How does feedback affect beliefs and decision making among students?

Symposium on Economic Experiments in Developing Countries (SEEDEC)

May 30th - 31st, 2019

Catalina Franco, Universidad del Rosario

Motivation

- People tend to overestimate their absolute and relative abilities:
  - In the lab (e.g., Hoelz & Rustichini 2005, Moore & Healy, 2008)
  - Surveys (e.g., Svenson 1981, Englmaier 2006)
- Do these biases affect real-life behavior?
- Can providing information to individuals correct their biased beliefs and affect their decisions?

Research questions

1. How does providing information about relative performance to students affect their:
   - Beliefs?
   - Academic investments?
   - Choices?
   - Performance?

2. Are beliefs elicited with an IC task coherent with the beliefs revealed by real-life behavior?

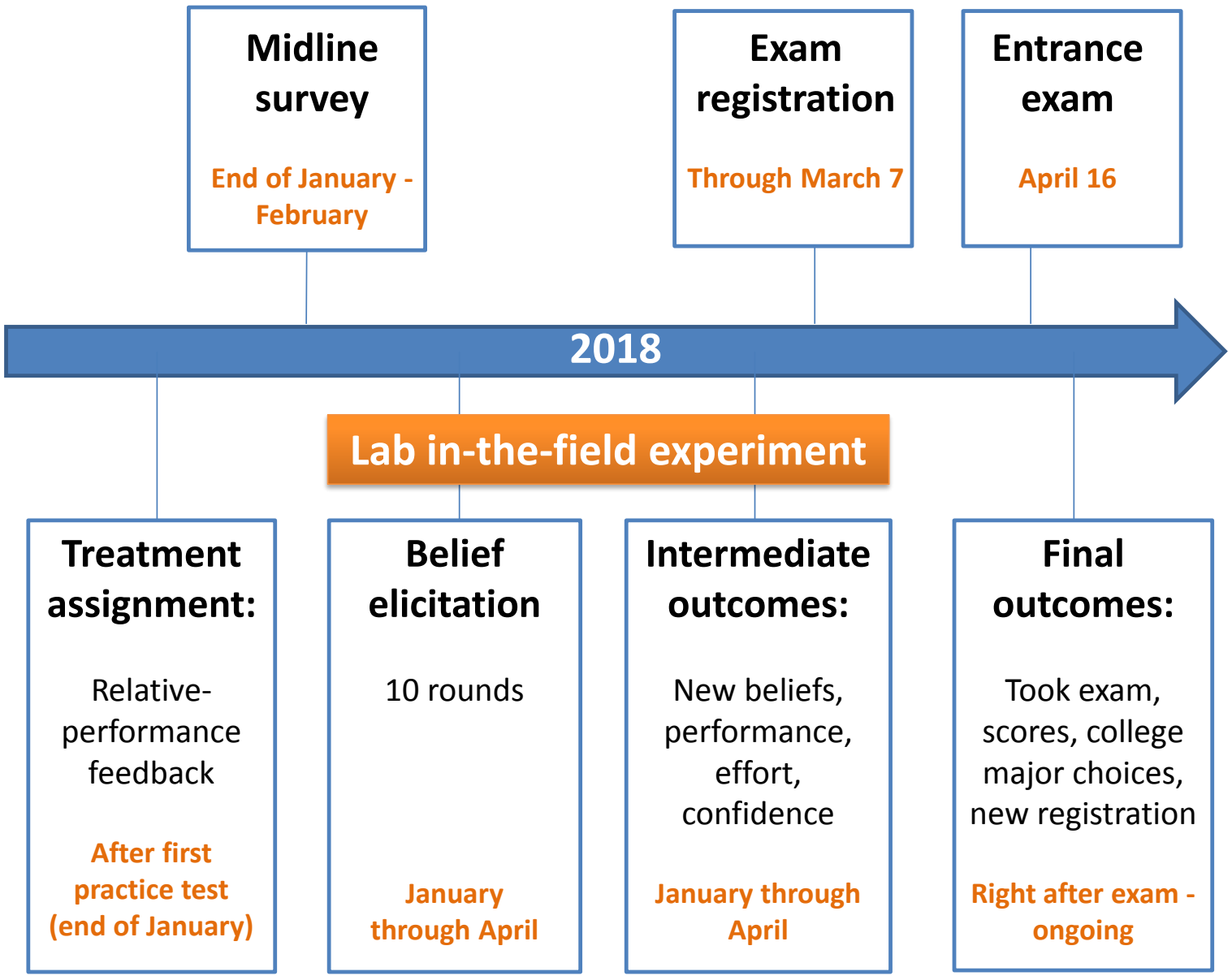
This paper

- I connect experimentally-measured relative performance beliefs with real-life behavior in a high-stakes education context
  - Examine how real-life decisions are consistent with behavior in the lab (Gill et al. 2016, Azmat & Irriberi 2010 and 2016, Eil & Rao 2011, Mobius et al. 2011, Ertac 2011)
  - Study other margins beyond grades (Azmat & Irriberi 2010 and 2016, Bandiera et al. 2015, Azmat et al. 2018, Murphy & Weindardt 2018)
  - Study intermediate adjustments to receiving feedback (Bobba & Frisancho 2016, Gonzalez 2017, Dizon-Ross 2018)

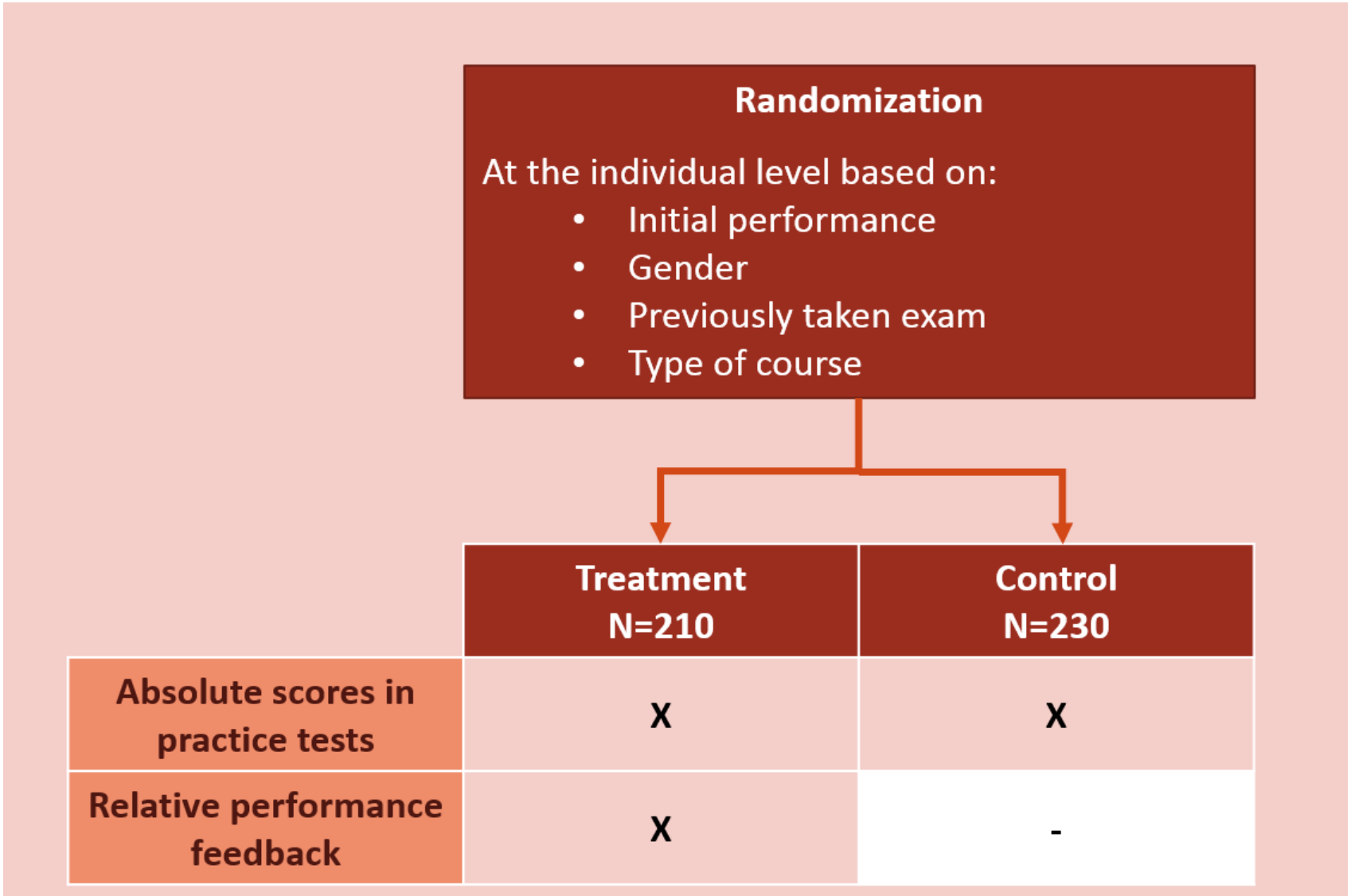
The setting

- Students enrolled in a test preparation center in Medellín, Colombia
- Preparing for entrance exam at Universidad de Antioquia:
  - High-stakes college entrance exam and very competitive admission
    Admission rates
  - Students choose up to two college major options *before* taking the exam
- Test preparation course:
  - Specific for this university exam
  - Weekly practice tests
  - I leverage institute's practice test performance report

Timeline


Midline survey
End of January - February
Exam registration
Through March 7
Entrance exam
April 16
2018
Lab in-the-field experiment
Treatment assignment:
Relative-performance feedback
After first practice test (end of January)
Belief elicitation
10 rounds
January through April
Intermediate outcomes:
New beliefs, performance, effort, confidence
January through April
Final outcomes:
Took exam, scores, college major choices, new registration
Right after exam - ongoing

Experimental design: field experiment



	Treatment N=210	Control N=230
Absolute scores in practice tests	X	X
Relative performance feedback	X	-

Feedback report

Regular report

Retroalimentación de desempeño relativo

Los siguientes gráficos muestran las **predicciones que hiciste el día del simulacro (encuesta 1)** junto con el cuartil en el cual quedó ubicado tu desempeño en el simulacro
El puntaje en el que quedaste sale de color VERDE.

Razonamento matemático:

CUARTIL 1	CUARTIL 2	CUARTIL 3	CUARTIL 4
Tu puntaje quedo en el cuartil 1	Asignaste	Asignaste	Asignaste
8	4	0	0

Competencia lectora:

CUARTIL 1
Tu puntaje quedo en el cuartil 1
6

Según tus asignaciones, pensaste que tu puntaje iba a quedar en un cuartil igual al que quedaste

Según tus asignaciones, pensaste que tu pu

Tu desempeño relativo fue mejor en competencia lectora que en razonamiento lógico matemático.

Información más detallada

Si quieres obtener información más detallada sobre tu desempeño por favor haz click **aquí.**

Balance of characteristics

By quartile | Attritors | Ext. validity

	Control	Treatment	P-value (T-C)	No. obs
Stratification variables				
Female	0.613	0.600	0.780	440
Previously taken entrance exam	0.795	0.810	0.699	439
AM course	0.426	0.414	0.803	440
PM course	0.357	0.372	0.746	440
Integrated UdeA - UNAL	0.043	0.042	0.975	440
Pre-medicine	0.148	0.148	0.995	440
Weekend course	0.026	0.024	0.879	440
Demographic variables				
Age	17.733	17.257	0.027	434
Single	0.973	0.976	0.787	433
Student	0.677	0.720	0.311	434
Residential strata	2.450	2.529	0.431	434
Urban	0.881	0.895	0.622	434
Academic variables				
Math no. correct (initial practice test)	11.579	11.811	0.553	439
Reading no. correct (initial practice test)	18.189	18.853	0.284	439
Avg. practice test score in classroom	38.067	38.143	0.762	440
Joint orthogonality test			0.2812	439

Empirical strategy

- Regression specification:

$$y_i = \beta_1 + \beta_2 T_i + \sum_{q=1}^{3} \alpha_q Q_i + \sum_{q=1}^{3} \tau_q Q_i * T_i + \rho strata_i + \mathbf{X_i}\gamma + \varepsilon_i$$

Where:
T_i: treatment assignment $\{0,1\}$
Q_i: quartile of initial practice test performance
$strata_i$: randomization strata
$\mathbf{X_i}$: baseline covariates

- Treatment effects:

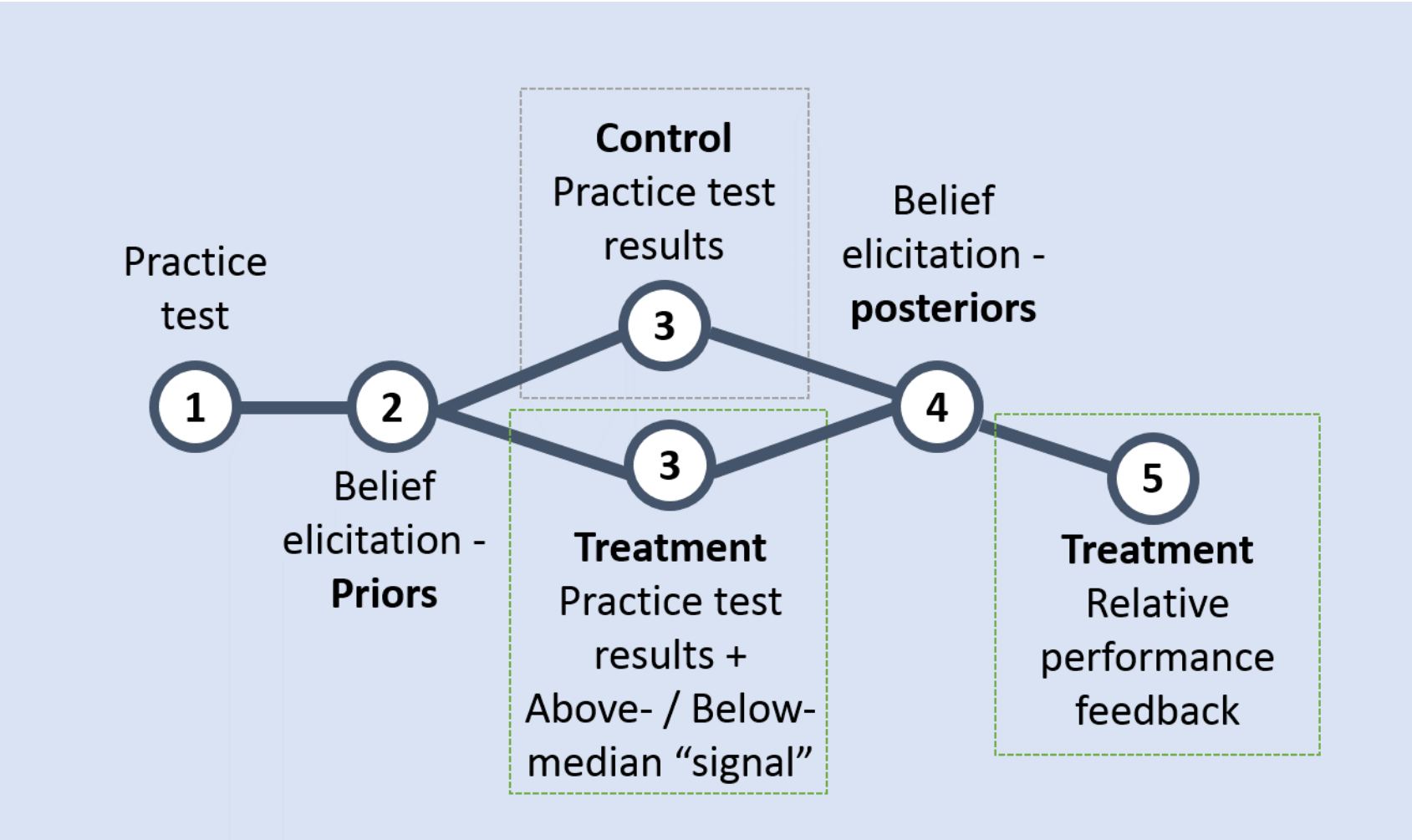
$$\mathbb{E}[y_i | T_i = 1, Q_i = q] - \mathbb{E}[y_i | T_i = 0, Q_i = q] = \beta_2 + \tau_q$$

Most students remain in the same or similar quartile relative to their initial performance Math

	Proportion of practice tests in reading quartile:			
	Q1=top	Q2	Q3	Q4=bottom
Q1 = top	0.089**	-0.056**	-0.043*	0.010
	(0.043)	(0.027)	(0.023)	(0.020)
Mean control	0.489	0.279	0.152	0.080
Q2	-0.071	0.015	0.054	0.002
	(0.055)	(0.037)	(0.035)	(0.031)
Mean control	0.364	0.270	0.217	0.149
Q3	-0.018	0.032	-0.004	-0.010
	(0.040)	(0.039)	(0.041)	(0.038)
Mean control	0.193	0.260	0.311	0.236
Q4 = bottom	-0.035	0.036	-0.012	0.011
	(0.042)	(0.042)	(0.041)	(0.059)
Mean control	0.151	0.241	0.313	0.295
N	3515	3515	3515	3515
N_clust	438	438	438	438

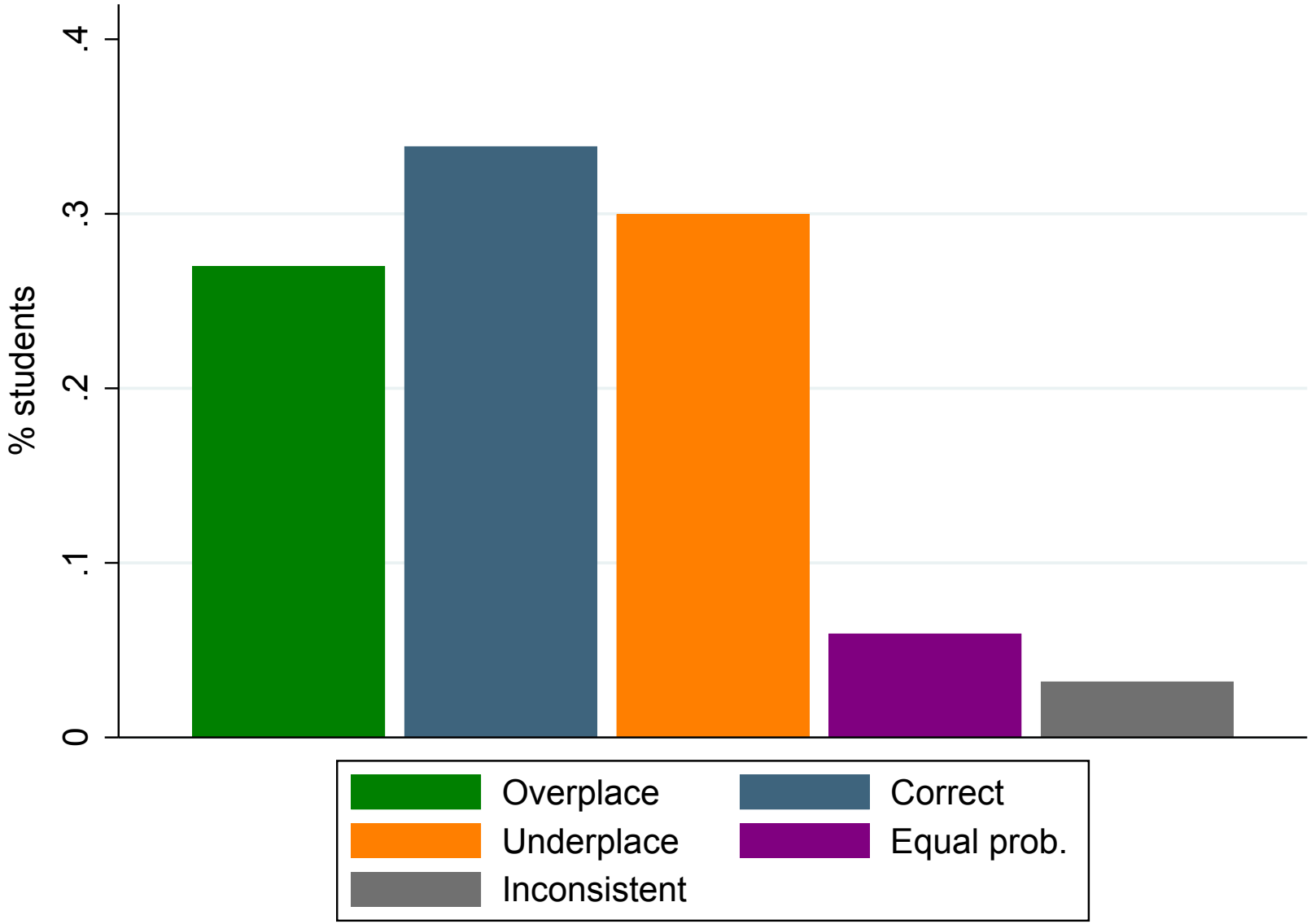
Biases in beliefs and effects of feedback on relative-performance beliefs

Lab-in-the-field: Timeline for rounds

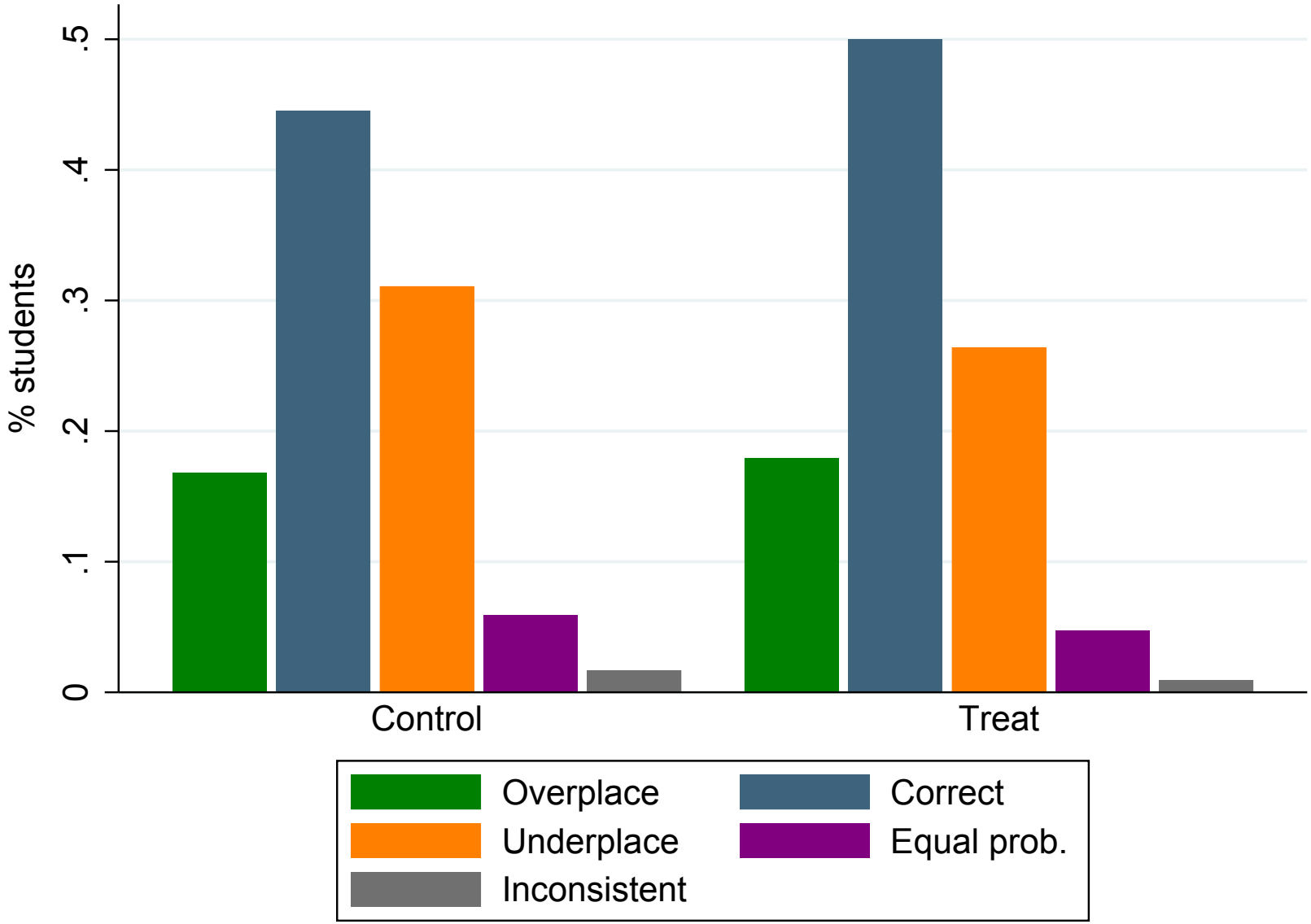


About 30% of students have correct prior beliefs

Rounds



Students become more correct in the posterior stage



Top performers are more likely to have correct reading priors when receiving feedback Math Posteriors

	Correct	Overplace	Underplace
Q1 = top	0.101**	-0.062*	-0.006
	(0.039)	(0.034)	(0.036)
Mean control	0.417	0.230	0.274
Q2	0.007	0.117**	-0.075
	(0.041)	(0.046)	(0.050)
Mean control	0.295	0.222	0.345
Q3	0.004	-0.008	-0.032
	(0.047)	(0.047)	(0.037)
Mean control	0.309	0.320	0.206
Q4 = bottom	0.044	-0.140*	0.048
	(0.060)	(0.071)	(0.059)
Mean control	0.288	0.365	0.160
N	2551	2551	2551
N_clust	433	433	433

Effects of relative performance feedback on academic investments and decisions

Poor-performing students receiving feedback are less likely to take practice tests

Rounds

	Takes practice tests	Math study hours	Reading study hours	Math correct answers	Reading correct answers
Q1 = top	0.011	0.819	0.278	0.696	0.490
	(0.011)	(0.594)	(0.569)	(0.736)	(0.518)
Mean control	0.953	5.018	4.449	21.688	22.856
Q2	0.010	-0.791	-0.114	-1.004	-1.290
	(0.019)	(0.856)	(0.792)	(0.875)	(0.793)
Mean control	0.926	6.179	5.348	18.640	20.831
Q3	0.011	-0.580	-0.291	0.391	-0.593
	(0.019)	(0.806)	(0.745)	(0.881)	(0.688)
Mean control	0.931	5.140	4.455	16.285	19.231
Q4 = bottom	-0.052***	-2.011*	-1.537*	-1.717*	-1.279
	(0.019)	(1.107)	(0.871)	(1.020)	(1.047)
Mean control	0.956	6.303	5.236	15.120	17.557
N	3645	2289	2285	3442	3442
N_clust	438	425	425	438	438

Top and bottom performers are less likely to take the exam

			ITT
	Did not take exam	**Never registered**	**Did not take exam**
Q1 = top	0.056**	0.059**	0.008
	(0.025)	(0.025)	(0.024)
Mean control	0.000	0.000	0.035
Q2	0.042	-0.000	-0.037
	(0.052)	(0.044)	(0.037)
Mean control	0.052	0.052	0.107
Q3	-0.016	-0.016	-0.041
	(0.024)	(0.024)	(0.025)
Mean control	0.021	0.021	0.062
Q4 = bottom	0.106*	0.104*	-0.025
	(0.057)	(0.056)	(0.036)
Mean control	0.000	0.000	0.091
N	438	438	985

Poor performers receiving feedback switch to easier majors

	Switched to harder major	Switched to easier major	Cutoff score first option	First option cutoff in top scores
Q1 = top	-0.030	0.024	0.899	0.077
	(0.099)	(0.071)	(1.587)	(0.071)
Mean control	0.235	0.088	80.267	0.438
Q2	-0.323**	-0.067	-1.927	-0.046
	(0.146)	(0.091)	(2.062)	(0.084)
Mean control	0.429	0.048	79.484	0.426
Q3	0.107	-0.066	0.041	0.104
	(0.141)	(0.103)	(1.963)	(0.086)
Mean control	0.150	0.100	78.918	0.298
Q4 = bottom	-0.242	0.257*	-1.986	-0.034
	(0.193)	(0.132)	(2.417)	(0.102)
Mean control	0.400	0.000	79.439	0.381
N	172	172	421	421

Effects of relative performance feedback on academic performance

Few differences in admission rates and exam performance

	Math score	Reading score	Total score	Admitted to first option	Admitted to second option
Q1 = top	1.632	-2.561	-0.375	-0.076	0.021
	(3.176)	(2.637)	(2.325)	(0.071)	(0.030)
Mean control	70.888	73.867	72.266	0.313	0.025
Q2	0.807	-1.725	-0.351	0.135	-0.045
	(4.766)	(4.539)	(3.450)	(0.084)	(0.031)
Mean control	60.644	63.163	61.849	0.130	0.037
Q3	1.052	-7.674	-5.221	0.004	-0.005
	(4.955)	(4.950)	(4.196)	(0.050)	(0.040)
Mean control	50.553	53.538	53.319	0.043	0.043
Q4 = bottom	-0.653	3.736	1.581	0.003	-0.014
	(5.862)	(6.120)	(4.974)	(0.068)	(0.020)
Mean control	42.377	46.339	44.360	0.071	0.024
N	421	421	421	421	421

No missings

Conclusion and discussion

- Information can discourage low-performing people
  - Reduce investments leading to take an important exam
  - Are less likely to take the exam
  - Consistent with "dropouts" behavior (Muller & Schotter, 2010)
- Elicited beliefs not consistent with beliefs revealed by behavior
  - How meaningful are belief elicitation mechanisms outside of the lab?
- Policy implications?
  - Efficient: Students who have higher chances of gaining admission will be competing for the slots
  - Effort vs. achievements?

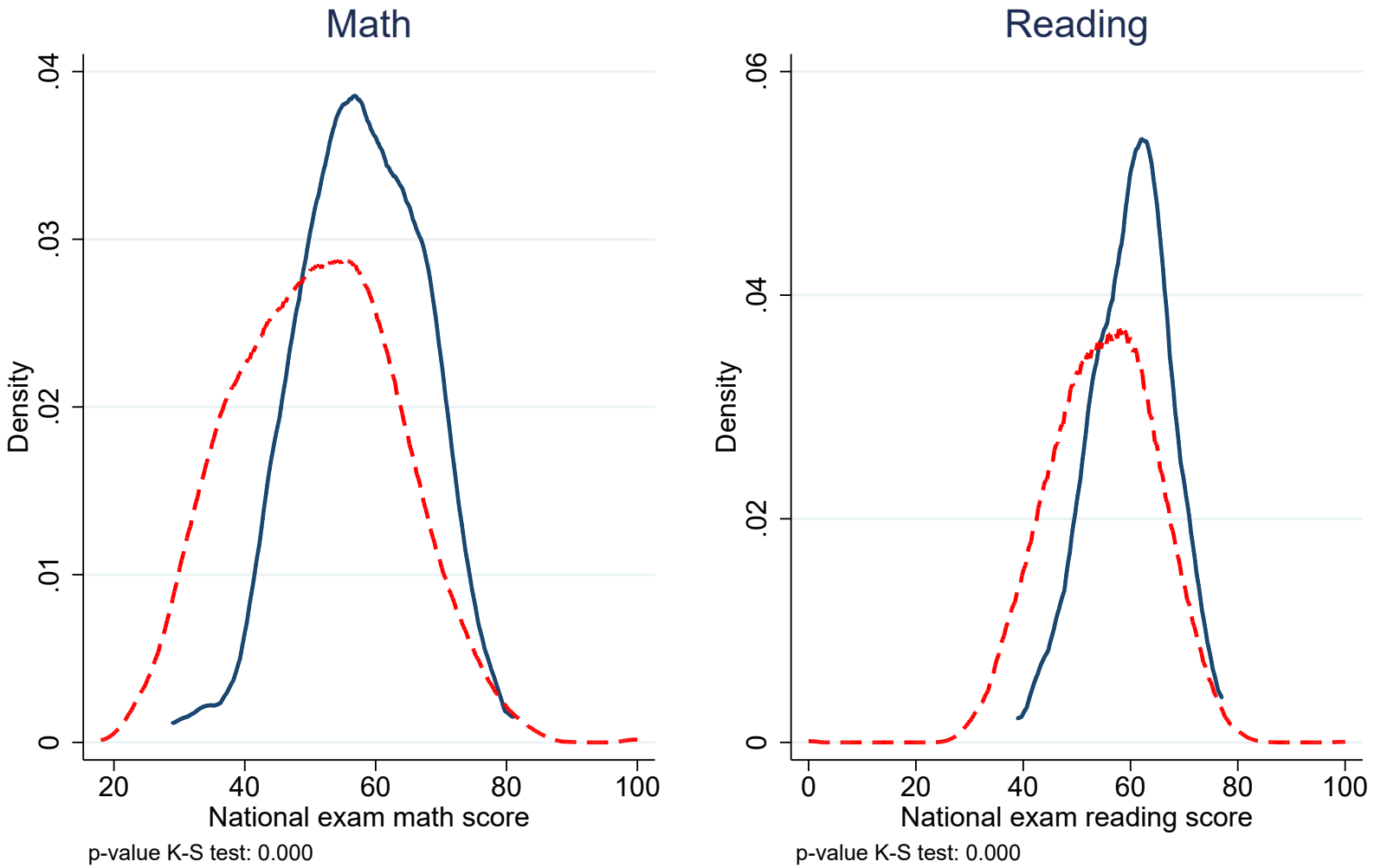
Admission rates of most competitive majors

Back

CARRERA	TOTAL INSCRITOS 1 Y 2 OPCIÓN	TOTAL ADMITIDOS	TASA DE ADMISION
PSICOLOGÍA	3,482	48	1.38%
MEDICINA	10,070	139	1.38%
ENFERMERÍA	3,974	58	1.46%
NUTRICIÓN Y DIETÉTICA	2,189	35	1.60%
INSTRUMENTACIÓN QUIRÚRGICA	1,979	33	1.67%
COMUNICACIÓN AUDIOVISUAL Y MULT.	1,214	25	2.06%
ODONTOLOGÍA	2,545	55	2.16%
MEDICINA VETERINARIA	2,816	62	2.20%
ADMINISTRACIÓN EN SALUD …	1,381	34	2.46%
TRADUCCIÓN INGLÉS-FRANCÉS-ESPAÑOL	1,620	41	2.53%
INGENIERÍA CIVIL	2,943	78	2.65%
ADMINISTRACIÓN DE EMPRESAS	3,225	91	2.82%
ENTRENAMIENTO DEPORTIVO	1,433	42	2.93%
LICENCIATURA EN LENGUAS EXTRANJERAS	1,433	42	2.93%
TRABAJO SOCIAL	1,835	59	3.22%
INGENIERÍA DE SISTEMAS	2,257	74	3.28%

Who are the students at the institute?

Back



Performance report control group

Back

EVALUACIÓN	FECHA	PUNTAJE
PRE-U DE A 2018-I	ENERO 23, 2018	2.5
NOMBRES - APELLIDOS	**CÓDIGO**	**# DE LISTA**
ESTUDIANTE 001 PRUEBA	1000001	1

TOTAL UdeA
25.0

RAZONAMIENTO LÓGICO MATEMÁTICO				
CATEGORÍA	# PREGUNTAS	CORRECTAS	INCORRECTAS	PUNTAJE
ANALISIS	10	2	8	2.0
TOTAL	10	2	8	2.0

COMPETENCIA LECTORA				
CATEGORÍA	# PREGUNTAS	CORRECTAS	INCORRECTAS	PUNTAJE
INTERPRETACION	10	3	7	3.0
TOTAL	10	3	7	3.0

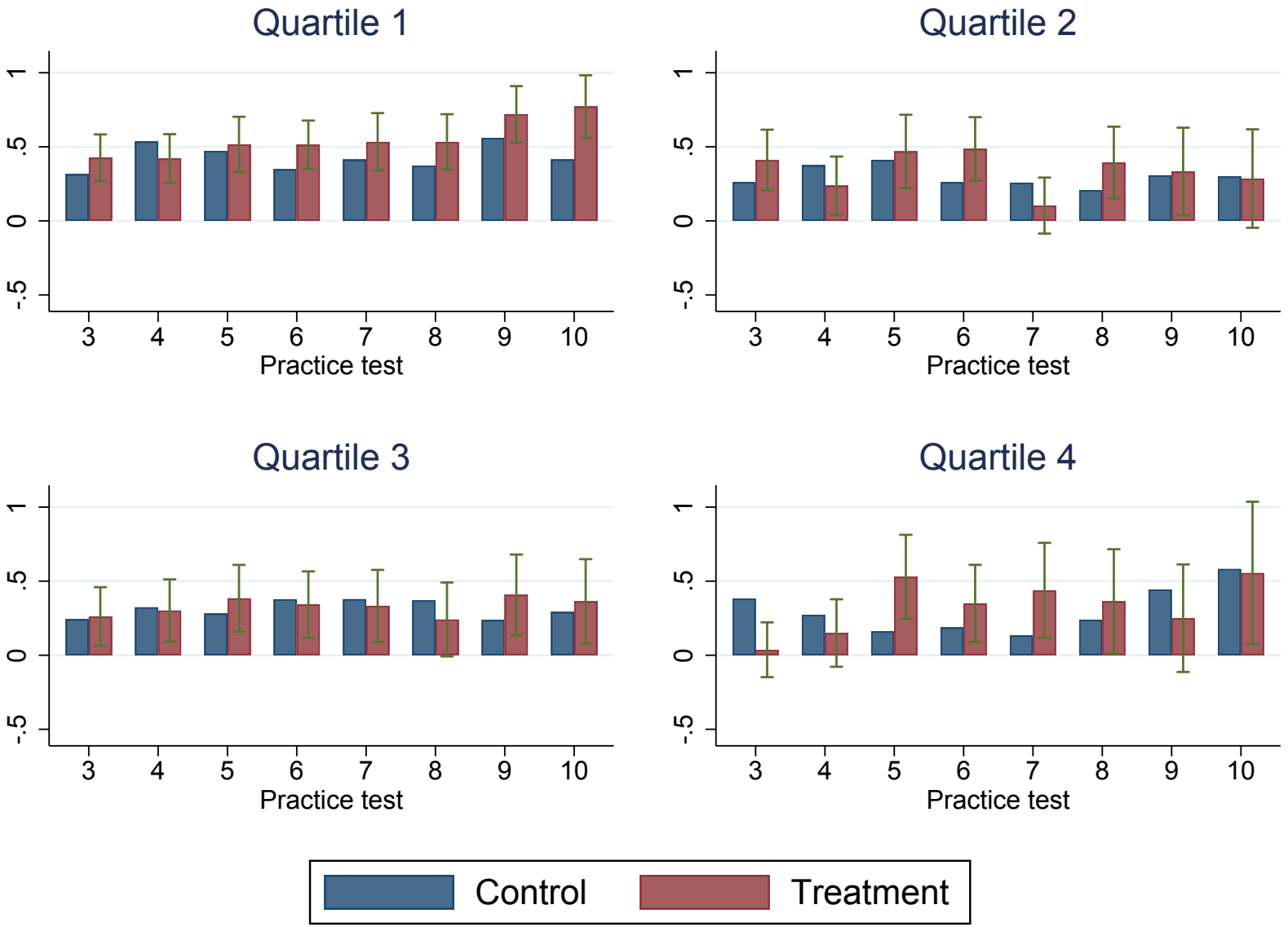
Most students remain in the same or similar quartile relative to their initial performance

Back

	Proportion of times in math quartile:			
	1 = top	2	3	4 = bottom
Panel A. Students in top quartile in initial practice test				
Treated	0.060	-0.021	-0.033	-0.007
	(0.055)	(0.036)	(0.025)	(0.022)
Constant	0.553***	0.237***	0.114***	0.096***
	(0.055)	(0.039)	(0.026)	(0.019)
Obs	1211	1211	1211	1211
No. students	168	168	168	168
Panel B. Students in bottom quartile in initial practice test				
Treated	-0.074	-0.046	0.063	0.057
	(0.045)	(0.058)	(0.049)	(0.071)
Constant	0.169***	0.337***	0.221***	0.273***
	(0.047)	(0.058)	(0.044)	(0.062)
Obs	510	510	510	510
No. students	75	75	75	75

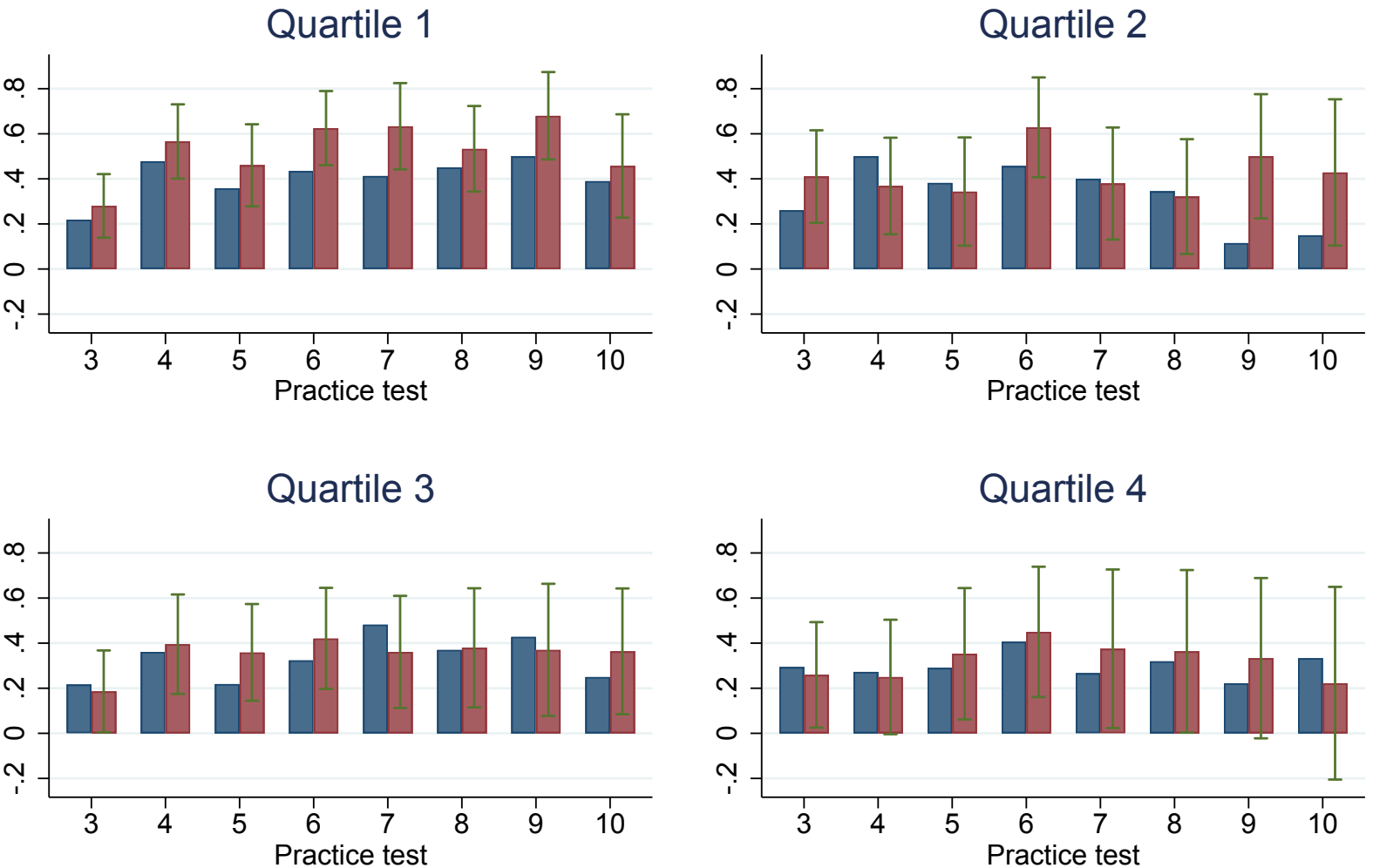
Correct prediction in reading by round and treatment

Back



Correct prediction in math by round and treatment

Back



Balance table by quartile

Back

	Q1 = top		Q2		Q3		Q4 = bottom	
	Control	Treat	Control	Treat	Control	Treat	Control	Treat
Female	0.652	0.562	0.604	0.592	0.652	0.683	0.527	0.563
Age	17.518	17.157	17.769	17.248	17.878	17.563	18.076	17.602
Single	0.974	0.975	1.000	1.000	1.000	0.981	0.951	1.000
Student	0.693	0.846**	0.763	0.667	0.649	0.714	0.793	0.760
Residential strata	2.627	2.866	2.485	2.336	2.319	2.540	2.608	2.304
Urban	0.910	0.892	0.891	0.888	0.877	0.938	0.876	0.896
Previously taken entrance exam	0.861	0.846	0.808	0.792	0.754	0.850	0.692	0.668
Math score (initial practice test)	3.915	3.842	3.061	3.108	2.763	2.751	2.023	1.992
Reading score (initial practice test)	6.383	6.555	5.189	5.242	4.171	4.286	2.821	2.444
Avg. practice test score in class	38.043	38.202	37.570	37.795	37.768	36.75*	36.190	36.435
AM course	0.925	0.874	0.902	0.928	0.820	0.922	0.863	0.914
PM course	0.303	0.361	0.382	0.416	0.446	0.468	0.459	0.480
Weekend course	0.009	0.008	0.013	0.000	0.008	0.015	0.009	0.012
Integrated UdeA - UNAL	0.007	0.02	0.021	0.027	0.051	0.012	0.043	0.058
Pre-medicine	0.166	0.191	0.145	0.138	0.107	0.122	0.199	0.101

Sampling frame and attrition

Back

	Q1 = top	Q2	Q3	Q4 = bottom	All
Panel A. Students who consented participation					
Assigned to control	149	132	135	107	523
Assigned to treatment	149	129	134	108	520
TOTAL	298	261	269	215	1,043
Fraction of all participants	28.6%	25.0%	25.8%	20.6%	
Panel B. Students who checked at least one performance report					
Assigned to control	80	58	48	42	228
Assigned to treatment	86	43	49	32	210
TOTAL	166	101	97	74	438
Fraction of all participants	37.9%	23.1%	22.1%	16.9%	
Fraction of participants in quartile	55.7%	38.7%	36.1%	34.4%	
Panel C. Statistics or report checking (conditional on checking at least one report)					
Average (out of 8)	2.70	2.42	2.35	2.04	2.45
Standard deviation	1.96	1.73	1.77	1.29	1.77
Minimum	1	1	1	1	1
Maximum	8	8	8	6	8
Average seconds spent in report	41.01	34.06	41.32	36.69	39.15

Balance of characteristics - attritors

Back

	Control	Treatment	P-value (T-C)	No. obs
Stratification variables				
Female	0.553	0.575	0.592	605
Previously taken entrance exam	0.797	0.793	0.910	604
AM course	0.447	0.461	0.733	605
PM course	0.237	0.242	0.894	605
Integrated UdeA - UNAL	0.058	0.062	0.849	605
Pre-medicine	0.061	0.064	0.859	605
Weekend course	0.197	0.171	0.417	605
Demographic variables				
Age	17.682	17.667	0.953	568
Single	0.969	0.974	0.734	568
Student	0.822	0.834	0.747	569
Residential strata	2.618	2.581	0.681	569
Urban	0.907	0.919	0.643	569
Academic variables				
Math no. correct (initial practice test)	11.060	11.019	0.894	604
Reading no. correct (initial practice test)	17.461	17.252	0.676	604
Avg. practice test score in classroom	37.607	37.872	0.220	604
Joint orthogonality test			0.9572	551

First stage IV

Back

	Checking report at least once	Checking intensity	Intensity (conditional on checking at least once)
Q1 = top	0.043	0.162	0.104
	(0.057)	(0.229)	(0.303)
Mean control	0.537	1.430	2.663
Q2	-0.106*	-0.163	0.230
	(0.060)	(0.197)	(0.354)
Mean control	0.439	1.015	2.310
Q3	0.008	-0.120	-0.368
	(0.058)	(0.187)	(0.356)
Mean control	0.356	0.904	2.542
Q4 = bottom	-0.098	-0.245	-0.158
	(0.065)	(0.167)	(0.312)
Mean control	0.393	0.822	2.095
N	1042	1042	438

Top performers are more likely to have correct math priors when receiving feedback

Back

	Correct	Overplace	Underplace
Q1 = top	0.124***	-0.070**	-0.016
	(0.042)	(0.031)	(0.043)
Mean control	0.395	0.179	0.341
Q2	0.003	0.078*	-0.013
	(0.053)	(0.047)	(0.058)
Mean control	0.363	0.154	0.323
Q3	0.014	-0.047	-0.033
	(0.051)	(0.055)	(0.045)
Mean control	0.328	0.280	0.232
Q4 = bottom	0.051	-0.154**	0.001
	(0.066)	(0.070)	(0.052)
Mean control	0.301	0.365	0.187
N	2551	2551	2551
N_clust	433	433	433

Top performers update more in reading when receiving the above- / below-median signal

Back

	Correct	Overplace	Underplace
Q1 = top	0.148***	-0.037	-0.114**
	(0.055)	(0.035)	(0.052)
Mean control	0.488	0.127	0.338
Q2	0.012	0.125**	-0.039
	(0.067)	(0.059)	(0.062)
Mean control	0.328	0.194	0.328
Q3	0.046	-0.007	-0.048
	(0.065)	(0.076)	(0.062)
Mean control	0.377	0.279	0.246
Q4 = bottom	0.033	-0.176**	-0.007
	(0.094)	(0.087)	(0.081)
Mean control	0.307	0.398	0.227
N	1072	1072	1072
N_clust	438	438	438

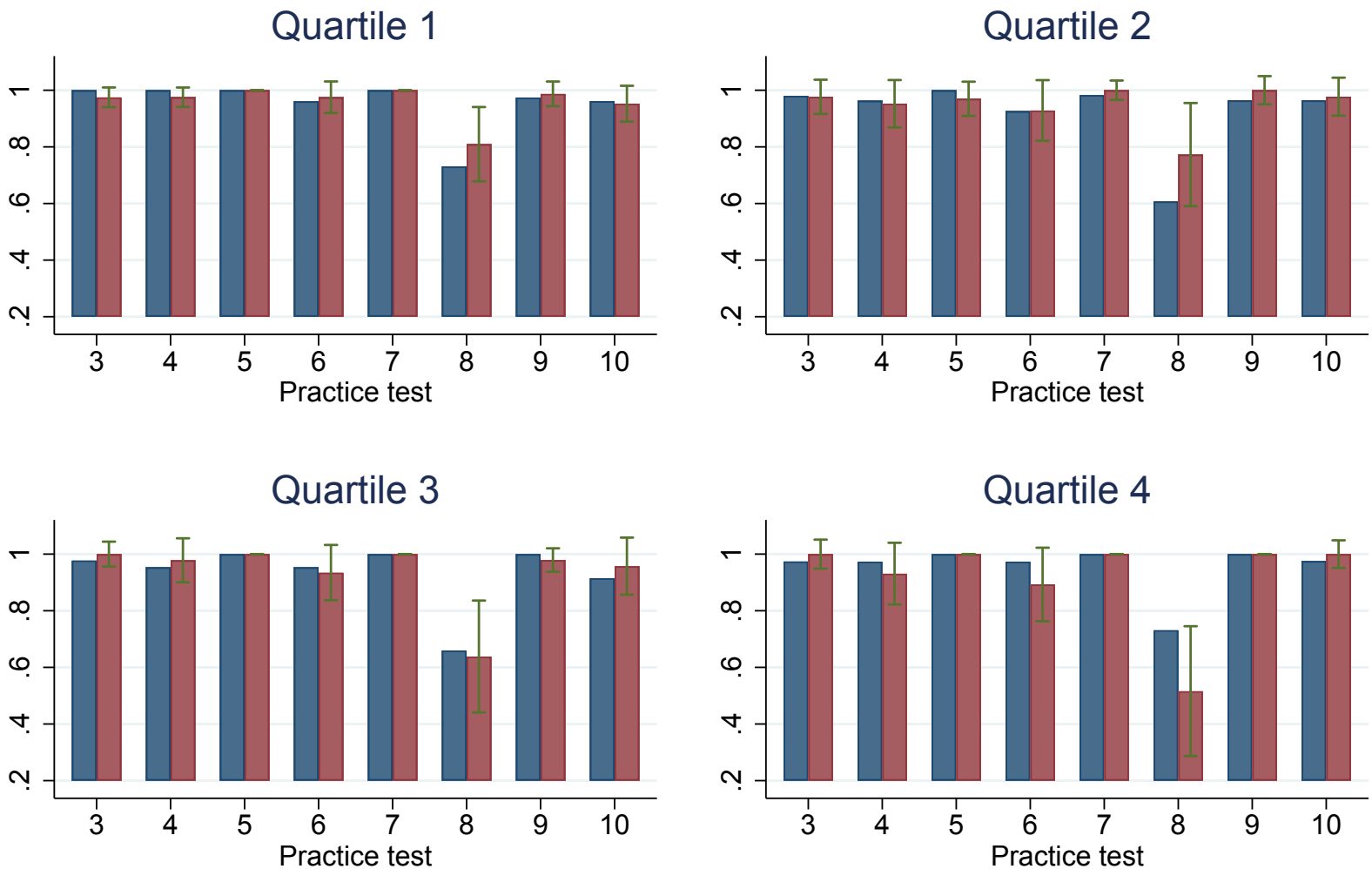
Top performers are more likely to have correct math posteriors when receiving feedback

Back

	Correct	Overplace	Underplace
Q1 = top	0.149***	0.017	-0.154***
	(0.051)	(0.034)	(0.051)
Mean control	0.490	0.096	0.351
Q2	0.040	0.014	0.076
	(0.073)	(0.053)	(0.070)
Mean control	0.422	0.141	0.273
Q3	0.101	-0.055	-0.059
	(0.071)	(0.066)	(0.074)
Mean control	0.391	0.227	0.300
Q4 = bottom	0.123	-0.163*	-0.097
	(0.098)	(0.095)	(0.077)
Mean control	0.289	0.361	0.253
N	1018	1018	1018
N_clust	419	419	419

Took practice test by round and treatment

Back



Similar results when replacing missings by zeros

Back

	Admitted to first option (zero if did not take exam)
Q1 = top	-0.088
	(0.070)
Mean control	0.313
Q2	0.124
	(0.079)
Mean control	0.121
Q3	0.008
	(0.050)
Mean control	0.042
Q4 = bottom	-0.011
	(0.064)
Mean control	0.071
N	438

Effects of relative performance feedback by gender

Bottom-performing treated men reduce investments

Back

	Takes practice tests		**Math study hours**		**Reading study hours**	
	Female	Male	Female	Male	Female	Male
Panel A. Students in top quartile in initial practice test						
Treated	0.018	0.002	0.848	0.696	0.077	0.495
	(0.014)	(0.017)	(0.829)	(0.801)	(0.805)	(0.720)
Mean control	0.956	0.946	5.312	4.456	4.763	3.852
DiD F vs. M	0.015		0.151		-0.418	
	(0.022)		(1.149)		(1.072)	
Panel D. Students in bottom quartile in initial practice test						
Treated	-0.038	-0.067**	-0.737	-3.727**	-1.476	-1.504
	(0.025)	(0.030)	(1.516)	(1.520)	(1.251)	(1.198)
Mean control	0.949	0.963	6.010	6.611	5.390	5.074
DiD F vs. M	0.029		2.989		0.028	
	(0.039)		(2.149)		(1.739)	

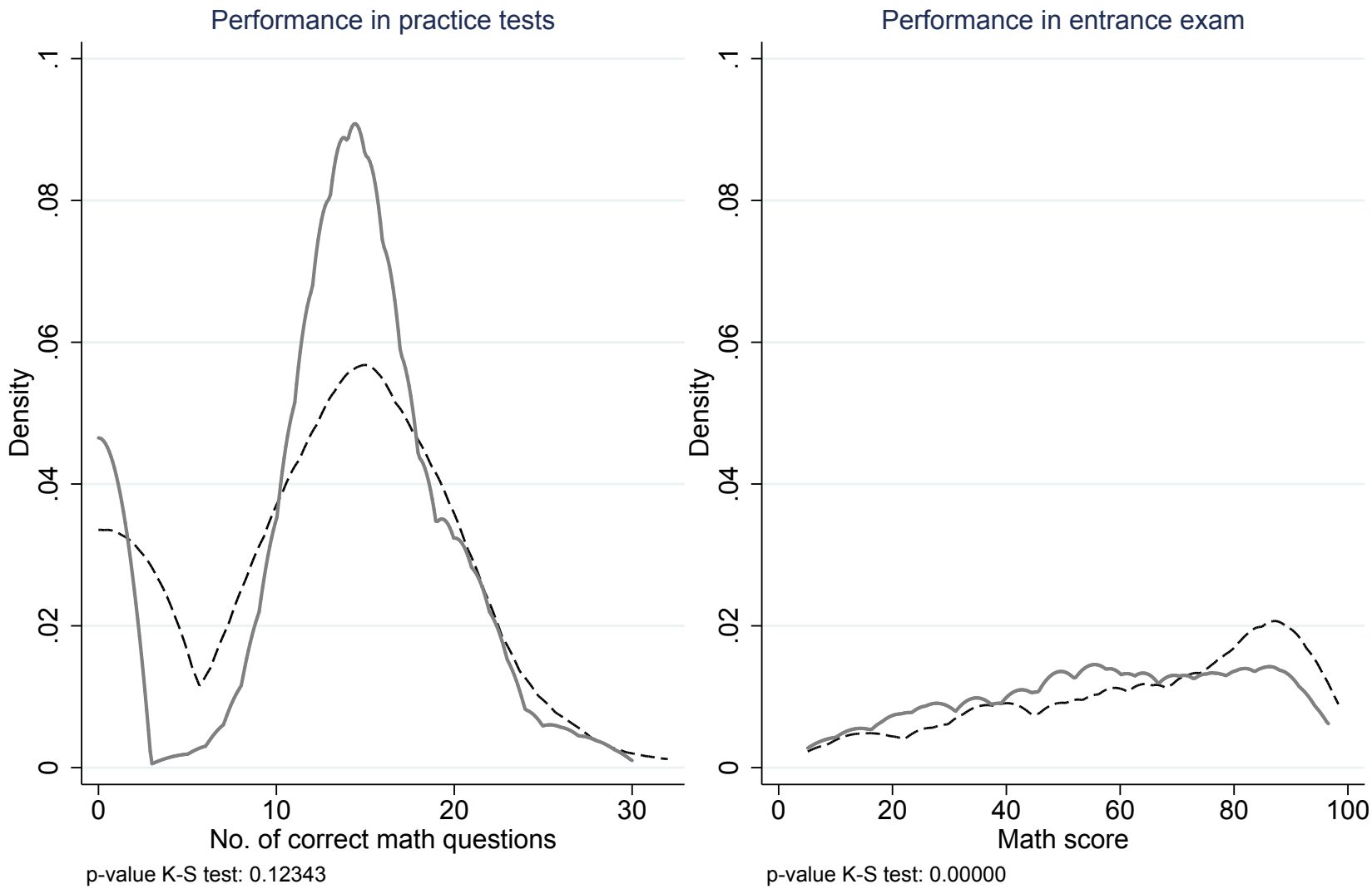
Treated women are less likely to take the exam

Back

	Did not take exam		**Never registered**	
	Female	Male	Female	Male
Panel A. Students in top quartile in initial practice test				
Treated	0.072***	0.035	0.075**	0.037
	(0.036)	(0.029)	-0.036	-0.029
Mean control	0.000	0.000	0.000	0.000
DiD F vs. M	0.037		0.038	
	(0.045)		(0.045)	
Panel B. Students in bottom quartile in initial practice test				
Treated	0.119**	0.091	0.123**	0.081
	(0.081)	(0.075)	(0.080)	(0.072)
Mean control	0.000	0.000	0.000	0.000
DiD F vs. M	0.029		0.042	
	(0.109)		(0.106)	

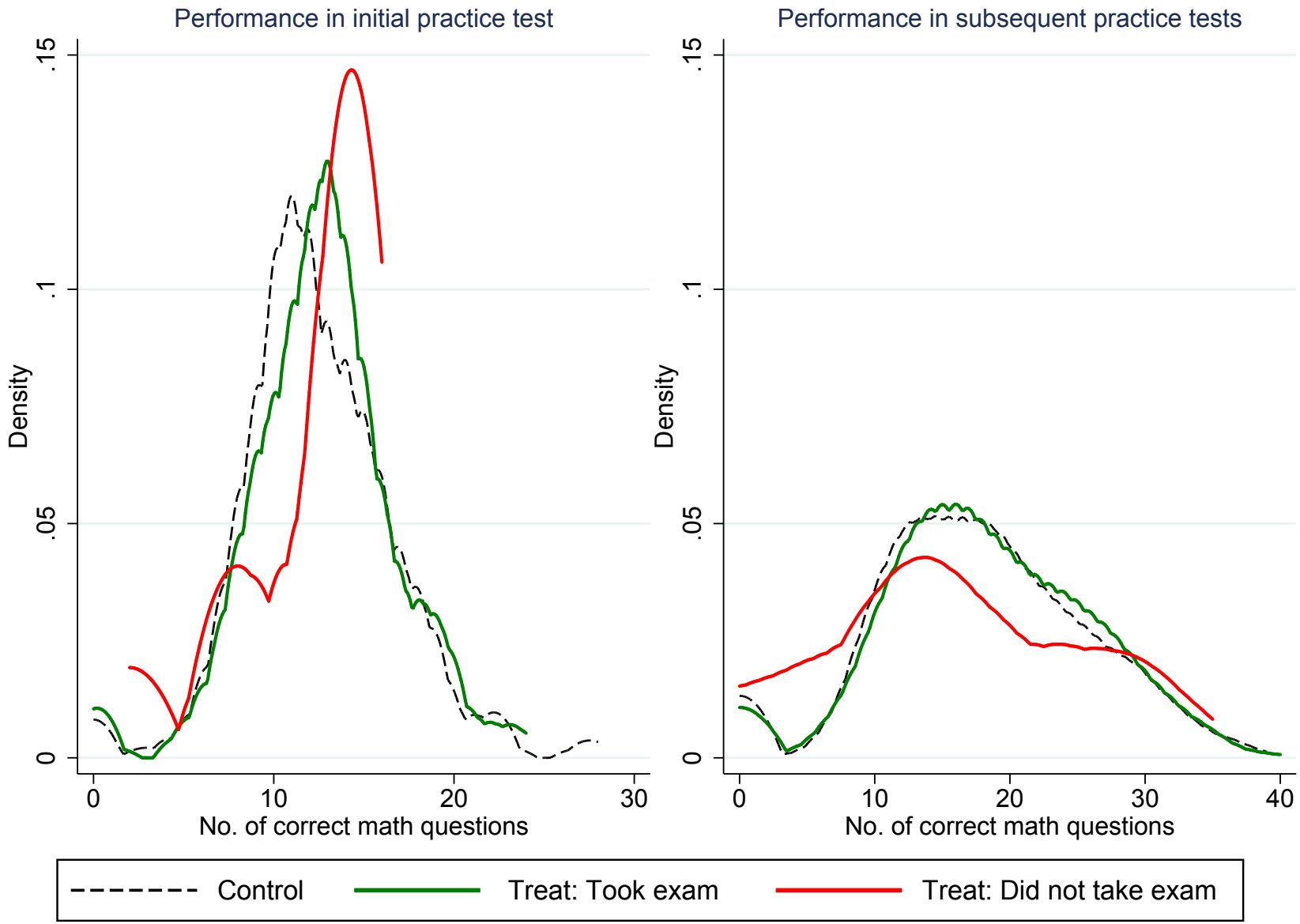
Men outperform women in the exam but not in prac. tests

Back



Explaining who is most responsive to relative performance feedback

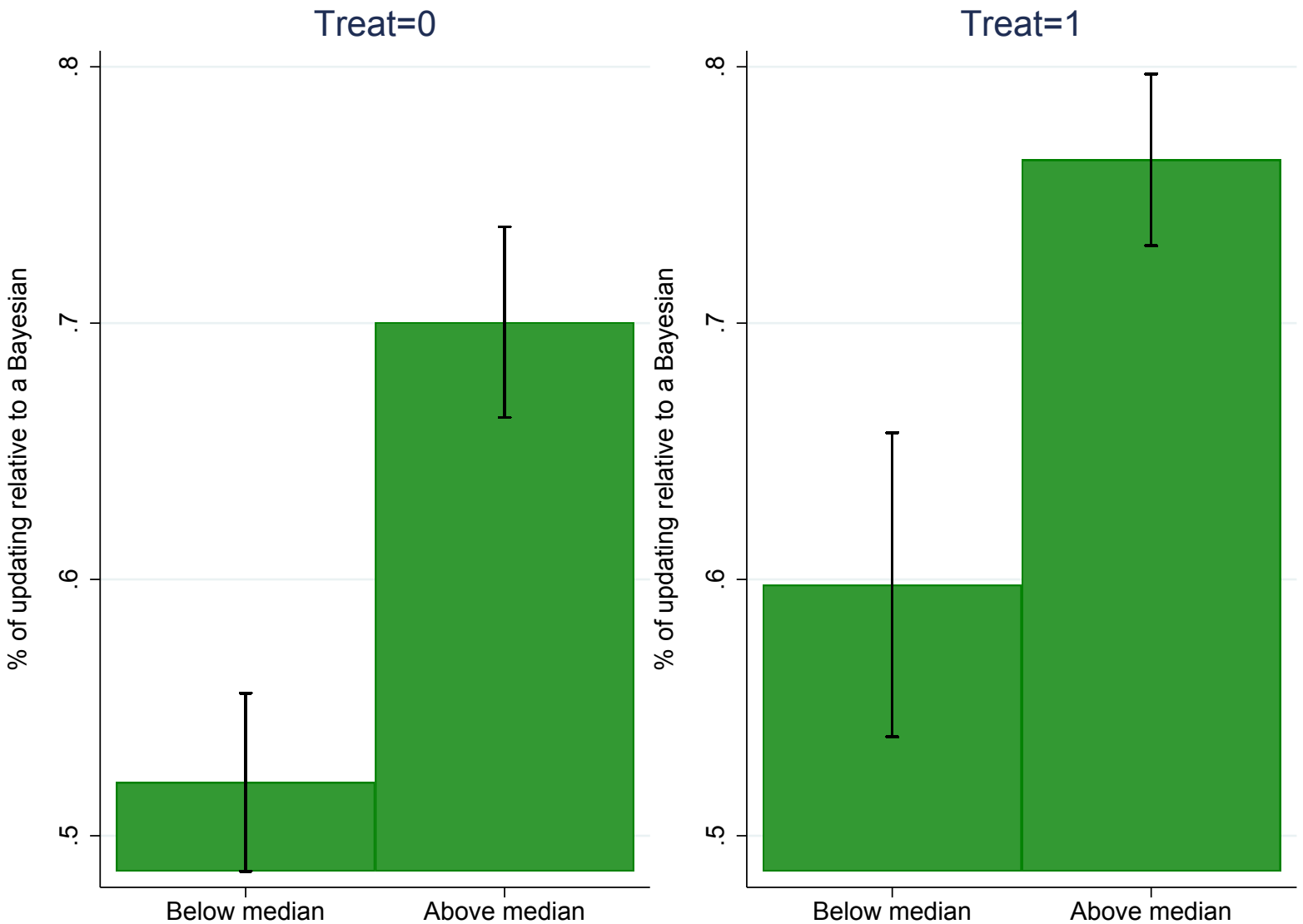
Who is dissuaded from taking the exam? Back



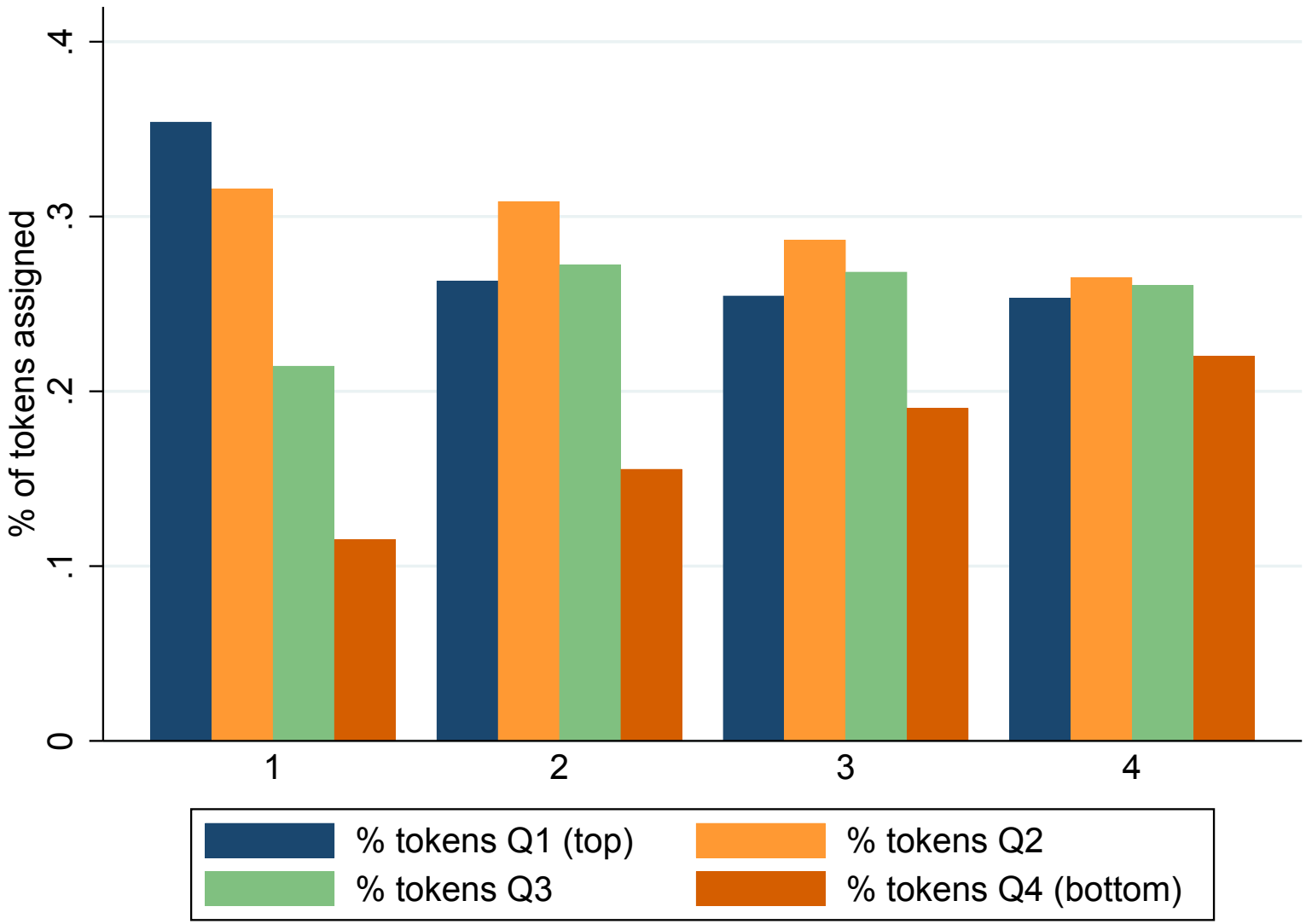
Why students' actions do not match their reported beliefs?

Students update like subjects in the lab

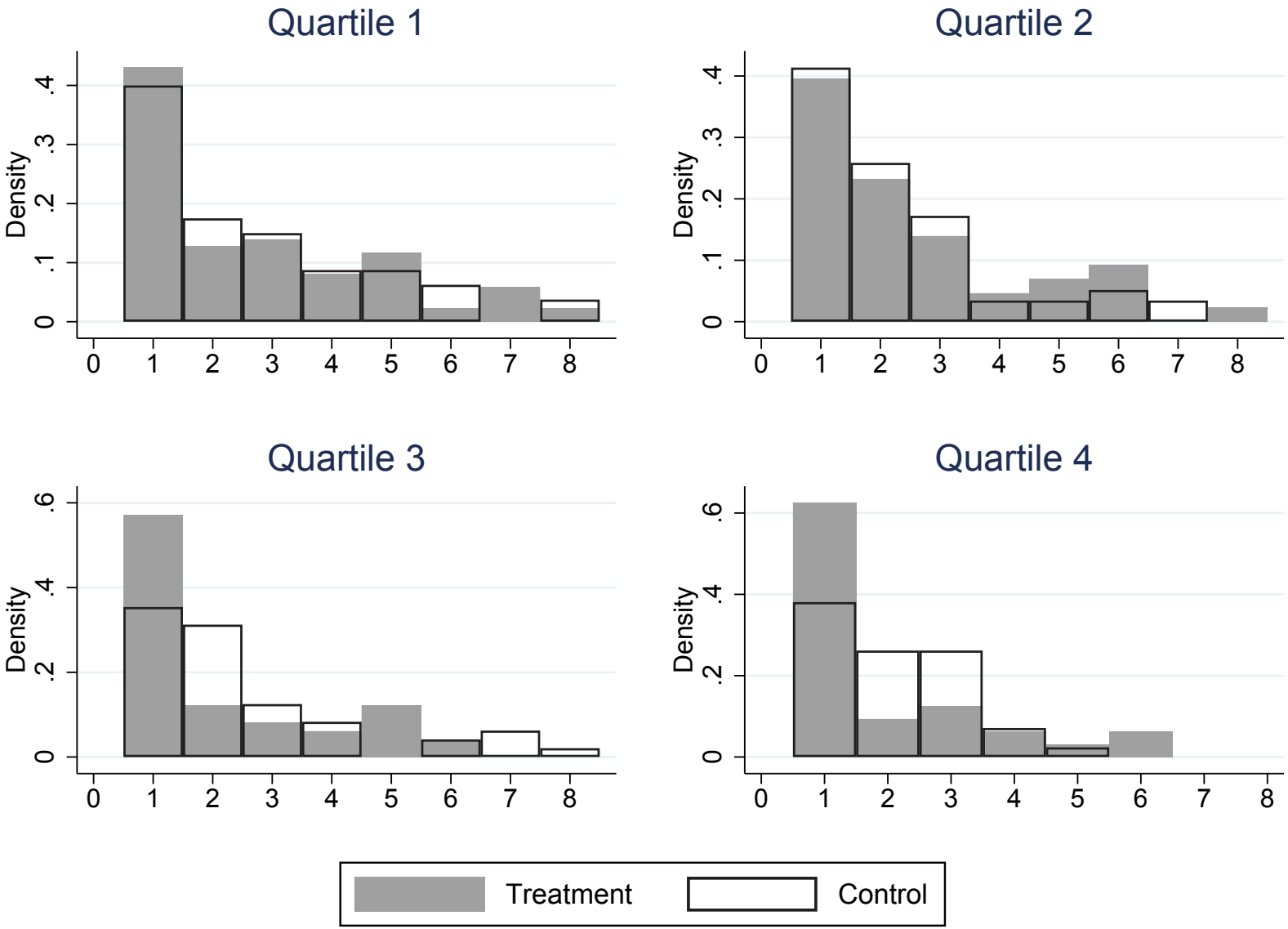
Conservatism and asymmetry



Hypothesis 1: Students do not understand the task or have limited attention



Hypothesis 2: Optimistic self-deception



Other hypotheses

- **Hypothesis 3**: High-stakes vs. low-stakes
  - Practice tests are meaningful
  - Lab-in-the-field task perceived as not important
- **Hypothesis 4**: Belief elicitation does not elicit the beliefs we want
  - Good approximation to turn latent into observable (Schotter & Trevino 2014)
  - Need more evidence on how meaningful outside of the lab

Behavioral theories on self-confidence

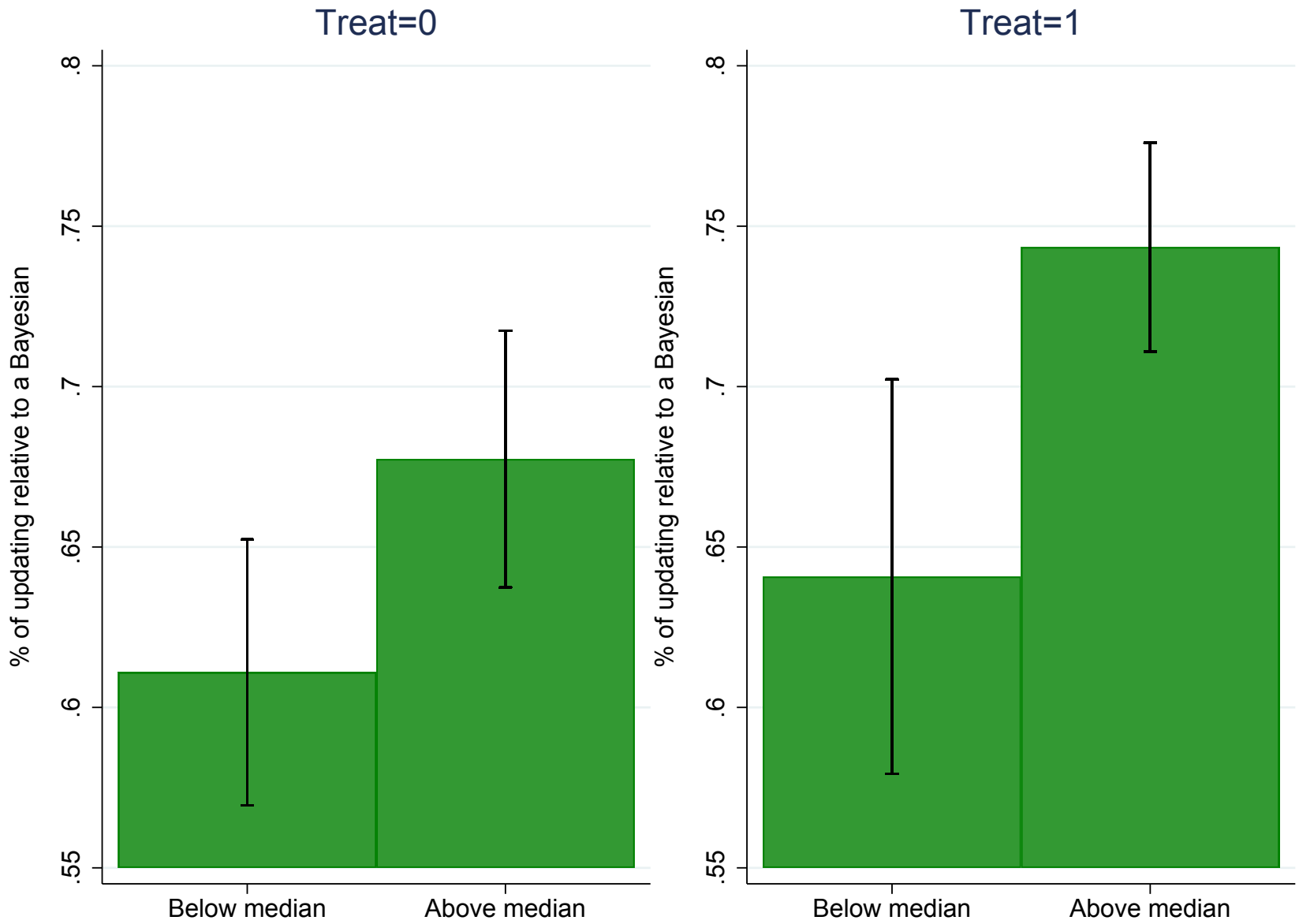
1. Biases in information processing
   - People do not update like Bayesians in the lab: conservatism and asymmetry
   - E.g., subjects update about 35% of what a Bayesian would and update more when receiving a positive rather than a negative signal (Mobius et al., 2014)
2. Self-relevance of beliefs (Koszegi, 2006; Weinberg, 2006)
   - Individuals derive utility from having a high belief about themselves
3. Confirmatory bias (Rabin & Schrag, 1999)
   - Individuals update more when receiving a signal confirming their prior than when receiving a disconfirming signal

Updating according to Bayes' rule

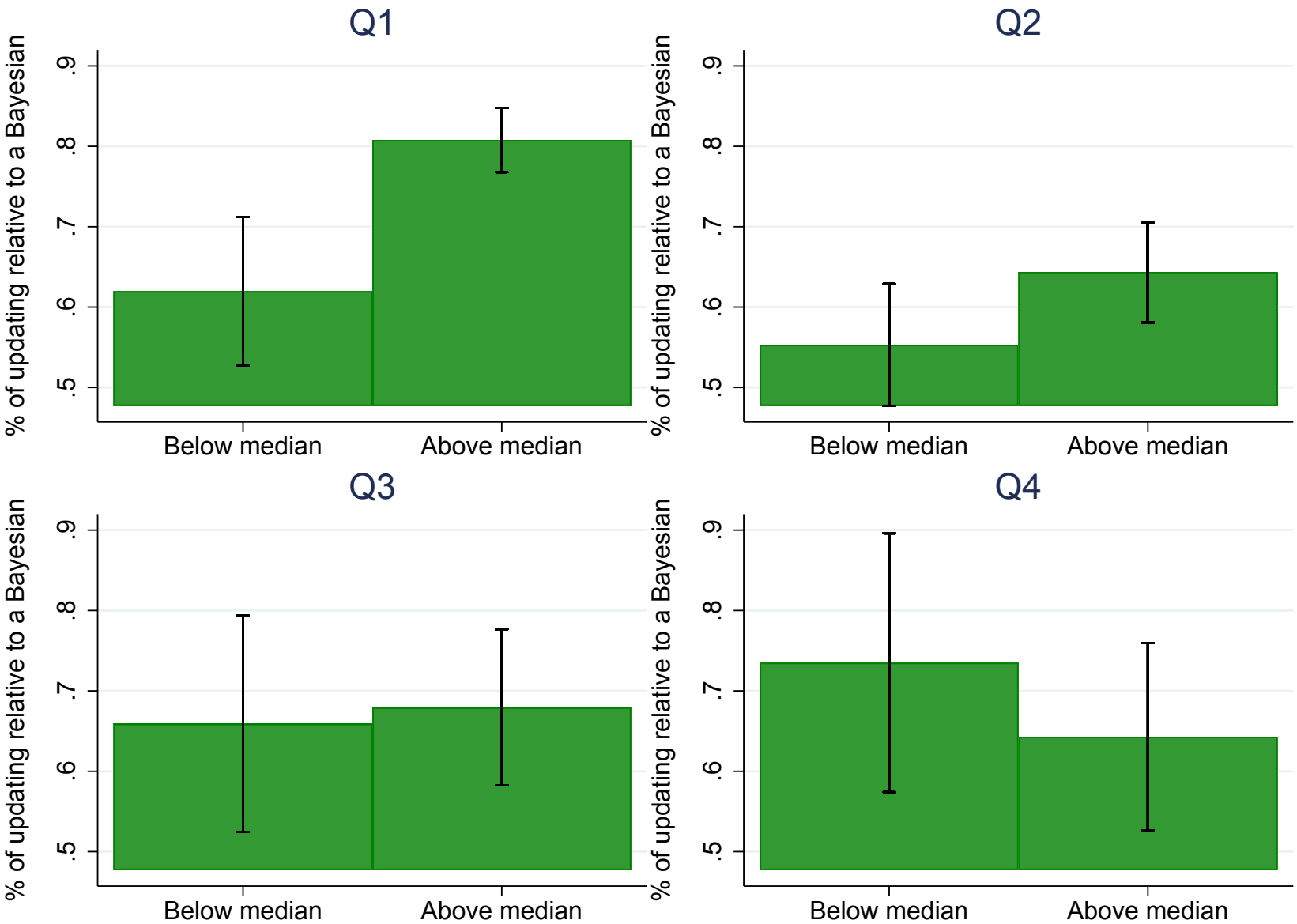
Back

- $$\mathbb{P}[Q1|S_i = \text{“}Top\text{”}] = \frac{\mathbb{P}[S_i = \text{“}Top\text{”}|Q1] \cdot \mathbb{P}[Q1]}{\mathbb{P}[S_i = \text{“}Top\text{”}|Q1] \cdot \mathbb{P}[Q1] + \mathbb{P}[S_i = \text{“}Top\text{”}|Q2] \cdot \mathbb{P}[Q2]}$$
- $$\mathbb{P}[Q2|S_i = \text{“}Top\text{”}] = \frac{\mathbb{P}[S_i = \text{“}Top\text{”}|Q2] \cdot \mathbb{P}[Q2]}{\mathbb{P}[S_i = \text{“}Top\text{”}|Q1] \cdot \mathbb{P}[Q1] + \mathbb{P}[S_i = \text{“}Top\text{”}|Q2] \cdot \mathbb{P}[Q2]}$$
- $$\mathbb{P}[Q3|S_i = \text{“}Bottom\text{”}] = \frac{\mathbb{P}[\text{“}Bottom\text{”}|Q3] \cdot \mathbb{P}[Q3]}{\mathbb{P}[\text{“}Bottom\text{”}|Q3] \cdot \mathbb{P}[Q3] + \mathbb{P}[\text{“}Bottom\text{”}|Q4] \cdot \mathbb{P}[Q4]}$$
- $$\mathbb{P}[Q4|S_i = \text{“}Bottom\text{”}] = \frac{\mathbb{P}[\text{“}Bottom\text{”}|Q4] \cdot \mathbb{P}[Q4]}{\mathbb{P}[\text{“}Bottom\text{”}|Q3] \cdot \mathbb{P}[Q3] + \mathbb{P}[\text{“}Bottom\text{”}|Q4] \cdot \mathbb{P}[Q4]}$$

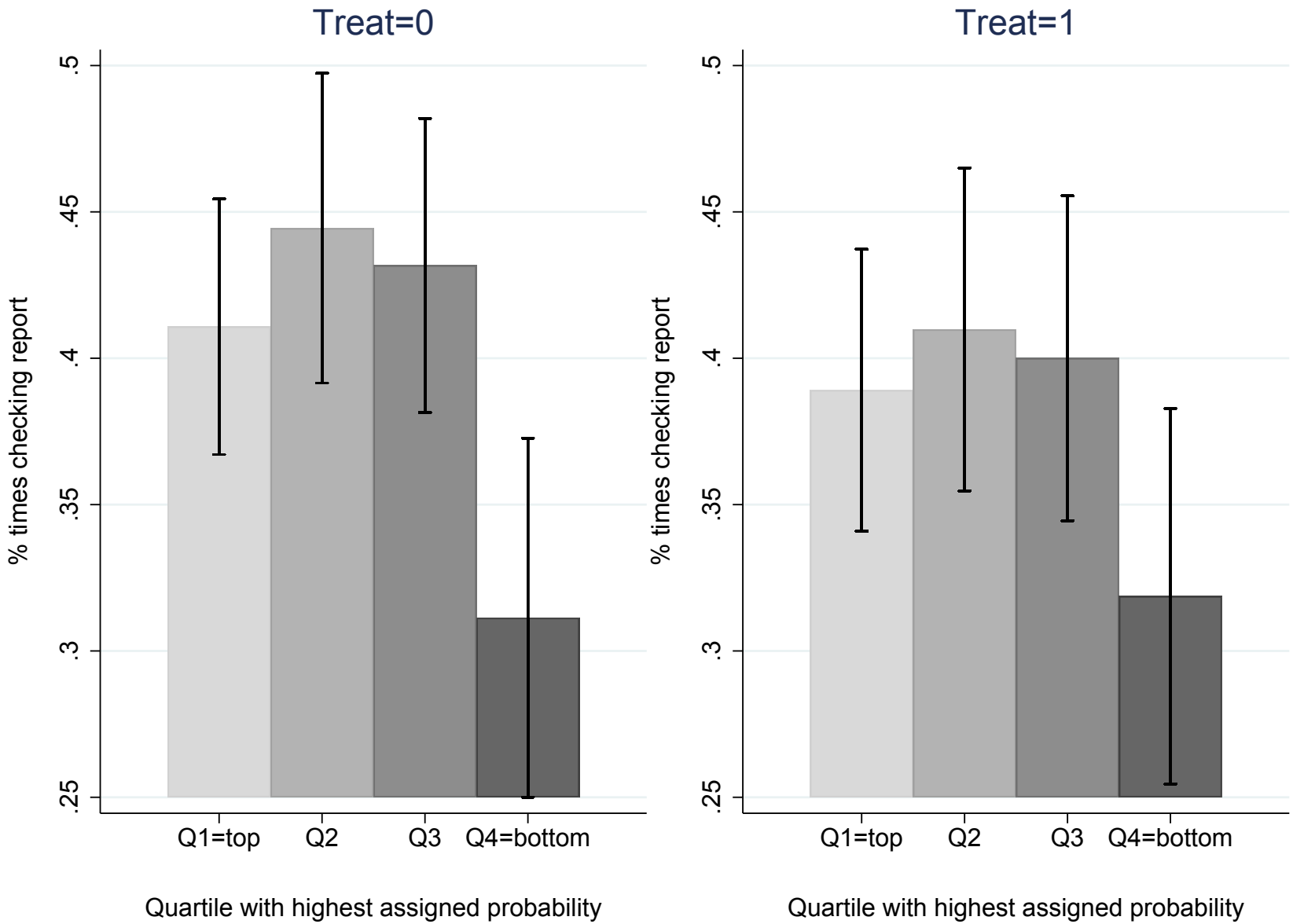
Conservatism and asymmetry - Math



Asymmetry varies by quartile in initial practice test - Math



Koszegi's model prediction does not hold - Math



Evidence for confirmatory bias - Math