Online Appendix for "Pulled-in and Crowded-out: Heterogeneous Outcomes of Merit-based School Choice"

Antonio Dalla-Zuanna Kai Liu Kjell G. Salvanes

A The Norwegian education system

The Norwegian education system consists of four levels, primary school (grades 1–7), middle school or lower secondary school (grades 8–10), high school or upper secondary school (three years), and then higher education. Norwegian compulsory education starts at age six, lasts for 10 years and consists of primary school and lower secondary school. Norwegian municipalities operate schools to provide compulsory education, and the vast majority (98%) of pupils attend public, local schools during compulsory schooling. At the elementary school level, all pupils are allocated to schools based on fixed school catchment areas within municipalities. With the exception of some religious schools and schools using specialized pedagogic principles, parents are not able to choose the school to which their children are sent (except by moving to a different neighborhood). There is a direct link between elementary school attendance and attendance at middle or lower secondary schools (ages 13–16/grades 8–10), in that elementary schools feed directly into lower secondary schools. In many cases, primary and lower secondary schools are also integrated.

The high schools have two main tracks, vocational and academic. High schools are administered at the county level (above the level of municipalities) and attendance is not mandatory, although since the early 1990s everybody graduating from middle schools has been guaranteed a slot in high school. Admissions procedures differ across counties for upper secondary schools. In some counties, pupils can freely choose schools, while in others children are allocated to schools based on well-defined catchment areas, or high school zones. Within schools, there is no systematic sorting of students into classes.

About 95% of students moving into high school enroll in the year they finish compulsory education. About 45% enroll in the academic track, which qualifies for higher education. The rest of the students enroll in the vocational track, and there are several subject fields for this track. There is an option also for students coming from the vocational track to enroll in university, but that requires some extra coursework. Admission at different universities and in different majors at universities is based on high school GPA. This is a combination of non-blind grading by local teachers and the results of the final-year exams, which are prepared centrally by the Directorate for Education (a branch of the Ministry of Education) and are subject to blind grading. The high school GPA is not normalized at the school level.

A.1 Teacher grading and exam grading at middle school (Middle school GPA)

At the end of middle school, students are evaluated both non-anonymously by their teachers for 11 subjects taught in school, and in addition anonymously in 2 nationally administered exit exams, which are graded by external examiners (who are not students' teachers). The subjects for the national exit exams are randomly selected for each student among the 11 subjects in which they are also evaluated by their teachers. The assessment for Norwegian and English consists of both oral and written exams. For the rest of the subjects, the assessment consists of only written exams. In each assessment, the grade ranges between 1 and 6. The final grade of a subject is determined by a simple average between the grades by the teacher and grades from the national exam in the final year, if present. The middle school GPA is not standardized at the school level so they are not grading to the curve.

The final-year middle school GPA used by post-reform high school admission is the sum of final grades across 11 subjects, for students who had grades in at least 3 subjects, hence it ranges between 3 and 66 (i.e. 6×11). Note that the algorithm for calculating the final-year GPA changed in the school year 2006/2007 (hence the last two years we observe). Instead of summing up the grades in 11 subjects, the final-year GPA is determined by first taking a simple average across all subjects and then multiplying by 10. This means that the maximum GPA is 60 for the last cohort under study. For this reason, when we use middle school GPA throughout the paper we generally refer to within-cohort GPA percentiles, assuming that the ranking of students in percentile bins is invariant to the change in the grading.

Grading principles are set by the Education Act of 1998 ("Opplæringslova"). In the Prescript to the Education Act of 1998 (Forskrift til opplæringslova) it is stated that teacher evaluations are to be based on the degree to which students have achieved the competence goals stated by the subject-specific centrally set "Learning goals," which are stated in each topic. For each subject, the final teacher evaluation grade is given in April and is set based on the performance in the final year of middle school. Notably, it is specifically stated that student behavior ("orden og oppførsel") is not to be reflected in grading, and (of course) that student background should not count in grading ("Prescript to Education Act"). Effort is allowed to be included in grading in gymnastics. Teacher grades are given *before* the grading of national exams, and hence teachers are not aware of the student's national exam score at the time when teacher assessment is given.

B The 2005 High School Admission Reform in Hordaland County: Details

The reform we analyze was passed by the government of Hordaland county in the autumn of 2004 and changed the enrollment system for students applying in the spring of 2005 and starting high school in August 2005. Students are assigned to high schools by the county's school administrative office, which, before the 2005–2006 academic year, practically used to enrol students in the schools which had spare capacity and were closer to their homes, to reduce travel time. Hence, although no fixed catchment areas were defined, in practice distance to each school determined the high school a person attended. It was still possible for very high-ability students to request to attend one specific school, but this caused only a small number of high-ability students who lived out of the city center to attend high-reputation schools in the city center. As these are students with high grades in middle school, we are comfortable considering them as always-takers (see Section III.B), hence not affecting our identification strategy.

The reform to the high school admission system was approved by the Hordaland county administration as a response to pressure from different interest groups. It established that starting from the following academic year (2005-2006) students were allowed to apply to different schools with no geographical restrictions. They could rank up to six schools, and assignments would have been based on preferences and, if a school was oversubscribed, the middle school GPA. Only the county's school administrative office, and not the schools, is then involved in the assignment procedure.

As students are assigned to school cohorts on the basis of year of birth (hence, students born in December are assigned to one cohort, while students born in January are assigned to the next cohort), the reform affected students born in 1989 onward. The reform changed the allocation of students to high schools in the whole county, but we focus on the municipality of Bergen and its neighboring municipalities (Os, Øygarden and Askøy) because students who live further away would have a very long commute to reach a high school that is not the one in their municipality. In addition, we focus only on the academic track. The reform only affected academic-track students, as vocational tracks are specific to one subject and often there is only one school offering that specific subject within the county. Thus, students who were willing to attend one specific vocational course were generally allowed to enroll in the only school offering that course both before and after the reform. In the period we consider, there were 16 academic high schools in the Bergen area.

Parents and pupils are well aware of the quality of each high school in our data period because the school rankings were provided by a publicly available website and extensively reported in the newspapers. Public information about school performance across high schools in Norway (i.e., league tables) became available in 2001.

C Proof of Proposition 1

In this section, we show how we can identify the effect of attending a competitive high school for crowded-out compliers (CC). The treatment effect for the pulled-in compliers (CP) can be identified following the same steps, focusing on students above the admission cutoff. Consider now the set of students below the admission cutoff ($C = 0$). The share of CC (π_{CC}), always-takers (π_{AT}) and never-takers (π_{NT}) can be identified from the conditional expectation of school choice (D) given $Post$:[55]

$$P(D = 1|Post = 0, C = 0) = \pi_{CC} + \pi_{AT}$$

$$P(D = 0|Post = 0, C = 0) = \pi_{NT}$$

$$P(D = 1|Post = 1, C = 0) = \pi_{AT}$$

Combining the first and the last data moments, we can identify the proportion of CC: $\pi_{CC} = P(D = 1|Post = 0, C = 0) - P(D = 1|Post = 1, C = 0)$.

The second step of the identification argument is to combine the population shares by compliance types with data on outcomes for subpopulations defined by the realized $Post$ and D. For instance, the subpopulation with $Post = 0, D = 1$ is a mixture of CC compliers and AT. Given Assumptions 1–4 in our framework, we can write the moments that characterize the mean outcome conditional on $Post$ and D as follows:

[55] Note that these compliance shares are conditional on $C = 0$ and not population shares; $\pi_{CC} + \pi_{AT} + \pi_{NT} = 1$.

(1) $E[Y|D=1, Post=0, C=0] = \frac{\pi_{CC}}{\pi_{AT}+\pi_{CC}} E[Y(1)|CC] + \frac{\pi_{AT}}{\pi_{AT}+\pi_{CC}} E[Y(1)|AT]$

(2) $E[Y|D=0, Post=0, C=0] = E[Y(0)|NT]$

(3) $E[Y|D=1, Post=1, C=0] = E[Y(1)|AT]$

(4) $E[Y|D=0, Post=1, C=0] = \frac{\pi_{NT}}{\pi_{NT}+\pi_{CC}} E[Y(0)|NT] + \frac{\pi_{CC}}{\pi_{NT}+\pi_{CC}} E[Y(0)|CC]$

Inserting (3) in (1) and re-arranging allows us to identify

$$E[Y(1)|CC] = \frac{\pi_{AT}+\pi_{CC}}{\pi_{CC}} E[Y|D=1, Post=0, C=0] - \frac{\pi_{AT}}{\pi_{CC}} E[Y|D=1, Post=1, C=0].$$

Similarly, inserting (2) in (4), we obtain

$$E[Y(0)|CC] = \frac{\pi_{NT}+\pi_{CC}}{\pi_{CC}} E[Y|D=0, Post=1, C=0] - \frac{\pi_{NT}}{\pi_{CC}} E[Y|D=0, Post=0, C=0].$$

Note that the above derivation relies on Assumption 2 ("no peer effects"). In particular, we require that the potential outcomes for AT and NT are independent of choices made by other students. To see the significance of this assumption, consider a counterexample where the potential outcome for AT is affected by choices made by all other students, $\tilde{D}(\boldsymbol{z})$. Then, (3) identifies $E[Y(1)|AT, \tilde{D}(\mathbf{1})]$, whereas (1) identifies the weighted average of $E[Y(1)|CC]$ and $E[Y(1)|AT, \tilde{D}(\mathbf{0})]$. In general, $\tilde{D}(\mathbf{0}) \neq \tilde{D}(\mathbf{1})$ because the reform changes the allocation of many students simultaneously, and hence $E[Y(1)|AT, \tilde{D}(\mathbf{1})] \neq E[Y(1)|AT, \tilde{D}(\mathbf{0})]$. In this case, the model is under-identified.

Finally, we can subtract the equations we derived for $E[Y^1|CC]$ and $E[Y^0|CC]$ to obtain the effect of attending a competitive school for CC in terms of observed data moments. Note that, because the compliance types are mutually exclusive, $(\pi_{NT} + \pi_{CC})$ is the probability that students do not attend a competitive school when they are below the cutoff post-reform, i.e. $P(D=0|Post=1, C=0) = \pi_{NT} + \pi_{CC}$, while π_{AT} is the probability that a student attends a competitive school below the cutoff post-reform i.e. $P(D=1|Post=1, C=0) = \pi_{AT}$. Hence,

$$(\pi_{NT}+\pi_{CC})E[Y|D=0, Post=1, C=0] + \pi_{AT}E[Y|D=1, Post=1, C=0] = E[Y|Post=1, C=0].$$

Similarly, it is easy to show that

$$(\pi_{AT}+\pi_{CC})E[Y|D=1, Post=0, C=0] + \pi_{NT}E[Y|D=0, Post=0, C=0] = E[Y|Post=0, C=0],$$

Using these results, the effect of attending a competitive school for CC is

$$E[Y(1)-Y(0)|CC] = \frac{E[Y|Post=1,C=0]-E[Y|Post=0,C=0]}{P(D=1|Post=1,C=0)-P(D=1|Post=0,C=0)}.$$

D Competitive Schools' Cutoffs

In the data, we do not have information on the cutoff used by oversubscribed schools to admit students in the post-reform period. However, we observe the middle school GPA for every student and we know which school they end up attending. We can identify the cutoff for any school (if it exists) following the threshold estimation literature, as in Hansen (2000). The same approach has been applied to school systems in Hoekstra (2009), Landaud, Ly, and Maurin (2020) and Bütikofer, Ginja, Løken, and Landaud (2023).

For every school and cohort, we exclude students who attend a middle school that is more than 30 kilometers away, as they are not likely to attend the school and won't likely contribute to determining the admission cutoff. We experimented with different distances, and we found that 10 or 8km changes the estimated cutoff only for one school (out of 16) every year compared to what we find using 30km. We then consider each school and each year separately.

For every value G_n of the middle school GPA distribution of the pool of potential applicants ($n \in [1, N]$ where N is the total number of values that GPA takes among potential applicants), we define a dummy $g_{n,i}$ that takes value 1 if student i scores above that specific value (i.e. $g_{n,i} = \mathbf{1}[GPA_i \geq G_n] \ \forall \, n \in [1, N]$). Next, we run one bivariate regression for each of these N values, where the dependent variable is a dummy for being admitted at the school ($s_i = 1$ if student i is admitted at the school we consider) and the independent variable is the dummy $g_{n,i}$ defined above:

$$s_i = \alpha + \beta g_{n,i} + \varepsilon_i$$

We select as the admission cutoff for a school in a specific year the value G_n of the GPA distribution for which the regression of the associated dummy has the highest R^2 among all the school-year-specific regressions, under the restriction that it estimates a significantly positive coefficient β. If the coefficient is negative or not significant, then no cutoff is assigned to that school. Using this procedure, we estimate a cutoff for five out of 16 high schools in 2005 and 2006, and four out of 16 high schools in 2007. The estimated cutoffs range between the 10th and the 75th percentiles of the middle school GPA in 2005, between the 11th and the 75th percentiles in 2006, and between the 10th and the 55th percentiles in 2007. In Figure A4 we show that the enrolment in competitive high schools exhibits a clear jump around the estimated admission cutoffs every year (the figure collapses together school admission around all the estimated cutoffs in the same year). Taken together, admission probability increases by 15 percentage points around the estimated cutoffs.

We then need to assign a cutoff for admission at selective schools to each student affected by the reform. We define it as the lowest cutoff that grants access to one of the competitive schools within eight kilometers from the middle school attended, in the year the student turned 16. As mentioned, 80% of the students attended a high school within this distance. Only a few students (about 2% of the sample) attended a competitive school beyond eight kilometers in the post-reform period, although they do have a competitive school in their eight kilometers neighborhood, hence they are assigned the "wrong" within-eight kilometers cutoff. We experimented with other definitions, changing the maximum distance between middle school and competitive high school considered to assign the "relevant" cutoffs. As mentioned in Section IV.A we generally find that the estimated share of pulled-in compliers declines if, for example, we expand this distance to 10km, which suggests that we do not gain more compliers by allowing individuals to have schools more than 8km away in their choice set.

E Empirical Specifications to Estimate the Aggregate Impact of the Reform

We describe the empirical specifications reported in Section V.A to estimate the aggregate effect of the reform.

Cross-markets Difference-in-Difference Model The first model compares changes in outcomes in Bergen with other education markets, where no reforms took place over the same period, in a difference-in-difference framework. In particular, we include in the sample students who enrolled in academic high schools between 2002 and 2007 in Bergen (and neighboring municipalities), and the four largest Norwegian municipalities Trondheim, Stavanger, Drammen and Kristiansand excluding Bergen and the capital Oslo. We estimate the following regression:

$$Y_i = \alpha_0 + \alpha_1 Bergen_i \times Post_i + \alpha_\mathbf{2}\mathbf{Munic_i} + \alpha_\mathbf{3}\mathbf{Cohort_i} + \alpha_\mathbf{4}\mathbf{Munic_i} \times t_i + \varepsilon_i, \tag{E.1}$$

where $Bergen_i$ is a dummy which is 1 if student i started an academic high school in Bergen, $Post_i$ is a dummy for being born after 1988 (affected by the reform if in Bergen), $Munic_i$ is a vector of dummies for enrolling in high school in each of the included municipalities, $Cohort_i$ is a vector of dummies for belonging to each cohort and t_i is a linear function of the month of birth (ranging from 1 to 72). This specification is thus a standard two-way fixed effect model, where we also control for municipality-specific linear trends in the month of birth, to avoid confounding pre-existing differences in the trends of outcomes between municipalities with the impact of the reform. The effect of the reform in Bergen is captured by the estimate of parameter α_1.

Cross-markets Difference-in-Difference Model, by GPA Quintiles We also analyze heterogeneity in the impact of the reform according to the grades obtained at middle school. In particular, we define five dummies, each taking value 1 if the student obtained a middle school GPA within each quintile of the middle school GPA distribution of the respective cohort (we rank students attending academic high schools in the 5 municipalities within the respective cohort, without separating between municipalities). We then estimate an equation similar to Equation (E.1), this time further multiplying the interaction between $Bergen_i$ and $Post_i$ with each of these five dummies, and controlling for quintile dummies' main effect and its interaction with all the controls (municipality fixed effect, cohort fixed effect and a linear trend in month of birth allowed to be heterogeneous between municipalities). Calling $Quintile_{q,i}$ the dummy for quintile q, we thus estimate the following model

$$\begin{aligned} Y_i = \alpha_0 &+ \sum_{q=1}^{5} \alpha_{q,post} Quintile_{q,i} \times Bergen_i \times Post_i + \alpha_\mathbf{2}\mathbf{Munic_i} + \alpha_\mathbf{3}\mathbf{Cohort_i} \\ &+ \alpha_\mathbf{4}\mathbf{Munic_i} \times t_i + \sum_{q=1}^{5} Quintile_{q,i} \times (\gamma_q + \gamma_\mathbf{q,1}\mathbf{Munic_i} + \gamma_\mathbf{q,2}\mathbf{Cohort_i} + \gamma_\mathbf{q,3}\mathbf{Munic_i} \times t_i) + \varepsilon_i. \end{aligned} \tag{E.2}$$

RDD Model Our second strategy relies on students attending an academic high school in the Bergen area and exploits an RDD framework which is similar to Equation (5), where the *Post* dummy appears on its own (not interacted with indicators for being above or below competitive high schools' admission cutoffs):

$$Y_i = \alpha_0 + \alpha_1 Post_i + Week_i \times (\alpha_2 + \alpha_3 Post_i) + \alpha_{\mathbf{4}} \mathbf{X_i} + \varepsilon_i, \tag{E.3}$$

where $Post_i$ is a dummy for being born after December 31^{st}, 1988, $Week_i$ is a linear function of the week of birth (normalized to 0 for those born in the first week of 1989) and X_i is a vector of 52 dummies for being born in each week of the year. Hence, parameter α_1 is the difference between the outcomes of students born just before and after December 31^{st}, 1988 controlling for constant seasonal effects.

RDD Model by GPA Quintiles Also for this specification, we analyze heterogeneity in the effect of the reform by interacting the $Post_i$ dummy with dummies for quintiles of the middle school GPA. For this last piece of analysis, we estimate the parameters of the following linear model:

$$Y_i = \alpha_0 + \sum_{q=1}^{5} \alpha_{q,post} Quintile_{q,i} \times Post_i + Week_i \times (\alpha_2 + \alpha_3 Post_i) + \alpha_{\mathbf{4}} \mathbf{X_i} + \sum_{q=1}^{5} \alpha_{5,q} Quintile_{q,i} + \varepsilon_i. \tag{E.4}$$

F Empirical Specifications for the Reduced-form Model

Reduced Form. The reduced form equation described in Section V.B is specified as follows:

$$\begin{aligned} Y_i = {} & \beta_0 + \beta_{CC} Post_i \times Below_i + \beta_{CP} Post_i \times Above_i + \beta_1 Below_i \\ & + Week_i \times (\beta_2 + \beta_3 Below_i + \beta_4 Post_i) + \beta_5 GPA_i + \beta_{\mathbf{6}} \mathbf{X_i} + \varepsilon_i, \end{aligned} \tag{F.1}$$

which corresponds to the regression of the outcome on the same variables as in the first stage equation (Equation (6)). Under assumption 3, the estimate of parameters β_{CC} and β_{CP} is the effect of the reform on low- and high-middle school GPA students, respectively.

Reduced Form exploiting distance-based variation. The reduced form equation estimated exploiting the additional variation offered by the distance between the middle school attended by the student and any competitive school is the following

$$Y_i = \beta_0 + (\beta_{CC}^{close} close_i + \beta_{CC}^{far} far_i) \times Post_i \times Below_i + (\beta_{CP}^{close} close_i + \beta_{CP}^{far} far_i) \times Post_i \times Above_i + Below_i \times (\beta_1 + \beta_2 close_i) + Week_i \times (\beta_3 + \beta_4 Below_i + \beta_5 Post_i) + \beta_6 GPA_i + \beta_{\mathbf{7}} \mathbf{X_i} + \varepsilon_i, \quad \text{(F.2)}$$

which corresponds to the regression of the outcome variable on the same variables as in the first stage equation (Equation (7)). The main effect of $close_i$ is also included in X_i together with the other controls as described in Section IV.B. Under assumption 3, the estimate of parameters β_{CC}^{close} (β_{CC}^{far}) is the effect of the reform for low-middle school GPA students whose middle school is located less than (more than) 3km away from any competitive high school, and β_{CP}^{close} (β_{CP}^{far}) is the effect of the reform for high-middle school GPA students whose middle school is located less than (more than) 3km away from any competitive high school.

G Robustness and Validation Checks: Additional Details

The econometric model displayed in Section III relies on Assumptions 1–4 (exclusion restriction, no peer effects, independence, and conditional monotonicity). In this section, we provide additional discussion and empirical evidence that in the framework we build these assumptions hold.

G.1 Robustness Checks

Table A8 reports the estimates for the first stage and the LATE, separately for CC and CP, from specifications where we change the functional forms and the definitions of the control variables.

In column (1) we report our baseline estimates. In column (2) we impose the same linear trend pre- and post-reform in terms of week of birth (the trend is allowed to be different for high- and low-GPA students). This corresponds to a before-after design where we compare the overall changes in outcomes before/after the reform controlling for a single linear trend which is allowed to differ for students with GPA above or below the cutoff but not to change after the reform. More specifically, we estimate the following first and second-stage equations:

$$
\begin{aligned}
Y_i = \gamma_0 + \gamma_{CC} D_i \times Below_i + \gamma_{CP} D_i \times Above_i + \gamma_1 Below_i \\
+ Week_i \times (\gamma_2 + \gamma_3 Below_i) + \gamma_5 GPA_i + \gamma_{\mathbf{6}} \mathbf{X_i} + \varepsilon_i,
\end{aligned} \tag{G.1}
$$

and

$$
\begin{aligned}
D_i \times p_i = a_0^p + a_{CC}^p Post_i \times Below_i + a_{CP}^p Post_i \times Above_i + a_1^p Below_i \\
+ Week_i \times (a_2^p + a_3^p Below_i) + a_5^p GPA_i + \mathbf{a_6^p X_i} + u_i^p, \quad \forall p \in \{Above, Below\}.
\end{aligned} \tag{G.2}
$$

where the coefficients γ_2 and γ_3 estimate linear trend in weeks of birth and its interaction with the *Below* dummy. X includes the same controls as in equation 5. Effectively, here we treat $Post_i$ as an instrument and assume that the timing of the reform is exogenous subject to a pre-existing linear trend which may vary by students' GPA. Column (2) shows that both the first stage and LATE results are very close to the ones in the main specification in terms of magnitude and precision.

In columns (3) and (4) we add more flexibility to the time trends by allowing the change in the post-reform trend to differ by GPA groups (column 3) and by allowing for a quadratic trend in the running variable (column 4).[56] Finally in column (5) we change the definition of the running variable, where we express all time variables (the linear trend and the seasonality dummies) in terms of the month of birth. Because some of these specifications are more demanding in terms of the number of parameters to estimate we lose some statistical power. Overall, the first stage coefficients and the point estimates of the treatment effects are broadly consistent with the results presented in Section V.B. They confirm the negative effect of the reform on crowded-out students and the small and insignificant return of attending competitive schools for students who are pulled into these schools as a consequence of the reform.

Finally, we conduct sensitivity checks on the way we categorize pre-reform students into high and low-GPA groups. We estimate the baseline model using alternative pre-reform cutoffs of the GPA distribution, where instead of the 28^{th} percentile we use the 25^{th} and the 35^{th}. Table A9 shows that our results are robust to these different specifications.

[56] For the estimates in column (3) we include the term $Week_i \times Post_i \times Below_i$ to Equation (5). For the estimates in column (4) we include $Week_i^2$ and its interaction with $Below_i$ and $Post_i$.

G.2 Exclusion Restriction and Independence

If the reform impacted school resources in terms of teacher quality, class size and facilities, this may result in a violation of the exclusion restriction (Assumption 1). In Norway, schools are centrally financed and resources do not depend on the quantity or quality of students, hence changes in the pool of students do not affect school finances. Class size is also regulated. In addition, because we only look at reform-induced changes in the first three years after the reform other characteristics, such as schools' location, facilities, or reputation, are unlikely to change so rapidly, even in the presence of big changes in students' characteristics. Although teachers can decide to move between schools, in Figure A5 we show that the characteristics of teachers (age and qualifications) in competitive and non-competitive schools did not change systematically with the reform.[57]

Assumption 3 instead requires that students affected by the reform are comparable to those not affected by the reform. In our approach, this translates to an identifying assumption that students who are born just after December 31st, 1988 are comparable to those born a few days before. We can thus analyze whether students born around this date are similar, by investigating whether the characteristics of individuals born before and after this date are significantly different. A test like this resembles tests that aim at assessing if randomization has been properly conducted in randomized experiments (Lee and Lemieux, 2010). In Figure A6 we plot the average characteristics of students born in each week around December 31st, 1988, in terms of gender and parental background. The figures show no clear jump in the first week of 1989. In addition, we estimate parameters of a regression of all these characteristics on a dummy for being born from 1989 onwards and linear time trends in the week of birth, allowed to differ before and after the last week of 1988. The estimated coefficients for the dummy at the discontinuity are reported on top of each figure, they are small and never statistically significant, providing evidence that the assumption of no differences between students affected and those not affected by the reform holds.

[57] Although average within-school teacher quality as measured by their observed qualifications does not change after the reform, this does not exclude the possibility that the same teachers adapt their teaching style to changing student composition. For example, teachers may be more or less effective in their classes when facing more homogeneous classrooms in terms of underlying ability (Duflo, Dupas, and Kremer, 2011). Teaching styles are not observed, and therefore we cannot rule out that teaching style may respond to class composition.

G.3 No Peer Effects

Assumption 2 imposes that the outcome of a student only depends on his/her school choices. In practice, the reform induced within-school changes in student compositions and one might be rightfully concerned that these changes in student composition may affect the schooling outcome of a student via the general peer effects.

We present two pieces of evidence to show that changing peer quality is unlikely to be driving our results. First, we show that there exists pre-reform differences between competitive and non-competitive schools beyond differences in student composition, hence changes in this composition alone should not fully explain the estimated effect of attending a more competitive school. We estimate high schools' value added (VA) for the years before the reform, by regressing final year high school test scores on school fixed effects and students' middle-school GPA (the school fixed effects identify VA).[58] The VA estimates tell us whether students in competitive schools have better outcomes once we control for initial student quality and sorting into these schools. We then estimate the same VA model including also controls for peers' background, to see how much of the difference between schools can be explained by the type of students enrolled in the school. In Figure A7 we show that three out of the five competitive schools are also among the ones with the highest VA, something which cannot be explained by peer composition. Although our definition of a competitive school is based upon post-reform admission cutoffs, this additional result reassures that there were prevailing differences in "productivity" between the competitive and non-competitive schools and that these are not fully explained by student composition.

Second, we can test whether the reform affects the outcomes of students who attend the same school regardless of the reform. In particular, Assumption 2 and our identification result will not hold if the outcomes of Always Takers and Never Takers are affected by the reform. We consider two subpopulations for which the reform is unlikely to affect their school choice: (1) students who have a middle school GPA above the 75^{th} percentile of the distribution (who are always above the cutoff) and who enrolled at a middle school within 1.5 kilometers of a competitive high school (and thus are

[58]Note that here we use only pre-reform data. Post-reform, we would be able only to observe high school GPA for students who complete high school, so any effect of the reform on high school completion would affect the distribution of high school GPA and would then be difficult to separate any change in VA due to the reform and the one due to changes in the probability of completing high school. For the same reason, we do not include high school GPA as one of the outcomes in our analysis.

likely to be within the catchment area of a competitive high school) have the possibility to attend a competitive high school irrespective of whether or not they are affected by the reform[59]; (2) students who have a middle school GPA below the 15^{th} percentile of the distribution and who enrolled at a middle school that is more than 5km away from any competitive high school, instead, are likely to never enroll in a competitive high school. We estimate the aggregate effect of the reform for each subpopulation separately, using an RDD framework.[60] Table A10 displays the results. In column (1) we start by estimating the change in the probability of attending a competitive school for the groups of students with middle school GPA above the 75^{th} percentile and of those below the 15^{th} percentile, without imposing the distance restriction. As expected, the probability of enrolling in one of these schools increases for the former and decreases for the latter. Column (2) shows that this effect is much smaller and not significant when we impose the distance limitation, hence suggesting that the reform does not impact high-GPA students living close to competitive schools and low-GPA students living far away by changing their high school choices. Columns (3) to (5) show that the estimates of the effect of the reform are generally small and never statistically significant. These results provide further support to the validity of Assumption 2.

G.4 Middle School GPA Distribution

Our empirical regression includes middle school GPA (in percentiles) as controls; middle school GPA also determines whether a student is assigned above or below an admission cutoff. Our strategy therefore requires that the middle school GPA is exogenous to the reform. For instance, if exams are not blindly graded, teachers' grading practice may change in response to the reform.[61] Middle school GPA can also be endogenous to the reform for other reasons, such as families selecting a different middle school to boost their children's grades or certain students exerting more effort to obtain access to their preferred school.

We argue that the institutional setting and the relatively short time span we consider alleviate some of these concerns. The cohorts under study already started middle school at the time of the reform,

[59] We selected 1.5km because, as shown in Table 5 even among students attending a middle school within 3km from a competitive school there is a significant fraction of pulled-in compliers.

[60] We estimate a linear regression where the outcome variables are regressed on a dummy for being born after December 1988 (and hence attending high school after the reform took place), a linear trend in week of birth that is allowed to differ before and after the reform, and seasonal dummies for the week of birth within the year (52 dummies).

[61] As an example, a teacher may inflate grades for certain students aiming for competitive schools.

making it unlikely that parents move their children to a different middle school in order to boost their GPA. As explained in Appendix A.1, although part of the GPA consists of teacher-graded exams that are not blindly graded, GPA is also determined by national exams which are blindly graded externally. Therefore, teachers only had control over part of the final GPA, with the rest determined by externally graded exams. Furthermore, the teacher assessment is given prior to the national exams, so teachers do not know the grade of the students in the blindly assessed part when giving their marks. In addition, the merit-based system was not in place before the reform, hence in the first few years of the reform teachers (and students) operate under lots of uncertainty, in terms of where the admission cutoff is and also which schools are going to be oversubscribed. Therefore, the lack of information at the time of teacher assessment means that teachers are very limited (if not impossible) to change their grading practice to affect the admission outcome for certain students.[62]

We perform two empirical tests to further justify our assumption that middle school GPA is exogenous to the reform. As our first test, we compare the distribution of middle school GPA for students in Bergen before and after the reform, relative to the distribution of grades at the national level. If teachers in Bergen manipulate grades, or if students exert more effort and there is an overall improvement in middle school grades, we should observe changes in the GPA distribution in Bergen relative to the national distribution. We thus plot the densities of the nationally-ranked middle school GPA in Bergen before and after 2005 in panel (a) in Figure A8. These distributions overlap and a Kolmogorov–Smirnov test excludes any significant difference. Similarly, in panels (b) and (c), we plot the same distributions for cohorts affected and not affected by the reform in Bergen, separately for individuals of different SES. Plots and formal tests exclude any difference between pre- and post-reform cohorts, suggesting that the reform had no systematic effect in boosting students' middle school GPAs.

Our second test provides more direct evidence showing that teachers do not change their grading style following the reform. Considering that grade inflation is only a potential issue with teacher assessment but not the national tests (as the latter is blindly graded), we estimate the relationship between national test scores and middle-school GPA and test whether the relationship between middle-school GPA and national test scores differs systematically after the reform. For instance, if teachers systematically inflate their grades for the best students after the reform, we should expect the gradient

[62]This stands in contrast to the case in Diamond and Persson (2016), where admission cutoff is known to the teacher at the time when the exams are graded.

of middle-school GPA with respect to national test score to become steeper. To this end, we regress the within-cohort rank of middle-school GPA on national test scores and interact post-reform dummy with national test scores. The results from this regression is reported in Table A11. The interaction term between post-reform dummy and the national test score is small and insignificantly different from zero, suggesting a stable relationship between the national test (blindly graded) and middle school GPA (which includes teacher assessment). This additional evidence provides further empirical support that there appears no systematic grade inflation following the reform.

H Tables and Figures

	Overall		Bottom Quintile of Middle School GPA		Top Quintile of Middle School GPA	
	(1) Pre-Reform	(2) Post-Reform	(3) Pre-Reform	(4) Post-Reform	(5) Pre-Reform	(6) Post-Reform
(a) Students Characteristics						
Parents with University	0.42	0.40	0.29	0.25	0.56	0.56
Parental Earnings	375,873 (216,036)	411,738 (252,177)	329,446 (187,460)	358,976 (191,265)	425,113 (254,167)	447,744 (244,720)
Female	0.53	0.55	0.36	0.41	0.69	0.65
(b) Academic Outcomes						
Admission at 5 Comp. Schools	0.38	0.36	0.35	0.09	0.44	0.59
High School Completion	0.90	0.90	0.73	0.73	0.96	0.96
University Enrolment	0.85	0.85	0.66	0.70	0.94	0.92
University Completion	0.82	0.83	0.57	0.62	0.95	0.95
Days Absent	6.60 (4.94)	7.98 (6.16)	8.12 (5.55)	10.52 (7.08)	4.89 (3.76)	5.63 (4.26)
Elite Uni Enrolment	0.15	0.15	0.03	0.02	0.38	0.36
Elite Uni Completion	0.11	0.09	0.02	0.01	0.27	0.21
STEM Enrolment	0.16	0.15	0.12	0.12	0.15	0.16
STEM Completion	0.10	0.10	0.07	0.07	0.09	0.12
N	3,675	4,285	711	842	749	880

Table A1: Descriptive statistics and post-high school outcomes.

Notes : Columns (1) and (2) include all students not affected and those affected by the reform, respectively. Columns (3) and (4) include students whose middle school GPA is below the 20^{th} percentile of the middle school GPA distribution of their cohort. Columns (5) and (6) include students whose middle school GPA is above the 80^{th} percentile of the middle school GPA distribution of their cohort. "Parents with university" is the proportion of students whose parents (both) completed a university degree. Parental earnings are expressed in 1998 NOK (6NOK $\approx$ 1USD). "Days Absent" is the yearly average of the number of days the student was absent from school in the three years of high school. "Elite Uni" refers to attending either one of two prestigious institutions (NHH and NTNU) or a medicine degree (see Section IV.A). "STEM" refers to attending any course in a STEM field. S.d. for continuous variables are in parentheses.

	Competitive Schools × Below		Competitive Schools × Above	
First Stage				
Post × Below	-0.282***		0.001	
	(0.034)		(0.019)	
Post × Above	-0.024***		0.153***	
	(0.005)		(0.024)	
	Crowded-out Compliers		Pulled-in Compliers	
Outcomes	Treatment Effect	Reduced Form	Treatment Effect	Reduced Form
Avg Days of Absence[a]	-0.518	0.127	2.122	0.327
	(2.760)	(0.637)	(2.094)	(0.301)
Elite Uni Enrolment	0.024	-0.007	0.056	0.008
	(0.070)	(0.020)	(0.124)	(0.019)
Elite Uni Completion	0.024	-0.007	-0.009	-0.002
	(0.056)	(0.016)	(0.111)	(0.017)
STEM Uni Enrolment	0.108	-0.030	0.054	0.006
	(0.112)	(0.032)	(0.128)	(0.019)
STEM Uni Completion	0.089	-0.025	-0.013	-0.004
	(0.085)	(0.024)	(0.096)	(0.015)
N		6,908		

[a]Average days of absence are observed only for students who completed high school, hence these results are conditional on high school completion.

Table A2: Treatment effects and Reduced Form for CC and CP, additional outcomes

Notes : "Avg Days of Absence" is the yearly average of the number of days the student was absent from school in the three years of high school. "Elite Uni" refers to attending either one of two prestigious institutions (NHH and NTNU) or a medicine degree (see Section IV.A). "STEM" refers to attending any course in a STEM field. Estimates of the Treatment effect come from a regression of the outcome on a dummy for attending a competitive school and having middle school GPA below the relevant cutoff for admission at a competitive school ("Competitive Schools × Below") and on another dummy for attending a competitive school and having middle school GPA above the relevant cutoff ("Competitive Schools × Above"). The former identifies the effect for Crowded-out Compliers, the latter for Pulled-in Compliers. Each of these two variables is instrumented with two other dummies, one for being born after December 31^{st}, 1988 and having middle school GPA below the cutoff ("Post × Below") and another for being born after December 31^{st}, 1988 and having a middle school GPA above the cutoff ("Post × Above"). In addition, each regression controls for a dummy for having middle school GPA below the relevant cutoff, a linear trend in week of birth (with weeks of birth normalized to 0 for those born in the first week of 1989 and the trend allowed to differ before and after the reform and above and below the relevant cutoff), a linear function of middle school GPA percentiles, seasonal (52 weeks) dummies, five dummies indicating whether each competitive school is located within 3km of the attended middle school interacted with the "Below" dummy, and one dummy indicating whether at least one competitive school is located within 8km of the attended middle school. S.e. are clustered at the week-of-birth level. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	(1)	(2)
	Panel (a): First Stage	
	Competitive Schools × Below	Competitive Schools × Above
Post × Below	-0.282***	0.001
	(0.034)	(0.019)
Post × Above	-0.024***	0.153***
	(0.005)	(0.024)
First-stage F	77.37	43.44
	Panel (b): Reduced Form	
Outcomes	Post × Below	Post × Above
School Completion	-0.075**	0.000
	(0.037)	(0.014)
University Enrolment	-0.079*	-0.020
	(0.042)	(0.017)
University Completion	-0.073	-0.012
	(0.049)	(0.018)
N	6,908	

Table A3: First stage and Reduced Form for CC and CP

Notes : Estimates of the Reduced Form (First Stages) come from a regression of the outcome (dummy for attending a competitive school interacted with dummies "Below" and "Above") on a dummy for being born after December 31^{st}, 1988 and having middle school GPA below the cutoff ("Post × Below") and another for being born after December 31^{st}, 1988 and having middle school GPA above the cutoff ("Post × Above"). In addition, each regression controls for a dummy for having a middle school GPA below the relevant cutoff, a linear trend in week of birth (with weeks of birth normalized to 0 for those born in the first week of 1989 and the trend allowed to differ before and after the reform and above and below the relevant cutoff), a linear function of middle school GPA percentiles, seasonal (52 weeks) dummies, five dummies indicating whether each competitive school is located within 3km of the attended middle school and one dummy indicating whether at least one competitive school is located within 8km of the attended middle school. Sanderson Windmejier F-statistics are reported. S.e. of are clustered at the week-of-birth level. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	(1) Crowded-out Compliers		(2) Pulled-in Compliers
Avg Days of	-1.083		1.087
Absence[a]	(2.018)		(2.073)
Overid. p-value		0.386	
Elite Uni	-0.042		0.003
Enrolment	(0.052)		(0.129)
Overid. p-value		0.293	
Elite Uni	-0.036		-0.131
Completion	(0.036)		(0.118)
Overid. p-value		0.078	
STEM Uni	0.081		0.168
Enrolment	(0.082)		(0.134)
Overid. p-value		0.305	
STEM Uni	0.082		0.071
Completion	(0.062)		(0.098)
Overid. p-value		0.394	
N		6,908	

[a] Average days of absence are observed only for students who completed high school, hence these results are conditional on high school completion.

Table A4: Treatment effects for CC and CP, additional outcomes, IV interacted with schools distance

Notes : "Avg Days of Absence" is the yearly average of the number of days the student was absent from school in the three years of high school. "Elite Uni" refers to attending either one of two prestigious institutions (NHH and NTNU) or a medicine degree, which have on average higher returns to completion (see Section IV.A). "STEM" refers to attending any course in a STEM field. Estimates of the Treatment effect come from a regression of the outcome on a dummy for attending a competitive school and having middle school GPA below the relevant cutoff for admission at a competitive school ("Competitive Schools × Below") and on another dummy for attending a competitive school and having middle school GPA above the relevant cutoff ("Competitive Schools × Above"). The former identifies the effect for Crowded-out Compliers, the latter for Pulled-in Compliers. Each of these two variables is instrumented with four other dummies; the interaction of the dummy for being born after December 31^{st}, 1988 and having middle school GPA below the cutoff ("Post × Below") and the one for being born after December 31^{st}, 1988 and having middle school GPA above the cutoff ("Post × Above") are again interacted with a dummy for having attended a middle school which has at least one competitive school within 3km ("Close") and another one for not having competitive high schools within 3km ("Far"). In addition, each regression controls for a dummy for having a middle school GPA below the relevant cutoff, a linear trend in week of birth (with weeks of birth normalized to 0 for those born in the first week of 1989 and the trend allowed to differ before and after the reform and above and below the relevant cutoff), a linear function of middle school GPA percentiles, seasonal (52 weeks) dummies, five dummies indicating whether each competitive school is located within 3km of the attended middle school, a dummy indicating whether at least one competitive school is located within 3km of the attended middle school interacted with the "Below" dummy, and one dummy indicating whether at least one competitive school is located within 8km of the attended middle school. p-values for the Hansen test are reported. S.e. are clustered at the week-of-birth level. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	(1)		(2)	
	Panel (a): First Stage			
	Competitive Schools × Below		Competitive Schools × Above	
Post × Below × Close	-0.471***		0.048***	
	(0.043)		(0.015)	
Post × Below × Far	-0.181***		-0.000	
	(0.035)		(0.014)	
Post × Above × Close	-0.001		0.087***	
	(0.004)		(0.026)	
Post × Above × Far	-0.018***		0.169***	
	(0.005)		(0.027)	
First-stage F	38.88		14.39	
	Panel (b): Reduced Form			
Outcomes	Post × Below		Post × Above	
	Close	Far	Close	Far
School	-0.136***	-0.035	-0.024	0.019
Completion	(0.050)	(0.041)	(0.015)	(0.016)
University	-0.144***	-0.037	-0.021	-0.021
Enrolment	(0.054)	(0.045)	(0.020)	(0.019)
University	-0.131**	-0.036	-0.007	-0.016
Completion	(0.060)	(0.052)	(0.021)	(0.021)
N	6,908			

Table A5: First stage and Reduced Form for CC and CP, Distance-based design

Notes : Estimates of the Reduced Form (First Stages) come from a regression of the outcome (dummy for attending a competitive school interacted with dummies "Below" and "Above") on four other dummies; the interaction of the dummy for being born after December 31^{st}, 1988 and having middle school GPA below the cutoff ("Post × Below") and the one for being born after December 31^{st}, 1988 and having middle school GPA above the cutoff ("Post × Above") are again interacted with a dummy for having attended a middle school which has at least one competitive school within 3km ("Close") and another one for not having competitive high schools within 3km ("Far"). In addition, each regression controls for a dummy for having a middle school GPA below the relevant cutoff, a linear trend in week of birth (with weeks of birth normalized to 0 for those born in the first week of 1989 and the trend allowed to differ before and after the reform and above and below the relevant cutoff), a linear function of middle school GPA percentiles, seasonal (52 weeks) dummies, five dummies indicating whether each competitive school is located within 3km of the attended middle school, a dummy indicating whether at least one competitive school is located within 3km of the attended middle school and one dummy indicating whether at least one competitive school is located within 8km of the attended middle school. Sanderson Windmejier F-statistics are reported. S.e. of are clustered at the week-of-birth level. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	Parental Background				Gender			
	Competitive Schools × Below		Competitive Schools × Above		Competitive Schools × Below		Competitive Schools × Above	
	Low SES	High SES	Low SES	High SES	Men	Women	Men	Women
First Stage								
Post × Below	-0.266***	-0.305***	0.002	-0.006	-0.282***	-0.285***	-0.013	0.020
	(0.027)	(0.040)	(0.012)	(0.013)	(0.030)	(0.032)	(0.011)	(0.014)
Post × Above	-0.017***	-0.008**	0.130***	0.142***	-0.017***	-0.006**	0.073***	0.189***
	(0.004)	(0.003)	(0.022)	(0.022)	(0.004)	(0.003)	(0.023)	(0.023)
First-Stage F	87.65	76.66	41.96	43.41	99.87	78.92	15.02	47.63
	Crowded-out Compliers		Pulled-in Compliers		Crowded-out Compliers		Pulled-in Compliers	
	Low SES	High SES	Low SES	High SES	Men	Women	Men	Women
Treatment Effect								
School	0.222	0.343**	0.062	0.028	0.178	0.387**	0.066	0.028
Completion	(0.154)	(0.164)	(0.109)	(0.105)	(0.142)	(0.165)	(0.187)	(0.086)
University	0.220	0.369**	0.091	-0.244*	0.250	0.279	0.096	-0.134
Enrolment	(0.173)	(0.169)	(0.142)	(0.136)	(0.153)	(0.186)	(0.236)	(0.103)
University	0.255	0.252	0.031	-0.091	0.117	0.476**	0.010	-0.057
Completion	(0.200)	(0.202)	(0.151)	(0.142)	(0.185)	(0.211)	(0.266)	(0.111)
N	6,908							

Table A6: First stage and Treatment effect for CC and CP by parental background and by gender

Notes : Estimates come from a regression of the outcome on a dummy for attending a competitive school and having middle school GPA below the relevant cutoff for admission at a competitive school ("Competitive Schools × Below") and on another dummy for attending a competitive school and having middle school GPA above the relevant cutoff ("Competitive Schools × Above"). The former identifies the effect for Crowded-out Compliers, the latter for Pulled-in Compliers. Each of these two variables is instrumented with two other dummies, one for being born after December 31^{st}, 1988 and having middle school GPA below the cutoff ("Post × Below") and another for being born after December 31^{st}, 1988 and having middle school GPA above the cutoff ("Post × Above"). We report the estimates of a regression where we interact each of these with a dummy for having two parents with a university degree and with a dummy for having at least one parent with no university degree ("Parental Background"), and one where the two treatment dummies are interacted with a dummy for being male and a dummy for being female ("Gender"). In addition, each regression controls for a dummy for having a middle school GPA below the relevant cutoff, a linear trend in week of birth (with weeks of birth normalized to 0 for those born in the first week of 1989 and the trend allowed to differ before and after the reform and above and below the relevant cutoff), a linear function of middle school GPA percentiles, seasonal (52 weeks) dummies, five dummies indicating whether each competitive school is located within 3km of the attended middle school, one dummy indicating whether at least one competitive school is located within 8km of the attended middle school and the main effect of the dummy for parental background and gender, respectively, interacted with the "Below" dummy. Sanderson Windmejier F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	Parental Background		Gender	
	Most Selective Schools × Above		Most Selective Schools × Above	
	Low SES	High SES	Men	Women
First Stage				
Post × Below	0.000	-0.002	-0.004	0.005
	(0.005)	(0.006)	(0.005)	(0.007)
Post × Above	0.187***	0.227***	0.181***	0.225***
	(0.026)	(0.030)	(0.031)	(0.029)
First-Stage F	53.18	59.46	40.67	63.63
Treatment Effect				
School Completion	0.121	0.069	0.138	0.048
	(0.101)	(0.082)	(0.112)	(0.081)
University Enrolment	-0.005	-0.175	-0.105	-0.132
	(0.140)	(0.108)	(0.146)	(0.108)
University Completion	0.048	0.054	0.030	0.045
	(0.135)	(0.099)	(0.135)	(0.108)
N	6,177			

Table A7: First stage and Treatment effect for attending the two most selective schools, by parental background and by gender

Notes : Estimates come from a regression of the outcome on a dummy for attending one of the two most selective schools and having middle school GPA below the relevant cutoff for admission at the most selective schools school and on another dummy for attending a competitive school and having middle school GPA above the relevant cutoff ("Most Selective Schools × Below"). The former identifies the effect for a mix of crowded out compliers and compliers who are crowded out from a most selective school or pulled into the selective school from a non-competitive school (not reported), the latter for compliers pulled-in the most selective schools. Each of these two variables is instrumented with two other dummies, one for being born after December 31^{st}, 1988 and having middle school GPA below the cutoff ("Post × Below") and another for being born after December 31^{st}, 1988 and having middle school GPA above the cutoff ("Post × Above"). We report the estimates of a regression where we interact each of these with a dummy for having two parents with a university degree and with a dummy for having at least one parent with no university degree ("Parental Background") and one where the two treatment dummies are interacted with a dummy for being male and a dummy for being female ("Gender"). In addition, each regression controls for a dummy for having middle school GPA below the relevant cutoff, a linear trend in the week of birth (with weeks of birth normalized to 0 for those born in the first week of 1989 and the trend allowed to differ before and after the reform and above and below the relevant cutoff), a linear function of middle school GPA percentiles, seasonal (52 weeks) dummies, two dummies indicating whether each most selective school is located within 3km of the attended middle school, one dummy indicating whether at least one of the two most selective schools is located within 8km of the attended middle school and the main effect of the dummy for parental background and gender, respectively, interacted with the "Below" dummy. The sample only includes post-reform students who attended a middle school which is within 8km from one of the two most selective schools. Sanderson Windmejier F-statistics are reported. S.e. of are clustered at the week-of-birth level. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	(1) Baseline	(2) Single Linear Time Trend	(3) Heterogenous Post Time Trend	(4) Week Quadratic	(5) Months
			(a) Crowded-out Compliers		
First Stage					
Post × Below	-0.282***	-0.282***	-0.279***	-0.295***	-0.279***
	(0.034)	(0.034)	(0.033)	(0.035)	(0.033)
Post × Above	-0.024***	-0.024***	-0.023***	-0.037***	-0.025***
	(0.005)	(0.005)	(0.005)	(0.010)	(0.005)
Treatment Effects					
School	0.268**	0.279**	0.207	0.182	0.276*
Completion	(0.136)	(0.133)	(0.143)	(0.141)	(0.150)
University	0.280*	0.288**	0.257*	0.228	0.283*
Enrolment	(0.147)	(0.145)	(0.151)	(0.145)	(0.168)
University	0.259	0.296*	0.187	0.095	0.268
Completion	(0.177)	(0.176)	(0.183)	(0.178)	(0.195)
			(b) Pulled-in Compliers		
First Stage					
Post × Below	0.001	-0.013	-0.013	-0.033	0.004
	(0.019)	(0.018)	(0.018)	(0.027)	(0.019)
Post × Above	0.153***	0.153***	0.153***	0.142***	0.149***
	(0.024)	(0.024)	(0.024)	(0.030)	(0.023)
Treatment Effects					
School	0.042	0.044	0.035	0.077	0.052
Completion	(0.095)	(0.094)	(0.094)	(0.128)	(0.087)
University	-0.087	-0.086	-0.090	-0.039	-0.075
Enrolment	(0.112)	(0.111)	(0.112)	(0.155)	(0.119)
University	-0.036	-0.031	-0.045	0.108	-0.018
Completion	(0.126)	(0.125)	(0.125)	(0.172)	(0.126)
N			6,908		

Table A8: Robustness checks for the main specification

Notes : Column (1) reports the baseline estimates from Table 4, column (2) imposes a linear trend in the week of birth that is invariant to the reform but allowed to differ between students above and below the cutoff), column (3) adds an interaction between *Post*, *Week* and *Below*, column (4) includes a quadratic function of the running variable *Week* and its interaction with *Below* and with *Post* and column (5) uses the month of birth instead of week of birth as the running variable. S.e. are clustered at the week-of-birth level apart from column (5) where they are clustered at the month-of-birth level. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	(1) Cutoff at 25th GPA Percentile		(2) Cutoff at 35th GPA Percentile	
	Panel (a): First Stage			
	Competitive Schools × Below	Competitive Schools × Above	Competitive Schools × Below	Competitive Schools × Above
Post × Below	-0.264***	-0.013	-0.291***	0.010
	(0.034)	(0.019)	(0.033)	(0.019)
Post × Above	-0.019***	0.143***	-0.029***	0.170***
	(0.005)	(0.023)	(0.006)	(0.024)
First-stage F	79.74	39.73	80.58	50.57
	Panel (b): Treatment Effect			
	Crowded-out Compliers	Pulled-in Compliers	Crowded-out Compliers	Pulled-in Compliers
School Completion	0.277*	0.039	0.318**	0.067
	(0.152)	(0.098)	(0.126)	(0.089)
University Enrolment	0.291*	-0.106	0.336***	-0.059
	(0.162)	(0.118)	(0.128)	(0.104)
University Completion	0.260	-0.045	0.314**	-0.016
	(0.195)	(0.134)	(0.155)	(0.115)

Table A9: First stage and Treatment effect for CC and CP changing pre-reform cutoff definition

Notes : Estimates come from the same specification as the one used to produce the estimates in Table 4. In column (1) we define the *Below* dummy as being equal to 1 for pre-reform students with a middle school GPA below the 25^{th} percentile of their cohort middle school GPA distribution. In column (2) we define the *Below* dummy as being equal to 1 for pre-reform students with a middle school GPA below the 35^{th} percentile of their cohort middle school GPA distribution. Sanderson Windmejier F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

	(1) Competitive HS Enrolment	(2) Competitive HS Enrolment	(3) HS Completion	(5) Uni Enrollment	(6) Uni Completion
	(a) Top Quartile Middle School GPA				
	Overall	Dist<1.5km	Dist<1.5km	Dist<1.5km	Dist<1.5km
Post	0.206***	0.098	-0.001	0.034	-0.033
	(0.041)	(0.070)	(0.057)	(0.064)	(0.044)
N	2,034	419	419	419	419
	(b) Bottom 15 Percentiles Middle School GPA				
	Overall	Dist>5km	Dist>5km	Dist>5km	Dist>5km
Post	-0.239***	0.006	-0.005	-0.006	-0.027
	(0.040)	(0.030)	(0.084)	(0.080)	(0.091)
N	1,155	515	515	515	515

Table A10: Estimates of the effect of the reform on specific subgroups of the population.

Notes : Panel (a) shows the estimates for students who have a middle school GPA above the 75th percentile, panel (b) for students who have a middle school GPA below the 15th percentile. “Dist” represents the distance between middle school and the closest of the five competitive high schools. “Post” are estimates of coefficients on a dummy for being born after December 31st, 1988 and thus for being affected by the reform. Regressions also control for linear trends in the week of birth, allowed to differ before and after the last week of 1988 (with weeks of birth normalized to 0 for those born in the first week of 1989), and for seasonal (52 weeks) dummies. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$.

	Middle School GPA
Blind Grade	19.16***
	(0.418)
Blind Grade ×	0.073
Post 2005	(0.572)
Year 2003	-3.602***
	(0.948)
Year 2004	-1.049
	(0.921)
Year 2005	-3.630
	(2.477)
Year 2006	-1.616
	(2.453)
Year 2007	-3.109
	(2.479)
Constant	-25.75***
	(1.780)
N	7,913

Table A11: Change in the correlation between middle school GPA and blinded grades, pre- and post-2005

Notes : Estimates of the parameters of a regression of the middle school GPA percentile (with students ranked among other students from the same cohort) on the grades obtained by the same student in the national middle school exam and the interaction between the national exam grades and a dummy taking value 1 if the student takes the exam in or after 2005. The regression controls for the year of graduation fixed effects. * $p < 0.10$, ** $p < 0.05$, ***$p < 0.01$

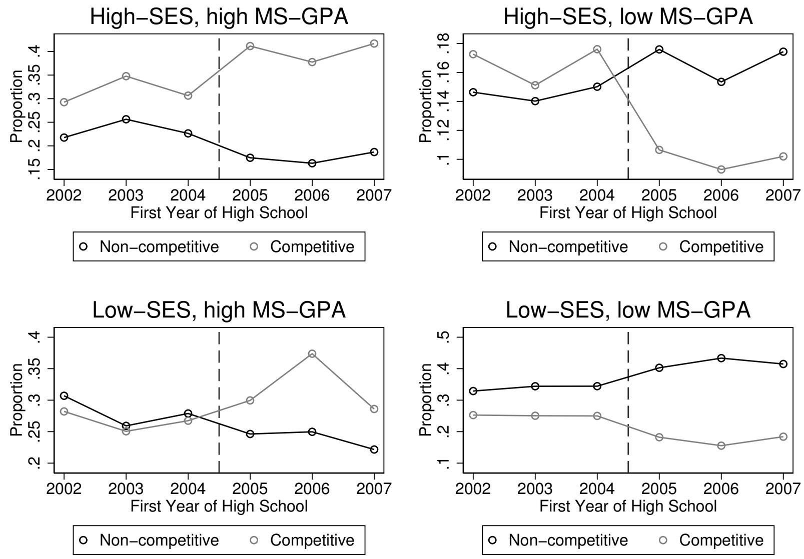


Figure A1: Proportion of students with different middle school GPA (MS-GPA) and different SES attending competitive and non-competitive high schools.

Notes : High-SES students have both parents with a university degree, Low-SES students have at least one parent with no university degree, see Section IV.A. High (Low) middle school-GPA students have a middle school GPA above (below) the median of the middle school GPA distribution in their cohort.

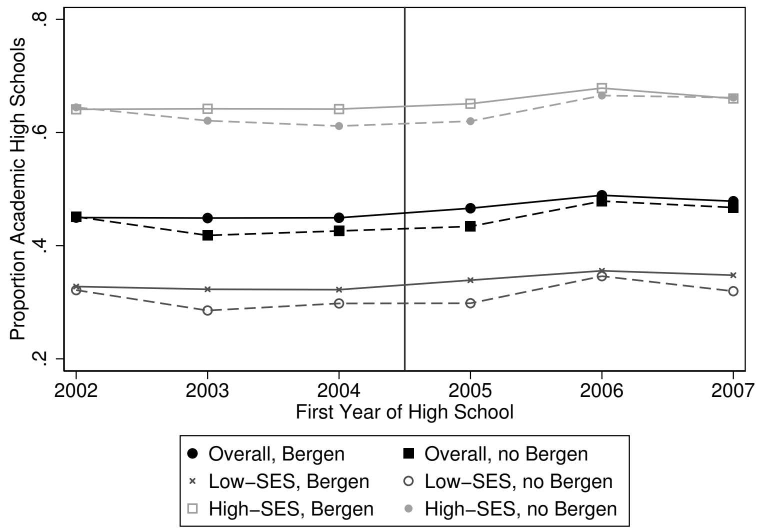


Figure A2: Proportion of students enrolling in academic high school, Bergen vs. rest of Norway.
Notes : The proportions are the ratio between the number of students enrolling in academic high school and the total number of students graduating from middle school every year. “Bergen county” includes all students graduating in Hordaland county (where Bergen is located). “Overall, no Bergen” includes students graduating in every county, apart from Hordaland. “High-SES” students’ parents (both) graduated from university, while at least one parent of “Low-SES” students did not complete university.

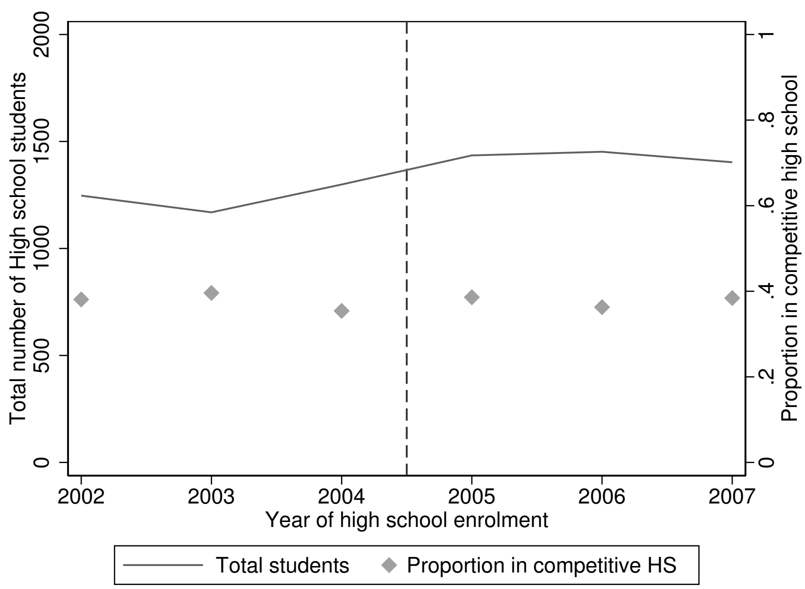
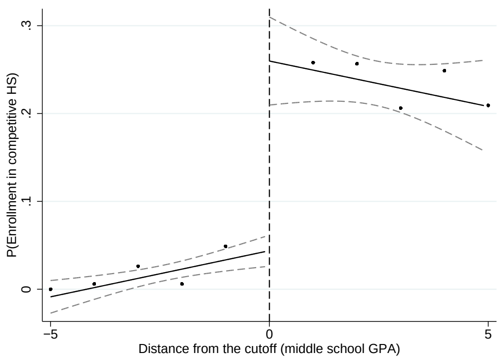
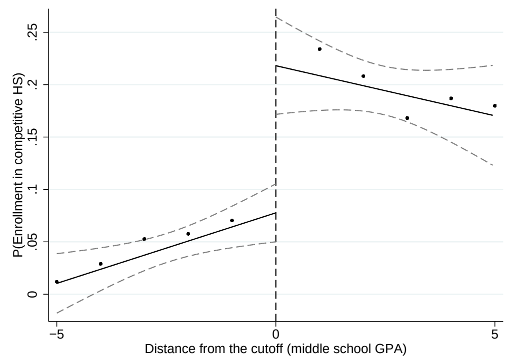


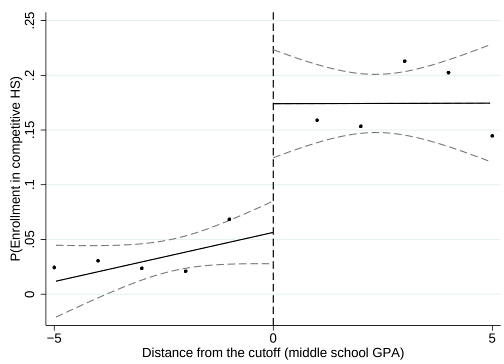
Figure A3: Number of students enrolling in academic high schools and, among these, proportion enrolling in competitive high schools every year.



(a) 2005



(b) 2006



(c) 2007

Figure A4: Probability of enrolling in a competitive high school around GPA admission cutoff in the three post-reform years.

Notes : These figures pull together all the different admission cutoffs in each post-reform year (hence, five admission cutoffs in 2005 and 2006 and four admission cutoffs in 2007, see Section IV.A). The plots show the proportion of students enrolling in a competitive high school for students whose GPA is between -5 and 5 GPA points from the cutoff (each point is the proportion for all students who are within that distance). Figures also display a linear fit and 95% confidence interval for this fit, allowed to be different above and below the admission cutoff. Details on how the thresholds are defined are provided in Section D.

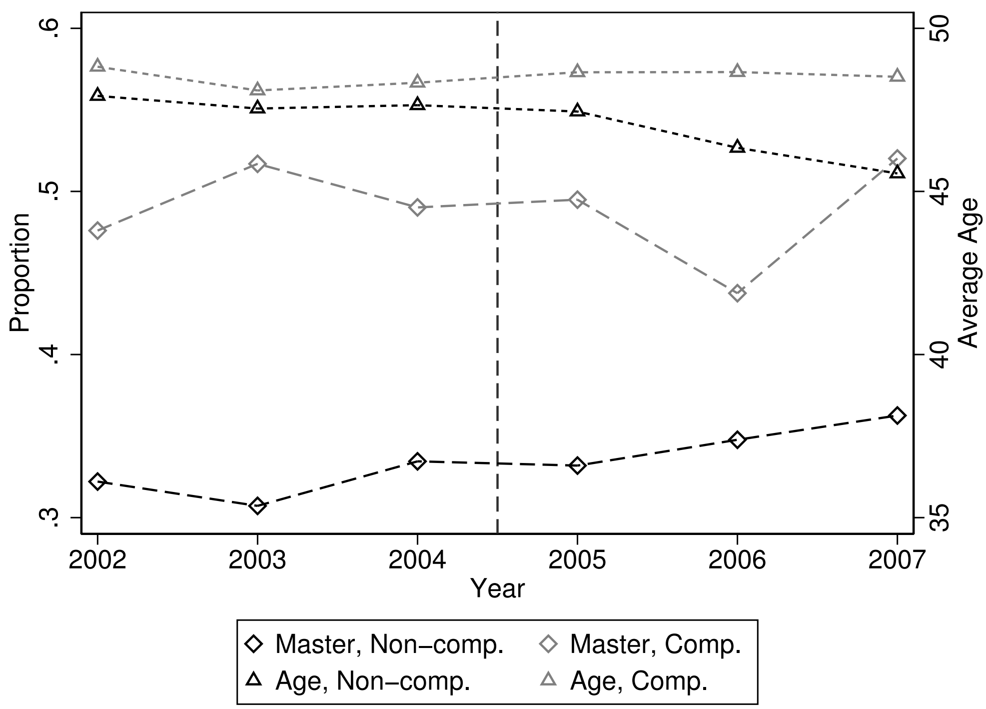
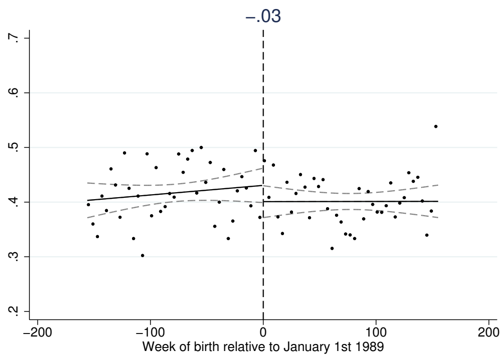
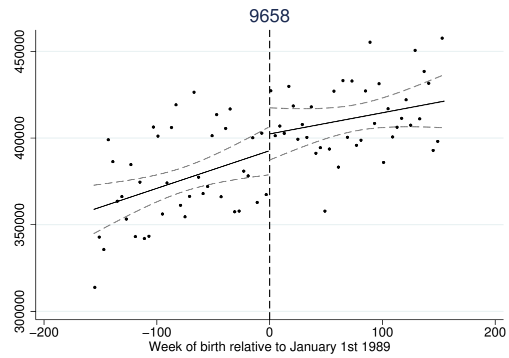


Figure A5: Teachers' average characteristics in every year, competitive and non-competitive schools.

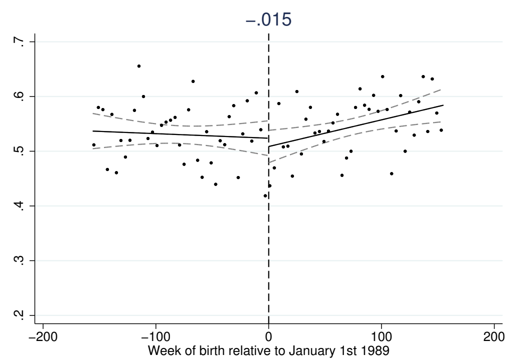
Notes : "Master" is the proportion of teachers holding a master's degree (left axis). "Age" is the average age of teachers (right axis). "Comp." is for competitive schools, "Non-comp." are the non-competitive schools.



(a) Parents with University



(b) Parental Earnings



(c) Female

Figure A6: Average characteristics of children born in different weeks around January 1^{st}, 1989.

Notes : The plots show the average value of each covariate (the proportion of students with both parents graduating from university (a), average annual parental earnings during the three years in high school (b), proportion of females (c)) for students born in the different weeks around the first week of 1989. Each dot represents the average for 4 weeks, while the linear fit (solid line) uses weekly data. The dashed line is the 95% confidence interval of such fit. The numbers on top of the figures are the estimates of the parameter on a dummy for being born after the threshold date, from a regression of the characteristics on the dummy for being born after the threshold (January 31^{st}, 1988) and a linear trend in the week distance from the threshold (allowed to differ before and after the threshold). Standard errors are clustered at the week-of-birth level and none of the reported coefficients is statistically significant. Parental earnings are expressed in 1998 NOK (6 NOK $\approx$ 1 USD).

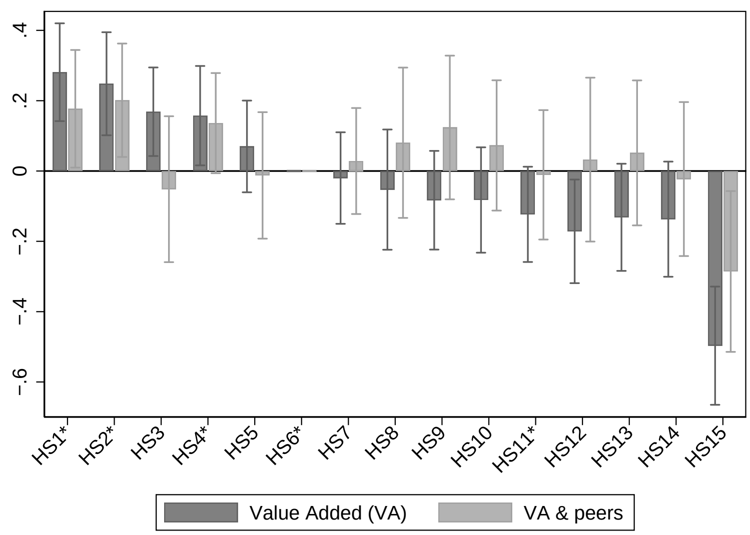
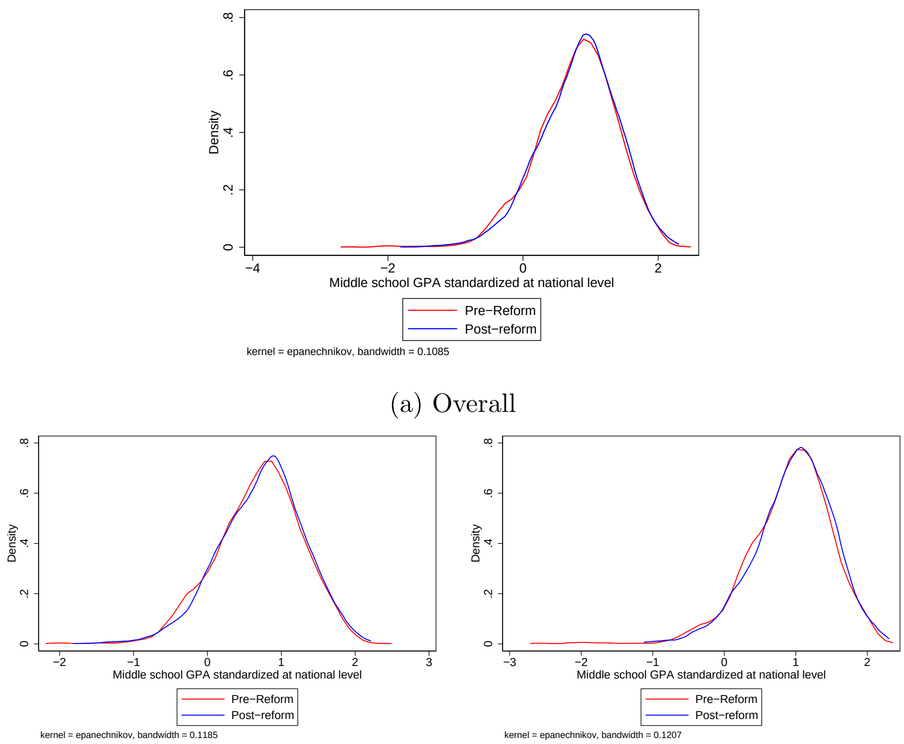


Figure A7: High school FE in Value-Added model for pre-reform cohorts.

Notes : The plotted estimates of the high school fixed effects (FE) parameters are derived from two different school Value Added (VA) models. The dark grey bars are OLS estimates of school FE from an individual-level regression of grades obtained at the end of high school (in the national exam in the Norwegian language) on school FE, middle school GPA (to have VA measures) and cohort FE (we include the three pre-reform cohorts of students). The light grey bars are estimates of the same parameters when we include as controls also the average characteristics of peers in the same school (the proportion of schoolmates' parents in different quartiles of the earnings distribution and the average education of schoolmates' parents). Given that we have multiple years available, we have variation in the quality of peers at the school level. The school FE in this second regression should be an estimate of the effect of the school beyond the impact of peers' composition. We have information on the grades in the national Norwegian exam for about 90% of the sample of pre-reform students. The lines around the bars are the 95% C.I. The starred schools (HS1, HS2, HS4, HS6 and HS11) are the competitive schools. Schools HS2 and HS4 are the most selective schools we define in Section VI. Baseline school for the school FE is HS6.



(a) Overall

(b) Low SES

(c) High SES

Figure A8: Robustness Checks for Middle School GPA Distribution.

Notes : In panel (a) we test whether the distribution of middle school GPA for cohorts pre- and post-reform in Bergen are statistically different, where the middle school GPA is normalized with respect to the national distribution of middle school GPA for the respective cohort (so that the position assigned to of each student reflect the relative position in the national middle school GPA distribution of their cohort). The p-value of the Kolmogorov–Smirnov test for the equality of the distributions is 0.238. In panels (b) and (c) we plot the same distribution, separately for individuals with at least one parent without a university degree (panel (b)) and for individuals whose parents (both) have a university degree (panel (c)). The p-value of the Kolmogorov–Smirnov test for the equality of the distributions is 0.364 (panel (b)) and 0.269 (panel (c)). Cohorts enrolling in high school in 2003 and 2004 are pre-reform, while those enrolling in 2005 and 2006 are post-reform.