When Excellence Backfires: The Impact of Superstars on Peers

September 10, 2024

Raul Mathias Leon
Center for The Economics of Human Development
and Harris School of Public Policy
University of Chicago

Rigissa Megalokonomou
Department of Economics,
Monash University
IZA and CESifo

Abstract

Recent literature shows that high-achieving peers mostly create positive externalities for their classmates. But how do these peer effects evolve with the use of stronger definitions that classify students with extremely high performance as *superstars*? We study the impact of being exposed to superstars in early high school on the remaining students' scholastic outcomes and university choices. We leverage the quasi-random assignment of students and teachers to classrooms in Greece to obtain identification. Using novel data from a representative sample of schools in Greece, we find that the presence of a superstar in the classroom is associated with lower educational outcomes for their classmates compared with classrooms without a superstar in the same school cohort. In particular, students exposed to superstars decreased their subsequent test scores by 0.06 standard deviations, possibly indicating that these students—including other high achievers—may become discouraged and reduce their effort. We also find that the adverse effects of superstars are significantly more pronounced for the top 3 to 5 students in the classroom. Students exposed to a superstar in a specific field are less likely to pursue that same study path themselves and tend to shift towards the alternative path. This suggests that classmates of superstars may reduce effort in the field in which the superstar excels and shift towards an alternative study or career path. Our effects are primarily driven by students in lower-ranked schools and schools located in low-income neighborhoods. These results suggest that classroom interventions that create externalities may be more impactful in disadvantaged environments.

JEL Codes: D5, I2

Keywords: superstars, peer effects, education outcomes, homophily, complementarity

[1]Raul Mathias Leon: Center for The Economics of Human Development and Harris School of Public Policy, University of Chicago (email:rleon@uchicago.edu), Rigissa Megalokonomou: Department of Economics, Monash Business School, CESifo, and IZA (email: rigissa.megalokonomou@monash.edu). The data that support the findings of this study are available from the Ministry of Education and School Authorities but we have no liberty to publish all data online.

1 Introduction

Many policies used by institutions in the education sector can drastically alter cohort compositions (Zimmerman, 2003). For example, affirmative action can enhance racial diversity, while selective admissions can increase intellectual diversity within cohorts. These policies, along with the status quo, can significantly impact students' educational achievements, choice of study, and matriculation rates. Conceptually, a Pareto-efficient allocation could exist that depends on the extent of peer-group externalities and the complementarity among students.

The first study to find evidence that peers and parents can have a greater impact than teachers or school characteristics on educational attainment is the Coleman Report (Coleman et al., 1966). Subsequent studies have supported this view, making it a conventional wisdom in the economics of education literature (Dearden et al., 2002; Carrell et al., 2009; Zimmerman, 2003; Carrell et al., 2013; Sacerdote, 2001). Research has increasingly focused on interactions between genders in the classroom (Lavy and Schlosser, 2011; Goulas et al., 2024a; Brenøe and Zölitz, 2020; Busso and Frisancho, 2021; Niederle and Vesterlund, 2007; Ahn et al., 2019; Zölitz and Feld, 2021), while other studies have explored the prevalence of peer effects within racial groups (Hoxby, 2000; Sacerdote, 2011; Fryer Jr and Torelli, 2010). More recent literature has examined the impact of class composition and high achievers on the tendency of women to stop studying STEM (science, technology, engineering, and math) degrees (Fischer, 2017; Card and Payne, 2021; Brenøe and Zölitz, 2020; Bostwick and Weinberg, 2022), as well as classroom peer effects of high and low-ability individuals (Busso and Frisancho, 2021; Cools et al., 2019; Lavy et al., 2012; Modena et al., 2022; Mouganie and Wang, 2020). Additionally, there is research on how teachers' gender role attitudes, effort and gender influence the gender gap across various outcomes (Rockoff, 2004; Lavy and Megalokonomou, 2024b; Carrell et al., 2010; Duflo et al., 2011; Bettinger and Long, 2005).

On the other hand, Brown (2011) identified a superstar effect, where the presence of a large ability gap in golf tournaments leads to reduced effort levels among players. However, the nature and channels of the superstar effect may differ significantly in a classroom environment, which is not solely a competitive space but also one of cooperation and gender role formation. This paper primarily addresses this question at the high school level. For instance, a strand of a literature suggest that personal rivalry or competition could be a mechanism driving peer effects, as it can result in increased effort and risk-taking independently of the context (Ager et al., 2022; Genakos and Pagliero, 2012; Gross, 2020). But what if the ability gap is too large? This is exactly why this study focuses on the impact of superstars, rather than general high achievers. For instance, the higher the ability of the superstar in a classroom, the harder it is to be ranked first. Creating incentives to reduce effort. This paper aims to provide descriptive evidence on the perverse effects of heterogeneous abilities and effort in a classroom environment by examining the superstar effect on the second-best student and its impact on the top 3 and 5 students. Our findings suggest that the superstar effect is more pronounced among high-achieving students.

Additionally, it builds upon existing research by exploring how superstar status, gender, and STEM outcomes interact in a school setting (Fischer, 2017; Mouganie and Wang, 2020; Goulas et al., 2024b). The literature suggests that this is important in our context, as women tend to avoid competition more than men and are more negatively affected by concerns about rank (Niederle and Vesterlund, 2010; Garratt et al., 2013; Tincani, 2017). To achieve this, high-achieving students will be classified into categories such as female, male, STEM, and Humanities superstars. This classification will help in studying the impact of different types of superstars on the short- and long-term scholastic outcomes of the rest of the students in the classroom separately by gender. Our findings suggest that the gender-based superstar distinction has significant implications for short-term outcomes such as test-scores, while the STEM-humanities distinction influences long-term outcomes.

Furthermore, while extensive research and policy efforts have addressed women's underrepresentation in STEM fields, there has been limited progress in reducing gender inequalities as measured by the number of baccalaureates awarded in recent decades. The situation is even more concerning when considering academic positions (Casad, Franks, Garasky, Kittleman, Roesler, Hall, and Petzel, 2021). Recent trends in test scores across developed countries suggest that sex differences are not purely biological, as evidenced by the narrowing and even reversal of the gender gap in math ability in some countries. This highlights the potential influence of strong environmental factors (Ceci, Ginther, Kahn, and Williams, 2014). There are two strands of the literature explaining this trend: 1) end of high school STEM preparedness and gender differences in college major choice (Payne and Asebedo, 2017; Aucejo and James, 2021) and 2) class composition and high achievers on the tendency of women to stop studying STEM degrees in university or STEM persistence (Fischer, 2017; Bostwick and Weinberg, 2022).

As a result, this paper will explore how exposure to a (fe)male, STEM or Humanities superstar can alter environmental influences within the classroom, potentially affecting students' high school track choice decisions and university STEM related choices. A plausible hypothesis is that students exposed to a STEM superstar might be discouraged from pursuing further studies in STEM. Conversely, exposure to a STEM superstar could provide a positive role model or benchmark that other students aspire to achieve. We find evidence supporting the former: students exposed to a STEM superstar are less likely to pursue STEM study paths, while those exposed to a Humanities superstar tend to shift away from Humanities and towards STEM. Our current results differ from the current state since we aim to integrate the two existing strands of literature by examining how class composition of superstars influence women's STEM *intake* decisions.

Lastly, our research contributes to the literature by addressing the effects of high-school classroom composition, an area with limited exploration due to the challenges of isolating causal impacts. This results from the fact that high-school classroom composition is often shaped by the influences of parents, teachers, and, most importantly, school administrators.

2 Institutional Details and Data

2.1 School and Classroom Assignment

In Greece, the high school academic year is divided into two semesters: semester 1, which starts at the beginning of September, and semester 2, which begins in early February of the following year. Each semester lasts 16 weeks. Students are assigned to a physical classroom where all lectures transpire.

This study leverages a unique system that takes advantage of the quasi-random assignment of students to classrooms in Greece. This system stems from the institutional practice of assigning students alphabetically or lexicographically at the start of 10th grade, as required by law. Balance tables will be provided to demonstrate that classroom assignment acted as a pseudo-randomization, successfully achieving balance and meeting the conditions for being considered as good as random. Moreover, teachers are randomly assigned to public schools, as required by law.

As a direct consequence of this law, students with last names starting with letters earlier in the alphabet are assigned to lower-numbered classrooms compared to students with last names that start with letters further down the alphabet. Once this classroom assignment is made, students cannot change classrooms. Greatly alleviating concerns of students self-selecting into classrooms.

It is important to note that students are assigned to public schools based on their residential address and proximity to a selected school. As a result, most students cannot change high schools after they have been assigned unless the family chooses to move or enroll in a private school. Since the data identifies students only at the school level, it is not possible to track students across different schools. However, this limitation is minimal because more than 92% of students in the Greek educational system attend public schools (Goulas, Megalokonomou, and Zhang, 2024a).

2.2 Data

We use panel data that consist of a sample of 134 high schools in Greece spanning the years 2001-2009. The sample corresponds to approximately 10% of public schools in Greece and 70,000 students (Goulas, Griselda, and Megalokonomou, 2022a). In addition, Goulas and Megalokonomou

(2021) has shown that the panel data used in this study is a representative sample of the entire Greek secondary educational system with regards to a myriad of important variables such as female share and track choice selection rates.

Each student has an individual identifier at the school level. Furthermore, there is information for all test scores across all high school grades (i.e., 10th-12th grade), detailed information on the student as well as some demographic information: age at 10th grade, gender, national exam performance, high school information (private vs. public), neighborhood income, track choice (STEM vs. Humanities), university admission, STEM application to university, etc.

There are five core or compulsory courses which are taken by all students in all grades. Of these, three are STEM-related: Algebra, Geometry, and Physics. Modern Greek and History are treated as Humanities. The test scores from these courses are observed throughout high school. The baseline control that measures the prior performance of a student before exposure to a superstar will be constructed using the 1st exam in Semester 1 of grade 10.

A limitation of using the baseline test scores is that they were not taken in the first few days of class before any type of interaction. Instead, they were taken within a two-month time frame. Their exact nature was a mid-semester exam in the 1st semester of grade 10. Nevertheless, one can still consider student performance in these early exams as a benchmark since they take place shortly after randomization. In addition, it is the measurement with the least interaction between peers and teachers at our disposal.

In Greece, test scores range from 0 to 20, but all grades used throughout the analysis will be transformed into z-scores to facilitate interpretation between subjects. These grades will be used to construct most of the subsequent control variables as well as for classifying students as superstars or not. For example, a student will be classified as a standard superstar if their average test score across the five core courses in the baseline exam is more than 2 standard deviations above the classroom mean. As mentioned before, we also have information regarding university admission, university application, and STEM application. [1]

2.3 High School Track and University Choices

In grade 11, students must choose a track based on their intended field of study at university. They can select from three tracks: Science, Information & Technology (IT), and Classics. For analytical

[1] The data defines a STEM degree as a degree offered at Science or Engineering and Technology departments. As such, degrees such as Economics or even Medicine are not considered. For the sake of precision, Science or Engineering and Technology departments include mathematics, engineering, physics, computer science, biology, chemistry, pharmacy, veterinary. Non-STEM university departments are those in liberal arts, literature, psychology, journalism, philosophy, education, Greek language, history, foreign languages, home economics, law, economics, statistics, business and management, accounting, political science, and European studies.

purposes, the Science and IT tracks are combined into a unified STEM track. Each track includes different elective courses. Students have the option to change tracks in grade 12, although this is quite rare.

Grade 12 is the final year of high school and is particularly important for university admissions. During this year, students must take school-level exams as well as a national exam that includes both track-specific subjects and an elective exam. Track-related subjects have a greater weight in the calculation of the admission grade. The exams are administered by the Ministry of Education between late May and early June, are blind, and are graded centrally. The national exam performance, or admission grade, is the sole criterion used for university admissions.

After the national exam results are released, students must create a ranked list of their preferred university degrees. The Ministry of Education then assigns top-ranking students to their preferred choices. This allocation process continues for lower-ranking students based on availability of spaces.

2.4 Defining a Superstar

Let i be an individual. Let $\mu_{s,c,t=10}$ and $\sigma_{s,c,t=10}$ be the average and standard deviation of the baseline test scores of students within a given class at each school. The standard superstar variable in the analysis is defined as an indicator function taking the value of 1 if a student has obtained average test scores over 2 standard deviations above the mean across all core courses in the baseline exam, and 0 otherwise. Specifically, observation i will be classified as a superstar if the following indicator variable takes the value of 1:

$$\text{standard\_superstar}_i = \mathbb{1}\{\text{average test scores of } i \text{ are over } \mu_{s,c,t=10} + 2\sigma_{s,c,t=10}\}$$

As shown in Figure 2, approximately 1% of students in the sample are categorized as superstars. This implies that about 15% of classrooms in the sample are exposed to a superstar.

The superstar group can be further subdivided by gender. A student is classified as a female superstar if she is female and meets the superstar criteria. Likewise, a student is classified as a male superstar if he is male and meets the superstar criteria. Figure A2 shows the dynamics by gender.

Additionally, students are classified as STEM superstars if their average test scores in STEM core subjects exceed two standard deviations above the mean, and as Humanities superstars if their scores in Humanities core subjects meet this threshold. Figure A1 in the appendix shows the dynamics of this other type of superstars over the years. We also define STEM supergirls and Humanities supergirls for female students who meet these criteria, and STEM superboys and Humanities superboys for male students who meet these criteria.

3 Empirical Strategy

3.1 Measuring the Impact of a Superstar on the Average Student

Since students are randomly assigned to superstars within school and classes, we structure the empirical design around this randomization. Therefore, the specification we use in a regression model includes a school-by-year fixed effect. Specifically, the description of the problem suggests the following initial econometric specification:

$$y_{icst} = \beta_{s,t} + X_i^T \delta + S_{cst}^T \pi + \omega D_{cst} + \varepsilon_{icst} \tag{1}$$

where y_{icst} is is the short- and longer-term scholastic outcomes for student i, in class c, in school s, in time t. These are a student's test scores in the final exams in grade 10, 11, 12 and university admission outcomes. The end-of-grade student performance is measuring overall performance across all subjects, only STEM, or only Humanities subjects. The longer-term outcomes include average test scores in grades 11 and 12, whether the student pursued a STEM or Humanities track in grade 11, national exam performance, university admission, admission to a top 10 university, application to STEM or Humanities programs, and admission to STEM or Humanities fields. $\beta_{s,t}$ is a school by year fixed effect; X_i is a matrix of student level controls; S_{cst}^T is a matrix of class level controls; D_{cst} is a binary indicator that takes the value of 1 if class c in school s in cohort t has a superstar and 0 otherwise; ε_{icst} is the error term that allows for any type of correlation within observations of the same school, time, individuals and class. Individual controls include whether a student is female, age in grade 10, whether they were born in the first quarter of the year, average overall test scores at baseline, and specific test scores at baseline if applicable. Class controls include the proportion of females, classroom size, and the leave-one-out mean overall test scores for the classroom. It is crucial to include school-by-year fixed effects to control for the endogenous sorting of students across schools in a given year.

The coefficient of interest is ω since it captures the effect of being exposed to a superstar on the students' outcomes compared to that of students in other classrooms with the same school cohort. The basic idea is to compare the test scores and study choices of students who are exposed to the same school environment and have similar characteristics, except for the fact that one classroom has a student whose academic performance really stand out compared with other classes in the same school-cohort for idiosyncratic reasons. We exclude superstars from the analysis, as the focus is on examining the impact of superstar exposure on their classmates. We then consider different types of superstars and examine their effects on the remaining students. Those include STEM superstar, Humanities superstar, Female Superstar and Male Superstar. Standard errors are clustered at the school by year level to allow for heteroskedasticity and serial correlation in the outcomes of students in the same classroom. This relies on work by Abadie, Athey, Imbens, and Wooldridge (2017), which suggests clustering at a higher level of aggregation than that of the randomization (in this case class-level) when restricted to finite samples.

3.2 Measuring the Impact of a Superstar on Top Performing Students

Another question of interest was the existence of the perverse effect of heterogeneous abilities in effort in a classroom environment. Thus, it is required to check the effect of a superstar on the second-best student in the classroom as well as its impact on the top 3 and 5 students as a whole. For clarification, the second-best student has been defined as either the best student after the superstar in a classroom with superstar or the best student in a classroom that hasn't been exposed. The top 3 and 5 students have been identified using the same logic.

For this we use two different methodologies. To measure the impact on the top 3 and 5 students as a whole, we need to adapt equation 1 by adding a top 3 and top 5 student dummy variable and its interaction with D_{cst}, accordingly. As a result 1 is modified into:

$$y_{icst} = \beta_{s,t} + X_i^T\delta + S_{cst}^T\pi + \omega D_{cst} + \pi T_{icst} + \rho D_{cst} \times T_{icst} + \varepsilon_{icst} \tag{2}$$

where T_{icst} is a dummy variable that takes the value of 1 if the student i is a top student in class c in school s in time t. It is important to notice that $\omega + \rho$ captures the effect of interest i.e. the superstar effect on a top student. In contrast, π just captures the average difference in achievement between a top student and non-top student. Finally, ρ just captures the difference in the superstar effect between a top and a non-top student.

To measure the exposure of a superstar on the second best we just take specification 1 and run a regression conditioned on the data points being the second best in the classroom (within-second best students empirical strategy). We didn't run specification 2 because in each class only one individual is defined as the second-best.

3.3 Evidence on the Quasi-random Assignment of Students to Classrooms

The estimation of peer effects is challenging, typically due to reflection bias, selection bias, and unobserved correlated effects (Brock and Durlauf, 2001; Manski, 1993). In our setting, high school students who attend the same school are assigned to classrooms based on their surnames in lexicographical order. This means that students with a surname starting with a letter earlier in the alphabet are given a smaller classroom number than those with a surname starting with a letter later in the alphabet. As such, the process of student allocation can be considered as good as random. Consequently, students cannot choose their classes in high school, which helps to alleviate concerns about selection bias (Heckman, 1990).[2]

To provide evidence that the lexicographic assignment to classrooms can be considered as good as random, we examine whether it successfully achieved balance. Specifically, in Table 2 we regress a binary indicator variable for having a Superstar in the classroom on each of the predetermined student characteristics and baseline performance. In columns (1)-(3), we regress the treatment variable separately on each of the student characteristics, and in column (4), we regress the treatment on all student-level characteristics simultaneously to capture any correlated effects. In A2 we do it separately for male and female students.

In addition, Table A3, in which we regress predetermined controls on the class number, provides evidence that the alphabetical assignment of students to classrooms is practically random. This shows that classroom number is not systematically associated with differences in students' baseline performance in core courses, the proportion of females in the classroom, average age, or the likelihood of being exposed to a Superstar. Due to multicollinearity, classroom number 6 is omitted from the regression and should be understood as the point of comparison.

4 Main Results

4.1 Short-run Impact

Table 3 presents the main result by using specification 1. The outcomes variable in columns (1-3) are a student's performance in all five core courses, STEM and Humanities at the end of grade 10, respectively. We include individual controls (baseline overall performance, baseline performance in the same subject, a student's age, gender, and a binary indicator for whether the student was born in the first quarter of the calendar year), school-by-year fixed effects as well as class controls (class size, leave-out proportion of females in the class, class's leave- out mean baseline performance in the same subject, class's leave-out mean age and class's leave-out share of students who were born in the first quarter). These variables absorb any unobserved heterogeneity at class level. Finally, the outcome is the standardized performance —that is, the grade transformed into z-scores to facilitate interpretation. Both coefficients are negative and of medium magnitude. For instance,

[2] Lavy and Megalokonomou (2024b); Dinerstein, Megalokonomou, and Yannelis (2022); Goulas, Griselda, and Megalokonomou (2022b,0); Goulas, Gunawardena, and Megalokonomou (2023); **?**); Goulas, Megalokonomou, and Zhang (2024b) also provide evidence that students are randomly assigned to classrooms within school cohorts in Greece. Evidence of the random assignment of teachers to classrooms in the same context can be found in Lavy and Megalokonomou (2024b) and Lavy and Megalokonomou (2024a).

students exposed to superstars' test scores in STEM decreased by 0.059 standard deviations in their end-of-the-year exam. Table A4 presents the results for individual courses.

While 3 provides baseline results for a distance of 2 standard deviations from the class mean, we further analyze smaller distances, such as 1.8, 1.6, 1.4, 1.2, and 1 standard deviation in Table A5. This approach helps us determine the point at which the results cease to be statistically significant, which occurs at 1 standard deviation. In the table, we include the total number of superstars and the share of students exposed at these new cutoff distances.

In order to estimate how the female and male populations each separately respond to being exposed to a superstar, we estimate several within-gender regressions. Table A6 presents the within-gender regressions of Table 3. The results remain similar, in terms of sign, across both genders.

Table A14 presents the estimates of exposure to multiple superstars. For this, we include two dummy variables: one representing classes exposed to a single superstar and another for classes exposed to 2-4 superstars. Finally, A7 replicates the analysis from 3, focusing exclusively on cohorts that include a superstar. The results remain consistent.

4.2 Long-term: University and Career Choices

Table 4 presents long-term outcomes such as final-year grades in the last year of high school, national exam performance, and various university-related outcomes. The negative effect on test scores persists and is also significant for national exam performance. The reduction of the national exam performance is intimately related with the university quality. Nonetheless, being exposed to a superstar does not appear to significantly influence the decision to pursue the STEM track in grade 11. There also appears to be an impact on the likelihood of applying to a STEM degree at university. For instance, on average, exposure to a superstar decreases the probability of applying to a STEM program by approximately 0.02 percentage points.

5 Mechanisms

5.1 Effects on the Top Students in the Classroom

Table 5 presents the main results using specification 2 for grade 10. Panel A does it for the top 5 students in the classroom, while Panel B does it for the top 3. The second estimate in each column is the interaction term and captures the difference in the superstar effect between a top student and a non-top student. The third estimate shows the sum of the first two coefficients, calculated using the delta method, to illustrate the superstar effect on a top student. We find that the effect on top students is significantly larger. The estimate of the superstar effect on the average student is attenuated, further supporting the presence of strong non-linearities. Table A8 reproduces the results from Table 5 for eleventh grade with similar findings. While Table A9 presents the estimates of the superstar effect on the second-highest performing student across all three years of high school.

5.2 By Type of Superstar

Table 7 shows the effect of being exposed to a variety of superstars. Girl superstars appear to drive the STEM effects in test scores, while boy superstars influence the humanities effects. Table A10 show the same exercise for long term outcomes. In the long run, it appears that STEM and Humanities superstars play the most significant roles. For example, the presence of a STEM superstar has a negative effect on the likelihood of students pursuing a STEM track, applying to STEM programs, and being admitted to STEM fields, while it positively influences the Humanities track. Conversely, exposure to a Humanities superstar has the opposite effect, positively impacting the STEM track and admissions but negatively affecting the Humanities track, admissions, and applications—a sort of crowding-out effect. Neither type of superstar improves the chances of attending university.

5.3 Gender-specific Homophily

Table 6 shows that female superstars are contributing to the negative effects in STEM courses across both genders. Similarly, it shows a similar pattern for male superstars, who affect Humanities across both genders. Table A15 and A17 presents the effects of female Humanities and STEM superstars, while A16 and A18 shows the effects of male Humanities and STEM superstars, respectively. The most interesting case is the impact of a Female Humanities Superstar, which decreases the end of grade 10 final performance but increases the the likelihood of students pursuing the STEM track in grade 11.

6 Heterogeneous Effects

We will now test for heterogeneous effects based on school quality and neighborhood income to examine whether exposure to a superstar impacts students differently across varying competitive environments and socioeconomic backgrounds, which are correlated with effort levels and parental investment. Table A11 presents the results, which suggest that the negative superstar effects observed previously are more pronounced in low-income neighborhoods and low-ranking schools.

This pattern may be explained by models of neighborhood choice, where students sort into selected neighbourhoods and schools to be able to interact with complementary peers (Eshaghnia et al., 2023).

In addition, Table A12 presents the results for schools with a low historical number of superstars compared to schools with many superstars. Although all the estimates remain negative, there is no clear pattern of significance. Lastly, A13 presents the analysis separately for students whose baseline performance is above and below the median

7 Falsification Exercise

To support the identification strategy and further validate the causal interpretation of our baseline results, we ran regressions with placebo treatments to confirm that our results are not driven by random chance or confounding factors. The general idea is that the falsification exercise should not yield results similar to the main findings.

7.1 Superstar Reassigned to Another Classroom within the School Cohort

In this falsification exercise, we randomly reassign superstar students to different classrooms within the same school cohort. Across 100 simulations, the resulting estimates are negligible and statistically insignificant (see Figure 3). This indicates that the superstar effect is not due to spurious correlations between superstar exposure and other factors within the school cohort. Instead, the results reflect genuine social interactions between students and their actual superstar classmates.

7.2 Measurement error

Building on Goulas, Gunawardena, and Megalokonomou (2023), we introduce three types of noise into the baseline performance used to identify superstars in each class: additive noise, which uniformly affects all individuals' baseline performance; multiplicative noise, whose impact varies with the level of baseline performance; and impulse noise, which introduces sudden random jumps to the baseline test scores of high-achieving male students. These distortions may result in misidentifying a student as the top performer, even though they are not truly the highest achiever. As such, these students, are less likely to influence their classmates in the same way as the actual superstar.

For instance, one may be concerned that we might not be correctly identifying male superstars, given evidence that males often exert less effort on exams unless the stakes are high (Montolio and

Taberner, 2021). Since the midterm exam (our baseline) is at most medium-stakes, we add some noise to the performance data of certain males to assess the robustness of our results. Introducing noise into the baseline performance of high-achieving male students could result in identifying a high-performing male student as a superstar, even though they may not truly be one. These high-performing students have a smaller ability gap, and based on the literature, we would not expect them to exert the same negative effect as an actual superstar. To investigate this, we perform simulations by adding impulse noise—sudden random jumps—to the baseline test scores of high-achieving male students (those 1 standard deviations above the mean) and re-estimate 1. After introducing noise into the simulated baseline scores, we identify a new group of superstar students. Consistent with our main analysis, we exclude the new superstars and focus on the effect of superstar exposure on the remaining students in the class. The simulation results in Figure A5 show that the superstar effect sign changes to positive, providing evidence that we are successfully distinguishing between male high-achievers and male superstars.

Figure A3 presents the simulated estimated effects under additive noise. As measurement error increases, we observe a downward attenuation bias, indicating that even small additional errors can significantly influence the results. Figure A4 replicates this analysis with multiplicative noise, where the noise has an even greater impact, substantially attenuating all estimates.

Overall, the patterns observed across the Monte Carlo simulations further suggest that our estimated effects accurately capture the true superstar effect, free from confounding factors.

8 Conclusions

Using data from 134 Greek high schools from 2001 to 2009, this study provides novel evidence on the unintended consequences of exposing students to superstars within classroom environments by exploiting the quasi-random assignment of teachers and students to classes. Contrary to conventional wisdom, which typically associates high-achieving peers with positive spillover effects, our findings suggest that superstars can have a detrimental impact on the academic outcomes of their classmates. Exposure to superstars resulted in a 0.06 standard deviation decrease in test scores for peers, indicating that a significant ability gap might lead to discouragement and reduced effort. The adverse effects are notably more pronounced for the top-performing students in the classroom.

Furthermore, we observe distinct long-term effects based on the superstar's academic focus. Students exposed to STEM superstars were less likely to pursue STEM pathways, while those exposed to Humanities superstars shifted toward STEM fields, suggesting that superstars can influence academic choices in unanticipated ways. These effects were more prevalent in lower-ranked schools and schools in economically disadvantaged areas, indicating that family, context and school environment play crucial roles in how peer effects manifest.

Future research could explore the effects of superstars in more competitive environments, such as top universities, and examine how superstars impact other superstars. Especially concerning creativity outcomes (Gross, 2020). Additionally, studying how superstars influence teaching effort and curricular difficulty may uncover indirect channels contributing to the superstar effect (Duflo, Dupas, and Kremer, 2011). Our findings suggest policy interventions, such as establishing schools specifically designed for gifted students, particularly for students in low-income neighborhoods. In these specialized schools, gifted students would not only minimize their impact on peers but also have an environment to flourish.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017). When Should You Adjust Standard Errors for Clustering? Technical report, National Bureau of Economic Research.

Ager, P., L. Bursztyn, L. Leucht, and H.-J. Voth (2022). Killer incentives: Rivalry, Performance and Risk-Taking Among German Fighter Pilots, 1939–45. *The Review of Economic Studies 89*(5), 2257–2292.

Ahn, T., P. Arcidiacono, A. Hopson, and J. R. Thomas (2019). Equilibrium Grade Inflation with Implications for Female Interest in STEM Majors. Technical report, National Bureau of Economic Research.

Aucejo, E. and J. James (2021). The Path to College Education: The Role of Math and Verbal Skills. *Journal of Political Economy 129*(10), 2905–2946.

Bettinger, E. P. and B. T. Long (2005). Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students. *American Economic Review 95*(2), 152–157.

Bostwick, V. K. and B. A. Weinberg (2022). Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs. *Journal of Labor Economics 40*(2), 397–436.

Brenøe, A. A. and U. Zölitz (2020). Exposure to More Female Peers Widens the Gender Gap in STEM Participation. *Journal of Labor Economics 38*(4), 1009–1054.

Brock, W. A. and S. N. Durlauf (2001). Interactions-Based Models. In *Handbook of econometrics*, Volume 5, pp. 3297–3380. Elsevier.

Brown, J. (2011). Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy 119*(5), 982–1013.

Busso, M. and V. Frisancho (2021). Good Peers have Asymmetric Gendered Effects on Female Educational Outcomes: Experimental Evidence from Mexico. *Journal of Economic Behavior & Organization 189*, 727–747.

Card, D. and A. A. Payne (2021). High School Choices and The Gender Gap in STEM. *Economic Inquiry 59*(1), 9–28.

Carrell, S. E., R. L. Fullerton, and J. E. West (2009). Does your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics 27*(3), 439–464.

Carrell, S. E., M. E. Page, and J. E. West (2010). Sex and science: How Professor Gender Perpetuates the Gender Gap. *The Quarterly Journal of Economics 125*(3), 1101–1144.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica 81*(3), 855–882.

Casad, B. J., J. E. Franks, C. E. Garasky, M. M. Kittleman, A. C. Roesler, D. Y. Hall, and Z. W. Petzel (2021). Gender Inequality in Academia: Problems and Solutions for Women Faculty in STEM. *Journal of Neuroscience Research 99*(1), 13–23.

Ceci, S. J., D. K. Ginther, S. Kahn, and W. M. Williams (2014). Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest 15*(3), 75–141.

Coleman, J. S., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld, and R. York (1966). The Coleman Report. *Equality of Educational Opportunity 3*(18), 1–32.

Cools, A., R. Fernández, and E. Patacchini (2019). Girls, Boys, and High Achievers. Technical report, National Bureau of Economic Research.

Dearden, L., J. Ferri, and C. Meghir (2002). The Effect of School Quality on Educational Attainment and Wages. *Review of Economics and Statistics 84*(1), 1–20.

Dinerstein, M., R. Megalokonomou, and C. Yannelis (2022, November). Human Capital Depreciation and Returns to Experience, journal = American Economic Review. *112*(11), 3725–62.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence From a Randomized Evaluation in Kenya. *American Economic Review 101*(5), 1739–1774.

Eshaghnia, S., J. J. Heckman, and G. Razavi (2023). Pricing Neighborhoods. Technical report, National Bureau of Economic Research.

Fischer, S. (2017). The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence. *Labour Economics 46*, 211–226.

Fryer Jr, R. G. and P. Torelli (2010). An Empirical Analysis of 'Acting White'. *Journal of Public Economics 94*(5-6), 380–396.

Garratt, R. J., C. Weinberger, and N. Johnson (2013). The State Street Mile: Age and Gender Differences in Competition Aversion in the Field. *Economic Inquiry 51*(1), 806–815.

Genakos, C. and M. Pagliero (2012). Interim Rank, Risk taking, and Performance in Dynamic Tournaments. *Journal of Political Economy 120*(4), 782–813.

Goulas, S., S. Griselda, and R. Megalokonomou (2022a). Comparative Advantage and Gender Gap in STEM. *Journal of Human Resources*.

Goulas, S., S. Griselda, and R. Megalokonomou (2022b). Comparative Advantage and Gender Gap in STEM. *Journal of Human Resources 0320-10781R2*.

Goulas, S., S. Griselda, and R. Megalokonomou (2023). Compulsory class attendance versus autonomy. *Journal of Economic Behavior & Organization 212*, 935–981.

Goulas, S., B. Gunawardena, and R. Megalokonomou (2023). Same-Gender Peer Role Models in Education. *Mimeo*.

Goulas, S. and R. Megalokonomou (2021). Knowing who you Actually Are: The Effect of Feedback on Short- and Longer-term Outcomes. *Journal of Economic Behavior & Organization 183*, 589–615.

Goulas, S., R. Megalokonomou, and Y. Zhang (2024a). Female Neighbors and Careers in Science. *Available at SSRN 4766917*.

Goulas, S., R. Megalokonomou, and Y. Zhang (2024b). Females, Classroom Disruption and STEM Careers. *Working Paper*.

Gross, D. P. (2020). Creativity Under Fire: The Effects of Competition On Creative Production. *Review of Economics and Statistics 102*(3), 583–599.

Heckman, J. J. (1990). Selection Bias and Self-Selection. In *Econometrics*, pp. 201–224. Springer.

Hoxby, C. M. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation.

Lavy, V. and R. Megalokonomou (2024a). Alternative Measures of Teachers' Value Added and Impact on Short and Long-Term Outcomes: Evidence From Random Assignment. *NBER Working Paper, No. 32671*.

Lavy, V. and R. Megalokonomou (2024b, April). The short- and the long-run impact of gender-biased teachers. *American Economic Journal: Applied Economics 16*(2), 176–218.

Lavy, V., M. D. Paserman, and A. Schlosser (2012). Inside the Black Box of Ability peer effects: Evidence from Variation in The Proportion of Low Achievers in The Classroom. *The Economic Journal 122*(559), 208–237.

Lavy, V. and A. Schlosser (2011). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics 3*(2), 1–33.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies 60*(3), 531–542.

Modena, F., E. Rettore, and G. M. Tanzi (2022). Asymmetries in The Gender Effect of High-Performing Peers: Evidence From Tertiary Education. *Labour Economics 78*, 102225.

Montolio, D. and P. A. Taberner (2021). Gender Differences Under Test Pressure and Their Impact on Academic Performance: A Quasi-Experimental Design. *Journal of Economic Behavior & Organization 191*, 1065–1090.

Mouganie, P. and Y. Wang (2020). High-Performing Peers and Female STEM choices in School. *Journal of Labor Economics 38*(3), 805–841.

Niederle, M. and L. Vesterlund (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics 122*(3), 1067–1101.

Niederle, M. and L. Vesterlund (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of economic perspectives 24*(2), 129–144.

Payne, P. and S. Asebedo (2017). Financial Self-Efficacy and the Financial Satisfaction of Credit-Card Users. In *2018 Academic Research Colloquium for Financial Planning and Related Disciplines*.

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence From Panel Data. *American economic review 94*(2), 247–252.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics 116*(2), 681–704.

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, HOW Big are They and How Much Do We Know Thus Far? In *Handbook of the Economics of Education*, Volume 3, pp. 249–277. Elsevier.

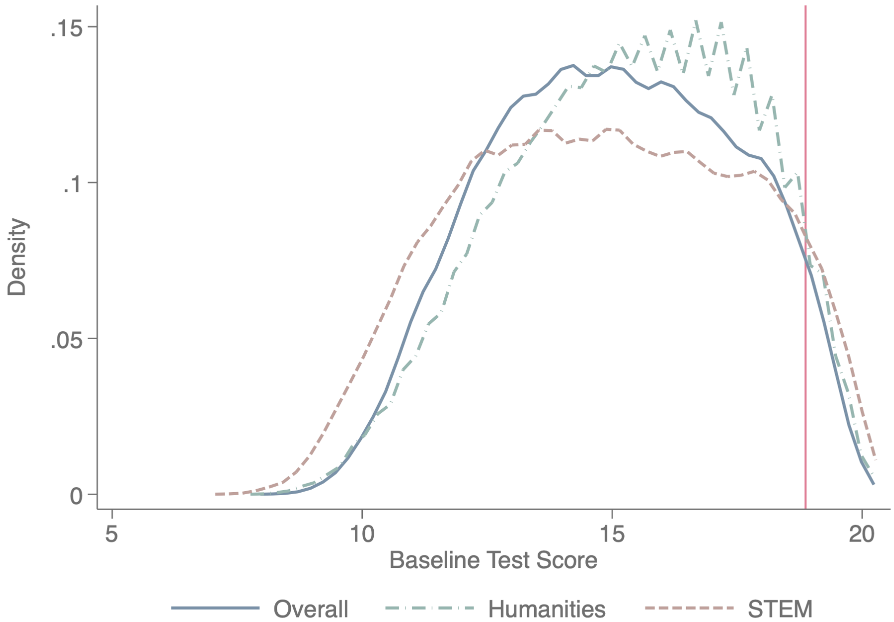
Tincani, M. (2017). Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence.

Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence From a Natural Experiment. *Review of Economics and statistics 85*(1), 9–23.

Zölitz, U. and J. Feld (2021). The Effect of Peer Gender on Major Choice in Business School. *Management Science 67*(11), 6963–6979.

Main Figures

Figure 1: Distribution of Baseline Test Scores



Note: The red vertical line indicates the average overall performance of superstars (i.e. 18.86).

Figure 2: Superstar Dynamics Over Time

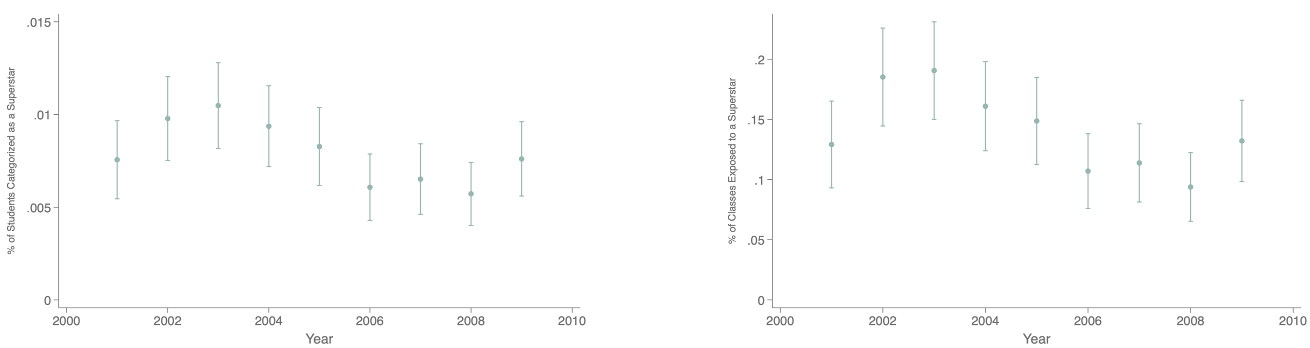
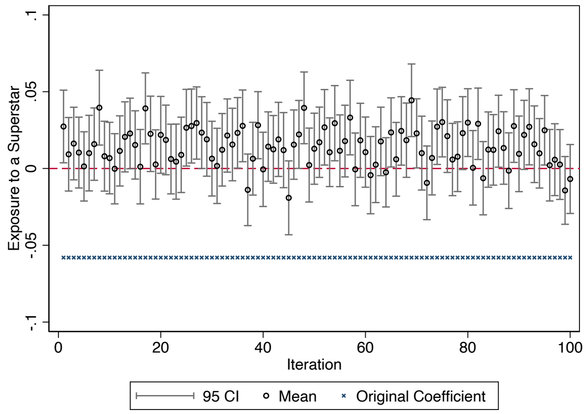
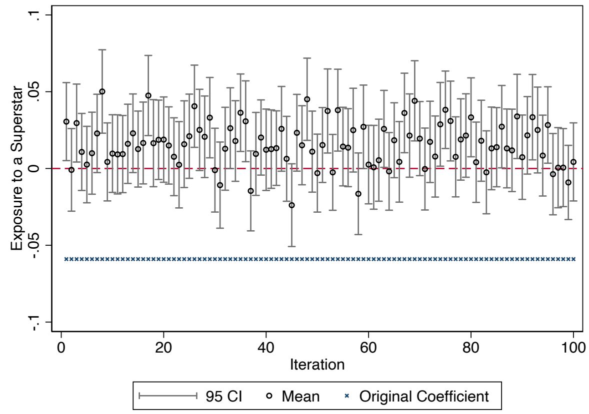


Figure 3: Superstar Reassigned to Another Classroom within the School Cohort

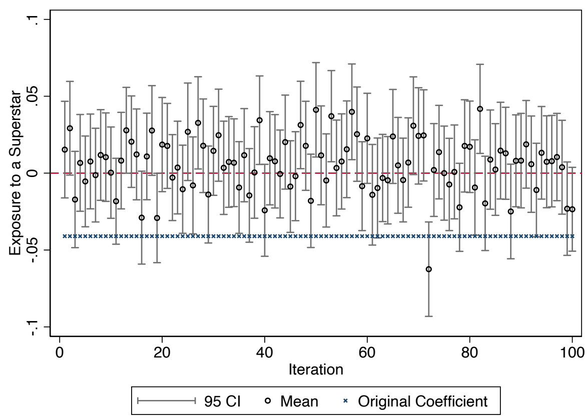
(a) Overall



(b) STEM



(c) Humanities



Main Tables

Table 1: Summary Statistics

	Mean	SD	Min	Max	N
Pre-treatment Characteristics					
Test Scores in Core Courses (Baseline)	15.02	2.40	8.20	20	64,691
Test Scores in STEM (Baseline)	14.80	2.76	7.33	20	64,691
Test Scores in Humanities (Baseline)	15.35	2.39	8	20	64,691
Female	0.56	0.50	0	1	64,691
Age at Grade 10	15.93	0.44	15	25	64,691
Born in 1st Quarter	0.12	0.33	0	1	64,691
Outcomes					
Test Scores in Core Courses (End of Grade 10)	11.88	3.95	0	20	64,691
Test Scores in STEM (End of Grade 10)	10.77	4.88	0	20	64,691
Test Scores in Humanities (End of Grade 10)	13.56	3.48	0	20	64,691
STEM Track in Grade 11	0.63	0.48	0	1	64,691
Humanities Track in Grade 11	0.37	0.48	0	1	64,691
National Exam Performance	13.06	3.93	1.17	19.84	56,246
Admitted to a University	0.83	0.38	0	1	56,246
STEM Application	0.62	0.49	0	1	56,246
STEM Admitted	0.11	0.31	0	1	46,571
Humanities Application	0.31	0.46	0	1	64,691
Humanities Admitted	0.28	0.45	0	1	46,571
Variable of Interest					
Superstar	0.01	0.09	0	1	64,691
Exposure to a Superstar	0.14	0.35	0	1	64,691

The table reports descriptive statistics for student pre-treatment characteristics and outcomes. We report those statistics for the full sample.

Table 2: Balance Test at the Individual Level

	In a Classroom with a Superstar				
	(1)	(2)	(3)	(4)	(5)
Test Scores in STEM (Baseline)	0.000				−0.000
	(0.000)				(0.000)
Test Scores in Humanities (Baseline)		0.000			0.000
		(0.000)			(0.000)
Age at Grade 10			−0.000		−0.000
			(0.000)		(0.000)
Born in 1st Quarter				−0.000	−0.000
				(0.000)	(0.000)
Mean	0.144	0.144	0.144	0.144	0.144
Observations	64691	64691	64691	64691	64691
F-Statistic	33.496	28.773	28.966	33.496	25.353
P-value for joint significance	0.000	0.000	0.000	0.000	0.000
School by year FE	✓	✓	✓	✓	✓
Class Controls	✓	✓	✓	✓	✓

Notes: The treatment variable is regressed on each control variable in a separate regression. In column (5) we include all control variables simultaneously in the regression and report the joint significance of those variables. Each estimate is generated from a different regression. All grades presented have been standardized. Class controls include the proportion of female peers, the class's leave-out mean in baseline test scores in STEM and humanities, the number of students in the class, leave-out mean age, leave-out percentage of students born in the first quarter. Robust standard errors clustered at the school by year level are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table 3: Estimates of the Effect of a Superstar

	Grade 10: Avg Test Scores in Final Exam		
	(1)	(2)	(3)
	Overall	STEM	Humanities
Exposure to a Superstar	-0.058^{***}	-0.059^{***}	-0.041^{***}
	(0.012)	(0.014)	(0.015)
Observations	64177	64177	64177
Adjusted R-squared	0.781	0.731	0.688
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table 4: Estimates of the Effect of a Superstar on Long-Term Outcomes

	Grade 11: Avg. Test Scores			Grade 12: Avg. Test Scores		
	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	STEM	Humanities	Overall	STEM	Humanities
Panel A:						
Exposure to a Superstar	-0.075^{***}	-0.065^{***}	-0.079^{***}	-0.056^{***}	-0.048^{***}	-0.067^{***}
	(0.013)	(0.014)	(0.016)	(0.013)	(0.014)	(0.018)
Observations	64177	64177	64177	64177	64177	64177
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
	STEM Track	National Exam	Admitted to University	Admitted to Top 10	Applied to STEM	Admitted to STEM
Panel B:						
Exposure to a Superstar	-0.007	-0.172^{***}	0.006	-0.014^{***}	-0.018^{**}	-0.003
	(0.007)	(0.053)	(0.006)	(0.003)	(0.008)	(0.005)
Observations	64177	55785	55785	64177	55785	46114
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table 5: Estimates of the Effect of a Superstar on a Top Student: Year 10

	Grade 10: Test Score in Final Exam		
	(1)	(2)	(3)
	Overall	STEM	Humanities
Panel A: Top 5			
Exposure to a Superstar	-0.038^{***}	-0.037^{**}	-0.030^{*}
	(0.013)	(0.014)	(0.017)
Exposure to a Superstar x Top 5	-0.111^{***}	-0.122^{***}	-0.065^{***}
	(0.013)	(0.014)	(0.015)
Effect on a Top Student	-0.149^{***}	-0.159^{***}	-0.095^{***}
	(0.016)	(0.018)	(0.017)
Observations	64177	64177	64177
School by year FE	✓	✓	✓
Controls	✓	✓	✓
Panel B: Top 3			
Exposure to a Superstar	-0.050^{***}	-0.050^{***}	-0.035^{**}
	(0.013)	(0.014)	(0.016)
Exposure to a Superstar x Top 3	-0.092^{***}	-0.099^{***}	-0.064^{***}
	(0.015)	(0.017)	(0.017)
Effect on a Top Student	-0.142^{***}	-0.149^{***}	-0.099^{***}
	(0.017)	(0.020)	(0.018)
Observations	64177	64177	64177
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course.

Table 6: Estimates of the Effect of a Girl and Boy Superstar within-gender

	Grade 10: Test Score in Final Exam			Grade 11:
	(1)	(2)	(3)	(4)
	Overall	STEM	Humanities	STEM track
Panel A: All				
Exposure to a Girl Superstar	−0.073***	−0.071***	−0.024	−0.007
	(0.016)	(0.017)	(0.018)	(0.008)
Exposure to a Boy Superstar	−0.055***	−0.009	−0.085***	−0.002
	(0.019)	(0.019)	(0.024)	(0.010)
School by year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Panel B: Within-Female				
Exposure to a Girl Superstar	−0.069***	−0.068***	−0.031	−0.003
	(0.019)	(0.021)	(0.021)	(0.012)
Exposure to a Boy Superstar	−0.058***	−0.002	−0.065**	−0.000
	(0.021)	(0.023)	(0.029)	(0.014)
School by year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Panel C: Within-Male				
Exposure to a Girl Superstar	−0.074***	−0.067***	−0.008	−0.010
	(0.019)	(0.019)	(0.022)	(0.011)
Exposure to a Boy Superstar	−0.052**	−0.018	−0.104***	−0.001
	(0.024)	(0.024)	(0.028)	(0.015)
School by year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓

Notes: The Supergirl and Superboy effects come from separate regressions. Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table 7: Estimates of the Effect of a Superstar by Type

	Grade 10: Test Score in Final Exam									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	STEM	Humanities	STEM	Humanities	STEM	Humanities	STEM	Humanities	STEM	Humanities
Exposure to a Superstar	−0.059***	−0.041***								
	(0.014)	(0.015)								
Exposure to a STEM Superstar			−0.038**	−0.028*						
			(0.015)	(0.016)						
Exposure to a Humanities Superstar					−0.079***	−0.031*				
					(0.016)	(0.018)				
Exposure to a Girl Superstar							−0.071***	−0.024		
							(0.017)	(0.018)		
Exposure to a Boy Superstar									−0.009	−0.085***
									(0.019)	(0.024)
Observations	64177	64177	64236	64236	64364	64364	64345	64345	64523	64523
Adjusted R-squared	0.731	0.688	0.731	0.689	0.734	0.689	0.733	0.689	0.734	0.692
School by year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Appendix:

Figure A1: Superstar Dynamics Over Time

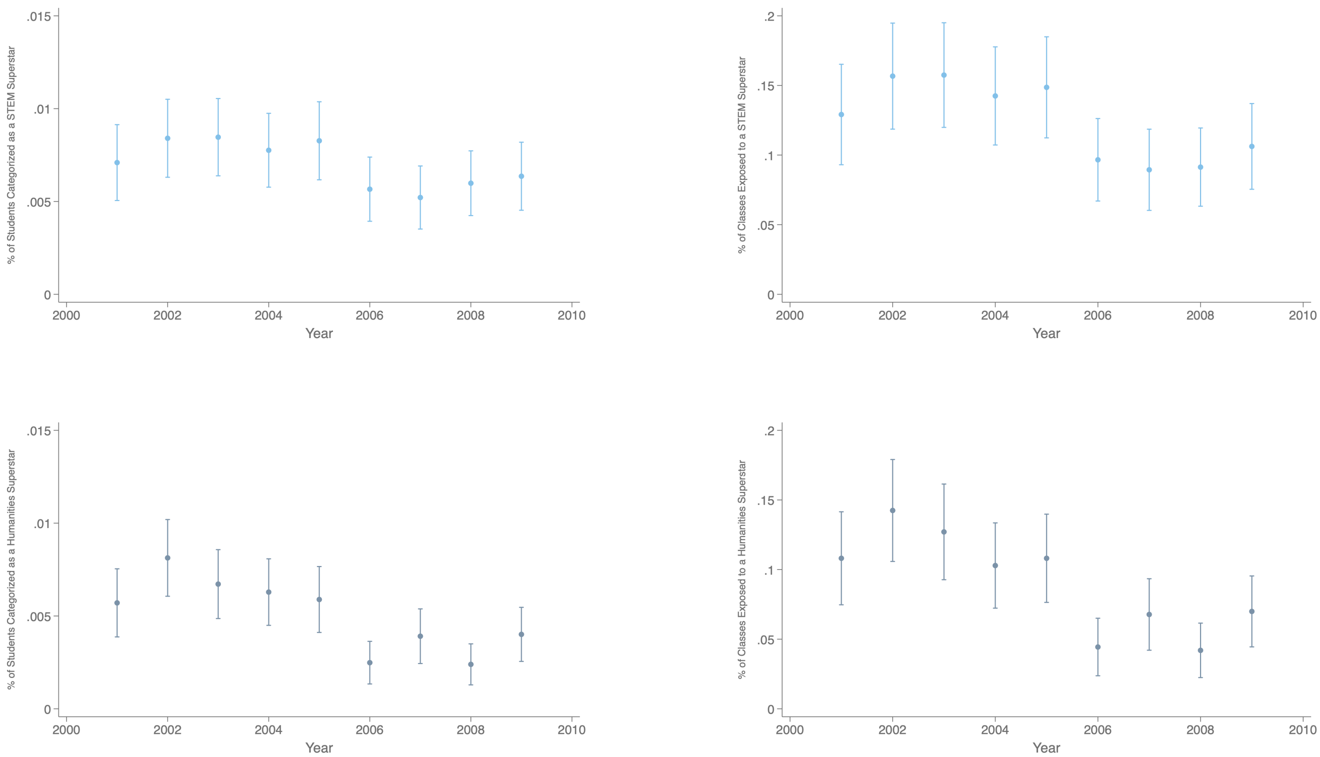


Figure A2: Superstar Dynamics by gender

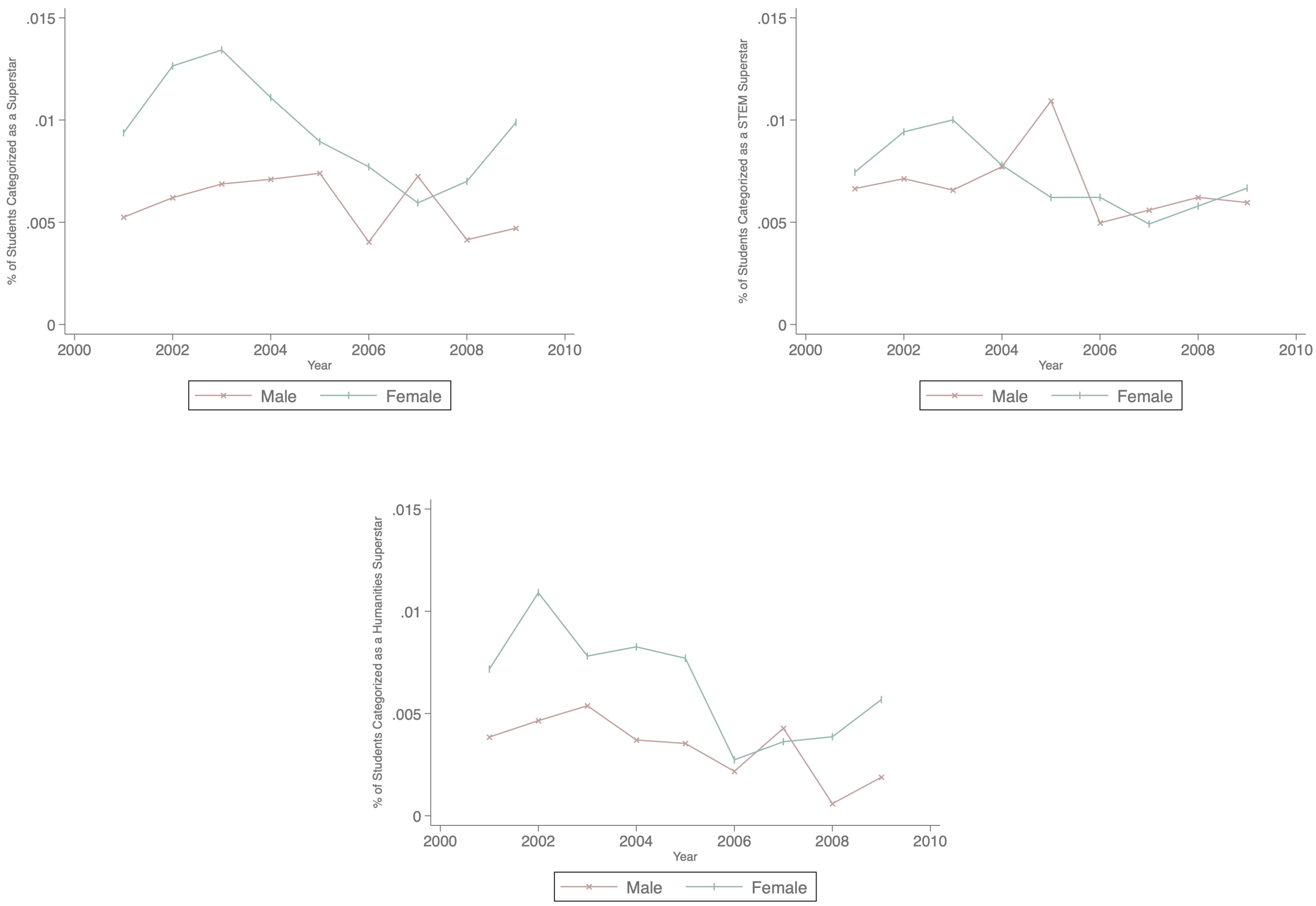
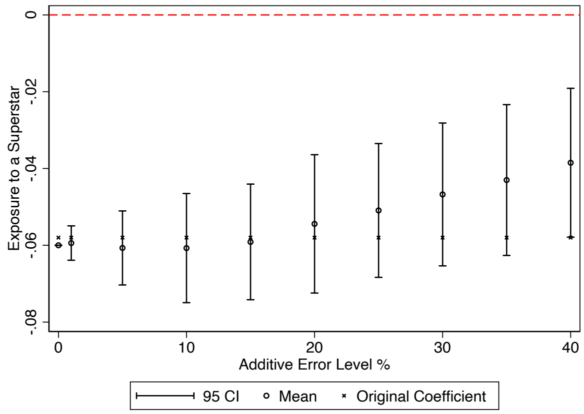
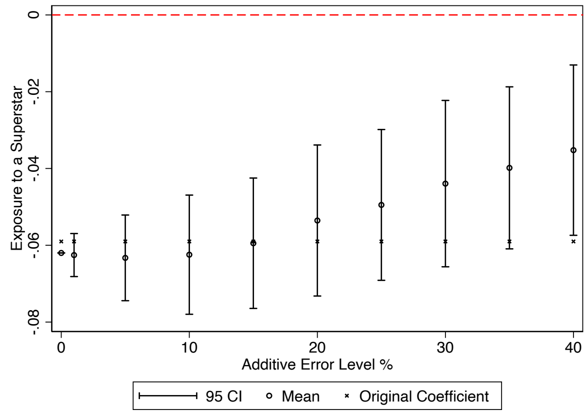


Figure A3: Additive Noise in Baseline Test Scores

(a) Overall



(b) STEM



(c) Humanities

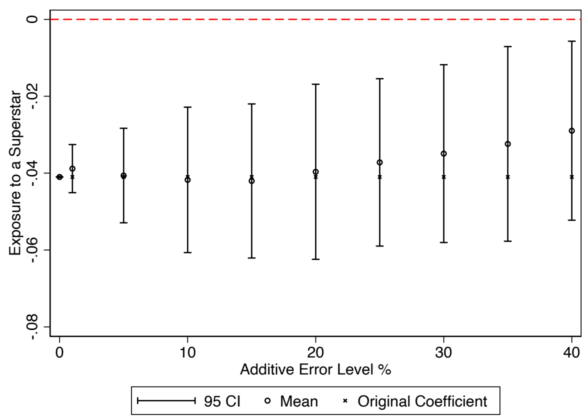
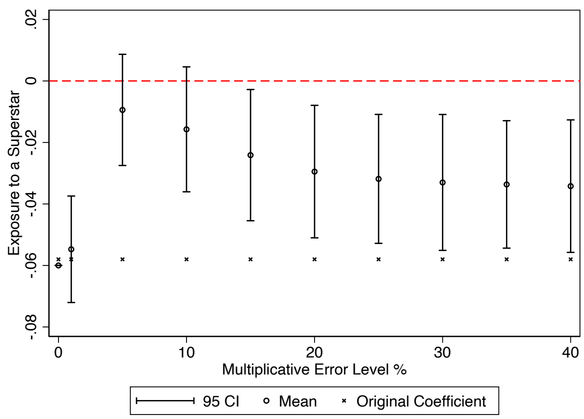
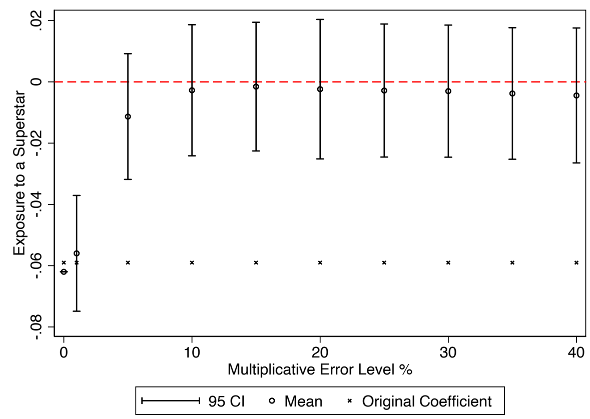


Figure A4: Multiplicative Noise in Baseline Test Scores

(a) Overall



(b) STEM



(c) Humanities

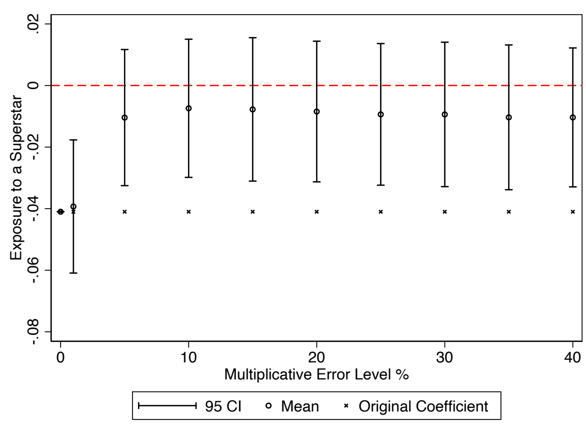
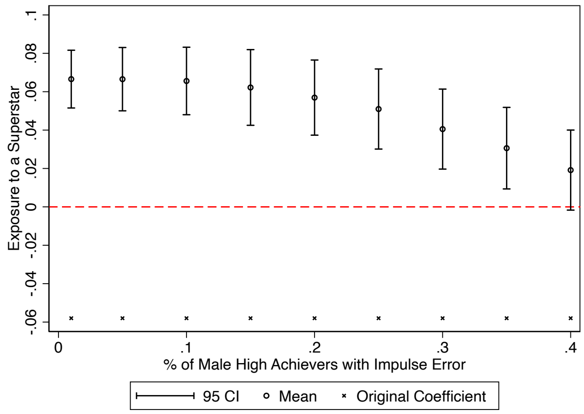
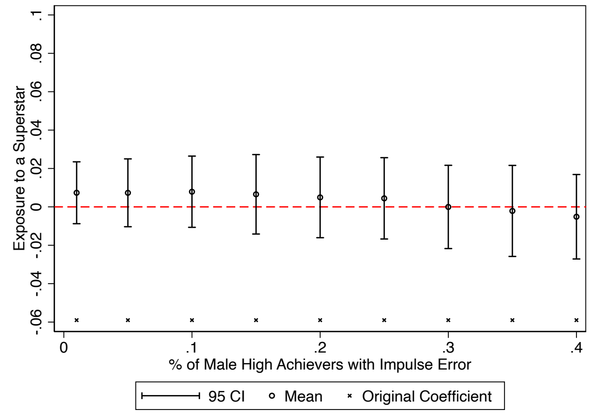


Figure A5: Impulse Noise in Male High Achievers Baseline Test Scores

(a) Overall



(b) STEM



(c) Humanities

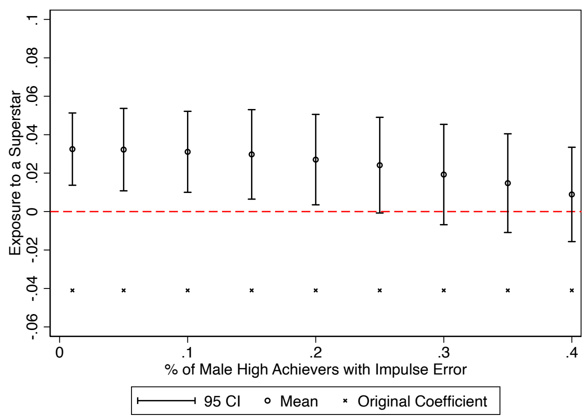


Table A1: Differences in Scores by Gender

	Male	Female	Diff. (Male - Female)	s.e.	Obs.
Algebra Test Score (Baseline)	14.6622	14.8659	-0.2037***	(0.0238)	64691
Modern Greek Test Score (Baseline)	14.3670	15.3292	-0.9622***	(0.0183)	64691
Physics Test Score (Baseline)	14.8913	14.8911	0.0001	(0.0239)	64691
Geometry Test Score (Baseline)	14.6751	14.8020	-0.1269***	(0.0242)	64691
History Test Score (Baseline)	15.4102	16.1065	-0.6963***	(0.0228)	64691

Table A2: Balance Test at the Individual Level Separately by Gender

	In a Classroom with a Superstar				
	(1)	(2)	(3)	(4)	(5)
Panel A: Females					
STEM Avg. Test Scores (Baseline)	−0.000				−0.001*
	(0.000)				(0.001)
Humanities Avg. Test Scores (Baseline)		0.000			0.001
		(0.001)			(0.001)
Age at Grade 10			0.001		0.000
			(0.002)		(0.003)
Born in 1st Quarter				−0.002	−0.001
				(0.004)	(0.005)
Mean	0.144	0.144	0.144	0.144	0.144
Observations	36093	36093	36093	36093	36093
F-Statistic	27.496	27.455	27.451	27.455	19.331
P-value for joint significance	0.000	0.000	0.000	0.000	0.000
Panel B: Males					
STEM Avg. Test Scores (Baseline)	0.000				0.001*
	(0.001)				(0.001)
Humanities Avg. Test Scores (Baseline)		−0.000			−0.001
		(0.001)			(0.001)
Age at Grade 10			−0.002		−0.001
			(0.003)		(0.006)
Born in 1st Quarter				0.003	0.001
				(0.004)	(0.008)
Mean of Outcome	0.144	0.144	0.144	0.144	0.144
Observations	28598	28598	28598	28598	28598
F-Statistic	26.154	26.147	26.159	26.158	18.370
P-value for joint significance	0.000	0.000	0.000	0.000	0.000

Notes: The treatment variable is regressed on each control variable in a separate regression. In column (5) we include all control variables simultaneously in the regression and report the joint significance of those variables. We show these estimates separately for male (upper panel) and female (lower panel) students. Each estimate is generated from a different regression. All grades presented have been standardized. All regressions include school by year FE and the following class controls: the proportion of female peers, the class's leave-out mean in baseline test scores in STEM and humanities, the number of students in the class, leave-out mean age, leave-out percentage of students born in the first quarter. Robust standard errors clustered at the school by year level are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A3: Balance Test at the Classroom Level

	Main variable	Baseline Class Avg. Grade in		Other			
	(1) Class with a Superstar	(2) STEM	(3) Humanities	(4) Class Size	(5) Avg. Class Age	(6) Proportion Born in Q1	(7) Number of Females
Class 1	−0.062	0.241	0.072	0.740	−0.022	0.002	0.274
	(0.071)	(0.612)	(0.472)	(0.479)	(0.017)	(0.010)	(0.698)
Class 2	−0.032	0.049	−0.033	0.634	−0.012	0.000	0.120
	(0.070)	(0.615)	(0.475)	(0.476)	(0.017)	(0.010)	(0.698)
Class 3	−0.045	0.185	−0.031	0.807*	−0.012	−0.002	0.356
	(0.071)	(0.614)	(0.473)	(0.484)	(0.017)	(0.010)	(0.685)
Class 4	−0.059	0.216	0.062	0.594	−0.016	0.004	0.823
	(0.071)	(0.609)	(0.477)	(0.494)	(0.016)	(0.010)	(0.701)
Class 5	−0.083	0.169	0.233	0.588	−0.011	0.003	0.436
	(0.070)	(0.643)	(0.483)	(0.533)	(0.017)	(0.010)	(0.740)
Observations	3338	3338	3338	3338	3338	3338	3338
Mean of Outcome	0.139	14.764	15.305	19.380	15.932	0.119	10.813
F-Statistic	0.981	0.388	0.367	0.895	0.892	0.358	2.081
P-value for joint significance	0.428	0.857	0.871	0.484	0.486	0.877	0.065
School by year FE	✓	✓	✓	✓	✓	✓	✓

Notes: The table shows estimated effects of the class number on a variety of outcomes. Class number 6 is omitted from the regression and thus should be interpreted as the reference category. The unit of observation is the class. Outcome variables are reported in the columns' headings and have been averaged at class level. In particular, we regress the classroom number on the treatment (column 1), average class baseline test scores in STEM (column 2), average class baseline test scores in humanities (column 3), the class size (column 4), the average class age (column 5), the average class proportion of students who are born in the first quarter of the calendar year (column 6), the class proportion of female students (column 7). F-Statistics for the joint significance of the regressors and the related P-value are also reported. They suggest that class numbers are not associated with differences in class-level outcomes. The mean of each outcome variable at class level is also reported. All regressions include a constant, and errors are clustered at unique class level. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A4: Estimates of the Effect of a Superstar by Individual Subject

	Grade 10: Test Score in Final Exam				
	(1)	(2)	(3)	(4)	(5)
	Algebra	Geometry	Physics	History	Modern Greek
Exposure to a Superstar	-0.051^{***}	-0.070^{***}	-0.056^{***}	-0.054^{***}	-0.001
	(0.016)	(0.017)	(0.016)	(0.018)	(0.020)
Observations	64177	64177	64177	64177	64177
Adjusted R-squared	0.604	0.589	0.623	0.561	0.619
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A5: Estimates of the Effect of a Superstar with Lower Thresholds

	Grade 10: Overall Avg Test Scores in Final Exam					
	(1)	(2)	(3)	(4)	(5)	(6)
	1 s.d.	1.2 s.d.	1.4 s.d.	1.6 s.d.	1.8 s.d.	2 s.d.
Exposure to a Superstar	-0.104	-0.111^{***}	-0.077^{***}	-0.061^{***}	-0.066^{***}	-0.058^{***}
	(0.136)	(0.024)	(0.012)	(0.009)	(0.010)	(0.012)
Number of Superstars	11967	8403	5245	2789	1237	514
Share Exposed to a Superstar	.998	.968	.837	.589	.311	.137
Observations	52724	56288	59446	61902	63454	64177
Adjusted R-squared	0.663	0.704	0.737	0.760	0.775	0.781
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A6: Estimates of the Effect of a Superstar By Gender

	Grade 10: Avg Test Scores in Final Exam		
	(1)	(2)	(3)
	Overall	STEM	Humanities
Panel A: Females			
Exposure to a Superstar	-0.052^{***}	-0.053^{***}	-0.035^{**}
	(0.015)	(0.017)	(0.017)
Observations	35747	35747	35747
Adjusted R-squared	0.795	0.744	0.701
School by year FE	✓	✓	✓
Controls	✓	✓	✓
Panel B: Males			
Exposure to a Superstar	-0.060^{***}	-0.061^{***}	-0.042^{**}
	(0.015)	(0.016)	(0.019)
Observations	28430	28430	28430
Adjusted R-squared	0.767	0.718	0.664
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A7: Estimates of the Effect of a Superstar With-in School-Cohorts that have a Superstar

	Grade 10: Avg Test Scores in Final Exam		
	(1)	(2)	(3)
	Overall	STEM	Humanities
Panel A: Overall			
Exposure to a Superstar	-0.060^{***}	-0.060^{***}	-0.045^{***}
	(0.013)	(0.016)	(0.016)
Observations	26603	26603	26603
Adjusted R-squared	0.762	0.711	0.665
School by year FE	✓	✓	✓
Controls	✓	✓	✓
	(1)	(2)	(3)
Panel B: Females			
Exposure to a Superstar	-0.050^{***}	-0.049^{***}	-0.040^{**}
	(0.016)	(0.019)	(0.018)
Observations	14776	14776	14776
Adjusted R-squared	0.779	0.727	0.679
School by year FE	✓	✓	✓
Controls	✓	✓	✓
Panel C: Males			
Exposure to a Superstar	-0.064^{***}	-0.065^{***}	-0.045^{**}
	(0.016)	(0.018)	(0.020)
Observations	11827	11827	11827
Adjusted R-squared	0.743	0.693	0.638
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A8: Estimates of the Effect of a Superstar on a Top Student: Year 11

	Grade 11: Avg. Test Scores		
	(1)	(2)	(3)
	Overall	STEM	Humanities
Panel A: Top 5			
Exposure to a Superstar	-0.056^{***}	-0.047^{***}	-0.062^{***}
	(0.013)	(0.014)	(0.017)
Exposure to a Superstar x Top 5	-0.108^{***}	-0.103^{***}	-0.088^{***}
	(0.014)	(0.016)	(0.015)
Effect on a Top Student	-0.164^{***}	-0.150^{***}	-0.150^{***}
	(0.017)	(0.019)	(0.019)
Observations	64177	64177	64177
School by year FE	✓	✓	✓
Controls	✓	✓	✓
Panel B: Top 3			
Exposure to a Superstar	-0.064^{***}	-0.054^{***}	-0.070^{***}
	(0.013)	(0.014)	(0.017)
Exposure to a Superstar x Top 3	-0.111^{***}	-0.107^{***}	-0.086^{***}
	(0.017)	(0.019)	(0.018)
Effect on a Top Student	-0.174^{***}	-0.162^{***}	-0.156^{***}
	(0.020)	(0.022)	(0.021)
Observations	64177	64177	64177
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A9: Estimates of the Effect of a Superstar on the Second Best

	(1) Overall	(2) STEM	(3) Humanities
Grade 10: Test Score in Final Exam			
Exposure to a Superstar	-0.033	-0.027	-0.042
	(0.034)	(0.039)	(0.037)
Observations	3338	3338	3338
School by year FE	✓	✓	✓
Controls	✓	✓	✓
	(1)	(2)	(3)
Grade 11: Avg. Test Scores			
Exposure to a Superstar	-0.092^{**}	-0.107^{**}	-0.041
	(0.038)	(0.042)	(0.043)
Observations	3338	3338	3338
School by year FE	✓	✓	✓
Controls	✓	✓	✓
	(1)	(2)	(3)
Grade 12: Avg. Test Scores			
Exposure to a Superstar	-0.071^{*}	-0.074^{*}	-0.059
	(0.039)	(0.044)	(0.042)
Observations	3338	3338	3338
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course.

Table A10: Long term outcomes by Type of Superstar

	(1) STEM Track	(2) STEM Admitted	(3) STEM Application	(4) Humanities Track	(5) Humanities Admitted	(6) Humanities Application	(7) Admitted
Panel A: Standard							
Exposure to a Superstar	−0.007	−0.018**	−0.003	0.007	0.002	0.003	0.006
	(0.007)	(0.008)	(0.005)	(0.007)	(0.007)	(0.008)	(0.006)
Observations	64177	55785	46114	64177	64177	46114	55785
School by year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Panel B: STEM							
Exposure to a STEM Superstar	−0.014*	−0.023***	−0.015***	0.014*	0.009	0.013	0.004
	(0.008)	(0.009)	(0.005)	(0.008)	(0.007)	(0.009)	(0.007)
Observations	64236	55835	46164	64236	64236	46164	55835
School by year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Panel C: Humanities							
Exposure to a Humanities Superstar	0.020**	0.022**	0.003	−0.020**	−0.016**	−0.038***	0.007
	(0.008)	(0.009)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Observations	64364	55958	46283	64364	64364	46283	55958
School by year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Panel D: Girl							
Exposure to a Girl Superstar	−0.007	−0.015*	0.000	0.007	0.003	0.009	0.003
	(0.008)	(0.009)	(0.006)	(0.008)	(0.008)	(0.010)	(0.008)
Observations	64345	55940	46266	64345	64345	46266	55940
School by year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Panel E: Boy							
Exposure to a Boy Superstar	−0.002	−0.016	−0.007	−0.007	0.002	−0.012	0.005
	(0.010)	(0.012)	(0.008)	(0.008)	(0.010)	(0.012)	(0.009)
Observations	64523	56091	46419	46419	64523	46419	56091
School by year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A11: Estimates of the Effect of a Superstar across School Quality and Neighborhood Income

	Low Ranking School			Top Ranking School		
	(1) Avg.Test Score in STEM	(2) Avg. Test Score in Humanities	(3) STEM Track in Grade 11	(4) Avg. Test Score in STEM	(5) Avg. Test Score in Humanities	(6) STEM Track in Grade 11
Exposure to a Superstar	-0.085^{***}	-0.047^{**}	-0.014	-0.027	-0.028	0.005
	(0.018)	(0.020)	(0.009)	(0.021)	(0.025)	(0.012)
Observations	32068	32068	32068	32109	32109	32109
Adjusted R-squared	0.722	0.678	0.191	0.740	0.697	0.179
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

	Low Income Nhbd.			High Income Nhbd.		
	(1) Avg.Test Score in STEM	(2) Avg. Test Score in Humanities	(3) STEM Track in Grade 11	(4) Avg. Test Score in STEM	(5) Avg. Test Score in Humanities	(6) STEM Track in Grade 11
Exposure to a Superstar	-0.068^{***}	-0.064^{***}	-0.029^{***}	-0.053^{***}	-0.026	0.007
	(0.022)	(0.025)	(0.011)	(0.018)	(0.020)	(0.009)
Observations	25868	25868	25868	38309	38309	38309
Adjusted R-squared	0.738	0.687	0.194	0.726	0.688	0.179
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized.

Table A12: Estimates of the Effect of a Superstar across Different Numbers of Superstars within a School

	School with Low Amount of Superstars			School with High Amount of Superstars		
	(1) Avg.Test Score in STEM	(2) Avg. Test Score in Humanities	(3) STEM Track in Grade 11	(4) Avg. Test Score in STEM	(5) Avg. Test Score in Humanities	(6) STEM Track in grade 11
Exposure to a Superstar	-0.037	-0.074^{**}	-0.010	-0.070^{***}	-0.035^{**}	-0.005
	(0.029)	(0.030)	(0.014)	(0.016)	(0.018)	(0.008)
Observations	29013	29013	29013	35164	35164	35164
Adjusted R-squared	0.747	0.702	0.190	0.718	0.676	0.182
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A13: Estimates of the Effect of a Superstar separtely for students whose baseline performance is above and below the median

	Students with Test Scores Below the Median			Students with Test Scores Above the Median		
	(1) Avg.Test Score Overall	(2) Avg.Test Score in STEM	(3) Avg. Test Score in Humanities	(4) Avg.Test Score Overall	(5) Avg. Test Score in STEM	(6) Avg. Test Score in Humanities
Exposure to a Superstar	-0.039^{***}	-0.043^{***}	-0.027	-0.014	-0.010	-0.017
	(0.013)	(0.015)	(0.019)	(0.016)	(0.019)	(0.017)
Observations	34131	34131	34131	30046	30046	30046
Adjusted R-squared	0.456	0.411	0.415	0.615	0.561	0.563
School by year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A14: Estimates of the Exposure to More Than One Superstar

	Grade 10: Avg Test Scores in Final Exam		
	(1)	(2)	(3)
	Overall	STEM	Humanities
Exposure to One Superstar	-0.056^{***}	-0.059^{***}	-0.035^{**}
	(0.013)	(0.015)	(0.016)
Exposure to Two or More Superstars	-0.077^{***}	-0.057^{*}	-0.091^{**}
	(0.030)	(0.034)	(0.046)
Observations	64177	64177	64177
Adjusted R-squared	0.781	0.731	0.688
School by year FE	✓	✓	✓
Controls	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A15: Estimates of the Effect of a Female Humanities Superstars

	Grade 10: Test Score in Final Exam			Grade 11:	
	(1)	(2)	(3)	(4)	(5)
	Overall	STEM	Humanities	STEM Track	Humanities Track
Panel A: All					
Exposure to a Female Humanities Superstar	-0.039^{**}	-0.064^{***}	-0.019	0.022^{**}	-0.022^{**}
	(0.018)	(0.019)	(0.020)	(0.009)	(0.009)
Observations	64459	64459	64459	64459	64459
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel B: Within-Female					
Exposure to a Female Humanities Superstar	-0.044^{**}	-0.061^{***}	-0.019	0.036^{***}	-0.036^{***}
	(0.021)	(0.021)	(0.024)	(0.014)	(0.014)
Observations	35861	35861	35861	35861	35861
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel C: Within-Male					
Exposure to a Female Humanities Superstar	-0.028	-0.060^{***}	-0.015	0.007	-0.007
	(0.022)	(0.023)	(0.024)	(0.012)	(0.012)
Observations	28598	28598	28598	28598	28598
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A16: Estimates of the Effect of a Male Humanities Superstars

	Grade 10: Test Score in Final Exam			Grade 11:	
	(1)	(2)	(3)	(4)	(5)
	Overall	STEM	Humanities	STEM Track	Humanities Track
Panel A: All					
Exposure to a Male Humanities Superstar	-0.056^{**}	-0.092^{***}	-0.044	0.014	-0.014
	(0.024)	(0.027)	(0.031)	(0.013)	(0.013)
Observations	64596	64596	64596	64596	64596
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel B: Within-Female					
Exposure to a Male Humanities Superstar	-0.057^{**}	-0.082^{***}	-0.032	0.010	-0.010
	(0.029)	(0.031)	(0.036)	(0.018)	(0.018)
Observations	36093	36093	36093	36093	36093
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel C: Within-Male					
Exposure to a Male Humanities Superstar	-0.057^{**}	-0.102^{***}	-0.065^{*}	0.018	-0.018
	(0.028)	(0.035)	(0.037)	(0.017)	(0.017)
Observations	28503	28503	28503	28503	28503
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A17: Estimates of the Effect of a Female STEM Superstars

	Grade 10: Test Score in Final Exam			Grade 11:	
	(1) Overall	(2) STEM	(3) Humanities	(4) STEM Track	(5) Humanities Track
Panel A: All					
Exposure to a Female STEM Superstar	−0.066***	−0.052***	−0.017	−0.013	0.013
	(0.018)	(0.019)	(0.020)	(0.010)	(0.010)
Observations	64432	64432	64432	64432	64432
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel B: Within-Female					
Exposure to a Female STEM Superstar	−0.070***	−0.047**	−0.046**	−0.016	0.016
	(0.021)	(0.023)	(0.023)	(0.014)	(0.014)
Observations	35834	35834	35834	35834	35834
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel C: Within-Male					
Exposure to a Female STEM Superstar	−0.063***	−0.055**	0.018	−0.012	0.012
	(0.022)	(0.022)	(0.025)	(0.012)	(0.012)
Observations	28598	28598	28598	28598	28598
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.

Table A18: Estimates of the Effect of a Male STEM Superstars

	Grade 10: Test Score in Final Exam			Grade 11:	
	(1)	(2)	(3)	(4)	(5)
	Overall	STEM	Humanities	STEM Track	Humanities Track
Panel A: All					
Exposure to a Male STEM Superstar	-0.063^{***}	-0.007	-0.049^{**}	-0.012	0.012
	(0.017)	(0.018)	(0.024)	(0.010)	(0.010)
Observations	64495	64495	64495	64495	64495
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel B: Within-Female					
Exposure to a Male STEM Superstar	-0.056^{***}	-0.002	-0.044	-0.013	0.013
	(0.019)	(0.021)	(0.027)	(0.015)	(0.015)
Observations	36093	36093	36093	36093	36093
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Panel C: Within-Male					
Exposure to a Male STEM Superstar	-0.075^{***}	-0.015	-0.057^{**}	-0.011	0.011
	(0.022)	(0.023)	(0.028)	(0.015)	(0.015)
Observations	28402	28402	28402	28402	28402
School by year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes: Regressions control for school by year fixed effects; student-level controls; student predetermined characteristics; and class-by-year characteristics. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Robust standard errors clustered at the school by year level are reported in parentheses. All grades presented have been standardized. The controls include the student's gender, age at grade 10, class size, proportion of females in the class, baseline GPA, average class GPA, and prior performance in the specific course. Nevertheless, due to presentation purposes it is not shown.