EVALUATION GUIDELINES - Take-home examination

DRE 70061
Panel Data/Microeconometrics

Department of Economics

Start date:	17.06.2019	Time 09:00
Finish date:	18.06.2019	Time 15:00

For more information about formalities, see examination paper.

Problem 1a.

The main points we want the students to appreciate are:

- This study documents a randomized controlled experiment (a field experiment) where a random number (0-3) of new entrants are assigned into a regulated market of grocery stores.
- There is imperfect compliance in the experiment, so the actual number of market entrants do not correspond to the number of entrants from the experimental design. The authors solve this by using an instrumental variable method, where the experimental design is used as an instrument for the actual entrants.

Problem 1b.

- An instrumental variable strategy (under some assumptions, particularly the monotonicity assumption) identify the LATE, the local average treatment effects or the average treatment effect among compliers.
- The LATE is equal to an average treatment effect if there is no treatment effect heterogeneity between groups (always-takers, never-takers, compliers).
- Instrumental variable strategies do not in general identify average treatment effects.

Problem 1c

- It is problematic to replace the endogenous covariate “number of entrants” with the variable “number of entrants >0”, because of the instrumental variable strategy. The exclusion restriction demands that the instrument only affects the outcome through the endogenous variable. It is quite likely that the instrument affects the endogenous variable at both extensive and intensive margins - and that both extensive and intensive margins of the treatment variable affect the outcome. In that case, the instrumental variable assumption fails with the endogenous variable “number of entrants >0” even if it holds with the original endogenous covariate “number of entrants”.
- Replication is ok with commands like (students do not need to reproduce these):
    - regress log_P_index Hgt0 if round==2 & sample1==1, cluster(mercado)
    - [with controls] regress log_P_index Hgt0 log_P_index_pret M_col_pre_treatment Q_index_pret provincia_* M_total_benef_pret M_percent_ben HH_income_dis DD_educ1 DD_educ2 DD_educ3 urbano if round==2 & sample1==1, cluster(mercado)
    - ivregress 2sls log_P_index (Hfgt0=Hgt0) if round==2 & sample1==1, cluster(mercado)
    - [with controls] ivregress 2sls log_P_index (Hfgt0=Hgt0) log_P_index_pret M_col_pre_treatment Q_index_pret provincia_* M_total_benef_pret M_percent_ben HH_income_dis DD_educ1 DD_educ2 DD_educ3 urbano if round==2 & sample1==1, cluster(mercado)
- Would need to do this with “sample2” and “sample3” in place of “sample1” to reproduce all the results.

- When we redo the results with multivalued endogenous variable (and in the final column, the multivalued instrument), we get:

	Original	Continuous treatment	And cont. instrument
Sample 1 - no controls	-4.0 (1.8)	-2.0 (1.1)	-1.2 (0.8)
Sample 2 - no controls	-5.6 (2.4)	-2.9 (1.3)	-1.9 (1.2)
Sample 3 - no controls	-6.0 (2.8)	-3.2 (1.6)	-1.9 (1.5)
Sample 1 - controls	-2.7 (1.5)	-1.4 (0.8)	-2.2 (0.8)
Sample 2 - controls	-5.2 (2.1)	-2.8 (1.2)	-4.1 (1.8)
Sample 3 - controls	-5.2 (2.3)	-2.8 (1.3)	-3.7 (1.8)

- it seems like the original results overstate the effect of one more entrant compared to the more natural and robust implementation in column 2.
- it is not always clear that results are significant.
- if “number of entrants in experimental design” is a valid instrument, then “number of entrants >0 in design” is also a valid instrument. It still makes sense to use all the variation in the instrument.
- column 3 gives weaker results than column 2 without controls but stronger results with controls.

Problem 1d

- Without controls, the F-stats are 10.65, 8.35, 6.51
- With controls, the F-stats are 14.99, 7.69, 6.51
- These F-stats suggest potentially unreliable 2sls inference.
- Yes, just use the reduced form / intention to treat column from table 5. It is fairly easy to see e.g. that all the results without controls are significant at the 1 pct level.

Problem 1e

- Taking y as dependent variable, x as endogenous as z as instrument, we need to search for “beta” such that the regression with y - beta*x on left hand side and z on right hand side gives a p-value of about 5 pct in a twosided test. (Higher p-value means that “beta” should be included in the confidence interval.)
- The table below should be correct to the reported number of digits.

- We do not expect students to employ any algorithm for doing the search.

	Original	New lower bound	New upper bound
Sample 1 - no controls	-4.0 (1.8)	-11	-1.1
Sample 2 - no controls	-5.6 (2.4)	-18	-1.7
Sample 3 - no controls	-6.0 (2.8)	-25	-1.6
Sample 1 - controls	-2.7 (1.5)	- 7	0.3
Sample 2 - controls	-5.2 (2.1)	-18	-1.7
Sample 3 - controls	-5.2 (2.3)	-21	-1.2

Problem 1f

- This is just a standard application of quantile regression. Note that we cannot simply apply the clustering. We do not aim for any good solution to this here and look only at the point estimates.

	lin regression	median reg	lower q regr	upper q regr
Sample 1 - no controls	-2.0	-2.2	-3.9	-2.4
Sample 2 - no controls	-1.4	-2.7	-3.6	-2.7
Sample 3 - no controls	-2.6	-2.8	-3.6	-2.1
Sample 1 - controls	-2.0	-2.3	-3.0	-0.4
Sample 2 - controls	-2.5	-2.1	-2.4	-1.3
Sample 3 - controls	-1.9	-2.1	-1.2	-0.6

- The point estimates suggest that most of the action is at the lower end of the price distribution. Maybe entrants compete more in the low-price segment of this market?

Problem 1g

- I do this only for sample1.
- This is just applying the formulas from Imbens and Rubin.

- Estimated shares of always-takers, never-takers and compliers are 0.306, 0.191 and 0.503.

- The estimated log price index for the four groups and the implied LATE are:

always-takers	-0.2476
never-takers	-0.2633
treated compliers	-0.2706
untreated compliers	-0.2308
LATE = difference	-0.0398

- The LATE computed in this manner should be identical (up to rounding errors) to the 2SLS estimate without controls (and it is).

Problem 2

The idea with subproblem 2a and 2b is that student get some credit if they are able to simulate data and do the analysis even if they are not able to systematically generate data and capture results.

We have left it somewhat open how the student can document how they are able to solve problems 2a and 2b without solving 2c and 2d. If they solve 2c and 2d, 2a and 2b are implicitly solved.

Problem 2 c.

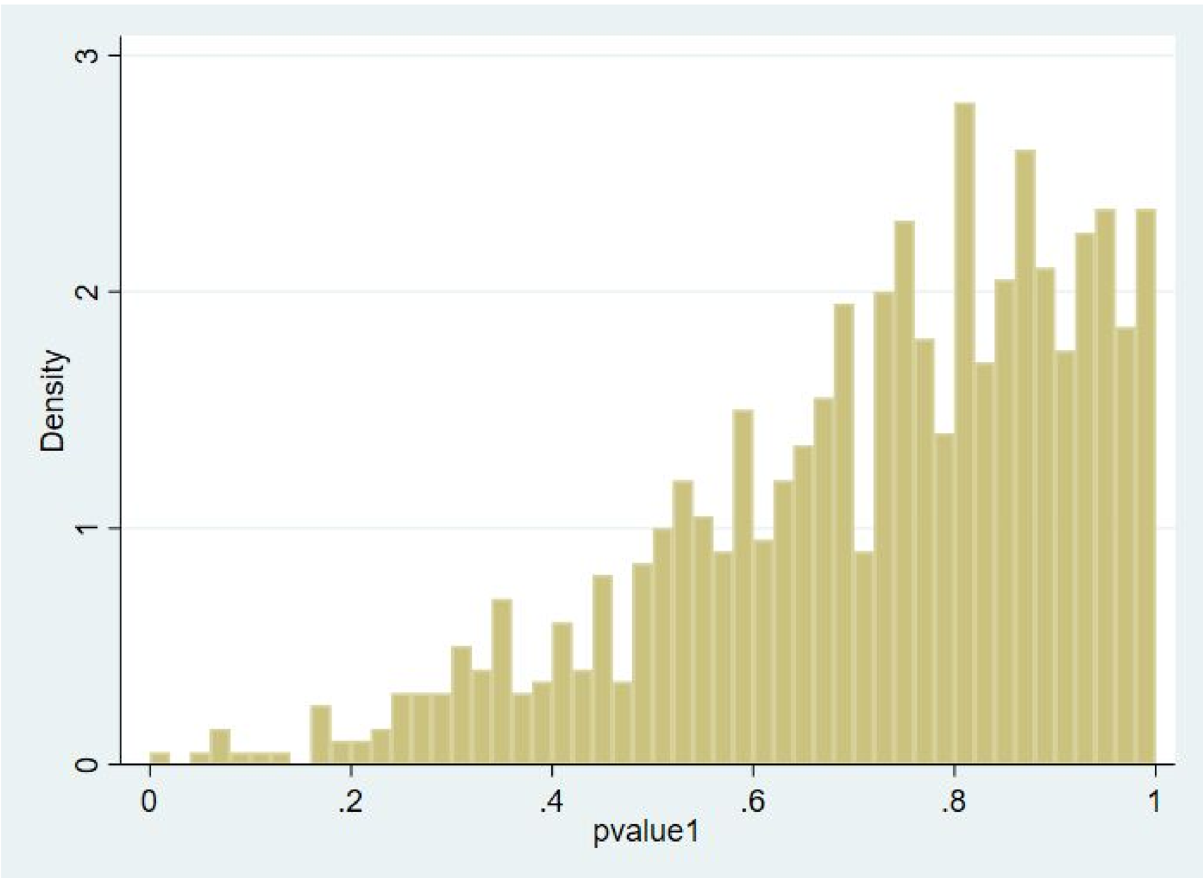
The rejection rate with significance level 5 percent should ideally be 0,05. The distribution of p-values should be uniform on the unit interval with this type of test. This distribution has mean 0.5 and standard deviation 0.289. It is useful to tabulate these statistics.

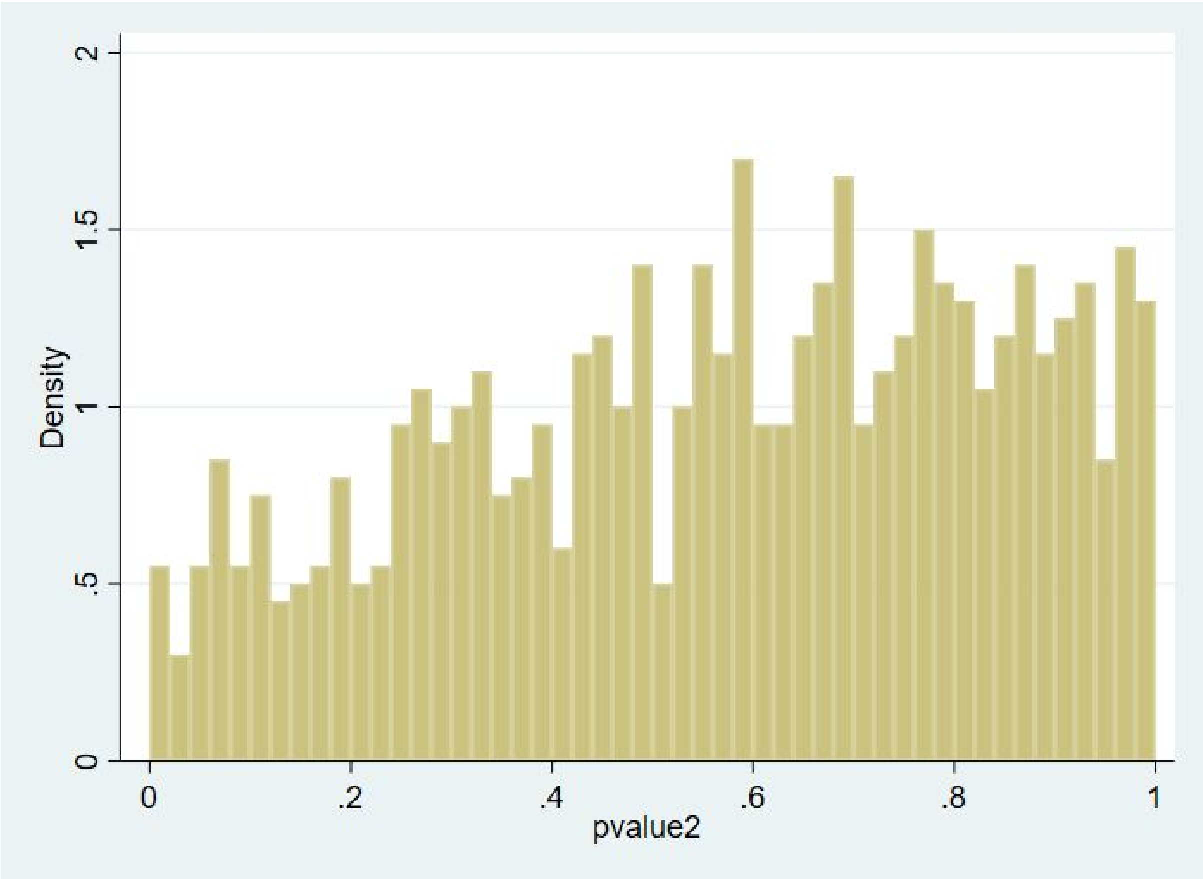
	Rejection rate (0.05)	Mean p-value	Std p-value
Test 1 (aggegation)	0.001	0.723	0.201
Test 2 (no fe, no clusters)	0.022	0.571	0.270
Test 3 (no fe, clustes)	0.066	0.482	0.298
Test 4 (fe, no clusters)	0.224	0.354	0.310
Test 5 (fe, clusters)	0.050	0.507	0.291

I get something like the graphs below for the p-value distributions:

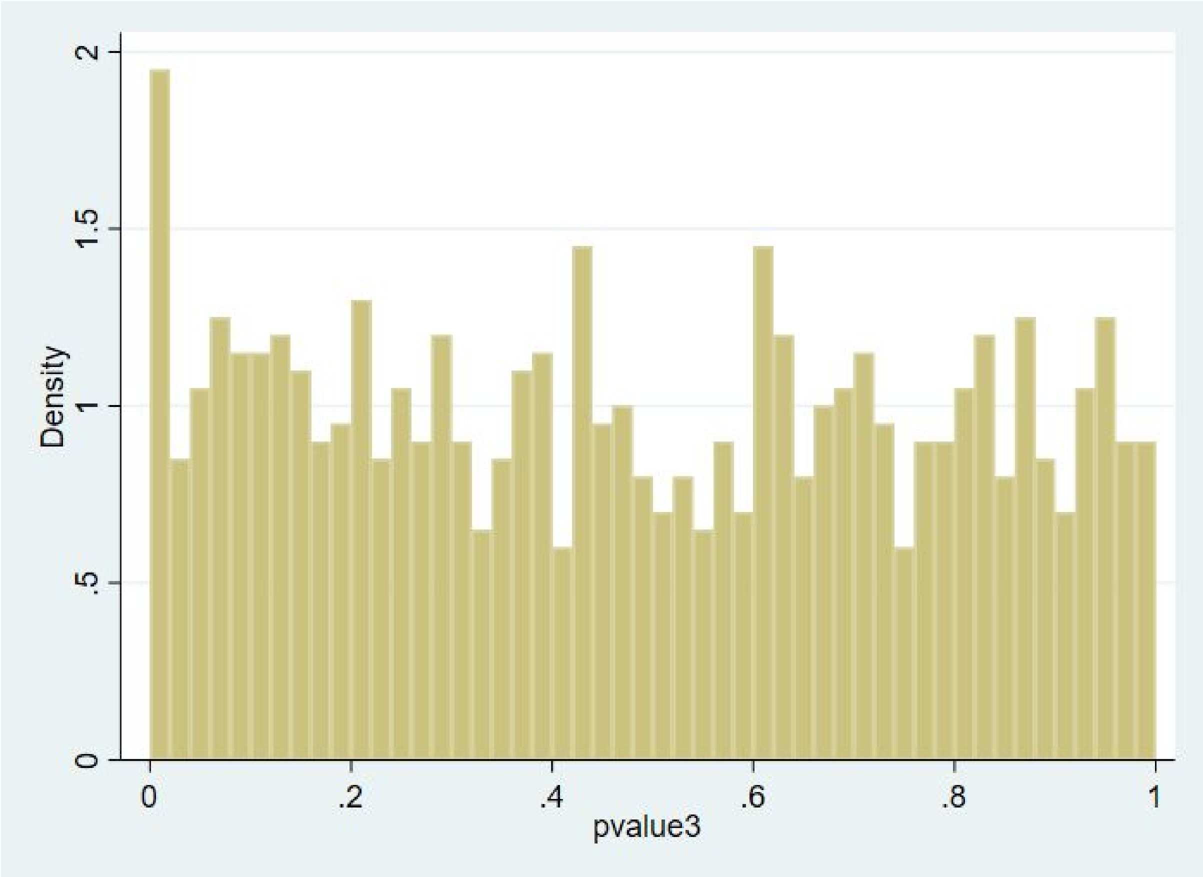
It is worthwhile to point out that Test 5 works fine - this follows both from the table and the p-value histogram below. Test 3 works reasonably fine, but has some slight overrejection driven by what looks like some overconcentration of p-values around 0.

Test 1 breaks unintentionally because of the dependence within clusters. This was not intentional. By taking differences before the regression is run, it is possible to get this right.

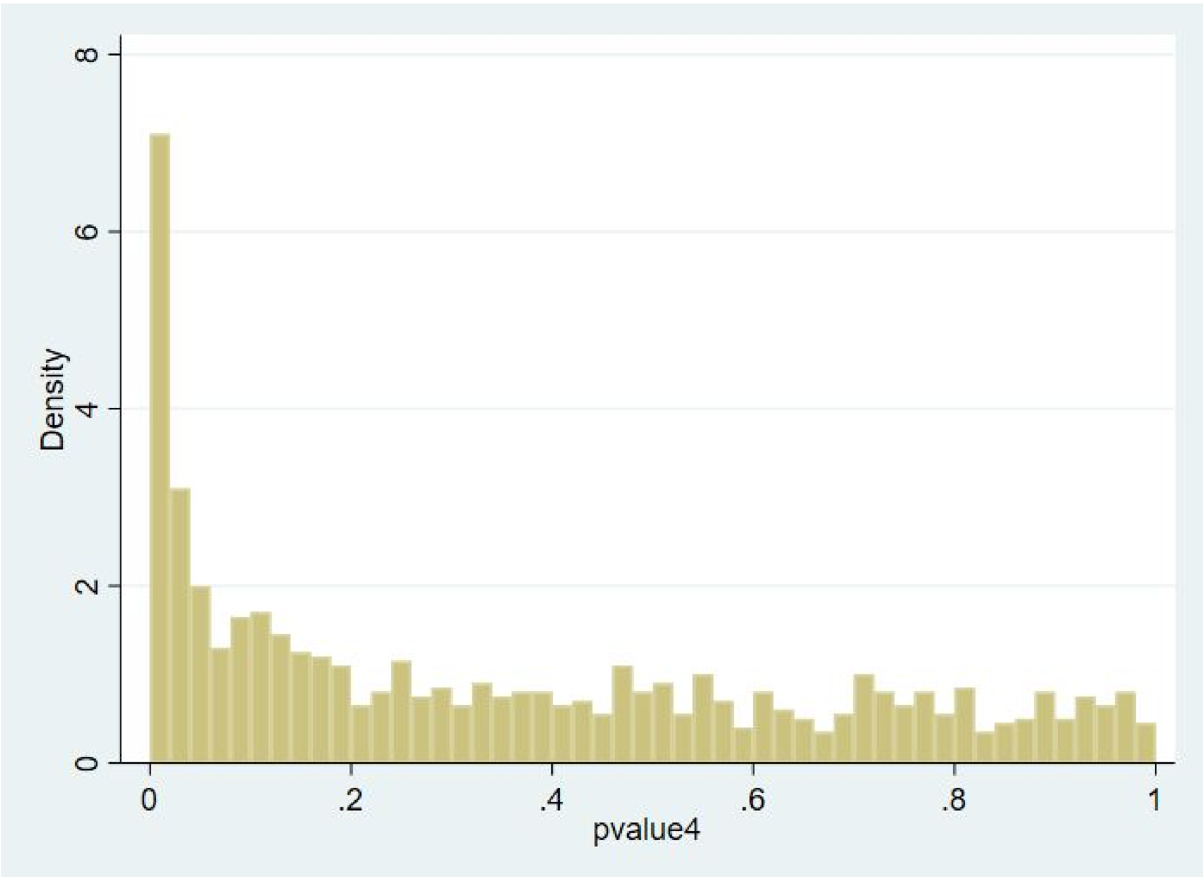




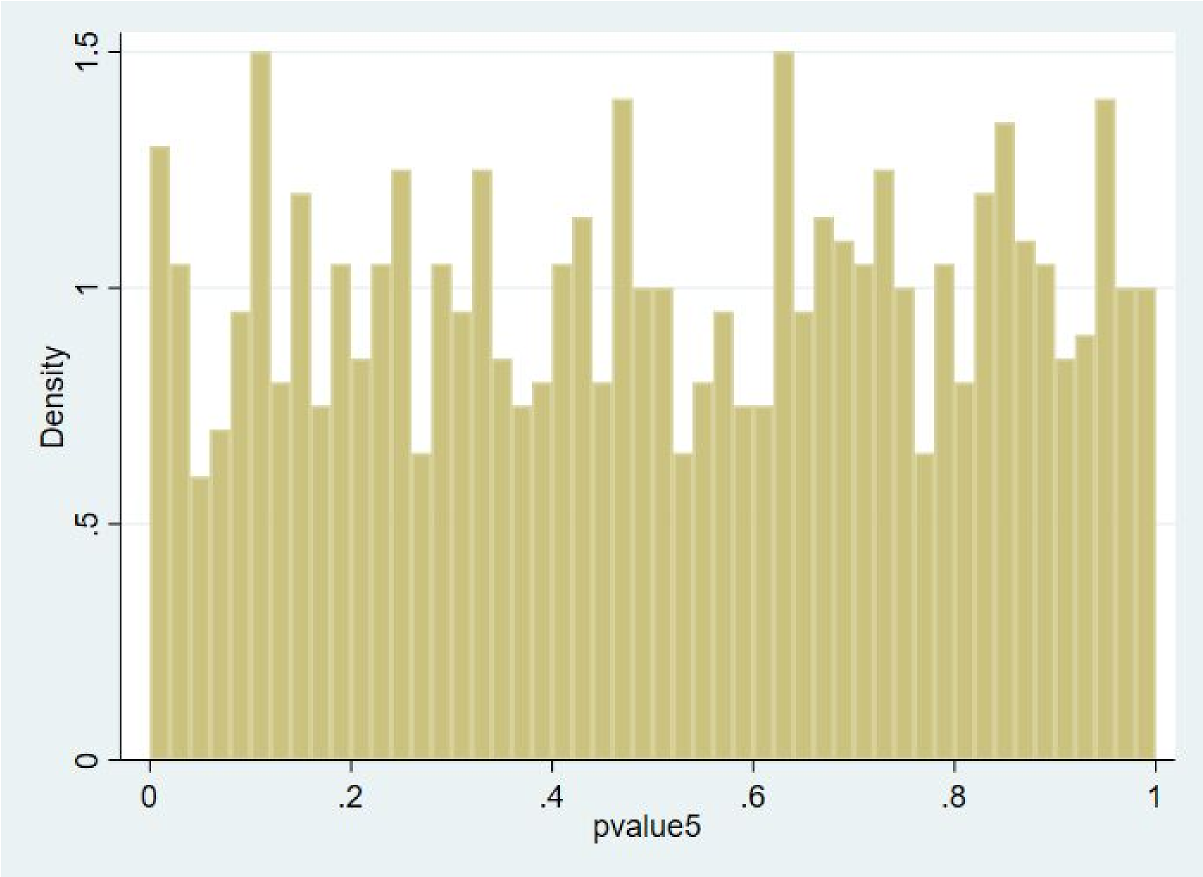
Density
2
1.5
1
.5
0
0
.2
.4
.6
.8
1
pvalue2



Density
0
.5
1
1.5
2
0
.2
.4
.6
.8
1
pvalue3



Density
0
2
4
6
8
0
.2
.4
.6
.8
1
pvalue4



I get the following results on 2d:

With only 6 groups

	Rejection rate (0.05)	Mean p-value	Std p-value
Test 1 (aggegation)	0.002	0.7236	0.206
Test 2 (no fe, no clusters)	0.031	0.557	0.279
Test 3 (no fe, clustes)	0.124	0.468	0.313
Test 4 (fe, no clusters)	0.217	0.359	0.312
Test 5 (fe, clusters)	0.103	0.489	0.308

With 24 groups

	Rejection rate (0.05)	Mean p-value	Std p-value
Test 1 (aggegation)	0.000	0.719	0.199
Test 2 (no fe, no clusters)	0.021	0.573	0.266
Test 3 (no fe, clustes)	0.069	0.485	0.291
Test 4 (fe, no clusters)	0.231	0.346	0.302
Test 5 (fe, clusters)	0.049	0.512	0.285

The important take-home message seems to be that with the parameters we use for this simulation, 6 units in the control group and 6 units in the treatment group seems to be enough. With half this number, things break down. In my numbers, it is not possible to see much of a difference in the performance between the analyses with 12 and 24 units. This may of course look a bit different in other simulations.

Some code

Some code used for solving the problems with STATA, I have just pasted in contents of do-files:

Problem 1c

```
ivregress 2sls log_P_index (Hfgt0=Hgt0) if round==2 & sample1==1, cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=Hgt0) if round==2 & sample1==1, cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=M_H_random) if round==2 & sample1==1,
cluster(mercado)

ivregress 2sls log_P_index (Hfgt0=Hgt0) if round==2 & sample2==1, cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=Hgt0) if round==2 & sample2==1, cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=M_H_random) if round==2 & sample2==1,
cluster(mercado)

ivregress 2sls log_P_index (Hfgt0=Hgt0) if round==2 & sample3==1, cluster(mercado)
```

```
ivregress 2sls log_P_index (M_H_obs=Hgt0) if round==2 & sample3==1, cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=M_H_random) if round==2 & sample3==1,
cluster(mercado)


global controls = "log_P_index_pret M_col_pre_treatment Q_index_pret provincia_*
M_total_benef_pret M_percent_ben HH_income_dis DD_educ1 DD_educ2 DD_educ3
urbano"

ivregress 2sls log_P_index (Hfgt0=Hgt0) $controls if round==2 & sample1==1,
cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=Hgt0) $controls if round==2 & sample1==1,
cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=M_H_random) $controls if round==2 & sample1==1,
cluster(mercado)

ivregress 2sls log_P_index (Hfgt0=Hgt0) $controls if round==2 & sample2==1,
cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=Hgt0) $controls if round==2 & sample2==1,
cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=M_H_random) $controls if round==2 & sample2==1,
cluster(mercado)

ivregress 2sls log_P_index (Hfgt0=Hgt0) $controls if round==2 & sample3==1,
cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=Hgt0) $controls if round==2 & sample3==1,
cluster(mercado)
ivregress 2sls log_P_index (M_H_obs=M_H_random) $controls if round==2 & sample3==1,
cluster(mercado)
```

---

problem 1d

```
regress Hfgt0 Hgt0 if round==2 & sample1==1, cluster(mercado)
test Hgt0==0
regress Hfgt0 Hgt0 $controls if round==2 & sample1==1, cluster(mercado)
test Hgt0==0

regress Hfgt0 Hgt0 if round==2 & sample2==1, cluster(mercado)
test Hgt0==0
regress Hfgt0 Hgt0 $controls if round==2 & sample2==1, cluster(mercado)
test Hgt0==0

regress Hfgt0 Hgt0 if round==2 & sample3==1, cluster(mercado)
test Hgt0==0
```

```
regress Hfgt0 Hgt0 $controls if round==2 & sample3==1, cluster(mercado)
test Hgt0==0
```

---

problem 1e

```
global controls = "log_P_index_pret M_col_pre_treatment Q_index_pret provincia_*
M_total_benef_pret M_percent_ben HH_income_dis DD_educ1 DD_educ2 DD_educ3
urbano"

* sample 1 without controls

*generate depvar = log_P_index
replace depvar = log_P_index
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.001 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.002 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.003 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.004 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.005 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.006 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.007 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.008 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.009 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.010 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.011 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
*pvalue is 0.05 about here
replace depvar = log_P_index + 0.012 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)

* lower bound of confidence interval

replace depvar = log_P_index + 0.08 * Hfgt0
```

```
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
* about here
replace depvar = log_P_index + 0.09 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.1 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.11 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)
*pvalue is 0.05 about here
replace depvar = log_P_index + 0.12 * Hfgt0
regress depvar Hgt0 if round==2 & sample1==1, cluster(mercado)


* sample 2 without controls

replace depvar = log_P_index
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)
replace depvar = log_P_index + 0.017 * Hfgt0
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)
replace depvar = log_P_index + 0.018 * Hfgt0
* between 17 and 18
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)
replace depvar = log_P_index + 0.019 * Hfgt0
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)


* lower bound of confidence interval

replace depvar = log_P_index + 0.17 * Hfgt0
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)
*here
replace depvar = log_P_index + 0.18 * Hfgt0
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)
replace depvar = log_P_index + 0.19 * Hfgt0
regress depvar Hgt0 if round==2 & sample2==1, cluster(mercado)

* sample 3 without controls

replace depvar = log_P_index
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.015 * Hfgt0
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.016 * Hfgt0
*between 16 and 17
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.017 * Hfgt0
```

regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.018 * Hfgt0
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.019 * Hfgt0
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)

* lower bound of confidence interval

replace depvar = log_P_index + 0.24 * Hfgt0
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
* between 24 and 25
replace depvar = log_P_index + 0.25 * Hfgt0
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.26 * Hfgt0
regress depvar Hgt0 if round==2 & sample3==1, cluster(mercado)

- and with controls:

global controls = "log_P_index_pret M_col_pre_treatment Q_index_pret provincia_* M_total_benef_pret M_percent_ben HH_income_dis DD_educ1 DD_educ2 DD_educ3 urbano"

* sample 1 with controls

*generate depvar = log_P_index
replace depvar = log_P_index
regress depvar Hgt0 $controls if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index - 0.002 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index - 0.003 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample1==1, cluster(mercado)

* lower bound of confidence interval

replace depvar = log_P_index + 0.06 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample1==1, cluster(mercado)
replace depvar = log_P_index + 0.07 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample1==1, cluster(mercado)

* sample 2 with controls

replace depvar = log_P_index
regress depvar Hgt0 $controls if round==2 & sample2==1, cluster(mercado)
replace depvar = log_P_index + 0.017 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample2==1, cluster(mercado)

```
replace depvar = log_P_index + 0.018 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample2==1, cluster(mercado)

* lower bound of confidence interval

replace depvar = log_P_index + 0.17 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample2==1, cluster(mercado)
replace depvar = log_P_index + 0.18 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample2==1, cluster(mercado)

* sample 3 with controls

replace depvar = log_P_index
regress depvar Hgt0 $controls if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.013 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.014 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample3==1, cluster(mercado)

* lower bound of confidence interval

replace depvar = log_P_index + 0.20 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample3==1, cluster(mercado)
replace depvar = log_P_index + 0.21 * Hfgt0
regress depvar Hgt0 $controls if round==2 & sample3==1, cluster(mercado)
```

---

problem 1f

```
global controls = "log_P_index_pret M_col_pre_treatment Q_index_pret provincia_*
M_total_benef_pret M_percent_ben HH_income_dis DD_educ1 DD_educ2 DD_educ3
urbano"

regress log_P_index Hgt0 if round==2 & sample1==1, cluster(mercado)
regress log_P_index Hgt0 if round==2 & sample2==1, cluster(mercado)
regress log_P_index Hgt0 if round==2 & sample3==1, cluster(mercado)
regress log_P_index Hgt0 $controls if round==2 & sample1==1, cluster(mercado)
regress log_P_index Hgt0 $controls if round==2 & sample2==1, cluster(mercado)
regress log_P_index Hgt0 $controls if round==2 & sample3==1, cluster(mercado)


qreg log_P_index Hgt0 if round==2 & sample1==1
qreg log_P_index Hgt0 if round==2 & sample2==1
qreg log_P_index Hgt0 if round==2 & sample3==1
```

```
qreg log_P_index Hgt0 $controls if round==2 & sample1==1
qreg log_P_index Hgt0 $controls if round==2 & sample2==1
qreg log_P_index Hgt0 $controls if round==2 & sample3==1

qreg log_P_index Hgt0 if round==2 & sample1==1, q(0.25)
qreg log_P_index Hgt0 if round==2 & sample2==1, q(0.25)
qreg log_P_index Hgt0 if round==2 & sample3==1, q(0.25)
qreg log_P_index Hgt0 $controls if round==2 & sample1==1, q(0.25)
qreg log_P_index Hgt0 $controls if round==2 & sample2==1, q(0.25)
qreg log_P_index Hgt0 $controls if round==2 & sample3==1, q(0.25)

qreg log_P_index Hgt0 if round==2 & sample1==1, q(0.75)
qreg log_P_index Hgt0 if round==2 & sample2==1, q(0.75)
qreg log_P_index Hgt0 if round==2 & sample3==1, q(0.75)
qreg log_P_index Hgt0 $controls if round==2 & sample1==1, q(0.75)
qreg log_P_index Hgt0 $controls if round==2 & sample2==1, q(0.75)
qreg log_P_index Hgt0 $controls if round==2 & sample3==1, q(0.75)
```

---

problem 1g

```
* I do this only for sample1

sum log_P_index if round==2 & sample1==1
* Gives estimate of the price of always-takers:
sum log_P_index if Hfgt0==1 & Hgt0==0 & round==2 & sample1==1
local pricea = r(mean)

* Gives estimate of the price of never-takers
sum log_P_index if Hfgt0==0 & Hgt0==1 & round==2 & sample1==1
local pricen = r(mean)

sum log_P_index if Hfgt0==1 & Hgt0==1 & round==2 & sample1==1
local pricex = r(mean)
sum log_P_index if Hfgt0==0 & Hgt0==0 & round==2 & sample1==1
local pricey = r(mean)

* The share of alwaystakers is estimated as
local pa = 34 / (34+77)
local pn =  55 / (55+233)
local pc =  1 - `pa' - `pn'
display `pa'
display `pn'
display `pc'
```

```
* The average price of treated compliers

local pricect = ((`pc'+`pa') / (`pc'))*`pricex' - (`pa' / `pc') * `pricea'
local pricecu = ((`pc'+`pn') / (`pc'))*`pricey' - (`pn' / `pc') * `pricen'


display `pricex'
display `pricey'

display `pricea'
display `pricen'

display `pricect'
display `pricecu'
```

---

problem 2

```
program define q2, rclass
	syntax [, obs(integer 12)]
	drop _all
	set obs `obs'
			gen person = _n
			gen y = rnormal(0,1)
			expand 6
			bysort person: generate year = _n
			xtset person year

			gen epsilon = rnormal(0,sqrt(0.21))

			forval i=2/6 {
			replace y = 0.9*l.y + epsilon if year == `i'
			}

			replace y = y + 1 if year>3 & person > (`obs'/2)

			generate treat = 1*(person > (`obs'/2))
			generate post = 1*(year>3)

			generate treatpost = treat*post

			generate yagg = (y + l.y + l2.y)/3 if year==3 | year==6
```

```
regress yagg treat post treatpost if year==3 | year==6, robust
scalar tvalue1 = (_b[treatpost]-1)/_se[treatpost]
scalar pvalue1 = tprob(2*`obs'-4,-abs(tvalue1))

regress y treat post treatpost, robust
scalar tvalue2 = (_b[treatpost]-1)/_se[treatpost]
scalar pvalue2 = tprob(6*`obs'-4,-abs(tvalue2))


regress y treat post treatpost, cluster(person)
scalar tvalue3 = (_b[treatpost]-1)/_se[treatpost]
scalar pvalue3 = tprob(6*`obs'-4,-abs(tvalue3))


regress y treat post treatpost i.person, robust
scalar tvalue4 = (_b[treatpost]-1)/_se[treatpost]
scalar pvalue4 = tprob(6*`obs'-15,-abs(tvalue4))

regress y treat post treatpost i.person, cluster(person)
scalar tvalue5 = (_b[treatpost]-1)/_se[treatpost]
scalar pvalue5 = tprob(6*`obs'-15,-abs(tvalue5))


end

clear all
do q2

simulate p1=pvalue1 p2=pvalue2 p3=pvalue3 p4=pvalue4 p5=pvalue5, reps(1000): q2,
obs(6)

generate rej1= p1<0.05
generate rej2= p2<0.05
generate rej3= p3<0.05
generate rej4= p4<0.05
generate rej5= p5<0.05

sum
drop rej1-rej5


simulate p1=pvalue1 p2=pvalue2 p3=pvalue3 p4=pvalue4 p5=pvalue5, reps(1000): q2,
obs(12)
generate rej1= p1<0.05
generate rej2= p2<0.05
generate rej3= p3<0.05
```

```
generate rej4= p4<0.05
generate rej5= p5<0.05

sum
drop rej1-rej5
histogram p1, width(0.02) start(0)
graph export "C:\Users\A1310280\Dropbox\undervisning\DRE 7006\STATA\Graph1a.png",
as(png) replace
histogram p2, width(0.02) start(0)
graph export "C:\Users\A1310280\Dropbox\undervisning\DRE 7006\STATA\Graph1b.png",
as(png) replace
histogram p3, width(0.02) start(0)
graph export "C:\Users\A1310280\Dropbox\undervisning\DRE 7006\STATA\Graph1c.png",
as(png) replace
histogram p4, width(0.02) start(0)
graph export "C:\Users\A1310280\Dropbox\undervisning\DRE 7006\STATA\Graph1d.png",
as(png) replace
histogram p5, width(0.02) start(0)
graph export "C:\Users\A1310280\Dropbox\undervisning\DRE 7006\STATA\Graph1e.png",
as(png) replace


simulate p1=pvalue1 p2=pvalue2 p3=pvalue3 p4=pvalue4 p5=pvalue5, reps(1000): q2,
obs(24)
generate rej1= p1<0.05
generate rej2= p2<0.05
generate rej3= p3<0.05
generate rej4= p4<0.05
generate rej5= p5<0.05

sum
drop rej1-rej5
```