Topics

Arnt O. Hopland and Ole Henning Nyhus*

Gender Differences in Competitiveness: Evidence from Educational Admission Reforms

DOI 10.1515/bejeap-2015-0008

Published online September 24, 2015

Abstract: This paper studies whether increased competition has adverse consequences for students' intrinsic motivation by studying an upper secondary admission reform in Norway. While earlier students were enrolled into their neighboring school, the new system introduces school choice, where admission is based on performance in lower secondary school. Using a difference-in-differences approach, we find that whereas the motivation for boys seems to be unaffected by the increased competition, there are adverse consequences on the motivation for girls.

Keywords: gender differences, competitiveness, student motivation, natural experiment

JEL classification: D02, D03, I20, I28, J16

1 Introduction

Experimental economists have found significant gender differences in preferences for competitive settings and how outcomes are affected by competition.¹ The general conclusions are that males are more in favor of competitive settings and perform better under competition than females. Such differences might have the potential to explain gender differences in outcomes in a variety of settings,

¹ For studies on willingness to compete, see, e.g., Hogarth, Karelaia, and Trujillo (2011), Niederle and Vesterlund (2007), or the literature review by Croson and Gneezy (2009). For studies on different outcomes between the genders in competitive settings, see, e.g., Gneezy, Niederle, and Rustichini (2003), and Gneezy and Rustichini (2004).

***Corresponding author: Ole Henning Nyhus**, Department of Economics, Norwegian University of Science and Technology, Dragvoll, N-7491 Trondheim, Norway, E-mail: ole.nyhus@sof.ntnu.no
Arnt O. Hopland, NoCeT and the Department of Business and Management Science, Norwegian School of Economics, Helleveien 30, N-5045 Bergen, Norway, E-mail: arnt.hopland@nhh.no

e.g., in the labor market. In recent years, this literature has been supplemented by field studies which in general provide more mixed results.² Hence, more empirical studies of gender differences in competitiveness are of great interest.

In Norway, the counties decide how students are allocated between upper secondary schools, and an important issue is whether admission is based on the neighborhood principle or performance-based ranking (school choice). Whereas the neighborhood rule involves little competition between students, performance-based ranking involves competition among lower secondary students because they can compete for enrollment in schools outside their defined neighborhood.³ In the empirical analysis, we use survey data on student motivation in lower secondary education collected by the Norwegian Directorate for Education and Training to analyze a reform in Oslo. The reform took place in 2009 and involved a change from an admission system based on the neighborhood rule to a system with performance-based ranking.

We analyze the reform using a difference-in-differences (DD) framework, with two different groups of untreated students as control. First, we compare 10th-grade students (last year of lower secondary school) in Oslo to 10th-grade students in counties that kept the neighborhood rule throughout the period. Second, we use Oslo students in the last year of primary school (7th grade) as an alternative control group. These students are not directly affected by the reform, since they will not apply for upper secondary until three years later. The second strategy is thus not sensitive for county heterogeneity. Importantly, the two strategies produce similar results, and both control groups pass formal placebo tests.

The results indicate interesting gender differences. Whereas boys' motivation is unaffected by the reform, there is a sizable negative effect on girls' motivation. Moreover, the differences in coefficients between girls and boys are in most cases statistically significant. Our findings are thus in line with findings in the experimental literature indicating that females are less in favor of competitive settings than males.

A potential mechanism can be that girls might be more skeptical to uncertainty of the final outcome, i.e., which upper secondary school they eventually end up in. We consider this to be consistent with the observation that females in

² See, e.g., Ors, Palomino, and Peyrache (2013); Jurajda and Munich (2011); Kleinjans (2009); Lavy 2013.

³ Haraldsvik (2012) studies one of the reforms we utilize and finds a significant effect on competitiveness, since there is an increase in students coming in from other schools than the neighboring ones. Second, using a DD approach similar to ours, she identifies a positive effect on lower secondary student achievement. Hence, the incentives implied by the increased competition are clearly real and of sufficient magnitude to affect student behavior.

general seem to be less willing to take risk (see, e.g., Byrnes, Miller, and Schafer 1999).

We also study a very similar reform that took place in the county Hordaland a few years earlier, effective from the school year 2005/2006. Because of better data, the results for Oslo are the most reliable, while the study of the reform in Hordaland acts as a valuable supplement and test of robustness. The results are similar to those obtained when studying the Oslo reform. It is reassuring that we observe very similar outcomes from two similar reforms in two different counties at two different points in time, since it is unlikely that the same spurious trend would occur in both cases.

The remainder of the paper is organized as follows: Section 2 provides a brief overview of the related literature. In Section 3 we present the relevant institutional background and discuss the reforms of interest. We present our data and empirical strategy in Section 4, before the main results are reported in Section 5. In Section 6 we present some specification tests, before we summarize our findings in Section 7.

2 Related Literature

Gender differences have received much attention in the experimental economics literature. The experimental data vary from controlled experiments on adults (e.g., Gneezy, Niederle, and Rustichini (2003); Niederle and Vesterlund (2007)), or children and youths (e.g., Gneezy and Rustichini (2004); Sutter and Glätzle-Rützler (2014)), to the use of data tapped from televised game shows (e.g., Hogarth, Karelaia, and Trujillo (2011)). The typical findings are that males are more in favor of competitive settings and perform better under competition than females.

This paper adds to the recent literature using field data, which reaches more mixed conclusions than the experimental studies.⁴ Kleinjans (2009) investigates whether gender differences in preferences for competition can explain some of the differences in career choices. Her findings suggest that women's greater distaste for competition decreases educational achievement. She also finds that different preferences for competition can explain part of the gender segregation in occupational fields. Jurajda and Munich (2011) study multiple university entrance exams taken by the same individuals and find that women do not

⁴ In a recent contribution, Filippin and Crosetto (2014) point out that even the experimental evidence of gender differences is less conclusive than what is commonly believed.

shy away from applying to more competitive institutions. Interestingly, men outperform women in entrance exams for the more prestigious institutions while no such differences exist for the exams for the less competitive schools. Ors, Palomino, and Peyrache (2013) find that males outperform females on high-competitive tests, while it is the other way around in tests with low competition. However, Lavy (2013) finds no gender-based difference in a rank-order tournament among high school teachers where prizes are based on their students' performance. Another strand of the literature suggests that males are more overconfident than women (e.g., Barber and Odean (2001); Correll (2001)), which may potentially explain why it seems that females dislike competition more than males.

This paper is also related to the school choice literature, which mainly analyzes how parental school choice works as an incentive for the school to improve productivity and quality of schooling. There are two main mechanisms outlined in this literature. First, students in below average schools get the chance to attend better schools, thus improving their achievement. Second, once students start to leave the neighborhood school in favor of a better school, schools will want to compete for the best students and therefore have incentives to provide the best possible education to the student consumer.⁵

3 Institutional Background and the Reforms

Since 1994, all who graduate from lower secondary education in Norway have a legal right to three years of upper secondary education. The legal right lasts for five years after completed lower secondary education, but expires prior to this if the student reaches 24 years of age.⁶ The legal right to upper secondary education relates to education program and not to school. Hence, even though the students can prioritize among schools, they are not entitled to admission to a specific school.

The upper secondary schools are owned and operated by the 19 Norwegian counties, and the admission systems vary from county to county, mainly between a neighborhood rule and a system based on performance-based ranking. When

⁵ Economics of school choice emerged from Friedman (1955) who argued that neighborhood rules are inefficient as they prohibit competition among schools. More recently Figlio and Hart (2014) examine how student test scores are influenced by school competition and give an overview of earlier studies.

⁶ No students fail any grade in Norwegian compulsory education. The 24 year rule is thus, in practice, a non-binding constraint.

subject to a neighborhood admission system for upper secondary schooling, the students have limited possibility to influence school admission apart from choosing education program. The neighborhood admission system is quite rigid compared to the performance-based system where the students can prioritize among all upper secondary schools in the county. The schools that have more applicants than study places then rank students based on their final grade point average (GPA) from lower secondary school.⁷

The reform in Oslo was announced in the fall of 2008. The reform involved that the existing system where 50% of the slots at each school were reserved for the students residing in the neighborhood of the school would be replaced by performance-based ranking for all slots across all schools within the county.⁸ The new system was effective from the following school year (2009/2010), i.e., the students graduating from lower secondary education in Oslo in 2009 were the first to be affected by this reform.

In early spring, around the same time as the survey is conducted, the lower secondary graduates rank their upper secondary programs and schools of choice. Upper secondary education is divided into two main tracks, academic and vocational. Almost 50% of the students attend an academic track, and because the costs associated with these programs are about 30% lower than vocational study tracks, they are offered at roughly half of the schools.⁹ The vocational track consists of nine different education programs, where no program singles out as the dominant one. Owing to the limited demand for these programs and the fact that they can be costly to offer, most vocational tracks are offered only at a few schools within each county. Students applying for these thus face a limited school choice even under the performance-based system.¹⁰ However, since the share of students that choose the academic track is quite evenly distributed between the genders, different treatment of boys and girls should not be a problem for our study. Students in the academic track have higher GPA than students in the vocational tracks, and as grades are correlated with socio-economic background, one might have that students from more wealthy backgrounds are more affected by the competition. We will return to this in the next section.

⁷ The GPA consists of teacher assessed grades and grades from final exit examinations, and all grades are weighted equally.

⁸ Prior to the reform in 2009/2010, Oslo was divided into four such neighborhoods.

⁹ The difference in expenditures is mainly due to higher wage expenditures (higher teacher/student ratio) in vocational study programs, and to some extent expenditures associated with work shops, labs, and materials.

¹⁰ Haraldsvik (2012) studies this formally and finds evidence of increased competition mainly in academic tracks when studying the admission reform in Hordaland county in 2005.

4 Data and Empirical Strategy

4.1 Survey Data

Our left-hand side variable is tapped from survey data collected by the Norwegian Directorate for Education and Training since the school year 2006/2007. The survey is sent out to 7th- and 10th-grade students during the period January–April.¹¹ We use survey data for the five years 2007–2011, giving us two years of pre-treatment data and three years of post-treatment data. In addition to data for Oslo, we also collect data for a control group consisting of 10th-grade students in the 10 Norwegian counties that had a neighborhood rule throughout the period studied.¹²

Among the questions is one that is well suited for studying motivation, “*How interested are you in learning at school?*” The response 1 indicates very low interest, while 5 indicates that students are highly interested in their school-work. The question we use is similar to the TIMSS questions about student attitudes towards schooling in German states studied by Jüriges and Schneider (2010) who find that, to quote the title of their paper, “Central exit examinations increase performance... but take the fun out of mathematics”. We interpret this as a measure of the students’ intrinsic motivation.

School-level results for each year are available through the directorate’s webpage. Owing to the aggregation to the school level, our left-hand side variable is continuous rather than an ordinal scale as it would be if we had access to individual responses. In order to secure anonymity, results for schools with fewer than 10 students in the cohort are not made public. This also applies for the separate male and female scores, where the score values are omitted when there are fewer than 10 boys/girls in the cohort. To secure that the treatment and control groups are consistent across specifications, we only use data for schools where data are published for both genders. Consequently, all schools with cohorts of fewer than 20 students are excluded, in addition to schools with a just 20 students whenever there are fewer than 10 of either gender. We are thus left with 48 out of a total of 62 schools in the treatment group and 322 out of a total of 560 in the control group.

¹¹ In the remainder we refer to the year the survey was conducted. For example, when referring to the school year 2006/07, we henceforth write only 2007. The full survey is available at <https://skoleporten.udir.no/>.

¹² The counties in the control group are Østfold, Hedmark, Buskerud, Telemark, Vest-Agder, Sogn og Fjordane, Sør-Trøndelag, Nord-Trøndelag, Nordland and Troms. Two local governments in Sør-Trøndelag (Trondheim and Klæbu) are omitted from the control group owing to a reform in these local governments within the sample period.

Table 1: Population and sample.

	Oslo (treated)		Control group: 10 counties		Control group: primary schools in Oslo	
	Population	Sample	Population	Sample	Population	Sample
Number of municipalities	1	1	232	181	1	1
Number of schools	62	48	560	322	110	96
Mean number of student	83.8	100.3	46.0	74.9	47.6	55.3
Mean number of girls	40.4	48.9	22.3	36.4	22.9	26.7
Mean number of boys	43.4	51.4	23.7	38.4	24.7	28.6

A noteworthy side-effect of the regulations for the publication of results is that they reduce the number of small schools, making schools in Oslo and the other counties more homogenous with regards to school size. This is illustrated in Table 1, which compares the full population of schools to what we have in our sample. The difference in cohort size between the treatment and control group consisting of 10 different counties is reduced by about 33% when comparing the population to the sample. We also observe that there are about twice as many lower secondary than upper secondary schools in Oslo. Since almost all students enroll directly into upper secondary schools, these schools have about double the number of students in each cohort compared to that of lower secondary schools.

In order to ease the interpretation of the results, we have standardized the responses into a z-score with zero mean and a standard deviation of unity over the period studied.¹³ Figure 1 shows the descriptive statistics for the variable of interest in Oslo and the control groups. Figures A–C show descriptive statistics when applying 10 different counties as control group. We observe that whereas the trend is flat throughout the period for the control group there is some variation in Oslo. The general picture is that students in Oslo had higher reported motivation than the control group the whole period, but the difference decreased after the reform. The drop is particularly clear when looking at the girls, while the curve for the boys in Oslo remains almost completely flat throughout the period. Almost all the variation is thus owing to changes in the motivation of girls.

¹³ The z-score is generated by standardizing the residuals when regressing (OLS) the raw score on motivation as the dependent variable and year dummies as independent variables. The motivation score is close to being normally distributed. See Appendix Table 10 for descriptive statistics on the raw score.

Figure 1: Descriptive statistics on motivation.

Weighted means using the number of all students, girls and boys, respectively, as weights. Standardized values.

Figure 1(D–F) shows the descriptive statistics when using 7th-grade students in Oslo as control group. 7th-grade students report consistently a higher motivation than 10th-grade students, on average the difference (see Appendix Table 10) is about 4.9%.¹⁴ We see that both the treatment and control group show a weak tendency to increase prior to the reform. After the reform, the trend continues for the control group, but is broken for the treatment group, which now seems to decline. When we study the genders separately in E and F, we see that this pattern is mostly driven by the girls. Whereas boys have a completely flat development for both groups throughout the period, there is a bit more going on when studying the girls. In the pre-treatment period, girls in both 7th and 10th grades have an almost completely flat development. In the post-reform period, the development for 10th-grade girls declines, while we observe a weak increase in motivation for girls in 7th grade. However, a joint test of a post-reform trend in the control group is rejected with a p -value of 0.34.¹⁵

¹⁴ On the national level the difference is a somewhat larger, about 7.5%.

¹⁵ In this test we apply the number of girls as weights, and the standard errors are clustered at the school level. This is the same procedure as in the regressions below.

4.2 Econometric Specification

We proceed by estimating various versions of

$$y_{it}^g = \beta^g (\text{Reform} \times \text{post} - \text{reform period})_{it} + \text{Controls}'_{it} \kappa^g + \alpha_i^g + \delta_t^g + \varepsilon_{it}^g \quad [1]$$

y_{it}^g is the reported motivation in school i in year t and $g = (\text{all, boys, girls})$. When estimating regressions where the genders are pooled, we weight the regressions by the average total cohort size. When estimating using the genders separately, we weight the regressions by the number of students of the specific gender in the cohort.

The variable $(\text{Reform} \times \text{post} - \text{reform period})$ is the interaction term between treated schools and the post-treatment dummy, or in some specifications, dummies for each year of the years in the post-treatment period. α_i and δ_t capture school and time (i.e., year) fixed effects, respectively. Importantly, these capture both differences between schools that are constant over time, and changes from one year to another that are equal for all schools. Hence, the coefficient β captures the difference in differences between schools that are subject to the reform and those that are not. ε_{it} is an error term.

The vector of control variables include the size of the cohort and a proxy for class size, the share of girls in each cohort, total resources spent on schools in the municipalities during the school year per student and some important socio-economic characteristics of the municipality.¹⁶ We can have that wealthy parents choose to move to areas with good schools, and one might thus expect that the reform will affect students from, e.g., wealthy families the most.¹⁷ Interestingly, a recent Swedish study (Söderström and Uusitalo 2010) concludes that admission reforms similar to the Norwegian increase segregation not only by ability but also by ethnicity. The socio-economic characteristics include the (log) average private income, unemployment rates for the age groups 16–24 years and 25–66 years and the share of single supporters, immigrants, people with high school and short and long university education.¹⁸

Finally, it is well known that pooled regressions may wrongfully estimate the standard errors, and thereby the t -values. We deal with this problem by reporting t -values based on clustered standard errors taking into account that

¹⁶ The proxy for class size is defined as the number of students at levels 8 to 10 multiplied by the average number of hours in class divided by total teaching hours. Budgets are weighted with 5/12 and 7/12 of the budget for the year of the first and second semester, respectively.

¹⁷ Interestingly, Fiva and Kirkebøen (2011) find only a short-run effect on house prices from publication of test scores.

¹⁸ Descriptive statistics for the control variables are reported in Appendix Table 11.

the error terms from the same schools are correlated, see Wooldridge (2003). Note that it is not obvious on what level standard errors should be clustered in DD analyses (Angrist and Pischke 2009). It can be argued that one should cluster on the level where the treatment occurs, i.e., the county level. However, owing to the low number of counties and since a part of the analysis is based only on schools within a single county, we use the school level throughout our analysis. We have run the regressions while clustering on the county level where this is possible and also used bootstrapped standard errors without this changing the qualitative results.

5 Main Results

The results from regressions where 10th-grade students in other counties act as control group are presented in Table 2. In Columns (A)–(C) we study the effect on girls. We start out in Columns (A) and (B) with a simple formulation where all years of post-treatment are pooled together. The difference between Columns (A) and (B) is that the vector of control variables is omitted in (A), while included in (B).

The sign of the treatment effect is negative, as expected when studying girls, but not significantly so in either Column (A) or Column (B). In order to investigate this further, in Column (C) we look at the effect year by year after the implementation of the reform. We observe that the effect is insignificant in the first year, but increases sharply and becomes significant at the 5% level in the following two years. The coefficients for the treatment effect in 2010 and 2011 are sizable, indicating that the intrinsic motivation is roughly 75% of a standard deviation lower two and three years after the reform implementation. The effect thus seems quite strong, but one should note that the standard deviation of the motivation variable is not very large. By using the means and standard deviations reported in Appendix Table 10, we find that a coefficient of -0.75 is equivalent to a 0.16 reduction in the survey scale points, which again translate to about 4% of the average value.

The observation that we do not find any immediate effects is not necessarily a surprise, since the reform was announced shortly before the survey was conducted for the first treated cohort. Hence, the students had very little time to understand the implications of the system change. Moreover, it is likely that the students in the first treated cohort would have had a limited understanding of the full implications of the reform anyway. The reason is that they had very little (if any) grounds for comparison between the two regimes, since they were the first to be affected by it. This also corresponds to the findings by Haraldsvik (2012), who finds significant effects on achievements, but not in the first year post-treatment.

Table 2: Estimations of student motivation.

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)
	Girls	Girls	Girls	Boys	Boys	Boys	All	(B)–(E)	(C)–(F)
Treatment	–0.186 (0.132)	–0.344 (0.230)		0.105 (0.101)	0.265 (0.176)			–0.609** (0.239)	
Treatment*09			–0.291 (0.230)			0.272 (0.183)	–0.0485 (0.212)		–0.564** (0.240)
Treatment*10			–0.734** (0.317)			0.163 (0.241)	–0.464 (0.284)		–0.897*** (0.337)
Treatment*11			–0.750** (0.334)			0.0589 (0.256)	–0.459 (0.298)		–0.809** (0.351)
Control variables	No	Yes	Yes	No	Yes	Yes	Yes		
Observations	1,556	1,556	1,556	1,556	1,556	1,556	1,556		
No. of schools	370	370	370	370	370	370	370		

Note: Robust standard errors (clustered on the school level) in parentheses. School fixed effects, a constant term, and time dummies are included in all regressions. Weighted regressions using the number of girls/boys/total students as weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

When we look at boys in Columns (D)–(F) the picture is very different. The coefficients for the treatment consistently come out as insignificantly positive. In Column (G) we pool boys and girls together, to see if there are any total effects. We observe that while the coefficients are negative every year after the reform, they are not significantly different from zero. The main take-away from Table 2 is that it confirms the impression from the descriptive statistics reported in Figure 1 (A)–(C). There seems to be a clear negative effect on the motivation of girls in Oslo from the reform, but no effect on the motivation of boys.

Finally, in Columns (H) and (I) we test whether the coefficients for girls and boys are significantly different from each other. Interestingly, we have that the coefficients for boys and girls are significantly different both when using the single treatment dummy specification [Column (H)], and when analyzing the treatment effect year by year [Column (I)].¹⁹ Hence, despite the fact that the treatment is not significantly negative for girls per se in the one dummy formulation and in 2009, girls’ motivation is still significantly reduced compared to that of boys.

The results from regressions where 7th-grade students in Oslo act as control group is reported in Table 3. Importantly, the results follow the same main patterns

Table 3: Estimations of student motivation using primary schools as control group.

	(A)	(B)	(C)	(D)	(E)	(F)	(G)
	Girls	Girls	Boys	Boys	All	(A)–(C)	(B)–(D)
Treatment	–0.357** (0.150)		–0.0543 (0.139)			–0.302* (0.175)	
Treatment*09		–0.159 (0.183)		–0.0307 (0.192)	–0.166 (0.191)		–0.128 (0.212)
Treatment*10		–0.364* (0.190)		–0.0310 (0.180)	–0.311* (0.185)		–0.333 (0.229)
Treatment*11		–0.550*** (0.187)		–0.101 (0.174)	–0.362** (0.171)		–0.449* (0.230)
Observations	654	654	654	654	654		
No. schools*level	144	144	144	144	144		

Note: Robust standard errors (clustered on the gender specific school*grade level) in parentheses. School*grade fixed effects, a constant term, the share of girls, class size, number of students and time dummies are included in all regressions. Weighted regressions using the number of girls/boys/total students as weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

¹⁹ The test involves estimating a model including the same variables as in the models with control variables in addition to gender-specific school effects, and interaction terms between a gender dummy and all other variables.

as in Table 2. Whereas girls seem to be negatively affected, the coefficients for boys are far from being significant.²⁰ We also observe that the gender difference is significant when using the one-dummy formulation. When splitting this year by year, however, the difference is only significant in 2011. The main difference from the results in Table 2 is that the effects are actually also significant when pooling boys and girls together, for two out of three post-treatment years. Our findings so far are in line with the earlier studies concluding that females seem to dislike competition more than males. See, e.g., Hogarth, Karelaia, and Trujillo (2011) and Niederle and Vesterlund (2007) or the literature review by Croson and Gneezy (2009).

Students living in central areas are likely to experience tough competition from the reform, since many of the most popular schools are located in this area.²¹ Further, schools located in wealthy areas with a low density of immigrants west in Oslo are probably also attractive for students living in other districts. Hence, it would be interesting to see whether the gender differences vary across districts. In order to investigate this, we have estimated district specific treatment effects of the admission reform. The results are presented in Table 4.

The first column in Table 4 presents the estimated gender differences using the 10 different counties as control group. A negative sign indicates that girls' motivation is reduced relative to that of boys following the reform. We find a negative coefficient in 13 out of 15 districts, and the two positive estimates are far from being significant. The negative coefficients are significant in seven districts, mostly for districts located near the central area or in the western parts of Oslo.

In the model where we use primary schools in Oslo as control group (third and fourth columns in Table 4), we find a similar pattern. Although there is a significant post-reform effect only in four districts, it is reassuring that we also find significant effects in these districts when using a different control group (first two columns).

Next, we estimate the models where the treatment indicator is introduced prior to the reform, known as placebo difference in differences. The reform effects are now estimated relative to 2007. The results are reported in Table 5.

In Columns (A)–(D) we test the results from the analysis using 10th-grade students in other counties as control group. The coefficient for 2008 (the placebo year) comes out as insignificant, and we can thus not reject a common pre-treatment

20 If we do not weight the regressions, the treatment effect for girls in 2010 is no longer statistically significant. However, the one-dummy formulation is still significant. The coefficients (std. errors) when regressing girls' motivation are -0.276 (0.192) and -0.537 (0.185) in 2010 and 2011, respectively. The results without weights are also mostly similar as in the weighted regressions when studying boys and the genders pooled.

21 As an example, in 2013, the four most popular schools (Elvebakken, Foss, Ullern, and Oslo Katedralskole) were all located in the central and western areas of Oslo.

Table 4: Gender differences of treatment in different districts in Oslo.

Location		District	Control group: 10 different counties		Control group: Primary schools in Oslo	
			Coef	SE	Coef	SE
Central	East	Gamle Oslo	−0.628**	(0.315)	−0.246	(0.253)
		Grünerløkka	−1.613***	(0.223)	−1.138***	(0.165)
		Sagene	0.161	(0.745)	0.583	(0.721)
	West	St. Hanshaugen	0.381	(0.432)	0.605	(0.448)
		Frogner	−0.950*	(0.514)	−0.689	(0.467)
West	Ullern		−1.959***	(0.242)	−1.576***	(0.183)
	Vestre Aker		−0.342	(0.227)	−0.0522	(0.154)
	Nordre Aker		−0.835**	(0.418)	−0.579	(0.407)
North-East	Bjerke		−0.955***	(0.330)	−0.622**	(0.247)
	Grorud		−0.893***	(0.340)	−0.566*	(0.297)
	Stovner		−0.553	(0.359)	−0.241	(0.299)
	Alna		−0.643	(0.637)	−0.334	(0.567)
South-East	Østensjø		−0.154	(0.368)	0.103	(0.346)
	Nordstrand		−0.530	(0.420)	−0.231	(0.369)
	Søndre Nordstrand		−0.680	(0.556)	−0.333	(0.517)
	Observations		3,112		1,308	
		No. of schools	740		288	

Note: Robust standard errors (clustered on the gender specific school*grade level) in parentheses. School*grade fixed effects, a constant term, control variables and time dummies are included in all regressions. Weighted regressions using the number of girls/boys as weights. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

trend. In Columns (E)–(H) we test the results from the analysis using 7th-grade students in Oslo as control group. The coefficients for the year 2008 (the placebo year) are always far from being significant at any conventional level of significance. We can thus not reject a common pre-treatment trend in this case either.

6 Robustness Analysis

6.1 Robustness using the Oslo Reform

When we study boys (girls) in our benchmark analyses, the control group also consists only of boys (girls). One might thus worry that the difference in

Table 5: Placebo tests.

Variables	Control group: 10 different counties				Control group: Primary schools in Oslo			
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
	Girls	Boys	All	(A)–(B)	Girls	Boys	All	(E)–(F)
Treatment*08 (Placebo treatment)	0.248 (0.210)	0.0711 (0.195)	0.227 (0.192)	0.176 (0.259)	0.0998 (0.231)	–0.206 (0.226)	–0.0511 (0.216)	0.305 (0.271)
Treatment*09	–0.0880 (0.276)	0.330 (0.256)	0.137 (0.267)	–0.418 (0.297)	–0.107 (0.211)	–0.137 (0.247)	–0.192 (0.228)	0.0295 (0.259)
Treatment*10	–0.513 (0.359)	0.226 (0.297)	–0.263 (0.329)	–0.739* (0.380)	–0.312 (0.212)	–0.137 (0.224)	–0.337 (0.216)	–0.176 (0.259)
Treatment*11	–0.490 (0.414)	0.133 (0.339)	–0.221 (0.377)	–0.623 (0.437)	–0.499** (0.230)	–0.206 (0.232)	–0.388* (0.218)	–0.293 (0.287)
Observations	1,556	1,556	1,556		654	654	654	
No. of schools	370	370	370		144	144	144	

Note: Robust standard errors (clustered on the gender-specific school*grade level) in parentheses. School*grade fixed effects, a constant term, control variables and time dummies are included in all regressions. Weighted regressions using the number of girls/boys/total students as weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the results between the genders might be driven by differences in the control groups. In order to investigate this, we have estimated the effect on boys and girls using the full sample of students in the control group. The results are very similar to those reported, and differences in the control group are thus not driving the difference in the results for boys and girls.²²

Next, the validity of our diff-in-diff analysis relies on a proper control group. In our main analyses, we have included all counties that had a neighborhood system throughout the period studied. The reasoning behind this is to compare admission systems that were almost identical prior to the reforms in Oslo. The Norwegian counties are, however, quite heterogenous with respect to geography, population size, and demography. A potential worry is thus that the results can be driven by such differences.

To test the robustness of our main results, we have excluded one county at a time from the control group. In order to save space, we only report the results for the genders separately. This gives us a total of 20 new regressions for each of the treated counties. The results are reported in Table 6.

22 The reform effects for girls in 2010 and 2011 are –0.50 and –0.63, respectively.

Table 6: Robustness test: excluding one county from the control group at a time.

County excluded	Østfold				Hedmark				Buskerud				Telemark				Vest-Agder			
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Treatment*09	-0.257 (0.242)	0.273 (0.190)	-0.284 (0.237)	0.267 (0.185)	-0.354 (0.244)	0.132 (0.187)	-0.303 (0.233)	0.280 (0.184)	-0.241 (0.237)	0.418** (0.185)										
Treatment*10	-0.696** (0.331)	0.100 (0.257)	-0.663** (0.327)	0.262 (0.242)	-0.894*** (0.335)	-0.025 (0.255)	-0.747** (0.320)	0.151 (0.243)	-0.744** (0.331)	0.350 (0.235)										
Treatment*11	-0.677* (0.350)	0.041 (0.272)	-0.770** (0.349)	0.105 (0.262)	-0.844** (0.356)	-0.173 (0.263)	-0.759** (0.336)	0.076 (0.256)	-0.715** (0.353)	0.307 (0.252)										
Observations	1,375	1,375	1,426	1,426	1,377	1,377	1,418	1,418	1,418	1,418										
No. of schools	332	332	340	340	327	327	335	335	340	340										
County excluded	Sogn og Fjordane				Sør-Trøndelag				Nord-Trøndelag				Nordland				Troms			
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Treatment*09	-0.333 (0.239)	0.276 (0.187)	-0.353 (0.237)	0.263 (0.191)	-0.294 (0.235)	0.247 (0.184)	-0.293 (0.237)	0.268 (0.188)	-0.258 (0.234)	0.267 (0.196)										
Treatment*10	-0.727** (0.328)	0.201 (0.248)	-0.778** (0.327)	0.134 (0.258)	-0.721** (0.322)	0.165 (0.244)	-0.795** (0.331)	0.145 (0.256)	-0.653** (0.327)	0.117 (0.272)										
Treatment*11	-0.792** (0.349)	0.075 (0.264)	-0.858** (0.343)	0.031 (0.269)	-0.751** (0.341)	0.039 (0.259)	-0.763** (0.345)	0.037 (0.269)	-0.673* (0.343)	0.022 (0.291)										
Observations	1,453	1,453	1,454	1,454	1,461	1,461	1,394	1,394	1,461	1,461										
No. of schools	341	341	347	347	347	347	323	323	346	346										

Note: Robust standard errors (clustered on the school level) in parentheses. School fixed effects, a constant term, control variables, and time dummies are included in all regressions. Weighted regressions using the number of girls/boys as weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We observe that the results are quite robust, and it thus seems that our findings are not driven by any particular counties.

Another issue is related to the survey data from the Norwegian Directorate for Education and Training. When working with survey data, it is always a worry that the subjects' responses do not necessarily measure exactly what they are meant to do. This could be the case if students who are dissatisfied (satisfied) simply give a poor (good) score on all questions without really considering them separately, or if the students misunderstand the questions. This will create additional noise in our data, but might, in worst case, also lead to estimation of spurious relationships.

In order to investigate this more closely, one might test if the model predicts a treatment effect on variables which the reform should not affect. Specifically, we have estimated the model using questions regarding the satisfaction with the physical work environment as dependent variable instead of student motivation. There are 10 such questions. Since the physical work environment should not be affected by the reform, a significantly negative result for these could give reason for caution, since it might indicate that our estimates for motivation capture a broader dissatisfaction. Again, we only report the results from regressions for the genders separately.

The results are reported in Table 7. Importantly, the coefficients are mostly far from being significant. Despite a few exceptions, the general impression is that the reported satisfaction with the physical work environment is, as expected, not related to the reform. These results increase our confidence in the survey's ability to measure what it intends to measure, since we do not find that the reform affects variables that it should not affect. Further, the signs are varying, some being negative and some being positive. Positive signs indicate that the negative link between the reform and motivation is not driven by a "general dissatisfaction". Moreover, it is hard to identify any clear gender differences in the tables. Hence, our conclusions about gender differences from the discussion of our main results should not be undermined.²³

23 Our procedure is similar to that suggested by Carlsen and Johansen (2004) who include a survey question that should not be affected by the explanatory variable of interest as a control variable capturing "general satisfaction". We have also tried this approach, using the average satisfaction across the different facility measures as a proxy for general satisfaction. The results show that satisfaction with facilities is strongly correlated with motivation, even when including school fixed effects, but more importantly that the coefficient for the treatment effect remains unchanged. This is consistent with the findings in the reported test, indicating that there is no relationship between the reform and satisfaction with facilities.

Table 7: Estimations of satisfaction with physical work environment.

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)
	Indoor air	Indoor temperature	Class- rooms	Teaching equipment	School library	Rest- rooms	Wardrobes & showers	School buildings	Cleaning	Outdoor area
Panel A: Girls										
Treatment*09	0.105 (0.186)	0.0355 (0.192)	0.155 (0.159)	0.390** (0.179)	0.184 (0.139)	0.109 (0.147)	0.203 (0.148)	0.110 (0.125)	0.294* (0.171)	-0.013 (0.162)
Treatment*10	-0.018 (0.223)	-0.102 (0.243)	0.281 (0.212)	-0.062 (0.220)	0.151 (0.193)	0.215 (0.213)	0.338 (0.234)	0.121 (0.178)	0.319 (0.227)	0.084 (0.213)
Treatment*11	-0.092 (0.259)	-0.219 (0.274)	0.069 (0.233)	-0.360 (0.239)	0.065 (0.215)	0.339 (0.271)	0.321 (0.258)	-0.065 (0.206)	0.299 (0.253)	0.046 (0.243)
Observations	1,532	1,533	1,537	1,537	1,537	1,531	1,524	1,536	1,537	1,538
No. of schools	368	367	368	368	368	368	366	368	368	368
Panel B: Boys										
Treatment*09	-0.103 (0.189)	-0.105 (0.196)	0.057 (0.175)	0.338 (0.211)	0.269* (0.152)	0.043 (0.146)	0.066 (0.168)	0.105 (0.146)	0.177 (0.210)	-0.205 (0.197)
Treatment*10	-0.105 (0.223)	-0.234 (0.229)	0.215 (0.185)	-0.041 (0.229)	0.236 (0.195)	0.099 (0.204)	0.154 (0.238)	0.173 (0.192)	0.211 (0.252)	-0.032 (0.244)
Treatment*11	-0.121 (0.257)	-0.291 (0.257)	0.0278 (0.222)	-0.257 (0.250)	0.167 (0.211)	0.269 (0.248)	0.173 (0.267)	0.056 (0.214)	0.192 (0.265)	-0.200 (0.279)
Observations	1,534	1,538	1,537	1,538	1,536	1,531	1,535	1,538	1,538	1,538
No. of schools	368	368	368	368	368	367	368	368	368	368

Note: Robust standard errors (clustered on the school level) in parentheses. School fixed effects, a constant term, control variables, and time dummies (not reported) included. Weighted regressions using the number of girls/boys as weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6.2 A Similar Reform in a Different County

As a final test of robustness, we turn the attention to the county Hordaland. In Hordaland, the county announced in the fall of 2004 that the existing neighborhood school system would be replaced by performance-based ranking with the new system being effective from the following school year (2005/2006), i.e., Hordaland implemented a reform very similar to the one in Oslo. Students graduating from lower secondary education in Hordaland in 2005 were the first to be affected by this reform. Since Hordaland has a population pattern that is more similar to the control group counties than Oslo, this works as a test as to whether the results are driven by such differences. The main drawback with the reform in Hordaland is the timing. In order to analyze the reform in Hordaland, we must combine the current survey (2007–2011), with its predecessor where data are available for 2004 and 2005.²⁴ The new survey differs from the old in that several new questions were added and some were removed, but 10th-grade students were asked about their motivation using an identical formulation in both.²⁵ Moreover, the scale also changed from 1–4 to 1–5. A 1–4 scale does not allow students to give a “neutral” answer, since 1 and 2 would suggest low motivation, while 3 and 4 indicate high motivation. If students that reported either low or high motivation are more likely to report a “neutral” answer in the new survey (i.e., 3), this might introduce a bias. However, for the bias to be systematic, the treatment and control groups must be affected differently by the change in scale. Moreover, it is not obvious in which direction this bias might go.

Since the Hordaland reform took place in 2005, the survey in 2004 provides the only data available for the pre-reform period. Hence, we cannot conduct a placebo test of the kind we do for Oslo in Table 5. Another drawback is that we have relatively few observations in the first year of the survey (i.e., 2004).²⁶ Consequently, the results for Oslo should thus be considered as the most reliable, while the study of the reform in Hordaland acts as a valuable supplement and test of robustness.

²⁴ The data for 2006 are not available from the directorate’s database. The directorate has informed us that the data for this year “was lost” when updating a server, owing to a technical and/or human error.

²⁵ Note that 7th-grade students were not asked the same question in the predecessor, making it impossible to use 7th-grade students in Hordaland as an alternative control as we did when analyzing the Oslo reform.

²⁶ Since the low number of respondents in this year can give rise to a worry about sample selection, we have also estimated the model using only schools that participated in 2004. The results are reported in Appendix Table 12.

As when studying Oslo, we have standardized the responses into a z-score with zero mean and a standard deviation of unity over the period studied.²⁷ Since there was a change in the survey's scale (from 1–4 to 1–5) in 2007, this is done by standardizing the residuals when regressing (OLS) the raw score on motivation as the dependent variable and year dummies as independent variables for the old survey (2004–2005), and the new survey (2007–2011), respectively.

Table 8: Population and sample.

	Hordaland (treated)		Control group: 10 counties	
	Population	Sample	Population	Sample
Number of municipalities	33	26	232	183
Number of schools	102	73	580	336
Mean number of students	63.9	88.9	45.3	74.8
Mean number of girls	30.9	43.2	22.0	36.4
Mean number of boys	33.0	45.6	23.3	38.4

Table 8 presents the population of schools and the sample remaining, following the same rules for publication as described in Section 4. The sample of schools is reduced to 73 from a total of 102 in the treatment group and 336 out of 580 in the control group when studying the Hordaland reform. As when studying Oslo, the exclusion of small schools gives that the two samples are more equal. The difference in cohort size is reduced by 24% when comparing the population to the sample.

Figure 2 shows descriptive statistics on motivation for Hordaland and the control group. Interestingly, all three graphs show that Hordaland students had higher motivation than the control group prior to the reform, but that the control group caught up with Hordaland after the reform. On first sight, the gender differences are thus less apparent in the Hordaland data than in the Oslo data (see Figure 1), even though the change in motivation after the reform does seem to be larger for girls than boys. Further, whereas there seems to be a drop in girls' motivation shortly after the reform, the trend is flat for boys in several years before dropping slightly toward the end of the sample. Hence, we also observe indications of gender differences in the descriptive statistics for the Hordaland reform.

²⁷ The motivation score is close to being normally distributed in both of our cases. See Appendix Table 10 for descriptive statistics on the raw score.

Figure 2: Descriptive statistics on motivation in Hordaland. Weighted means using the number of all students, girls and boys, respectively, as weights. Standardized values.

The results from the analysis using the Hordaland reform are reported in Table 9. The year prior to the announcement of the reform (2004) acts as the pre-treatment comparison. When looking at the girls in Columns (A)–(C), the first thing we note is that the effect seems to be stronger, both in terms of coefficient size and statistical significance than for girls in Oslo. In particular, the coefficients for the treatment effect are also significantly negative when pooling all post-treatment together in Columns (A) and (B).

From Column (C) we observe that, as in Oslo, the treatment effect comes out as insignificantly negative in the first year. The effect three years after the reform (2007) is very similar to the effect we find 2–3 years after the reform in Oslo (2010 and 2011), about 56% of a standard deviation. Interestingly, the effect is a bit weaker, and falls short of significance in 2008, before the coefficients increase in absolute value in the remaining years of the sample. Although the coefficient in 2008 is non-significant, it is not significantly different from the other estimated treatment effects.

In the three last years, the treatment effect is on average around 65% of a standard deviation and is significant even at the 1% level in all three years. This is equivalent to about a 0.15 reduction in the survey scale points after the change of the scaling in 2007 (see Appendix Table 10).²⁸ This translates into

²⁸ Again, the results are not sensitive to the use of weights. If we do not use weights when regressing girls' motivation, the coefficients (std. errors) are -0.636 (0.251), -0.528 (0.228), -0.825 (0.246), -0.730 (0.218), and -0.802 (0.220) in 2007 to 2011, respectively. The results without weights are also similar as in the weighted regressions when studying boys and the genders pooled.

Table 9: Estimations of student motivation using data from Hordaland.

Variables	(A)		(B)		(C)		(D)		(E)		(F)		(G)		(H)		(I)	
	Girls	No	Girls	Yes	Girls	Yes	Boys	No	Boys	Yes	Boys	Yes	All		(B)–(E)		(C)–(F)	
Treatment	–0.516*** (0.191)		–0.417** (0.205)				–0.112 (0.205)	No	–0.0547 (0.218)	Yes	0.107 (0.266)	Yes	0.112 (0.275)		–0.363 (0.260)		–0.159 (0.333)	
Treatment*05					–0.0523 (0.273)						–0.0160 (0.277)		–0.305 (0.275)				–0.542* (0.307)	
Treatment*07					–0.558** (0.237)						–0.101 (0.236)		–0.259 (0.241)				–0.263 (0.289)	
Treatment*08					–0.364 (0.229)						–0.0247 (0.242)		–0.371 (0.259)				–0.645** (0.281)	
Treatment*09					–0.670*** (0.233)						–0.309 (0.243)		–0.642*** (0.240)				–0.311 (0.281)	
Treatment*10					–0.620*** (0.204)						–0.338 (0.252)		–0.631** (0.261)				–0.388 (0.274)	
Treatment*11					–0.725*** (0.223)													
Control variables	No		Yes		Yes		No		Yes		Yes		Yes					
Observations	2,017		2,017		2,017		2,017		2,017		2,017		2,017					
No. of schools	409		409		409		409		409		409		409					

Note: Robust standard errors (clustered on the school level) in parentheses. School fixed effects, a constant term, and time dummies are included in all regressions. Weighted regressions using the number of girls/boys/total students as weights.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4% of the average value in this period, i.e., exactly the same effect as we found when studying the reform in Oslo in Table 2.

As in the study of the Oslo reform, the picture is very different when studying the boys in Columns (D)–(F). Even though the coefficients for the treatment effect mostly come out as negative, as for girls, they are never close to being significant at any conventional level of significance. This is also reflected in Column (G) where the genders are pooled, where only the two last years come out as significant.

Finally, in Columns (H) and (I) we test whether the coefficients for girls and boys are statistically different. Despite the clear significance for girls, and the insignificant estimates for boys, the test for different effects between the genders provides slightly less clear-cut conclusions than when studying the Oslo reform. In the one-dummy formulation [Column (H)] the t-test for the difference is about 1.4 in absolute value. However, the difference is significant in the years 2007 and 2009 [Column (I)]. The reason for the less clear-cut results is that there are some weak indications in Table 9 that boys in Hordaland also experience a reduction in motivation following the reform, as we also noticed when studying Figure 2.

Still, the main take-away from Table 9 is that it mirrors the observations from Figure 2 since the motivation of girls is reduced following the reform. Further, while the sign of the treatment effect for boys is negative, reflecting the small negative trend in Figure 2, there is no significant treatment effect. The results for Hordaland thus support the findings from the study of the Oslo reform in Tables 2 and 3.

7 Concluding Remarks

This paper studies gender differences in competitiveness using data from upper secondary admission reforms in Norway and survey data on student motivation. In 2005, Oslo changed its admission system in upper secondary education from a neighborhood admission system to a system based on performance in lower secondary education (school choice). This introduces high stake competition for students in lower secondary education.

We analyze the reform using a DD framework, using two different groups of untreated students as control group, and the results indicate interesting gender differences. Whereas boys' motivation is unaffected by the reform, there is a sizable negative effect on girls' motivation. Moreover, the differences in coefficients between girls and boys are in most cases statistically significant. Our

findings are thus in line with findings in the experimental literature indicating that females are less in favor of competitive settings than males.

Among the tests of robustness, we also study a very similar reform that took place in the county Hordaland a few years earlier and find very similar results. We find it reassuring that we observe very similar outcomes of two similar reforms in two different counties at two different points in time, since it is unlikely that the same spurious trend would occur in both cases.

Acknowledgments: Valuable comments from two anonymous referees, Fredrik Carlsen, Torberg Falch, Marianne Haraldsvik, Bjarne Strøm, and participants at the 2013 EALE Meeting (Turin), the 2013 EPCS Meeting (Zurich), 2013 Annual Meeting of the Norwegian Economic Association (Stavanger), FIBE 2013 (Bergen), workshops in Lisbon and Catanzaro (IWAEE), and seminar participants in Trondheim are gratefully acknowledged. All remaining errors are our own.

Appendix

Table 10: Descriptive statistics, motivation.

Sample	2004–2005 (old survey)			2007–2011 (new survey)			
	Control			Control			Control
	Treatment	10 counties	Full	Treatment	10 counties	Full	7th grade
Panel A: Oslo reform							
Girls	–	–	–	4.17 (0.19)	4.10 (0.22)	4.12 (0.21)	4.39 (0.24)
Boys	–	–	–	4.07 (0.21)	3.90 (0.24)	3.93 (0.25)	4.27 (0.25)
All	–	–	–	4.12 (0.17)	4.00 (0.19)	4.02 (0.19)	4.32 (0.19)
Panel B: Hordaland reform							
Girls	3.45 (0.17)	3.34 (0.20)	3.36 (0.20)	4.09 (0.23)	4.10 (0.24)	4.10 (0.24)	–
Boys	3.23 (0.18)	3.16 (0.23)	3.17 (0.22)	3.92 (0.23)	3.90 (0.24)	3.90 (0.24)	–
All	3.34 (0.13)	3.25 (0.17)	3.26 (0.17)	4.00 (0.17)	4.00 (0.19)	4.00 (0.18)	–

Note: Mean values and standard deviations (in parentheses) are weighted with the number of girls/boys/total students. Data for 2006 are missing in the directorate's database.

Table 11: Descriptive statistics, control variables.

Variables	Oslo study			Hordaland study	
	Oslo	Control group		Hordaland	Control group
		10 counties	7th grade		
Share of girls	0.49 (0.05)	0.49 (0.06)	0.48 (0.07)	0.49 (0.06)	0.49 (0.06)
Class size	17.79 (2.78)	15.43 (2.36)	15.91 (2.84)	16.44 (2.90)	15.59 (13.04)
Municipal spending per student (1,000 NOK)	47.20 (1.93)	50.42 (5.53)	–	48.50 (6.54)	49.54 (5.80)
Log (mean gross income)	12.85 (0.06)	12.63 (0.10)	–	12.67 (0.13)	12.57 (0.13)
Unemployment rate, 16–24, percent	1.60 (0.37)	2.05 (0.84)	–	2.10 (0.96)	2.42 (1.06)
Unemployment rate, 25–66, percent	2.34 (0.49)	1.73 (0.57)	–	1.97 (0.73)	2.00 (0.78)
Single supporters, percent	1.78 (0.21)	2.29 (0.53)	–	2.33 (0.60)	2.50 (0.71)
Share of immigrants, 0–16, percent	30.81 (0.73)	7.73 (5.01)	–	6.11 (2.84)	7.12 (4.90)
Education, share with high school, percent	34.23 (0.67)	44.82 (3.24)	–	44.50 (4.39)	45.03 (3.25)
Education, share with BSc, percent	28.48 (0.40)	18.17 (3.47)	–	20.29 (4.07)	17.59 (3.56)
Education, share with MSc or PhD, percent	14.59 (0.86)	4.16 (1.95)	–	6.18 (3.29)	3.95 (1.90)

Note: Mean values and standard deviations (in parentheses) are weighted with the number of students.

Table 12: Estimations of student motivation using data from Hordaland.

Variables	(A)		(B)		(C)		(D)		(E)		(F)		(G)		(H)		(I)	
	Girls		Girls		Girls		Boys		Boys		Boys		All		(B)–(E)		(C)–(F)	
Treatment	–0.517*** (0.192)		–0.520** (0.239)		–0.0957 (0.417)		–0.113 (0.206)		–0.0820 (0.246)		0.0267 (0.367)		–0.0222 (0.418)		–0.438* (0.264)		–0.122 (0.459)	
Treatment*05																		
Treatment*07					–0.486 (0.361)						–0.00665 (0.409)		–0.195 (0.422)				–0.479 (0.430)	
Treatment*08					–0.655** (0.306)						–0.131 (0.270)		–0.447* (0.268)				–0.524 (0.392)	
Treatment*09					–0.600** (0.302)						–0.0197 (0.281)		–0.305 (0.322)				–0.581* (0.331)	
Treatment*10					–0.842*** (0.264)						–0.268 (0.255)		–0.572** (0.272)				–0.574* (0.335)	
Treatment*11					–0.957*** (0.278)						–0.328 (0.307)		–0.637* (0.336)				–0.629** (0.267)	
Control variables	No		Yes		Yes		No		Yes		Yes		Yes					
Observations	871		871		871		871		871		871		871					
No. of schools	146		146		146		146		146		146		146					

Note: Includes only schools participating in the 2004-survey. Robust standard errors (clustered on the school level) in parentheses. School fixed effects, a constant term, and time dummies are included in all regressions. Weighted regressions using the number of girls/boys/total students as weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

References

- Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Barber, B. M., and T. Odean. 2001. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics* 116:261–92.
- Byrnes, J. P., D. C. Miller, and W. D. Schafer. 1999. "Gender Differences in Risk Taking: A Meta-Analysis." *Psychological Bulletin* 125:367–83.
- Carlsen, F., and K. Johansen. 2004. "Subjective Measures of Employment Opportunities and Interregional Migration." *Labour* 18:563–89.
- Correll, S. J. 2001. "Gender and the Career Choice Process: The Role of Biased Self-Assessments." *American Journal of Sociology* 106:1691–730.
- Crosen, R., and U. Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47:448–74.
- Figlio, D., and C. M. D. Hart. 2014. "Competitive Effects of Means-Tested School Vouchers." *American Economic Journal: Applied Economics* 6:133–56.
- Filippin, A., and P. Crosetto. 2014. "A Reconsideration of Gender Differences in Risk Attitudes." IZA Discussion Papers 8184, Institute for the Study of Labor (IZA).
- Fiva, J. H., and L. J. Kirkebøen. 2011. "Information Shocks and the Dynamics of the Housing Market." *Scandinavian Journal of Economics* 113:525–52.
- Friedman, M. 1955. "The Role of Government in Education." In *Economics and the Public Interest*, edited by R. A. Solow, 123–44. New Brunswick, NJ: Rutgers University Press.
- Gneezy, U., M. Niederle, and A. Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics* 118:1049–74.
- Gneezy, U., and A. Rustichini. 2004. "Gender and Competition at a Young Age." *American Economic Review (Papers and Proceedings)* 94:377–81.
- Haraldsvik, M. 2012. "Does Performance-Based School Choice Affect Student Achievement?" Chapter 3 in *Influences on educational outcomes: three essays on the role of parents, peers and choice*. PhD diss. 2012:346, Norwegian University of Science and Technology.
- Hogarth, R. M., N. Karelaia, and C. A. Trujillo. 2011. "When Should I Quit? Gender Differences in Exiting Competitions." *Journal of Economic Behavior & Organization* 83:136–50.
- Jurajda, Š., and D. Münich. 2011. "Gender Gap in Admission Performance under Competitive Pressure: Admissions to Czech Universities." *American Economic Review (Papers and Proceedings)* 101:514–18.
- Jürges, H., and K. Schneider. 2010. "Central Exit Examinations Increase Performance... but Take the Fun Out Of Mathematics." *Journal of Population Economics* 23:497–517.
- Kleinjans, K. J. 2009. "Do Gender Differences in Preferences for Competition Matter for Occupational Expectations?" *Journal of Economic Psychology* 30:701–10.
- Lavy, V. 2013. "Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance-Based Pay Tournaments among Teachers." *The Economic Journal* 123:540–73.
- Niederle, M., and L. Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122:1067–101.
- Ors, E., F. Palomino, and E. Peyrache. 2013. "Performance Gender-Gap: Does Competition Matter?" *Journal of Labor Economics* 31:443–99.

- Sutter, M., and D. Glätzle-Rützler. 2014. "Gender Differences in the Willingness to Compete Emerge Early in Life and Persist." *Management Science*. <http://dx.doi.org/10.1287/mnsc.2014.1981>, forthcoming.
- Söderström, M., and R. Uusitalo. 2010. "School Choice and Segregation: Evidence From an Admission Reform." *Scandinavian Journal of Economics* 122:55–76.
- Wooldridge, J. M. 2003. "Cluster-Sample Methods in Applied Econometrics." *American Economic Review (Papers and Proceedings)* 93:133–8.