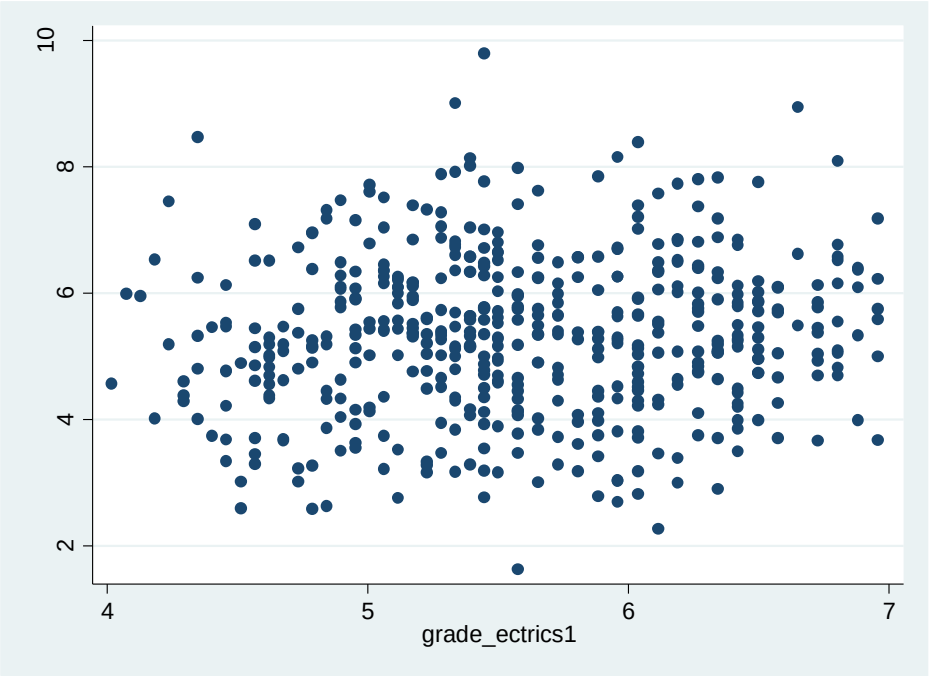
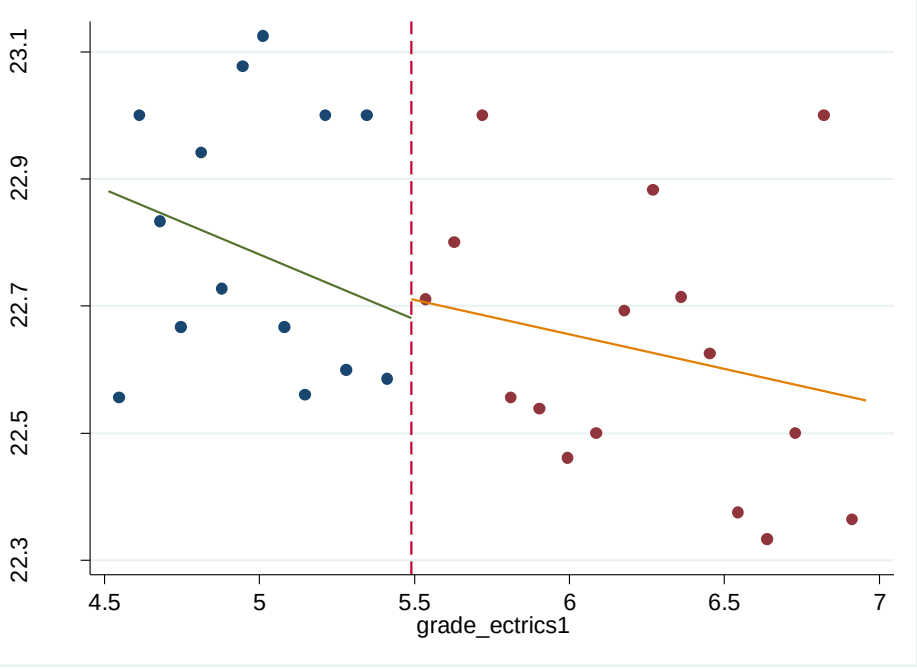
Econometrics Assignment 3a

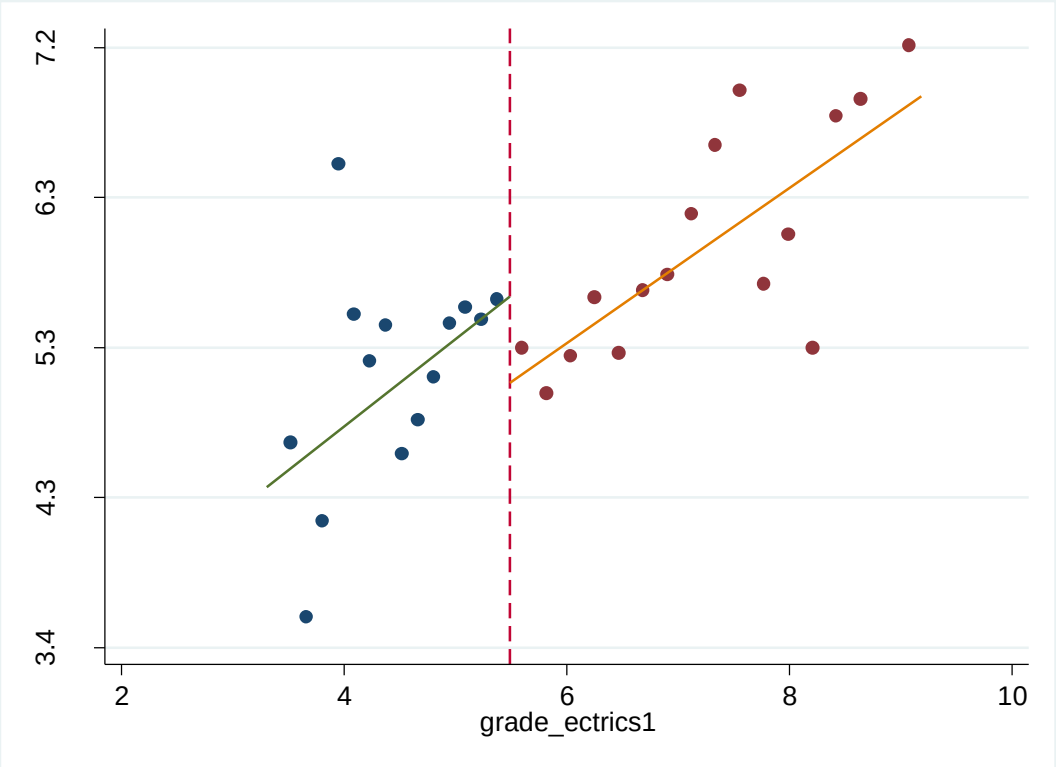
Joost Bouten, SNR: 1265889
Twan Vissers, SNR: 1266283
Fons Strik, SNR: 1257943

III.

(a)

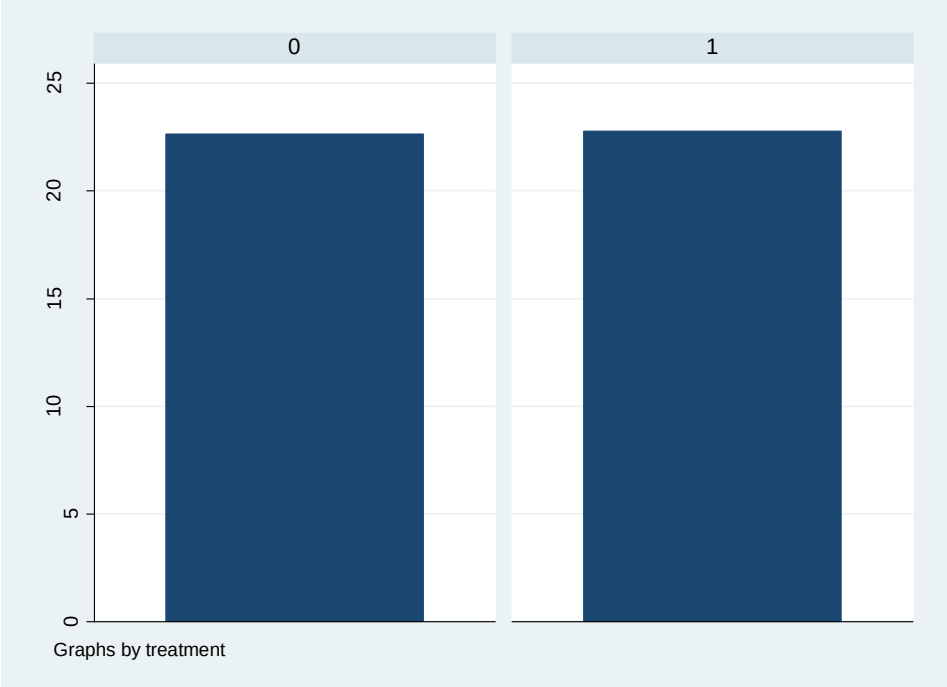


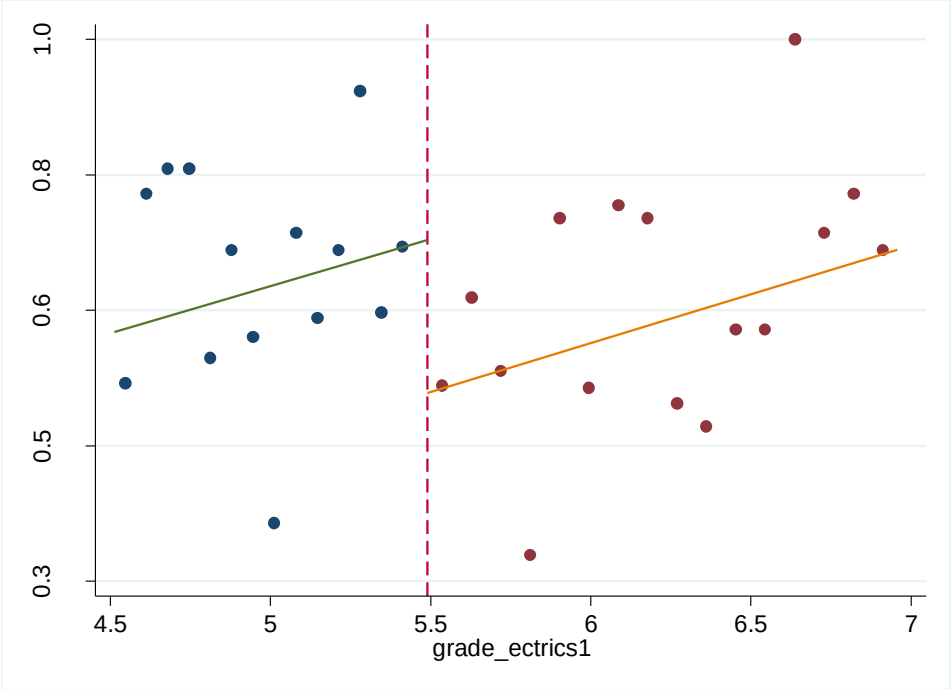


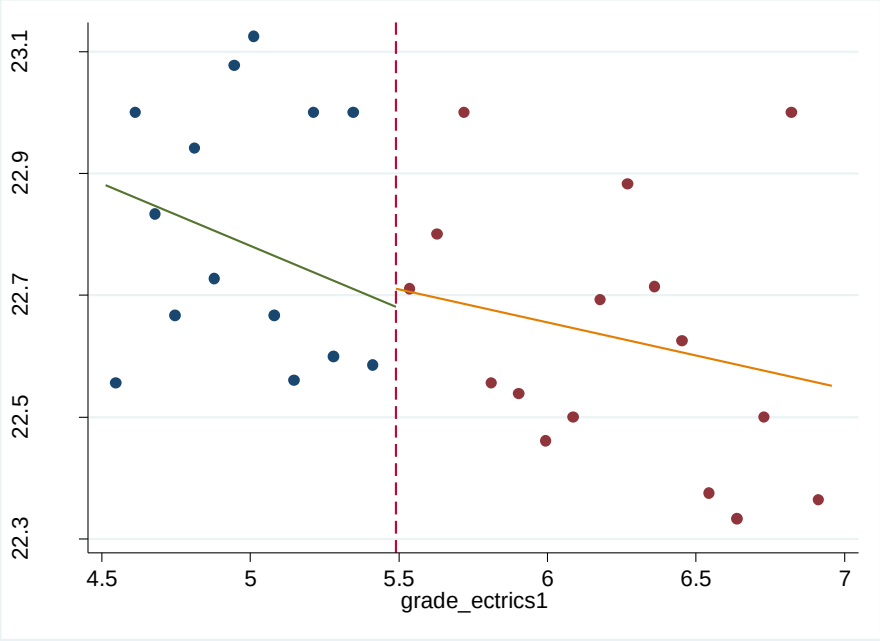


The relationship between the grades for Econometrics 1 and Econometrics 2 is positive, this can be seen by observing the positively sloped linear regression lines.
The observed relationship makes sense because people that perform relatively well on the first Econometrics course will generally be more capable in the field of econometrics.
The treatment effect is negative, the regression line jumps downwards from 5.62 to 5.08 at the cutoff point. Being treated leads to a $\sim 9.7\%$ fall in the grade for Econometrics 2.

(b)


0
1
25
20
15
10
5
0
Graphs by treatment


1.0
0.8
0.6
0.5
0.3
4.5
5
5.5
6
6.5
7
grade_ectrics1



For the variable *white* there exists an apparent discontinuity, as the regression lines do not touch each other. This is an indication of selection into treatment that can violate the manipulation assumption and thus this is something we should worry about.
For the variable *age* there is no such discontinuity, therefore we do not have to worry about bias caused due to manipulation.

(c) $Y_i = \alpha + \beta cy_i + \rho D_i + \varepsilon_i$
The estimated coefficient is statistically significantly different from zero at the 1% significance level. The effect is rather large, at a treatment effect of $\sim 10{,}0\%$.

(d) Because intuitively the covariates should not affect random selection into treatment through manipulation. This can be checked by running falsification tests. There is no problem when leaving out covariates as long as they change continuously and smoothly.

(e) When running the naïve regression, *grade_ectrics1* is not included in the regression. Using *treatment* as the only independent variable (apart from the covariates), we omit the variable *grade_ectrics1* leading to omitted variable bias. The effect is now underestimated and not statistically significantly different from 0.

```
 1 . do "M:\Master\Methods Econometrics I\Do-file CA3a.do"

 2 . * Computer Assignment 3a
 3 . * I (a)
 4 . use "C:\Users\u1266283\Downloads\ca3a_2017(1).dta"

 5 .
 6 . * II (a, b & c)
 7 . sum age
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	**541**	**22.67837**	**.7201298**	**21**	**26**

```
 8 . histogram grade_ectrics2
   (bin=23, start=1.631716, width=.35514279)

 9 . sum grade_ectrics2 if grade_ectrics1~=.&grade_ectrics2~=.
```

Variable	Obs	Mean	Std. Dev.	Min	Max
grade_ectr~2	**547**	**5.410306**	**1.277653**	**1.631716**	**9.8**

```
10 .
11 . * III
12 . * (a)
13 . graph twoway (scatter grade_ectrics2 grade_ectrics1) if grade_ectrics1>4&grade_ectrics1<7

14 . graph twoway (scatter grade_ectrics2 grade_ectrics1) (lfit grade_ectrics2 grade_ectrics1 if g
   > rade_ectrics1>4&grade_ectrics1<7

15 .
16 . sysdir set PLUS "C:\Users\u1266283"

17 . sysdir set PERSONAL "C:\Users\u1266283"

18 . ssc install cmogram
   checking cmogram consistency and verifying not already installed...
   all files already exist and are up to date.

19 .
20 . cmogram grade_ectrics2 grade_ectrics1, scatter lfit line(5.49) cutpoint(5.49)

   Plotting mean of grade_ectrics2, conditional on grade_ectrics1.

   n = 547

   Bin #1: [3.303360462188721,3.446239535013835] (n = 0) (mean = .)
   Bin #2: (3.446239535013835,3.589118607838949] (n = 1) (mean = 4.701108455657959)
   Bin #3: (3.589118607838949,3.731997680664063] (n = 2) (mean = 3.594998240470886)
   Bin #4: (3.731997680664063,3.874876753489177] (n = 4) (mean = 4.202601790428162)
   Bin #5: (3.874876753489177,4.017755826314291] (n = 2) (mean = 6.463948965072632)
   Bin #6: (4.017755826314291,4.160634899139405] (n = 3) (mean = 5.509772459665935)
   Bin #7: (4.160634899139405,4.303513971964518] (n = 7) (mean = 5.213524273463658)
   Bin #8: (4.303513971964518,4.446393044789632] (n = 7) (mean = 5.439592906406948)
   Bin #9: (4.446393044789632,4.589272117614746] (n = 20) (mean = 4.628305351734161)
   Bin #10: (4.589272117614746,4.732151190439859] (n = 22) (mean = 4.839850100603971)
   Bin #11: (4.732151190439859,4.875030263264973] (n = 17) (mean = 5.113263775320614)
   Bin #12: (4.875030263264973,5.017909336090087] (n = 32) (mean = 5.452022477984428)
   Bin #13: (5.017909336090087,5.160788408915201] (n = 25) (mean = 5.556419038772583)
   Bin #14: (5.160788408915201,5.303667481740315] (n = 39) (mean = 5.476571810551179)
   Bin #15: (5.303667481740315,5.44654655456543] (n = 58) (mean = 5.607839041742786)
   Bin #1: [5.49,5.707277681687299] (n = 54) (mean = 5.301325334442987)
   Bin #2: (5.707277681687299,5.924555363374598] (n = 38) (mean = 5.011795181977122)
   Bin #3: (5.924555363374598,6.141833045061897] (n = 54) (mean = 5.246747396610402)
   Bin #4: (6.141833045061897,6.359110726749196] (n = 44) (mean = 5.618155100128868)
   Bin #5: (6.359110726749196,6.576388408436495] (n = 40) (mean = 5.266262167692185)
   Bin #6: (6.576388408436495,6.793666090123794] (n = 12) (mean = 5.665901601314545)
   Bin #7: (6.793666090123794,7.010943771811093] (n = 21) (mean = 5.76173126129877)
   Bin #8: (7.010943771811093,7.228221453498392] (n = 18) (mean = 6.144279638926188)
   Bin #9: (7.228221453498392,7.445499135185691] (n = 10) (mean = 6.58453803062439)
```

```
Bin #10: (7.445499135185691,7.66277681687299] (n = 3) (mean = 6.928637981414795)
Bin #11: (7.66277681687299,7.880054498560289] (n = 5) (mean = 5.705449676513672)
Bin #12: (7.880054498560289,8.097332180247589] (n = 3) (mean = 6.018300533294678)
Bin #13: (8.097332180247589,8.314609861934889] (n = 2) (mean = 5.297590255737305)
Bin #14: (8.314609861934889,8.531887543622188] (n = 1) (mean = 6.76594352722168)
Bin #15: (8.531887543622188,8.749165225309488] (n = 1) (mean = 6.873388767242432)
Bin #16: (8.749165225309488,8.966442906996788] (n = 0) (mean = .)
Bin #17: (8.966442906996788,9.183720588684082] (n = 1) (mean = 7.213979721069336)

21 .
22 . reg grade_ectrics2 grade_ectrics1 if grade_ectrics1<5.5

      Source |       SS           df       MS      Number of obs   =       240
-------------+----------------------------------   F(1, 238)       =      8.42
       Model |  13.7629648         1  13.7629648   Prob > F        =    0.0041
    Residual |   389.10437       238  1.63489231   R-squared       =    0.0342
-------------+----------------------------------   Adj R-squared   =    0.0301
       Total |  402.867334       239  1.68563738   Root MSE        =    1.2786

------------------------------------------------------------------------------
grade_ectrics2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
grade_ectrics1 |   .5521149   .1902909     2.90   0.004     .1772453    .9269846
         _cons |   2.591561   .9447063     2.74   0.007     .7305069    4.452615
------------------------------------------------------------------------------

23 . display 2.591561 + 5.5*.5521149
   5.6281929

24 . reg grade_ectrics2 grade_ectrics1 if grade_ectrics1>=5.5

      Source |       SS           df       MS      Number of obs   =       307
-------------+----------------------------------   F(1, 305)       =     20.58
       Model |  30.6664982         1  30.6664982   Prob > F        =    0.0000
    Residual |  454.425515       305  1.48991972   R-squared       =    0.0632
-------------+----------------------------------   Adj R-squared   =    0.0601
       Total |  485.092013       306  1.58526802   Root MSE        =    1.2206

------------------------------------------------------------------------------
grade_ectrics2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
grade_ectrics1 |   .4919093   .1084262     4.54   0.000     .2785511    .7052674
         _cons |   2.375611   .6876465     3.45   0.001     1.022479    3.728743
------------------------------------------------------------------------------

25 . display 2.375611 + 5.5*.4919093
   5.0811122

26 . display (5.6281929-5.0811122)/5.6281929
   .09720362

27 .
28 . * (b)
29 . gen treatment=(grade_ectrics1<5.5)

30 . graph bar age if grade_ectrics1>=4.5&grade_ectrics1<=7, by(treatment)

31 . cmogram white grade_ectrics1 if grade_ectrics1>=4.5&grade_ectrics1<=7, scatter lfit line(5.49

   Plotting mean of white, conditional on grade_ectrics1.

   n = 463
```

```
Bin #1: [4.51233720779419,4.579066446849279] (n = 9) (mean = .5555555555555556)
Bin #2: (4.579066446849279,4.645795685904368] (n = 10) (mean = .8)
Bin #3: (4.645795685904368,4.712524924959456] (n = 6) (mean = .8333333333333334)
Bin #4: (4.712524924959456,4.779254164014544] (n = 6) (mean = .8333333333333334)
Bin #5: (4.779254164014544,4.845983403069632] (n = 17) (mean = .5882352941176471)
Bin #6: (4.845983403069632,4.912712642124721] (n = 11) (mean = .7272727272727273)
Bin #7: (4.912712642124721,4.97944188117981] (n = 13) (mean = .6153846153846154)
Bin #8: (4.97944188117981,5.046171120234898] (n = 8) (mean = .375)
Bin #9: (5.046171120234898,5.112900359289986] (n = 12) (mean = .75)
Bin #10: (5.112900359289986,5.179629598345074] (n = 25) (mean = .64)
Bin #11: (5.179629598345074,5.246358837400162] (n = 11) (mean = .7272727272727273)
Bin #12: (5.246358837400162,5.313088076455251] (n = 15) (mean = .9333333333333334)
Bin #13: (5.313088076455251,5.379817315510339] (n = 17) (mean = .6470588235294118)
Bin #14: (5.379817315510339,5.44654655456543] (n = 41) (mean = .7317073170731707)
Bin #1: [5.49,5.58175871372223] (n = 38) (mean = .5526315789473685)
Bin #2: (5.58175871372223,5.67351742744446] (n = 15) (mean = .6666666666666666)
Bin #3: (5.67351742744446,5.765276141166689] (n = 14) (mean = .5714285714285714)
Bin #4: (5.765276141166689,5.857034854888918] (n = 9) (mean = .3333333333333333)
Bin #5: (5.857034854888918,5.948793568611148] (n = 13) (mean = .7692307692307693)
Bin #6: (5.948793568611148,6.040552282333377] (n = 40) (mean = .55)
Bin #7: (6.040552282333377,6.132310996055606] (n = 14) (mean = .7857142857142857)
Bin #8: (6.132310996055606,6.224069709777836] (n = 13) (mean = .7692307692307693)
Bin #9: (6.224069709777836,6.315828423500065] (n = 17) (mean = .5294117647058824)
Bin #10: (6.315828423500065,6.407587137222294] (n = 14) (mean = .5)
Bin #11: (6.407587137222294,6.499345850944524] (n = 32) (mean = .625)
Bin #12: (6.499345850944524,6.591104564666753] (n = 8) (mean = .625)
Bin #13: (6.591104564666753,6.682863278388982] (n = 3) (mean = 1)
Bin #14: (6.682863278388982,6.774621992111212] (n = 8) (mean = .75)
Bin #15: (6.774621992111212,6.866380705833441] (n = 10) (mean = .8)
Bin #16: (6.866380705833441,6.958139419555664] (n = 11) (mean = .7272727272727273)

32 . cmogram age grade_ectrics1 if grade_ectrics1>=4.5&grade_ectrics1<=7, scatter lfit line(5.49)

Plotting mean of age, conditional on grade_ectrics1.

n = 462

Bin #1: [4.51233720779419,4.579066446849279] (n = 9) (mean = 22.55555555555556)
Bin #2: (4.579066446849279,4.645795685904368] (n = 10) (mean = 23)
Bin #3: (4.645795685904368,4.712524924959456] (n = 6) (mean = 22.83333333333333)
Bin #4: (4.712524924959456,4.779254164014544] (n = 6) (mean = 22.66666666666667)
Bin #5: (4.779254164014544,4.845983403069632] (n = 17) (mean = 22.94117647058824)
Bin #6: (4.845983403069632,4.912712642124721] (n = 11) (mean = 22.72727272727273)
Bin #7: (4.912712642124721,4.97944188117981] (n = 13) (mean = 23.07692307692308)
Bin #8: (4.97944188117981,5.046171120234898] (n = 8) (mean = 23.125)
Bin #9: (5.046171120234898,5.112900359289986] (n = 12) (mean = 22.66666666666667)
Bin #10: (5.112900359289986,5.179629598345074] (n = 25) (mean = 22.56)
Bin #11: (5.179629598345074,5.246358837400162] (n = 11) (mean = 23)
Bin #12: (5.246358837400162,5.313088076455251] (n = 15) (mean = 22.6)
Bin #13: (5.313088076455251,5.379817315510339] (n = 17) (mean = 23)
Bin #14: (5.379817315510339,5.44654655456543] (n = 41) (mean = 22.58536585365854)
Bin #1: [5.49,5.58175871372223] (n = 38) (mean = 22.71052631578947)
Bin #2: (5.58175871372223,5.67351742744446] (n = 15) (mean = 22.8)
Bin #3: (5.67351742744446,5.765276141166689] (n = 14) (mean = 23)
Bin #4: (5.765276141166689,5.857034854888918] (n = 9) (mean = 22.55555555555556)
Bin #5: (5.857034854888918,5.948793568611148] (n = 13) (mean = 22.53846153846154)
Bin #6: (5.948793568611148,6.040552282333377] (n = 39) (mean = 22.46153846153846)
Bin #7: (6.040552282333377,6.132310996055606] (n = 14) (mean = 22.5)
Bin #8: (6.132310996055606,6.224069709777836] (n = 13) (mean = 22.69230769230769)
Bin #9: (6.224069709777836,6.315828423500065] (n = 17) (mean = 22.88235294117647)
Bin #10: (6.315828423500065,6.407587137222294] (n = 14) (mean = 22.71428571428572)
Bin #11: (6.407587137222294,6.499345850944524] (n = 32) (mean = 22.625)
Bin #12: (6.499345850944524,6.591104564666753] (n = 8) (mean = 22.375)
Bin #13: (6.591104564666753,6.682863278388982] (n = 3) (mean = 22.33333333333333)
Bin #14: (6.682863278388982,6.774621992111212] (n = 8) (mean = 22.5)
Bin #15: (6.774621992111212,6.866380705833441] (n = 10) (mean = 23)
Bin #16: (6.866380705833441,6.958139419555664] (n = 11) (mean = 22.36363636363636)
```

```
33 .
34 . * (c)
35 . reg grade_ectrics2 treatment grade_ectrics1 if grade_ectrics1>=4.5&grade_ectrics1<=7
```

Source	SS	df	MS		
				Number of obs	= 468
				F(2, 465)	= 4.83
Model	14.9419557	2	7.47097786	Prob > F	= 0.0084
Residual	719.46643	465	1.54723963	R-squared	= 0.0203
				Adj R-squared	= 0.0161
Total	734.408386	467	1.57260896	Root MSE	= 1.2439

grade_ectrics2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treatment	.5338346	.2026563	2.63	0.009	.1355991	.93207
grade_ectrics1	.501468	.1615665	3.10	0.002	.1839772	.8189588
_cons	2.285306	.9901721	2.31	0.021	.3395402	4.231072

```
36 . margins, eydx(treatment)
```

Average marginal effects Number of obs = 468
Model VCE : **OLS**

Expression : **Linear prediction, predict()**
ey/dx w.r.t. : **treatment**

	ey/dx	Delta-method Std. Err.	t	P>|t|	[95% Conf.	Interval]
treatment	.0997596	.0379475	2.63	0.009	.0251897	.1743294

```
37 .
38 . * (d)
39 . reg grade_ectrics2 treatment grade_ectrics1 age gender sped if grade_ectrics1>=4.5&grade_ectr
```

Source	SS	df	MS		
				Number of obs	= 437
				F(5, 431)	= 1.92
Model	15.3870193	5	3.07740386	Prob > F	= 0.0902
Residual	691.67331	431	1.60481046	R-squared	= 0.0218
				Adj R-squared	= 0.0104
Total	707.060329	436	1.621698	Root MSE	= 1.2668

grade_ectrics2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treatment	.496477	.2133675	2.33	0.020	.0771067	.9158473
grade_ectrics1	.4887964	.1704715	2.87	0.004	.1537374	.8238554
age	.0316026	.083573	0.38	0.706	-.1326588	.195864
gender	.0733211	.1237246	0.59	0.554	-.1698576	.3164998
sped	.1477029	.2464363	0.60	0.549	-.3366635	.6320693
_cons	1.614782	2.223335	0.73	0.468	-2.755146	5.984711

```
40 . reg grade_ectrics2 treatment grade_ectrics1 if age~=.&gender~=.&grade_ectrics1>=4.5&grade_ect
```

Source	SS	df	MS		
				Number of obs	= 462
				F(2, 459)	= 4.71
Model	14.6279096	2	7.31395479	Prob > F	= 0.0095
Residual	713.307636	459	1.55404714	R-squared	= 0.0201
				Adj R-squared	= 0.0158
Total	727.935546	461	1.57903589	Root MSE	= 1.2466

grade_ectrics2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treatment	.5350275	.2045928	2.62	0.009	.1329727	.9370822
grade_ectrics1	.4992402	.1629908	3.06	0.002	.1789396	.8195408
_cons	2.303326	.9991973	2.31	0.022	.3397574	4.266894

```
41 .
42 . * (e)
43 . reg grade_ectrics2 treatment age gender sped if grade_ectrics1>=4.5&grade_ectrics1<=7
```

Source	SS	df	MS		
Model	2.19303472	4	.548258679	Number of obs	= 437
Residual	704.867295	432	1.63163726	F(4, 432)	= 0.34
				Prob > F	= 0.8537
Total	707.060329	436	1.621698	R-squared	= 0.0031
				Adj R-squared	= -0.0061
				Root MSE	= 1.2774

grade_ectr~2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treatment	-.0020276	.12472	-0.02	0.987	-.2471611	.2431058
age	.0131487	.0840184	0.16	0.876	-.151987	.1782844
gender	.1193672	.1236991	0.96	0.335	-.1237598	.3624942
sped	.1206229	.248305	0.49	0.627	-.3674133	.608659
_cons	4.996597	1.900404	2.63	0.009	1.26141	8.731785

```
44 .
   end of do-file

45 .
```