

# Midterm

Adriana Sham

11/5/2019

```
pacman::p_load(msm, dplyr, ggplot2, broom, purrr, tidyverse, stargazer)
math_sample = as.data.frame(read.csv("Final_sample.csv", header = TRUE, sep = ","))
math_sample[is.na(math_sample)] = 0

estimates =
  math_sample %>%
  nest(-background) %>%
  mutate(fit = map(data, ~ lm(pvlmath ~ female, data = .)), results
= map(fit, coefficients)) %>%
  unnest(results) %>%
  filter(row_number() %% 2 == 0)

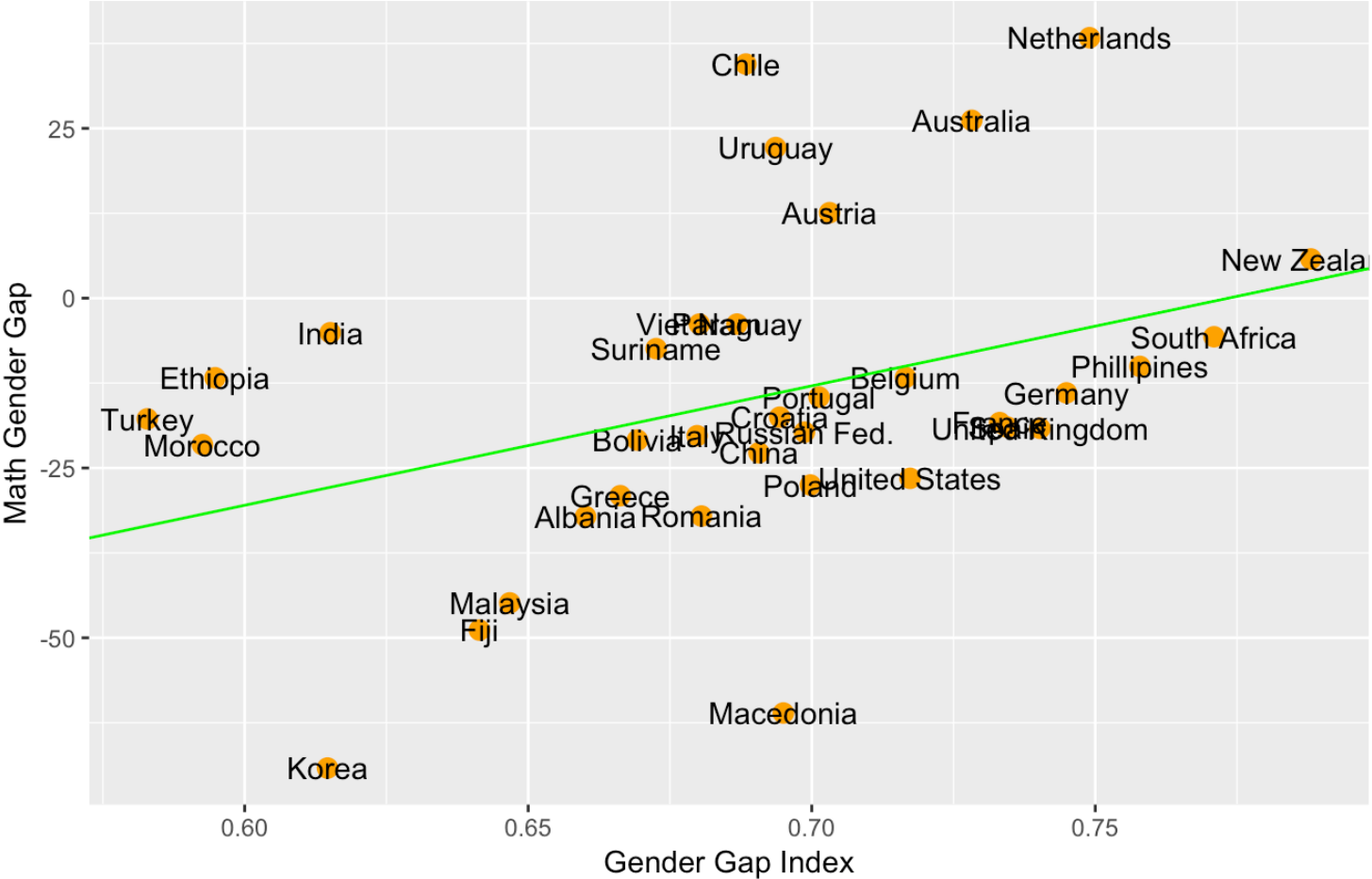
ggi = unique(math_sample$ggi)
final_data = cbind.data.frame(estimates, ggi)

mod = lm(results ~ ggi, final_data)
b = coef(mod)
```

## Including Plots

```
ggplot(final_data, aes(x = ggi, y = results)) +
  geom_point(colour = "orange", size = 3) +
  geom_text(aes(label = background)) +
  geom_abline(intercept = b[1], slope = b[2], col = "green") +
  labs(title = "Gender Gap Index vs Math Gender Gap", x = "Gender Gap Index", y = "Math Gender Gap", caption = "Source: 2003, 2006, 2009, and 2012 PISA datasets")
```

Gender Gap Index vs Math Gender Gap



Source: 2003, 2006, 2009, and 2012 PISA datasets

You can also embed plots, for example:

```
rm(list = ls())
math_sample = as.data.frame(read.csv("Final_sample.csv", header = TRUE, sep = ",") )
names(math_sample)
```

##	[1]	"year"	"background"	"cnt"	"country"
##	[5]	"female"	"age"	"diffgrade"	"fiscd"
##	[9]	"miscd"	"momwork"	"dadwork"	"hisei"
##	[13]	"homepos"	"pcgirls"	"private"	"metropolis"
##	[17]	"scoremath"	"pv1math"	"pv2math"	"pv3math"
##	[21]	"pv4math"	"pv5math"	"stweight"	"ggi"
##	[25]	"parliament90"	"parliament97"	"flp_90"	"gdppc"
##	[29]	"hdi"	"lgdppc"	"obs"	"norigin"
##	[33]	"stratum2003"	"hostregion"	"stratum2006"	"stratum2009"
##	[37]	"stratum2012"			

```
math_sample[is.na(math_sample)] = 0
str(math_sample)
```

```
## 'data.frame':    11527 obs. of  37 variables:
## $ year          : int  2012 2009 2009 2012 2009 2009 2012 2012 201
2 2012 ...
## $ background    : Factor w/ 35 levels "Albania","Australia",...: 5
5 5 5 5 5 5 5 5 5 ...
## $ cnt           : Factor w/ 9 levels "ARG","AUS","AUT",...: 1 1 1 1
1 1 1 1 1 1 ...
## $ country       : Factor w/ 9 levels "Argentina","Australia",...: 1
1 1 1 1 1 1 1 1 1 ...
## $ female        : int  0 1 0 1 1 1 1 0 1 1 ...
## $ age           : num  15.8 15.3 15.9 16 16.2 ...
## $ diffgrade     : int  0 1 1 0 0 1 0 0 0 0 ...
## $ fisced        : num  6 5 0 2 5 5 0 1 1 1 ...
## $ misced        : num  6 0 0 4 5 0 4 4 4 1 ...
## $ momwork       : num  0 0 1 1 0 0 0 0 0 1 ...
## $ dadwork       : num  1 1 1 1 1 1 1 0 1 1 ...
## $ hisei         : num  21.2 28 45 71.4 56 ...
## $ homepos       : num  -1.54 -1.92 -1.92 -0.23 -1.6 ...
## $ pcgirls       : num  0.043 0.577 0.249 0.182 0.287 ...
## $ private       : num  0 0 0 0 0 0 0 0 0 0 ...
## $ metropolis    : num  1 0 1 0 0 0 1 1 0 0 ...
## $ scoremath     : num  385 345 455 457 289 ...
## $ pvlmath       : num  415 364 505 467 291 ...
## $ pv2math       : num  355 319 411 474 227 ...
## $ pv3math       : num  432 349 426 442 289 ...
## $ pv4math       : num  355 333 493 466 334 ...
## $ pv5math       : num  369 359 438 437 303 ...
## $ stweight      : num  31.8 96 63.4 16.2 81.3 ...
## $ ggi           : num  0.669 0.669 0.669 0.669 0.669 ...
## $ parliament90  : num  0.09 0.09 0.09 0.09 0.09 ...
## $ parliament97  : num  7 7 7 7 7 7 7 7 7 7 ...
## $ flp_90        : num  50.3 50.3 50.3 50.3 50.3 ...
## $ gdppc         : num  3792 3792 3792 3792 3792 ...
## $ hdi           : num  0.659 0.659 0.659 0.659 0.659 ...
## $ lgdppc        : num  8.24 8.24 8.24 8.24 8.24 ...
## $ obs           : int  131 131 131 131 131 131 131 131 131 131 ...
## $ norigin       : int  4 4 4 4 4 4 4 4 4 4 ...
## $ stratum2003   : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
## $ hostregion : int 3202 3203 3201 3202 3201 3201 3202 3202 320
2 3205 ...
## $ stratum2006 : num 0 0 0 0 0 0 0 0 0 0 ...
## $ stratum2009 : num 0 3203 3201 0 3201 ...
## $ stratum2012 : Factor w/ 101 levels "", "ARG0001", "ARG0003", ...:
7 1 1 7 1 1 7 7 7 5 ...
```

```
lm1 = lm(pv1math ~ female * (age + diffgrade + hostregion + country)
+ country + I(ggi*female) + year + background, data = math_sample)
```

```
lm2 = lm(pv1math ~ female * (age + diffgrade + hostregion + country
) + country + I(ggi*female) + I(gdppc*female) + year + background,
data = math_sample)
```

```
lm3 = lm(pv1math ~ female * (age + diffgrade + hostregion + country
+ gdppc + ggi) + country + year + background, data = math_sample)
```

```
lm4 = lm(pv1math ~ female * (age + diffgrade + hostregion + country
+ misced + fisced) + country + I(ggi*female) + I(gdppc*female) + ye
ar + background, data = math_sample)
```

```
lm5 = lm(pv1math ~ female * (age + diffgrade + hostregion + country
+ misced + fisced + homepos + momwork + dadwork) + country + I(ggi*
female) + I(gdppc*female) + year + background, data = math_sample)
```

```
lm6 = lm(pv1math ~ female * (age + diffgrade + hostregion + country
+ misced + fisced + homepos + momwork + dadwork + pcgirls + private
+ metropolis) + country + I(ggi*female) + I(gdppc*female) + year +
background, data = math_sample)
```

```
output = stargazer(lm1, lm2, lm3, lm4, lm5, lm6, type = "text", titl
e="Gender Equality and the Math Gender Gap", omit = c("country", "ba
ckground"))
```

```
##
## Gender Equality and the Math Gender Gap
## =====
## =====
##
Dependent variable:
```

```

## -----
##
##
pvlmath
##              (1)              (2)
(3)              (4)              (5)
(6)
## -----
## -----
## female              -3,585.174              -4,016.866
-4,016.866              -3,827.305              -4,759.112
-4,874.282
##              (5,201.052)              (5,238.883)
(5,238.883)              (5,144.177)              (5,079.134)
(5,104.935)
##
## age              8.253**              8.277**
8.277**              9.541**              8.772**
8.409**
##              (4.079)              (4.080)
(4.080)              (3.996)              (3.945)
(3.939)
##
## diffgrade              -1.701              -1.723
-1.723              -1.739              -1.611
-1.543
##              (2.508)              (2.508)
(2.508)              (2.456)              (2.424)
(2.420)
##
## hostregion              -2.580**              -2.644**
-2.644**              -2.002*              -1.888*
-1.621
##              (1.175)              (1.178)
(1.178)              (1.156)              (1.142)
(1.148)
##
## misced
3.185***              2.133***              2.111***
##

```

(0.761)	(0.762)	(0.761)
##		
## fished		
7.300***	5.650***	5.530***
##		
(0.705)	(0.718)	(0.717)
##		
## homepos		
12.039***	11.723***	
##		
(1.477)	(1.478)	
##		
## momwork		
14.184***	13.952***	
##		
(2.974)	(2.969)	
##		
## dadwork		
16.133***	16.232***	
##		
(3.534)	(3.530)	
##		
## pcgirls		
7.550		
##		
(6.792)		
##		
## private		
7.026*		
##		
(3.698)		
##		
## metropolis		
18.055***		
##		
(3.883)		
##		
## I(ggi * female)	66.887*	87.192*
80.399*	89.839*	87.367*
##	(36.975)	(47.310)
(46.711)	(46.324)	(46.279)
##		

## I(gdppc * female)		-0.0002
-0.0002	-0.0002	-0.0002
##		(0.0002)
(0.0002)	(0.0002)	(0.0002)
##		
## gdppc		
-0.006***		
##		
(0.001)		
##		
## ggi		
5,391.734***		
##		
(482.186)		
##		
## year	-0.205	-0.204
-0.204	-0.559**	-0.146
-0.101		
##	(0.269)	(0.269)
(0.269)	(0.264)	(0.262)
(0.265)		
##		
## female:age	3.086	3.042
3.042	2.451	3.374
3.227		
##	(5.723)	(5.724)
(5.724)	(5.606)	(5.534)
(5.527)		
##		
## female:diffgrade	-3.091	-3.011
-3.011	-3.155	-2.231
-2.317		
##	(3.538)	(3.540)
(3.540)	(3.467)	(3.423)
(3.417)		
##		
## female:hostregion	1.088	1.219
1.219	1.163	1.450
1.487		
##	(1.625)	(1.636)
(1.636)	(1.606)	(1.586)
(1.594)		

```

##
## female:gdppc
-0.0002
##
(0.0002)
##
## female:ggi
87.192*
##
(47.310)
##
## female:miscd
1.547                1.136                1.192
##
(1.060)              (1.063)              (1.062)
##
## female:fiscd
-0.531               -1.074               -1.050
##
(0.992)              (1.014)              (1.014)
##
## female:homepos
5.171**              5.685***
##
(2.132)              (2.133)
##
## female:momwork
0.363                0.600
##
(4.208)              (4.201)
##
## female:dadwork
-2.397               -2.539
##
(4.936)              (4.930)
##
## female:pcgirls
18.198**
##
(9.083)
##
## female:private

```



```

-16.476***
##
(4.970)
##
## female:metropolis
-15.534***
##
(5.351)
##
## Constant          8,931.780**          9,134.453**
5,456.406          7,747.723**          6,561.530*
5,603.657
##          (3,792.137)          (3,803.648)
(3,813.827)          (3,732.912)          (3,685.495)
(3,706.379)
##
## -----
-----
-----
## Observations          11,527          11,527
11,527          11,527          11,527
11,527
## R2          0.233          0.233
0.233          0.265          0.284
0.287
## Adjusted R2          0.229          0.229
0.229          0.261          0.280
0.283
## Residual Std. Error    87.234 (df = 11468)    87.236 (df = 11
467)    87.236 (df = 11467)    85.437 (df = 11463)    84
.314 (df = 11457)    84.168 (df = 11451)
## F Statistic    60.128*** (df = 58; 11468) 59.114*** (df = 59
; 11467) 59.114*** (df = 59; 11467) 65.528*** (df = 63; 11463) 65.97
5*** (df = 69; 11457) 61.522*** (df = 75; 11451)
## =====
=====
=====
## Note:
*p<0.1; **p<0.05; ***p<0.01

```