen ra = .
. foreach v in k 1 2 3{
2. quietly gen tss`v' = treadss`v' + tmathss`v'
3. quietly replace sc = cltype`v' == 1
4. quietly replace ra = cltype`v' == 3
5. quietly regress tss`v' sc ra, robust
6. quietly eststo a_reg_`v'
7. }
```

The regression model is

$$Y_i = \beta_0 + \beta_1 sc_i + \beta_2 ra_i + u_i$$

Where  $Y_i$  is the outcome variable done for 4 different grades (Kindergarden though 3rd),  $sc_i$  is a binary variable for being randomized into a small class, and  $ra_i$  is a binary variable for being randomized into a class with an assistant.

Therefore, the constant term is the average total score of the control group (i.e. no assistant assigned and reduction in class size), and the coefficients of  $sc_i$  and  $ra_i$  are the change in total test scores for those groups respectively.

As can be seen in the output table, being randomized into a small class has an effect across all classes and is largest in 1st grade.

```

. label variable tssk "K"
. label variable tss1 "1st"
. label variable tss2 "2nd"
. label variable tss3 "3rd"
. label variable sc "Small Class"
. label variable ra "Regular + Ass"
.
. esttab a_reg_k a_reg_1 a_reg_2 a_reg_3, label se ar2

```

	(1) K	(2) 1st	(3) 2nd	(4) 3rd
Small Class	13.90*** (2.454)	29.78*** (2.831)	19.39*** (2.712)	15.59*** (2.396)
Regular + Ass	0.314 (2.271)	11.96*** (2.652)	3.479 (2.545)	-0.291 (2.273)
Constant	918.0*** (1.633)	1039.4*** (1.785)	1157.8*** (1.815)	1228.5*** (1.680)
Observations	5786	6379	6049	5967
Adjusted R-squared	0.007	0.017	0.009	0.010

Standard errors in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

## (b) Including controls

```
. quietly gen totex = .
. quietly gen ses = .
. foreach v in k 1 2 3{
2. quietly replace tss`v' = treadss`v' + tmathss`v'
3. quietly replace sc = cltype`v' == 1
4. quietly replace ra = cltype`v' == 3
5. quietly replace totex = totex`v'
6. quietly replace ses = ses`v'
7. quietly regress tss`v' sc ra totex ssex srace ses, robust
8. quietly eststo b_reg_`v'
9. }
```

The regression model is the same as before with added demographic controls for teacher experience and the gender race and socio-economic status of the student.

$$Y_i = \beta_0 + \beta_1 sc_i + \beta_2 ra_i + \beta_3 totex_i + \beta_4 ssex_i + \beta_5 srace_i + \beta_6 ses_i + u_i$$

Comparing the size of the coefficients of small class to the coefficients of different demographics shows that the effects aren't very big. Take socio-economic status for instance this is measured by whether the student is entitled to a free lunch or not, the coefficients on this are systematically larger and more than double the effect of small class.

The standard errors on  $sc_i$  went down for all coefficients, this is b/c by adding more controls the model is getting more exact. Also the adjusted R-Squared has gone up by quite a lot.

```
. label variable tssk "K"
. label variable tss1 "1st"
. label variable tss2 "2nd"
. label variable tss3 "3rd"
. label variable sc "Small Class"
. label variable ra "Regular + Ass"
. label variable totex "Exp Teacher"
. label variable ssex "Gender"
. label variable srace "Race"
. label variable ses "SE status"
.
. esttab b_reg_k b_reg_1 b_reg_2 b_reg_3, label se ar2
```

	(1) K	(2) 1st	(3) 2nd	(4) 3rd
Small Class	14.17*** (2.354)	24.80*** (2.625)	16.78*** (2.544)	11.99*** (2.277)
Regular + Ass	0.776 (2.156)	8.250*** (2.459)	4.811* (2.385)	-0.524 (2.146)
Exp Teacher	1.117*** (0.169)	0.235* (0.112)	0.124 (0.113)	0.0784 (0.105)
Gender	13.98***	11.91***	10.47***	7.654***

	(1.850)	(2.090)	(2.020)	(1.783)
Race	-10.36*** (2.063)	-27.97*** (2.789)	-31.55*** (3.145)	-20.06*** (2.616)
SE status	34.95*** (2.013)	56.98*** (2.431)	46.00*** (2.511)	42.02*** (2.151)
Constant	847.5*** (6.044)	974.5*** (7.459)	1115.0*** (8.124)	1181.2*** (6.940)
Observations	5748	6227	5707	5733
Adjusted R-squared	0.101	0.184	0.171	0.144

Standard errors in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

(c) Test for randomization

```
. foreach v in k 1 2 3{
2. quietly replace sc = cltype`v' == 1
3. quietly replace ra = cltype`v' == 3
4. quietly replace ses = ses`v'
5. quietly regress totexp`v' sc ra, robust
6. quietly eststo c_reg_`v'_exp
7. quietly regress ssex sc ra, robust
8. quietly eststo c_reg_`v'_ssex
9. quietly regress srace sc ra, robust
10. quietly eststo c_reg_`v'_srace
11. quietly regress ses sc ra, robust
12. quietly eststo c_reg_`v'_ses
13. }
```

The following tables check the distribution of observables among different test groups. It seems that while some of the randomizations (like gender) did actually randomize across demographics there were some demographics which weren't split randomly (specifically race and teacher experience).

```
. label variable totexpk "Exp K"
. label variable totexp1 "Exp 1st"
. label variable totexp2 "Exp 2nd"
. label variable totexp3 "Exp 3rd"
. esttab c_reg_k_exp c_reg_1_exp c_reg_2_exp c_reg_3_exp, label se
```

	(1) Exp K	(2) Exp 1st	(3) Exp 2nd	(4) Exp 3rd
Small Class	-0.149 (0.181)	1.899*** (0.262)	0.311 (0.268)	-0.172 (0.265)
Regular + Ass	0.670*** (0.175)	2.353*** (0.257)	1.102*** (0.250)	1.335*** (0.254)
Constant	9.068*** (0.122)	10.30*** (0.171)	12.66*** (0.182)	13.49*** (0.189)
Observations	6304	6810	6738	6751

Standard errors in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

```
. esttab c_reg_k_ssex c_reg_1_ssex c_reg_2_ssex c_reg_3_ssex, label se
```

	(1) Gender	(2) Gender	(3) Gender	(4) Gender
Small Class	0.0217 (0.0128)	0.0176 (0.0128)	0.0250* (0.0126)	0.0232 (0.0123)
Regular + Ass	0.0187 (0.0121)	0.00230 (0.0119)	0.0108 (0.0116)	0.00287 (0.0116)
Constant	1.464*** (0.00578)	1.468*** (0.00582)	1.465*** (0.00593)	1.466*** (0.00602)
Observations	11578	11578	11578	11578

Standard errors in parentheses

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

. esttab c\_reg\_k\_srace c\_reg\_1\_srace c\_reg\_2\_srace c\_reg\_3\_srace, label se

	(1) Race	(2) Race	(3) Race	(4) Race
Small Class	-0.0803*** (0.0139)	-0.0895*** (0.0139)	-0.0665*** (0.0135)	-0.0902*** (0.0130)
Regular + Ass	-0.0633*** (0.0128)	-0.116*** (0.0128)	-0.0386** (0.0128)	-0.0577*** (0.0123)
Constant	1.415*** (0.00658)	1.427*** (0.00656)	1.409*** (0.00671)	1.419*** (0.00702)
Observations	11453	11453	11453	11453

Standard errors in parentheses

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

. esttab c\_reg\_k\_ses c\_reg\_1\_ses c\_reg\_2\_ses c\_reg\_3\_ses, label se

	(1) SE status	(2) SE status	(3) SE status	(4) SE status
Small Class	0.00622 (0.0157)	0.0418** (0.0152)	0.0451** (0.0155)	0.0334* (0.0157)
Regular + Ass	-0.0256 (0.0151)	0.00609 (0.0145)	0.0130 (0.0148)	-0.000146 (0.0151)
Constant	1.523*** (0.0107)	1.470*** (0.00992)	1.468*** (0.0106)	1.484*** (0.0113)
Observations	6302	6651	6496	6520

Standard errors in parentheses

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001