

STAR Example Intent-to-treat Report

William Murrah

February 16, 2014

Contents

1	Introduction	2
2	Descriptive Statistics	3
3	Intent-to-treat Analyses	6
3.1	Pretest-posttest Regressions	6
	Appendices	8
A	Data Preparation Code	8
A.1	starRmake.R	8
A.2	prePost.R	8
A.3	stdPrePost.R	9
B	R Session Information	10

1 Introduction

This is an example report of the *STAR* data. The data consists of 100 cases and 10 variables. This example is a simplified version of the types of internal lab reports that can be generated using R and the `knitr` package to produce reproducible analyses.

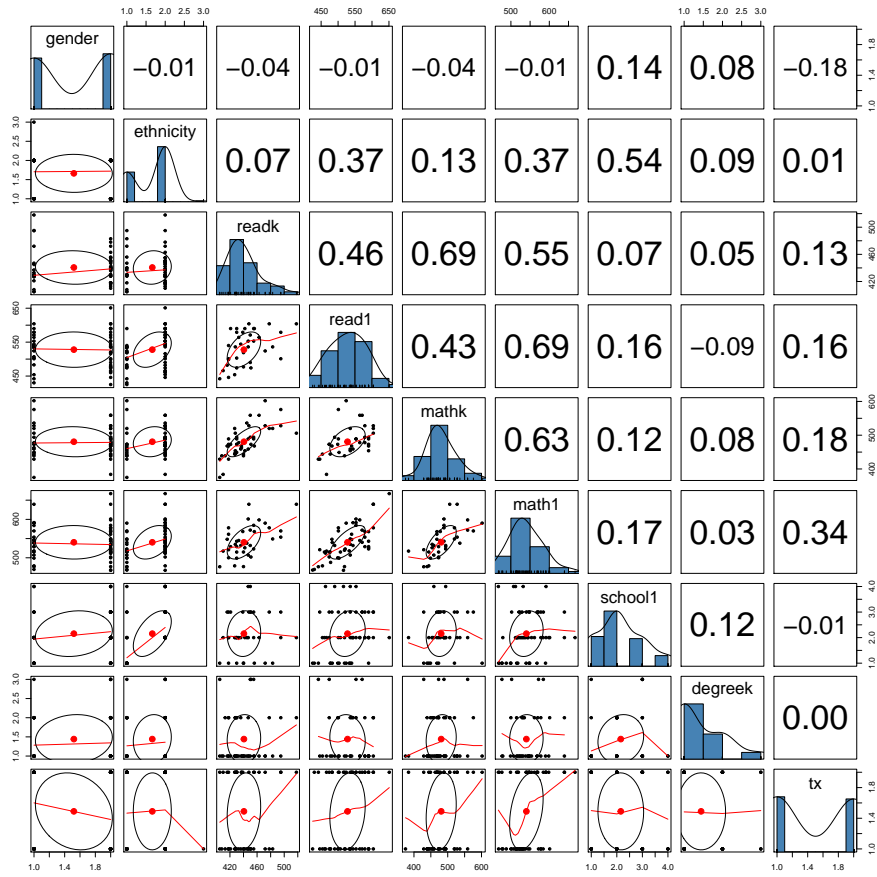
2 Descriptive Statistics

Variable	Levels	n _{control}	% _{control}	\sum % _{control}	n _{treatment}	% _{treatment}	\sum % _{treatment}	n _{all}	% _{all}	\sum % _{all}
gender	female	20	39.2	39.2	28	57.1	57.1	48	48.0	48.0
	male	31	60.8	100.0	21	42.9	100.0	52	52.0	100.0
$p = 0.11$	all	51	100.0		49	100.0		100	100.0	
ethnicity	afam	18	36.0	36.0	16	33.3	33.3	34	34.7	34.7
	cauc	31	62.0	98.0	32	66.7	100.0	63	64.3	99.0
	hispanic	1	2.0	100.0	0	0.0	100.0	1	1.0	100.0
$p = 0.58$	all	50	100.0		48	100.0		98	100.0	
school1	inner-city	7	23.3	23.3	7	25.0	25.0	14	24.1	24.1
	rural	14	46.7	70.0	12	42.9	67.9	26	44.8	69.0
	suburban	6	20.0	90.0	7	25.0	92.9	13	22.4	91.4
	urban	3	10.0	100.0	2	7.1	100.0	5	8.6	100.0
$p = 0.95$	all	30	100.0		28	100.0		58	100.0	
degreek	bachelor	17	63.0	63.0	16	64.0	64.0	33	63.5	63.5
	master	8	29.6	92.6	7	28.0	92.0	15	28.9	92.3
	master+	2	7.4	100.0	2	8.0	100.0	4	7.7	100.0
$p = 0.99$	all	27	100.0		25	100.0		52	100.0	

Table 1: Descriptive Statistics for Qualitative Variables

Variable	Levels	n	Min	q ₁	\tilde{x}	\bar{x}	q ₃	Max	s	IQR	#NA
readk	control	26	405.0	422.5	438.0	437.8	449.2	483.0	20.3	26.8	25
	treatment	22	405.0	428.0	433.5	444.0	455.8	518.0	29.2	27.8	27
$p = 0.40$	all	48	405.0	427.0	436.5	440.6	450.2	518.0	24.7	23.2	52
readl	control	27	425.0	477.5	521.0	520.0	558.5	604.0	50.3	81.0	24
	treatment	26	430.0	501.5	540.0	536.1	577.0	651.0	53.8	75.5	23
$p = 0.27$	all	53	425.0	483.0	536.0	527.9	571.0	651.0	52.2	88.0	47
mathk	control	26	375.0	445.2	473.0	473.2	494.0	559.0	39.5	48.8	25
	treatment	22	384.0	468.0	475.5	489.5	516.5	602.0	49.3	48.5	27
$p = 0.21$	all	48	375.0	452.8	473.0	480.7	506.0	602.0	44.6	53.2	52
mathl	control	30	468.0	512.0	523.0	526.8	537.2	601.0	31.9	25.2	21
	treatment	26	467.4	524.6	551.0	555.9	588.8	667.0	50.3	64.2	23
$p = 0.01$	all	56	467.4	515.0	533.5	540.3	568.7	667.0	43.6	53.8	44

Table 2: Descriptive Statistics for Quantitative Variables



3 Intent-to-treat Analyses

3.1 Pretest-posttest Regressions

	Reading	Math
(Intercept)	108.84 (143.35)	245.24*** (53.38)
txtreatment	-0.71 (16.46)	26.96** (9.35)
pretest	0.91** (0.32)	0.53*** (0.11)
gendermale	4.21 (17.58)	16.90 (9.94)
ethnicitycauc	40.31 (29.33)	34.05 (17.04)
degreekmaster	3.68 (20.79)	16.59 (11.82)
degreekmaster+	-27.08 (27.05)	-18.30 (15.69)
school1rural	-5.41 (35.17)	-11.29 (20.29)
school1suburban	-7.13 (34.81)	-3.54 (20.31)
school1urban	-26.16 (39.73)	-3.22 (23.01)
R ²	0.37	0.67
Adj. R ²	0.16	0.56
Num. obs.	36	37

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 3: Unstandardized ITT Models

	Reading	Math
(Intercept)	−0.35 (0.43)	−0.94** (0.29)
txtreatment	−0.01 (0.32)	0.62** (0.21)
scale(pretest)	0.43** (0.15)	0.54*** (0.11)
gendermale	0.08 (0.34)	0.39 (0.23)
ethnicitycauc	0.77 (0.56)	0.78 (0.39)
degreekmaster	0.07 (0.40)	0.38 (0.27)
degreekmaster+	−0.52 (0.52)	−0.42 (0.36)
school1rural	−0.10 (0.67)	−0.26 (0.46)
school1suburban	−0.14 (0.67)	−0.08 (0.47)
school1urban	−0.50 (0.76)	−0.07 (0.53)
R ²	0.37	0.67
Adj. R ²	0.16	0.56
Num. obs.	36	37

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 4: Standardized ITT Models

Appendices

A Data Preparation Code

A.1 starRmake.R

```
# starRmake.R -----
star <- read.csv(file='data/star.csv')
names(star)[1] <- 'id'
# Create treatment variable. Always set seed values when using functions
# utilize random number generation, so your numbers can be exactly
# replicated.
set.seed(111)
tx <- rbinom(n=100, size=1, prob=.50)
star$tx <- factor(tx, labels=c('control', 'treatment'))
mathEffect <- tx*(13 + rnorm(100, 0, 1))
star$math1 <- star$math1 + mathEffect
# END -----
```

A.2 prePost.R

```
*****
# prePost.R -----
source('data/starRmake.R')
# Regression formula for
fml <- 'posttest ~ tx + pretest + gender + ethnicity + degreek + school1'
df <- star
# reading -----
posttest <- star$read1
pretest <- star$readk
readmod.pp <- lm(fml,df)

# math -----
posttest <- star$math1
pretest <- star$mathk
mathmod.pp <- lm(fml,df)

# tables -----
# library(texreg)
# screenreg(list(readmod.pp, mathmod.pp))
# END -----
```


A.3 stdPrePost.R

```
*****
# stdPrePost.R -----
source('data/starRmake.R')
fml <- 'scale(posttest) ~ tx + scale(pretest) + gender + ethnicity +
degreek + school1'
df <- star
# reading -----
posttest <- star$read1
pretest <- star$readk
readmod.spp <- lm(fml,df)

# math -----
posttest <- star$math1
pretest <- star$mathk
mathmod.spp <- lm(fml,df)

# tables -----
# library(texreg)
# screenreg(list(readmod.spp, mathmod.spp))
# END -----
```

B R Session Information

```
R version 3.0.2 (2013-09-25)
Platform: x86_64-pc-linux-gnu (64-bit)

locale:
 [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
 [9] LC_ADDRESS=C             LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods    base

other attached packages:
[1] texreg_1.31      psych_1.4.2.3    reporttools_1.1.1 xtable_1.7-1
[5] knitr_1.5

loaded via a namespace (and not attached):
[1] evaluate_0.5.1 formatR_0.10     stringr_0.6.2    tools_3.0.2
```