Non-monetary Awards and Social Comparison in Mathematical Olympiad: Evidence from a Regression Discontinuity Design*

ROBERTO ASMAT†

November, 2021

Access to the latest version here

Abstract

I study how non-monetary awards (diplomas) impact future participation in competitions by using administrative data on participants in district Mathematical Olympiads in Slovakia. To identify the impact of diplomas, I use a regression discontinuity design in which competitors who achieve a certain score receive a diploma. I find that receiving a diploma increases the likelihood of participation in the following grade by 3 p.p. (10% with respect to the baseline). Furthermore, by exploiting results' transparency and idiosyncratic variation in cohort composition within districts, I study how the value of the diploma for the marginal recipient is affected by social comparison. I find that the diploma effect is weaker when the fraction of recipients in the district is larger. Moreover, I find that the diploma effect is weaker when surrounded by higher proportions of high-performing peers but not affected when exposed to higher proportions of low-performing peers in the district. These results suggest that the value of non-monetary awards is shaped by status concerns and learning channels.

*I am grateful to the Slovak Mathematical Olympiad Committee for their support. In particular, to Monika Dillingerová and Ivan Cimrak for sharing their knowledge about the Olympiad's organisation in Slovakia. I would like to thank Roland Rathelot, Manuel Bagues and Ben Lockwood for their guidance and support. This paper also benefited from feedback given by Dan Anderberg, Michela Redoano, Andreas Stegmann, Christine Braun, Stefano Caria, Luis Candelaria, Dita Eckardt, Amrita Kulka, and participants at the CAGE-AMES Workshop in the University of Warwick. Finally, I would like thank Diana Tadeo for her excellent research assistance.

†Department of Economics, Warwick University (email: r.asmat-belleza@warwick.ac.uk)

1 Introduction

When studying how principals incentivise agents, economists have mainly focused on the use of monetary incentives (Ashraf, Bandiera and Jack, 2014). Although non-monetary (pure symbolic) awards are also widely used by principals (Frey, 2007), much less attention has been paid to such awards as means to increase the motivation and effort of agents (Gallus, 2017). For instance, employers decorate productive workers as employees of the month, universities confer special titles to professors for their contributions, and the military bestows medals to soldiers who risk their lives in battles. Empirical research in economics has only recently devoted efforts to study the important role played by non-monetary awards in fostering individuals' motivation (Cotofan, 2021). However, questions about factors shaping the effectiveness of non-monetary awards abound.

From a theoretical point of view, the effectiveness of non-monetary awards on individuals' motivation can be affected by two different forces. First, non-monetary awards are used by principals to motivate agents by exploiting a preference for status (Besley and Ghatak, 2008). If non-monetary awards convey status on their recipients, one might expect that their value should depend on how scarce awards are: the higher (lower) the fraction of award recipients, the lower (higher) the status conferred by awards. Second, when agents observe a principal assigning rewards among them, they might interpret such awards as signals to learn about their own abilities and prospect in the organisation (Benabou and Tirole, 2003). As a consequence, one might expect that the perceived signal's strength by award recipients is affected by how good their peers performed: if they learn that there is a larger (smaller) ability gap with respect to the top performer, they might lower (raise) the value of the award as a signal about their abilities. However, testing these theoretical implications is confronted with two main problems: a credible identification of the effect of non-monetary awards on individuals' motivation, and the ability of individuals to value such awards via social comparison.

This paper contributes to our understanding on how non-monetary awards affect individuals' willingness to continue performing a task in the context of Mathematical Olympiads. I use administrative records on over 56,000 results from 5^{th} to 9^{th} grade participants at the district Mathematical Olympiads in Slovakia which designs a test for each of these grades. I exploit a discrete threshold that determines the provision of diplomas to students in each grade. Specifically, only students who score 9 points or more (in a scale from 0 to 18) receive recognition (a diploma) as "successful" participants. By linking data on participation of individuals in each grade over time, I use the diploma rule to construct regression discontinuity (RD) estimates of the effect of receiving a diploma on the voluntary decision to participate in the following grade. The advantage of Mathematical Olympiads in quantifying mathematical abilities, the use of cutoffs to determine the provision of diplomas, and the inability of either participants or graders to manipulate scores allows for a unique quasi-experiment to test whether non-monetary awards are influencing future participation.

The central finding of this paper is that receiving a diploma increases the likelihood of competitors' participation in the following category by 3 percentage points (10% with respect to the baseline). This suggests that non-monetary awards strongly encourage children to continue training in mathematics. This effect on students' participation might be due not only to responses of students themselves but also of their teachers, and/or parents. Disentangling these reactions is out of the scope of this paper as the current data set does not offer a chance to explore reactions of agents involved in the decision to participate in the Olympiads.

Furthermore, this paper studies how social comparisons affect the value of non-monetary awards by exploiting the transparency of results within districts. In these competitions, results are disseminated such that individuals are not only informed about their own performance but also on the performance of their peers at the district level. This feature allows us to investigate whether the diploma effect on subsequent participation is affected by social comparisons for the marginal award recipient. By exploiting idiosyncratic variation in cohort composition within districts over time, I provide two main findings. First, I find that the diploma effect for the marginal award recipient is weaker in districts with a larger fraction of award recipients. This result supports the theoretical prediction that the value of non-monetary awards comes from their scarcity. Second, I find that when the marginal award recipient is exposed to larger proportions of high-performing peers (e.g., those achieving the maximum score possible), the diploma effect is weaker. However, this effect remains unchanged in districts with larger proportions of low-performing peers (e.g., those achieving the minimum score possible). These results support the hypothesis that negative comparisons matter more than positive comparisons when assessing non-monetary awards as signals of abilities, and are consistent with prior research in labour contexts showing that job satisfaction depends on relative pay comparisons, and this relationship is nonlinear (see Hamermesh (2001) and Card et al. (2012)).

Finally, in light of the literature on gender differences in self-confidence and interest in mathematical competitions (e.g., Niederle and Vesterlund (2007) and Croson and Gneezy (2009)), I test whether there are gender differences in the effect of non-monetary awards on future participation in mathematical competitions. I find no evidence that boys and girls react differently to receiving the diploma. Moreover, in this setting, I find that girls enter into competition as much as boys do. These findings are at odds with the literature on gender differences in competitive traits, usually based on adolescents and adults, and suggest that gender differences are not relevant either for children or the selected sample of highly accomplished students.

This paper contributes to two strands in the literature of Economics of Education. First, the central finding of this paper builds on an extensive literature studying the effects of feedback on educational outcomes. A comprehensive review is available in Villeval (2020) who shows that the literature has focused on studying absolute and relative feedback. This paper provides evidence on a different type of feedback which has a less objective nature, and has been less explored in educational settings. To the best of my knowledge, only two papers have partially addressed the effects of symbolic awards in educational contexts. First, Bedard, Dodd and Lundberg (2021) evaluate whether positive feedback on performance in first economics courses at university increases the probability of majoring in Economics. The limitation of this field experiment is that the difference between the treatment ("strong performance in Economics 1") and control ("successfully completing Economics 1") feedback is a very light touch treatment in comparison to the setting of this study where the contrast is salient (positive feedback versus no feedback). Second, Hoogveld and Zubanov (2017) study the effect of recognition on performance involving also first-year university students and find that the recipients of recognition did not do better, while the non-recipients significantly improved their performance. Although designed as a field experiment, their main results are based on an RD approach with a restricted sample size, which poses concerns on their external validity. In contrast, this paper provides credible field-based evidence of the importance of recognition for educational outcomes and its results are robust to all sorts of different specifications.

This study also contributes to recent literature that studies the effect of failure in competitions on future participation. These studies conduct RD estimations relying on score cutoffs that determine failure at advancing to further rounds. Buser and Yuan (2019) use data on Dutch mathematical Olympiads to estimate the effect of losing relative to winning (round 1 of national Olympiads) on subsequent participation in the competition. Ellison and Swanson (2021) address the same research question in the context of the American Mathematical Competitions (AMC). Both studies find evidence that competitors react negatively to losing relative to winning. However, despite the appeal of an RD design, the interpretation of the estimates of these studies depends critically on the treatment given to students who narrowly pass the score cutoff. In particular, in both settings, there is a multi-stage structure such that students who reach a certain score in a given round will proceed to a next round for which they will train more, gain more competition experience, and even additional recognition if passing a certain score in later rounds. This raises the question: to what extent is the effect driven by failure/recognition? In my setting, recognition (diploma) is isolated from additional experiences since Olympiads for early grades have no multi-stage structure and, therefore, results can be attributed to it.

As results suggest that children infer their capacity to do mathematics not only from their absolute and relative performance but also from recognition, this study poses the following question for policy matters: shall we design a non-monetary awards structure that maximizes interest and participation in mathematics for all students, or one that focuses on talented students and relegates the less skilled at early ages? It seems that the design of praising participants at early ages might be unnecessary as there are no further rounds in the competition. Moreover, such institutional design might have massive consequences on children's encouragement since their non-cognitive traits such as grit and conscientiousness, relevant for this matter, are less developed than those of adolescents and adults (Mike et al., 2015) and, therefore, are more prone to be affected public recognition. However, prescriptions in this regard deserve further lines of research and normative analysis. The institutional design of non-monetary awards for talented young students has more complex consequences as there is evidence that the development of their mathematical skills expands the knowledge frontier later in life (Agarwal and Gaule, 2020).

The paper is structured as follows. Section 2 describes the Mathematical Olympiads and data used. Section 3 presents the econometric method, main results and heterogeneous effect analysis by gender of participants. Section 4 addresses score comparisons. Section 5 concludes.

2 Institutional Background and Data

2.1 Background

Mathematical Olympiads are competitions usually held on the basis of regional and national rounds within countries whose ultimate goal is to select the best students to represent them at the International Mathematics Olympiad (IMO). Olympiads are proof-based contests consisting of few problems (at the IMO, 6 questions to be solved in two days) drawn from geometry, number theory, algebra, and combinatorics. There is strong evidence the mathematical Olympiads are reliably capturing math abilities of students (Ellison and Swanson, 2010) and that IMO scores are highly predictive of math publications and citations twenty years in the future (Agarwal and Gaule, 2020).

Although these competitions are aimed at high school students for the reasons mentioned, Mathematical Olympiads in Slovakia target students at primary school levels as well. This special feature in Slovakia makes competitions at this level suitable for studying non-monetary awards effects. Indeed, from the 5^{th} to 8^{th} grades, students only compete in a district round, as opposed to students from the 9^{th} to 13^{th} grades, where additional rounds (regional and national) are added. As already discussed in the introduction, a multi-stage structure complicates the analysis of the effect of scoring above the cutoff as it implies that competitors not only receive the diploma but also gain additional training for competition in later rounds. Therefore, this analysis focuses on competitions designed for 5^{th} to 8^{th} grade students. Olympiads for elementary grades are categorised by adding the prefix Z to the corresponding grade. For instance, the Olympiad category for 5^{th} grade students is called "Z5".

Contests for primary levels are organised by the Slovak Committee of Mathematical

Olympiad (SKMO) every year at the district level. To do so, district level committees are formed to manage the competitions locally. Tests for 5^{th} to 8^{th} grade students consist of 3 questions (6 points each), while the 9^{th} grade test involves 4 questions (6 points each). Crucially, score thresholds are established by the SKMO in each category. For Z5, Z6, Z7 and Z8, students who score at least 9 points are recognised as "successful". According to the SKMO, diplomas are given to these students, while no recognition is given to students who score 8 points or less[1]. For Z9, the same recognition is given at 12 points. These thresholds have been constant during the period 2011-2018. Finally, it is important to note that district committees provide descriptive (total score), comparative (rank within the district) and non-monetary awards (diplomas stating a "successful" performance) to all participants within the district. This implies that a given participant is not only informed about his/her own score/rank/award but also about the score/rank/award of each participant in the district.

2.2 Data and Analysis Samples

I collected data on participants at the district mathematical Olympiads between 2011 and 2017 in categories Z5 to Z8. The panel structure allows us to track those competitors and observe whether they participate in the next corresponding categories (Z6 to Z9) during 2012-2018. To have a longitudinal structure, I built unique identifiers for students and schools based on names provided and consistency over time. This data includes grade (5^{th} to 9^{th}), gender[2], language (Slovak or Hungarian)[3] and type of school (regular or grammar school)[4]. For each participant in categories Z5 to Z8, I also observe the score obtained and the rank within the district for the correspondent category. For example, I can track competitor i in category c with score s_i and rank r_i at district d in year t, and observe whether he/she participates in the following category in year $t + 1$.

Table 1 shows characteristics of all participants (column 1), as well as characteristics of participants in the RD analysis samples (columns 2–3). Panel A shows characteristics for all participants in any category. As shown in column 1, for all 46,968 tests considering all categories (Z5 to Z8), 37 percent of students participated already in a previous category on average, almost half of participants are girls, 8 percent of students are taught in Hungarian language, and 14 percent of students are enrolled in a grammar school. Panels B, C, D, and E correspond to samples for categories Z5, Z6, Z7, and Z8, respectively. Two important differences across categories are worth mentioning. First, it can be seen that the higher the category, the lower the amount of total observations. For instance, I observe 16,334

[1]SKMO informed that the feedback provided to students who score 8 points or less depend on the tutor. As a general recommendation, teachers do not say to students that they were "unsuccessful". Instead, some teachers might deliver verbal messages encouraging these competitors to participate the next year.

[2]I inferred the sex of participants based on their names.

[3]Students who are taught in Hungarian language are identified by their school description.

[4]Like the language classification, students enrolled in regular or grammar schools are identified by their school description.

participants in the Z5 category, and 8,566 participants at Z8 category. Second, the higher the category, the higher the proportion of students with previous experience. For instance, in Z5, no student has previous experience in Olympiads as they only start in 5^{th} grade. In Z8 category, 66% of participants attended the Olympiads in previous categories at least once. However, it is interesting to note that most of the characteristics remain the same in each category. Interestingly, the proportion of girls participating represents around half of total participants in any category.

I construct two RD samples for analysis. Based on the cutoff score that determines the provision of diplomas, I select students who score within 2 points (column 2) and 3 points (column 3) of the cutoff. For instance, for column 2 we have 11,182 competitors who scored from 7 to 10 points. It can be seen that for any of these two windows, students do not vary along characteristics (past participation, proportion of girls, students taught in Hungarian language, and students enrolled in grammar schools). Moreover, students of the RD samples are comparable to students of the whole sample along these characteristics. Given the discrete nature of the running variable, an RD sample based on competitors scoring 8 or 9 points does not allow us to conduct the analysis since the treatment and points are perfectly correlated. In other words, it is not feasible to attribute changes in the outcome to the treatment.

3 RD-based Analysis of Non-monetary Awards

3.1 Research Design

To evaluate the effects of receiving a diploma on subsequent participation at the mathematical Olympiads, I adopt a sharp regression discontinuity (RD) design around the score cutoff of 9 points that determines the provision of diplomas. In particular, we estimate the following model:

$$Y_{i,c+1} = \alpha_1 \mathbf{1}(S_{ic} \geq 9) + f(S_{ic}) + \mathbf{1}(S_{ic} \geq 9) \times f(S_{ic}) + X_i'\gamma + \theta_{dtc} + \theta_s + u_{i,c+1}, \tag{1}$$

where $Y_{i,c+1}$ is an indicator variable for participating the next category $c+1$. $\mathbf{1}(S_{ic} \geq 9)$ is an indicator variable for scoring equal to or above the threshold in category c. $f(S)$ is a polynomial function of the number of points scored S. X_i is a vector of controls for gender, past participation, rank within the district in category c. θ_{dtc} is a district-by-year-category fixed effect that controls for any unobserved shocks common to all competitors within a district, year and category. Finally θ_s is a school fixed effect.

The parameter of interest α_1 provides an estimate of the causal effect of receiving a diploma on future participation averaged across all categories. Under the identification assumption that u_i does not change discontinuously at the threshold, this is an unbiased estimate even if controls for observable factors X are not included. Intuitively, the assumption is that the score at a given year is smoothly related to characteristics that affect participation the year after. Formally, $f(S)$ is constant in a neighbourhood around the threshold. Under this assumption, students who just scored below 9 points are a suitable control group for individuals with scores just above 9 points, and any difference in their participation next year can be attributed to the fact that they receive the diploma. Formally, $f(S)$ is nonparametrically identified at $S = 9$ (Hahn, Todd and der Klaauw, 2001).

To estimate the model in equation (1), I use two different bandwidths of 2 and 3 points to the left and the right of the threshold, and I adopt two different approaches. First, I assume that shocks in districts, years, and categories work independently and include separate fixed effects for each level (θ_d, θ_t and θ_c). Second, I relax that assumption and allow for shocks to be specific for a given district at a particular year in a specific category (θ_{dtc}). Although this model rules out all confounders at the district-year-category level, is more demanding because 1,787 fixed effects need to be estimated, compared to 78 district, 8 year, and 4 category fixed effects in a model with separate fixed effects.

3.2 Validity of RD Design

In our context, the ability of competitors or scorers to manipulate which side of the score threshold would fall might be a concern. In this regard, it must be assumed that it is random for a competitor to have a score either just below or just above the score cutoff in order to establish identification. Put differently, I assume that some competitors are randomly lucky, obtaining a score barely below the threshold, while other competitors are randomly unlucky and obtain a score barely above the threshold that determines receiving the diploma.

There are a number of reasons that this assumption is likely reasonable. The first is the level of accuracy of mathematical Olympiads to measure mathematical abilities (Agarwal and Gaule, 2020). This implies that students with small score differences are virtually equally talented. Second, students do not know the scoring criteria in advance, meaning that it is implausible for them to manipulate their answers and choose whether to get the diploma. In fact, it is fair to assume that all students are aiming to score the highest points. Third, scorers follow a national grading rule and do not have incentives to manipulate the score of students who are just below the threshold that determines the provision of diplomas. Indeed, district rounds are not followed by other rounds and, therefore, grading participants below or above the cutoff do not involve costs for the organisation. Moreover, scorers do not tutor participants so there is no incentive to encourage particular participants via score manipulation.

Figure 1 contains a histogram displaying the number of observations in each possible score at the mathematical Olympiads. It includes results from all categories (Z5 to Z8) during 2011–2017. The distribution of scores shows no evidence of endogenous sorting to one side of either of the threshold studied. Figure 2 shows the same analysis by category. Likewise, there are no concerns of score manipulation for any category.

It is interesting to observe heaps at 0, 6, 12 and 18 points in Figure 1. The fact that a good proportion of competitors scored 0 points might indicate that they were pushed by their parents without preparing for the examination. However, the other heaps might be explained by the fact that some competitors only aim at answering a question for which they know the complete proof and are not interested in providing wrong or incomplete proofs. This raises another type of concern by which competitors around the threshold might have different strategies to solve the test: some competitors might provide only complete proofs while others might be willing to provide incomplete proofs in order to obtain a higher score.

Finally, I conduct an additional test of sorting by examining regression models based on equation (1) with the students' characteristics used as the dependent variables which should remain unchanged at the thresholds. The participants' demographics I examine are gender and language of instruction (Slovak or Hungarian). Table 2 shows estimates of the effect of receiving the diploma on these predetermined characteristics. The regression models estimated employ bandwidths of 2 and 3 points away of the cutoff (9 points). For each student's demographics, all fixed effects (district, year, category and school) are included in the regressions.

Table 2 shows that overall the predetermined individuals' characteristics (gender and past participation in the Olympiad) are not statistically related to receiving the diploma. With the exemption of column (2), these characteristics are balanced between control and treatment groups. Further graphic evidence for the previous analysis is presented in Figure 3, which shows the relationship between students' score and predetermined students' characteristics. It shows bins of predetermined characteristics and corresponding fitted regression lines based on equation (1) which should remain unchanged across the scoring thresholds. Figure 3 shows that demographic factors are stable across the threshold. The stability of predetermined characteristics gives additional credibility that the regression discontinuity design can deliver unbiased estimates in this context.

3.3 Impact on Future Participation

In this subsection I examine the effect of receiving a diploma on future participation in the mathematical Olympiads. A virtue of the RD design is that it provides a graphical depiction showing how the diploma effect is identified. Thus, I begin with a graphical analysis of the diploma effect before turning to a more detailed regression-based analysis.

Figure 4 plots the fraction of participants in year t who participate in year $t+1$ by score obtained in t, and predicted participation rates based on simple regression models for all participants on each side of the cutoff score. Each observation is the proportion of competitors participating the year after in score bins. The running variable, score obtained in year t, has been normalized so the cutoff (nine points) is displayed at zero. Thus, the black lines represent the fitted regressions in the intervals -9 to -1, and 0 to 9.

Before focusing on the discontinuity, it is worth noting the strong relationship between the score obtained in a given year and the likelihood of participating in the following category as shown in Figure 4. The higher the score obtained, the higher the likelihood of subsequent participation. For instance, around 20 percent of students who scored 0 points (or -9 points below the cutoff) participate in the following category, while around 70 percent of students who scored 18 points (or 9 points above the cutoff) do so. Moreover, it can also be seen that the higher the points, the larger confidence intervals of the mean of future participation. This is explained by the fact that students scoring high points are fewer as shown in Figure 1.

Figure 4 reveals a sharp jump in the fraction of students participating in the following category at the cutoff, with competition participation rising from 41 percent to 51 percent. This graph provides strong evidence that the diploma has large effects on students' willingness to compete again. While the regression lines illustrate this relationship at the cutoff score, the unrestricted fraction means indicate the underlying noise in the data. As can be seen, on each side of the cutoff score, the relationship between score at t and participation at $t+1$ is smooth providing strong evidence that no other factor is affecting participation at the cutoff apart from the diploma. As a reminder, these students are all labelled as "successful" depending on which side they are from the cutoff.

Figure 5 contains the same plot for each of the four transitions (from Z5 to Z6, from Z6 to Z7, from Z7 to Z8, and from Z8 to Z9) separately. It can be seen that in all transitions there is a discontinuous jump in the probability of participating the next year when competitors fall just short of the cutoff, except for in the Z6 category. However, Figure 4 and Figure 5 should be taken carefully as they show the unconditional subsequent participation and do not include any control. In particular, they do not take into account that some students have participated before the period of analysis which might influence future participation for students around the threshold. For instance, in Panel B (transition Z6 to Z7), as shown in Table 1, 46 percent of these competitors participated in the Z5 category. Only Panel A of Figure 5 does not suffer from this issue: for the transition from Z5 to Z6, no student has previous experience in the Olympiads as they do not exist for lower grades.

Having shown the raw patterns of future participation around the score cutoff, I present regression-based estimates. Table 3 shows the estimated effect of receiving the diploma on future participation. I present estimates for two bandwidths as described in Table 1: +/-2 points; +/-3 points away from the cutoff. The discontinuity estimates are between 3.4 p.p. or around 10% with respect to the baseline (column 2) and 5.3 p.p. or around 15% (column 1). This means that competitors react positively to non-monetary awards. Table 3 also shows the estimated effect based on regressions with separate fixed effects (columns 1 and 3) and district-by-year-by-category fixed effects (columns 2 and 4). It can be seen that the estimated effect in the latter is slightly lower and consistent in all RD samples. Because of its consistency and virtue to account for all non-observable factors that might affect future participation at the district-year-category level, this model is my preferred specification.

This study has two important limitations. First, the data set does not allow us to disentangle the reactions of students from the responses of their tutors and parents. Indeed, the decision to participate is voluntary and might be based on interaction between students, tutors and parents. For instance, tutors might be more encouraging, offer better training and/or set higher expectations for students who receive the diploma, while parents might reward them and/or push their children to continue training. Understanding these mechanisms is important in designing non-monetary award policies. Second, I cannot identify how tutors manage the feedback given to participants who achieved a score just below the threshold as this information is not available. In this regard, tutors might either provide verbal encouragement or negative feedback to these students. With respect to the latter, some studies suggest that negative feedback may increase motivation and future performance (Cianci, Klein and Seijts, 2010) and, therefore, one could expect that the estimated effect in this study may be a lower bound.

3.4 Gender Differences in Reaction to Non-monetary Awards

In this subsection, I investigate whether there are gender differences in the reaction to receiving the diploma. Thus, the main interest is not the discontinuity itself (the effect of receiving the diploma on future participation), but whether there are gender differences in the discontinuity. In other words, whether boys and girls react differently to receiving the diploma. Therefore, I estimate the following equation:

$$(2) \quad Y_i = \alpha_1 T_i + f(S_i) + f(S_i) \times T_i + \alpha_2 F_i + \alpha_3 T_i \times F + f(S_i) \times F + f(S_i) \times F \times T_i + u_i,$$

where Y_i is an indicator variable for participating the next year, T_i is analogous to the indicator variable $\mathbf{1}(S_i \geq 9)$ in equation (1) which indicates whether participant i scored equal or above the threshold, $f(S)$ is a polynomial function of the number of points scored S. The parameter of interest in this analysis is α_3, which estimates the gender difference in reaction to receiving the diploma. Following Buser and Yuan (2019), I include two important interactions allowing for different slopes for each gender: the polynomial $f(S)$ with the participant's gender F; and the triple interaction between the polynomial $f(S)$, the treatment T, and the participant's gender F. Again, I run this analysis using separate fixed effects and district-by-year-by-category fixed effects.

Like the general analysis, I first provide graphical evidence of gender differences of the impact of the diploma on future participation in mathematical Olympiads. Figure 6 plots the fraction of participants in year t who participates in year $t+1$ by score obtained in t for boys and girls. First of all, it is remarkable that conditional on the score obtained in a given year, there is no evidence of gender differences in subsequent participation. To the left (less skilled) and the right (more skilled) of the cutoff, both boys and girls participate in Olympiads the year after. Regarding the gender difference around the cutoff, Figure 6 shows no graphical evidence of different reactions to receiving the diploma.

Next, I test gender differences in reaction to receiving the diploma by estimating equation (2). 5^{th} row in Table 4 shows the estimates of the difference between boys and girls in the effect of the diploma on future participation. It presents the estimates for two different ranges of score around the cutoff: 2 points (columns 1 and 2) and 3 points (columns 3 and 4). The sixth row in Table 4 reveals that the difference is not statistically significant except for one bandwidth (+/- 2 points) at the 10% level when allowing for district-by-year-by-category fixed effects. The results taken together suggest that there are no gender differences in reaction to receiving the diploma.

The evidence of gender differences from this study is striking. First of all, the fact that the sample is gender balanced goes in contrast with the consensus on the lower willingness of girls to compete in mathematical subjects (Niederle and Vesterlund, 2010). As shown in Table 1, 47 percent of all tests were taken by girls. Second, it shows that there is no evidence that boys and girls react differently to receiving the diploma, which also goes in contrast against similar studies (see Buser and Yuan (2019) and Ellison and Swanson (2021)). As previously explained in the introduction, although the setting of this study (one stage structure) is not comparable to theirs (multi-stage structure), this study provides evidence that such differences are not important in the context of high achievers and/or children.

4 Non-monetary Awards and Score Comparison

In this section I study whether the diploma effect is affected by social comparison. As explained in section 2.1, the SKMO disseminates results such that everyone is able to observe the score of everyone else in a district. By exploiting variation in the ability distribution of cohorts within districts over time, I investigate whether this transparency can affect the diploma effect on subsequent participation through score comparison in two different ways. First, I assess whether the diploma effect depends on its scarcity (the fraction of diploma recipients in the district). Second, I test whether the diploma effect is affected by the ability gap in the district (the performance of peers who are at the top and the bottom of the score distribution). I interpret these results as evidence that non-monetary awards motivate individuals via status concerns and inform individuals about their abilities and prospect in the competition, respectively.

4.1 Status Concerns

The first approach studies the relationship between the diploma effect for the marginal recipient in a given grade and the diploma scarcity in the district. To account for observed and unobserved characteristics of districts and competitors that might be correlated with the fraction of recipients, I exploit variation in the score distribution within district-category over time. By doing so, I compare the effect of receiving the diploma on competitors that face the same environment, except for the fact that in certain years they are exposed to a higher/lower fraction of diploma recipients as a result of variation in cohort composition.

I test additional gains or losses on the diploma effect for individuals exposed to different fractions of diploma recipients as follows:

$$\begin{aligned} Y_{id,c+1,t+1} = {} & \alpha_1 T_{idct} + \alpha_2 S_{idct} + \alpha_3 T_{idct} \times S_{idct} + \beta_1 F_{dct} + \beta_2 T_{idct} \times F_{dct} \\ & + X_i'\gamma + \theta_{dc} + \theta_t + \theta_s + u_{id,c+1,t+1}, \end{aligned} \tag{3}$$

where $Y_{id,c+1,t+1}$ is an indicator variable for individual i in district d participating in the next category $c+1$ in year $t+1$. $T_{idct} = \mathbf{1}(S_{idct} \geq 9)$ is an indicator variable for scoring equal or above the threshold in category c in year t. S_{idct} is the number of points scored. F_{dct} is the fraction of diploma recipients in district d at category c in year t. X_i is a vector of controls for gender, past participation, and rank within the district d in category c in year t. Finally, I include district-by-category, year, and school fixed effects. For this analysis, β_2 is the coefficient of interest as it captures the gain/loss of the diploma effect on subsequent participation when exposed to a higher fraction of diploma recipients.

Table 5 shows the results of estimating equation (3) using two RD samples. For instance, column (1) shows the estimations for the sample using a bandwidth of $+/-2$ points and separate district, year and category fixed effects. The interpretation for the interaction between treatment and the fraction of diploma recipients is as follows: an increase from 0 to 1 in the fraction of diploma recipients in the district is associated with a reduction of 18.9 percentage points of the diploma effect on subsequent participation (row 5). Columns (2), (3) and (4) show the same results. This evidence suggests that individuals value the diploma depending on the scarcity of the diploma in the district: the less (more) scarce, the lower (higher) its value.

4.2 Informational Channels

Since individuals might interpret the diploma as a signal of their abilities, in this section I inspect how the strength of this signal for the marginal recipient can be affected by his peers' performance. The main idea of this exercise is the following: if the diploma delivers information about abilities, its strength as a signal should be affected by the performance of peers at the top of the score distribution while it should not be affected by the performance of peers who are at the bottom of the ability distribution. The intuition behind this is that the marginal recipient might learn about his potential in these competitions by knowing how far is from the top performer but should not learn in this regard from the worst performer in the district. In contrast, if the diploma effect is affected by the performance of peers who are at the bottom of the score distribution, this would be evidence that the diploma is associated with relative utility comparisons rather than informational aspects.

This section starts by providing a simple graphical analysis of the effect of receiving the diploma on subsequent participation in districts with different maximum and minimum scores achieved. Since characteristics of the score distribution in a given district might be correlated with the probability of receiving the diploma, I then propose two causal approaches to inspect how peers' performance can affect the value of receiving the non-monetary award on encouraging competitors by exploiting idiosyncratic variation in the score distribution within districts.

To illustrate, consider two districts "A" and "B" with different score ranges. In district "A", the maximum score achieved is 12 points, while in district "B", is 18 points. The question is whether the effect of receiving the diploma is different in districts "A" and "B". The hypothesis is that the effect of the diploma is less strong in district "B" than in "A", as students who barely passed the cutoff might feel that the recognition is not too valuable, as there is one competitor who did extremely well in comparison to them. Analogously, if in district "A" the minimum score achieved is 0 points, while in district "B" it is 7 points, one can expect the effect of receiving the diploma to be less strong in district "B" than in "A" as competitors who barely passed the cutoff might feel that the recognition is not too valuable as there is least talented competitor almost obtained the same recognition.

I first provide graphical evidence by grouping districts with respect to the maximum and minimum scores achieved and show the effect of receiving the diploma separately. To simplify the analysis, I set different breaking points for maximum scores (13, 14, and 15 points) and minimum scores (3, 4 and 5) to classify districts. For the maximum score analysis (13 points breaking point), I compare the diploma effect among competitors from districts where the maximum score achieved was 12 or less, versus its effect among competitors living in districts where the maximum score achieved was 13 or more. Likewise, for the minimum score analysis (3 points breaking point), I compare the diploma effect among competitors in districts where the minimum score achieved was 2 or less, versus its effect among competitors from districts where the minimum score achieved was 3 or more. According to our reasoning, for both analyses, the diploma should have a stronger effect among competitors from the first group of districts than from the second group. Table 6 and 7 describe the samples used for the analysis.

Panels (a) and (b) of Figure 7 show the diploma effect in districts where the maximum score was 12 or less (equivalently, distance from the cutoff being 3 or less) versus its effect in districts where the maximum score was 13 or more (equivalently, distance from the cutoff being 4 or more), respectively. Naturally, in panel (a), the fitted line to the right of the cutoff is based on only 4 points, while in panel (b) there is no such restriction. The same applies in panel pairs (c) and (d), and (e) and (f) for 13 and 14 points as breaking points to group districts, respectively. Figure 7 suggests that the effect of receiving a diploma in districts where there are relatively less talented competitors is around twice that in districts where there are relatively more talented competitors. This suggests that non-monetary awards are affected negatively when recipients are surrounded by high-achievers.

Panels (a) and (b) of Figure 8 show the diploma effect in districts where the minimum score was 2 or less (equivalently, distance from the cutoff being -7 or less) versus its effect in districts where the minimum score was 3 or more (equivalently, distance from the cutoff being -6 or more), respectively. By construction, in panel (b), the fitted line to the left of the cutoff is based on only 6 points, while in panel (a) there is no such restriction. The same applies in panel pairs (c) and (d), and (e) and (f) for 4 and 5 points as breaking points to classify districts, respectively. Figure 8 suggests that there is no difference between the effects of receiving a diploma in both groups. As described in Table 7, there are few districts where the minimum score achieved was, for instance, 5 points or more (comprising only 1,085) which explains the noise shown in panel (f). Although Figure 8 suggests that the diploma effect is not affected when surrounded by low-achievers, I proceed to test it using a regression model.

The limitation of this analysis is that the score composition within a district can itself affect a competitor's performance and their likelihood to receive the diploma. In other words, competitors from districts with different score distributions are not comparable. For instance, Table 6 shows that 46% of competitors in districts where the maximum score achieved is 12 points have some previous experience at the Olympiads while the analogous figure for competitors in districts where the maximum score achieve is 13 points or more, is 35%. The same issue applies for the division of competitors based on different minimum score achieved in their districts.

More generally, the challenge of identifying the effect of being exposed to high- and low-performing peers on the value of the diploma relates to the fact that competitors are self-selected to live in districts based on their quality of schools and, therefore, the value of being recognised in mathematical Olympiad is mechanically related to the characteristics of their peers. To overcome this issue, I develop two different but related causal approaches to hold peer abilities constant and focus on "as good as random" differences in being exposed to low and high points of reference.

Exposure to Minimum and Maximum Scores

The first approach consists of capturing "as good as random" differences in minimum and maximum levels of scores achieved across districts. To account for observed and unobserved characteristics of districts and competitors that might be correlated with minimum and maximum scores achieved, I exploit idiosyncratic variation in the tails of score distributions within districts. By doing so, I compare the effect of receiving the diploma on competitors that face the same environment, except for the fact that in certain years they are exposed to a higher/lower maximum and minimum scores achieved as a result of idiosyncratic variation.

The key identifying assumption in this approach is that changes in the maximum and minimum scores achieved in a district are uncorrelated with observed and unobserved factors that could themselves affect the likelihood of receiving the diploma. To test for this, I check whether the diploma given to competitors around the threshold explains different levels of maximum and minimum scores achieved in their districts. Table 8 shows the results of conducting this analysis as in equation (1). For instance, in columns (1) and (2), the maximum and the minimum scores achieved in the district are the dependent variables using the +/- 2 points bandwidth, respectively. Table 8 shows that there is no evidence that the competitors above the threshold are more or less exposed to different levels of maximum and minimum scores than competitors below the threshold.

I study additional gains or losses on the diploma effect for individuals in districts with different minimum and maximum scores achieved as follows:

$$\begin{aligned} Y_{id,c+1} &= \alpha_1 T_{idc} + f(S_{idc}) + T_{ic} \times f(S_{idc}) \\ &\quad + \beta_1 S^{max}_{dc} + \beta_2 S^{min}_{dc} + \beta_3 T_{idc} \times S^{max}_{dc} + \beta_4 T_{idc} \times S^{min}_{dc} + X_i' \gamma + u_{id,c+1}, \end{aligned} \tag{4}$$

where $Y_{id,c+1}$ is an indicator variable for individual i in district d participating in the next category $c+1$. $T_{idc} = \mathbf{1}(S_{idc} \geq 9)$ is an indicator variable for scoring equal or above the threshold in category c. S_{idc} is the number of points scored. S^{max}_{dc} is the maximum score achieved in district d at category c. S^{min}_{dc} is the minimum score achieved in district d at category c. Finally, X_i is a vector of controls for gender, past participation, rank within the district in category c. Finally, all regressions are estimated under two models: using separate fixed effects and district-by-category, year and school fixed effects. For this analysis, β_3 and β_4 are the coefficients of interest as they capture the gain/loss of the diploma effect on subsequent participation when surrounded by better peers below and above the cutoff, respectively.

Table 9 shows the results of estimating equation (3) using the two RD samples used throughout this paper. For instance, column (1) shows the estimations for the sample using a bandwidth of $+/-2$ points and separate district, year and category fixed effects. The interpretation for the interaction between treatment and the maximum score achieved in the district is as follows: an increase of 1 point in the maximum score achieved in the district is associated with a reduction of 1.1 percentage points of the diploma effect on subsequent participation (row 6). However, I find that an increase in one point in the minimum score achieved in the district does not affect the effect of receiving the diploma (row 7). Rows 6 and 7 in columns (3) and (4) show the same results for the +/-3 points bandwidth. This evidence suggests that receiving the diploma is associated with a lower effect on subsequent participation when surrounded by better peers above the cutoff and no gain/loss when surrounded by better peers below the cutoff.

Exposure to High- and Low-performing Peers

The second approach is similar to Mouganie and Wang (2020) and accounts for observed and unobserved characteristics of districts and competitors that might be correlated with high- and low-performing peer composition by exploiting idiosyncratic variation in the distribution of test scores within districts over time. Thus, I compare the effect of receiving the diploma on competitors that face the same environment, except for the fact that in years they are exposed to a higher proportion of high-performing or low-performing peers as a result of idiosyncratic variation.

I conduct this analysis using three different definitions of high- and low-performing peers. In the most stringent definition, a high-performing peer is the competitor who scored 18 points (the maximum possible score) while a low-performing peer is defined as the peer who scored 0 points (the minimum possible score). The alternative pair of definitions for high- and low-performing peers correspond to 17 and 1 points, and 16 and 2 points, respectively.

Similarly, I first check whether the likelihood of obtaining the diploma is related to being exposed to different proportions of high- and low-performing peers. Tables 10, 11 and 12 show the results of conducting this analysis as in equation (1) for all the three definitions of high- and low-performing peers. For instance, in columns (1) and (2) of Table 9, the proportion of high-performing peers (those who obtained 18 points) and the proportion of low-performing peers (those who obtained 0 points) in the district are the dependent variables using the +/- 2 points bandwidth, respectively. Tables 10, 11, and 12 show that there is no evidence that the competitors above the threshold are exposed to different proportions of high- and low-performing peers than competitors below the threshold.

Thus, by exploiting random exposure to high- and low-performing peers, I estimate the following equation:

$$\begin{aligned}Y_{id,c+1} = {} & \alpha_1 T_{idc} + f(S_{idc}) + T_{ic} \times f(S_{idc}) \\ & + \beta_1 H_{dc} + \beta_2 L_{dc} + \beta_3 T_{idc} \times H_{dc} + \beta_4 T_{idc} \times L_{dc} + X_i'\gamma + u_{id,c+1},\end{aligned} \tag{5}$$

where $Y_{id,c+1}$ is an indicator variable for individual i in district d participating in the next category $c+1$. $T_{idc} = \mathbf{1}(S_{idc} \geq 9)$ is an indicator variable for scoring equal or above the threshold in category c. S_{idc} is the number of points scored. H_{dc} and L_{dc} are the proportion of high-performing and low-performing peers in category c at district c, respectively. Finally, X_i is a vector of controls for gender, past participation, rank within the district in category c. Similar to the previous strategy, I include district-by-category, year and school fixed effects to account for factors that would impact all competitors at the district-category level.

Table 13 shows the results on how the effect of receiving the diploma is affected by being exposed to a higher proportion of high- and low-performing peers. Columns (1) and (2) show estimates from the most stringent definition of peers' type of performance using 2 and 3 points as bandwidths, respectively. These estimates indicate that exposure to more high-performing peers decreases the diploma effect while exposure to low-performing peers has no effect. In column (2), the estimate of -0.027 in the 6$^{\text{th}}$ row indicates that a 1 standard deviation increase in the share of high-performing peers decreases the diploma effect on subsequent participation by 2.7 percentage points. In contrast, the estimate of -0.008 in the 7$^{\text{th}}$ indicates that the diploma effect is unaffected by being exposed to a larger number of low-performing peers. The estimate for the interaction between receiving the diploma and the exposure to high-performing peers is statistically significant at the 5% level for the 2 points bandwidth, and at the 1% level for the 3 points bandwidth. These results are robust to all definitions of high- and low-performing competitors, and the inclusion of separate district and year fixed effects (see Table 14).

These results support the hypothesis that negative comparisons (with respect to larger maximum scores or higher proportion of high-performing peers) matter more than positive comparisons (with respect to lower minimum scores or higher proportion of low-performing peers) for a competitor's encouragement. This asymmetric relation is consistent with prior research in labour contexts showing that job satisfaction depends on relative pay comparisons, and this relationship is nonlinear (see Hamermesh (2001) and Card et al. (2012)).

5 Conclusions

I study the effect of non-monetary awards on future participation in mathematical Olympiads. By using a regression discontinuity design based on a strict cutoff that determines the provision of diplomas, I credibly isolate the effect of receiving the diploma from other types of feedback (absolute test score and rank), and observable and non-observable factors influencing future participation. I find that recognising children with symbolic awards positively affects their subsequent participation in contests.

Furthermore, I provide evidence that the magnitude of the diploma effect is affected by social comparisons. By exploiting variation in the ability distribution of districts over time, I provide two main findings. First, I find that the effect of receiving the diploma is lower in districts with a larger fraction of diploma recipients, which suggests that the diploma effect is shaped by status concerns. Second, I find that the diploma effect is weaker when exposed to higher maximum scores and higher proportions of high-performing peers while it does not experience gain or loss in districts with lower minimum scores or higher proportions of low-performing peers. These results suggest that competitors interpret the diploma as a signal of their abilities as they only compare themselves with their peers who are at the top but not with the ones at the bottom of the score distribution.

Finally, I find no gender differences in the effect of receiving the diploma on future competition, meaning that both boys and girls are equally affected by non-monetary awards. Moreover, descriptive evidence shows that boys and girls almost equally participate in mathematical competitions, and they dropout at the same rates conditional on past performance. These results are in contrast with the vast literature pointing out gender differences in competitive traits. I speculate that one potential explanation for this divergence lies on the fact that this study is based on children whose traits are not yet affected by gender norms, while most of this literature is based on studies with adolescents and adults.

Although the use of non-monetary awards is ubiquitous in many educational settings, it has been relatively unexplored in the literature of Economics of Education. The results of this study show that non-monetary awards have a large effect on engaging children's interest in mathematics and indicate these policies deserves further research. Moreover, this study also opens new research avenues about the design of non-monetary awards given in schools.

References

Agarwal, Ruchir and Patrick Gaule. 2020. "Invisible Geniuses: Could the Knowledge Frontier Advance Faster?" *American Economic Review: Insights* 2(4):409–424.

Ashraf, Nava, Oriana Bandiera and Kelsey Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics* 120:1–17.

Bedard, Kelly, Jacqueline Dodd and Shelly Lundberg. 2021. "Can Positive Feedback Encourage Female and Minority Undergraduates into Economics?" *AEA Papers and Proceedings* 111:128–132.

Benabou, Roland and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70(3):489–520.

Besley, Timothy and Maitreesh Ghatak. 2008. "Status Incentives." *American Economic Review: Papers and Proceedings* 98(2):206–211.

Buser, Thomas and Huaiping Yuan. 2019. "Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad." *American Economic Journal: Applied Economics* 11(3):225–252.

Card, David, Alexandre Mas, Enrico Moretti and Emmanuel Saez. 2012. "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." *American Economic Review* 102(6):2981–3003.

Cianci, Anna M., Howard J. Klein and Gerard H. Seijts. 2010. "The effect of negative feedback on tension and subsequent performance: The main and interactive effects of goal content and conscientiousness." *Journal of Applied Psychology* 95(4):618–630.

Cotofan, Maria. 2021. "Learning From Praise: Evidence From a Field Experiment With Teachers." *Journal of Public Economics* forthcoming:155–194.

Croson, Rachel and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2):448–474.

Ellison, Glenn and Ashley Swanson. 2010. "The gender gap in secondary school mathematics at high achievement levels: Evidence from the American mathematics competitions." *Journal of Economic Perspectives* 24(2):109–128.

Ellison, Glenn and Ashley Swanson. 2021. "Dynamics of the Gender Gap in High Math Achievement." *The Journal of Human Resources* Forthcoming.

Frey, Bruno. 2007. "Awards as Compensation." *European Management Review* 4:6–14.

Gallus, Jana. 2017. "Fostering Public Good Contributions with Symbolic Awards: A Large-Scale Natural Field Experiment at Wikipedia." *Management Science* 63(12):3999–4015.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69(1):201–209.

Hamermesh, Daniel S. 2001. "The Changing Distribution of Job Satisfaction." *Journal of Human Resources* 36(1):2981–3003.

Hoogveld, Nicky and Nick Zubanov. 2017. "The power of (no) recognition: Experimental evidence from the university classroom." *Journal of Behavioral and Experimental Economics* 67:75–84.

Mike, Anissa, Kelci Harris, Brent Roberts and Joshua Jackson. 2015. "Conscientiousness." *International Encyclopedia of the Social and Behavioral Sciences (Second Edition)* pp. 658–665.

Mouganie, Pierre and Yaojing Wang. 2020. "High-Performing Peers and Female STEM Choices in School." *Journal of Labor Economics* 38(3):805–841.

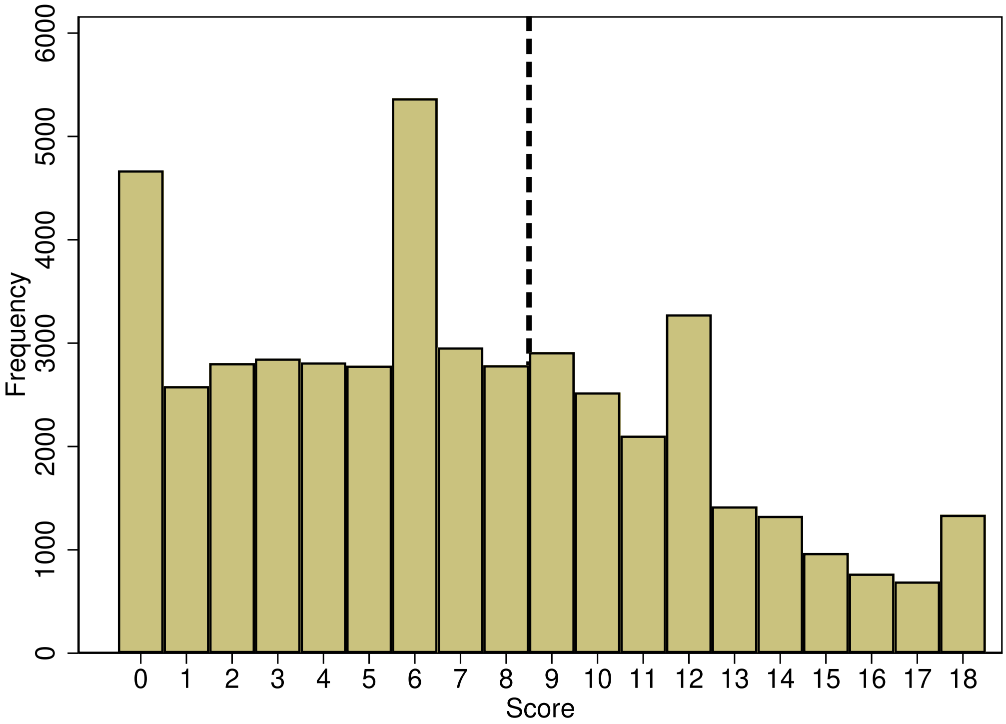
Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122(3):1067–1101.

Niederle, Muriel and Lise Vesterlund. 2010. "Explaining the Gender Gap in Math Test Scores: The Role of Competition." *Journal of Economic Perspectives* 24(2):129–144.

Villeval, Marie Claire. 2020. "Feedback policies and peer effects at work." *Handbook of Labor, Human Resources and Population Economics* .

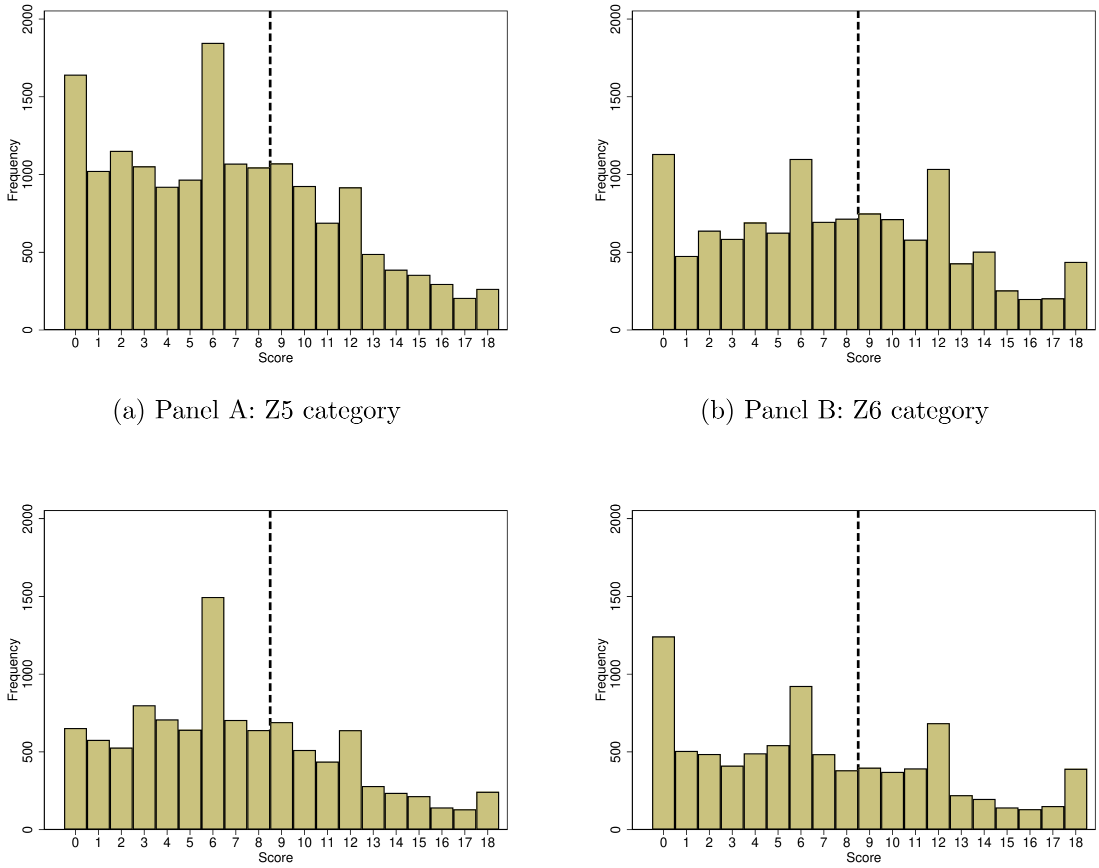
Figures

Figure 1: Histogram of Running Variable for RD Analysis



Notes: Sample is all student-test results from categories Z5 to Z8 between 2011-2017 (as in Table 1, column 1). Bars reported within bins of width 1. Sample size is 46,968.

Figure 2: Histogram of Running Variable for RD Analysis by categories



(a) Panel A: Z5 category

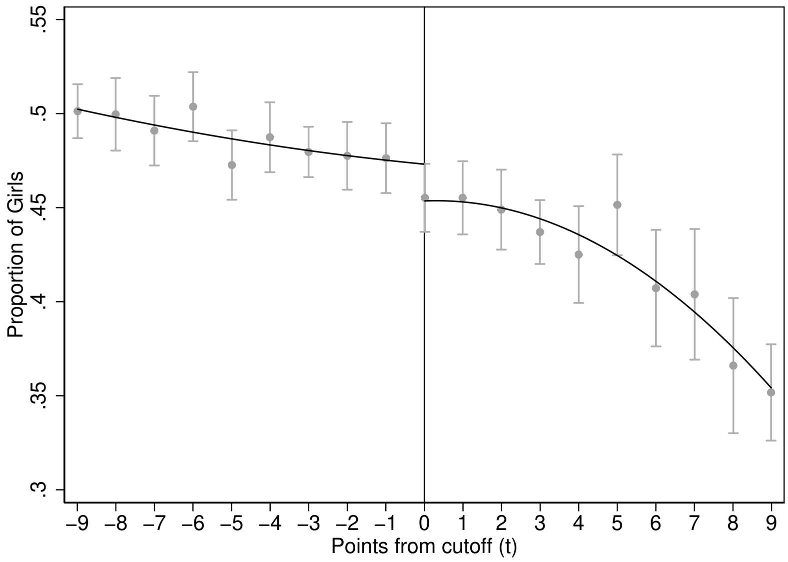
(b) Panel B: Z6 category

(c) Panel C: Z7 category

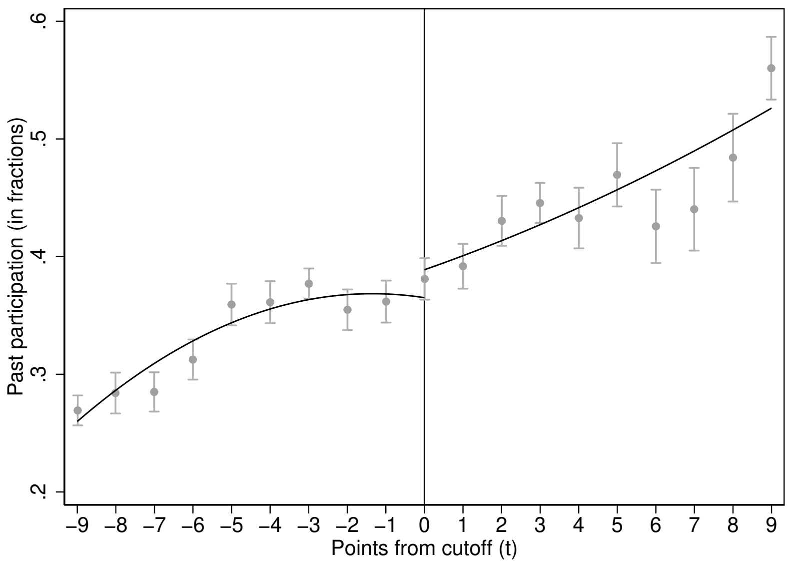
(d) Panel D: Z8 category

Notes: Samples is all student-test results for each category (Z5 to Z8) between 2011-2017 (as in Table 1, Panels B, C, D, and E). Bars reported within bins of width 1. Sample sizes for categories Z5, Z6, Z7, and Z8 are 16,334, 11,777, 10,291, and 8,566, respectively.

Figure 3: Predetermined Covariates by Points from Cutoff



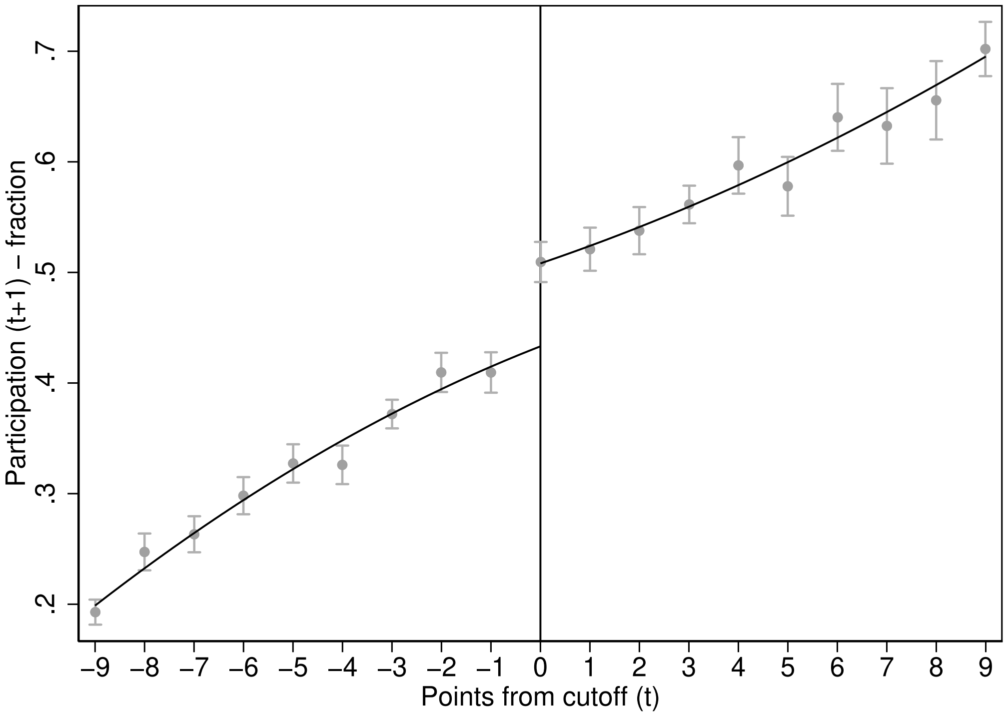
(a) Gender



(b) Past participation

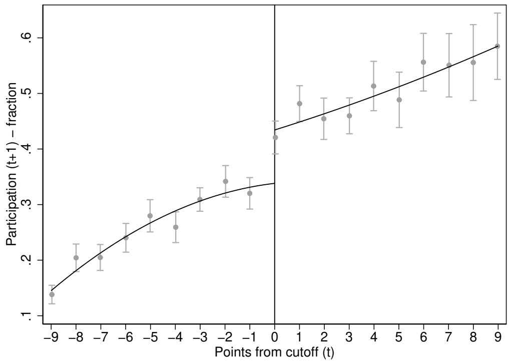
Notes: Vertical axis shows the average fraction of girls (Panel (a)), and fraction of students with previous experience (Panel (b)) who participated in the Olympiad. Horizontal axis shows normalised score S_i to the diploma cutoff, with $S_i \geq 0$ indicating the provision of diplomas and $S_i \leq 0$ indicating no diploma was awarded. Mean and 95% confidence intervals are shown for predetermined covariates within each bin (score). The figure also shows the estimated polynomial in points margin allowing for a discontinuity at the 0 margin. Sample is all student-test results in categories Z5 to Z8 between 2011-2017 (as in Table 1, column 2). Sample size is 46,968.

Figure 4: Diploma Effect on Future Participation

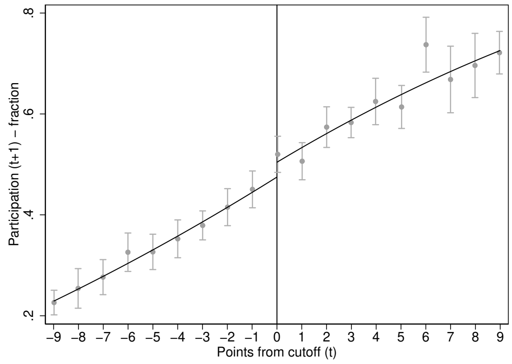


Notes: Vertical axis shows the fraction of students who participate in the Olympiads the year after. Horizontal axis shows normalised score S_i to the diploma cutoff, with $S_i \geq 0$ indicating the provision of diplomas and $S_i \leq 0$ indicating no diploma was awarded.. Mean and 95% confidence intervals are shown for the outcome within each bin (score). The figure also shows the estimated polynomial in points margin allowing for a discontinuity at the 0 margin. Sample is all student-test results in categories Z5 to Z8 between 2011-2017 (as in Table 1, column 2). Sample size is 46,968.

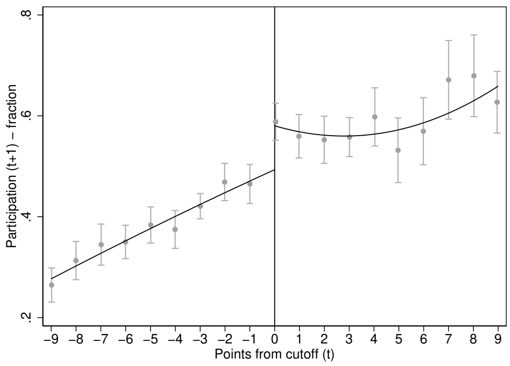
Figure 5: Diploma Effect on Future Participation by categories



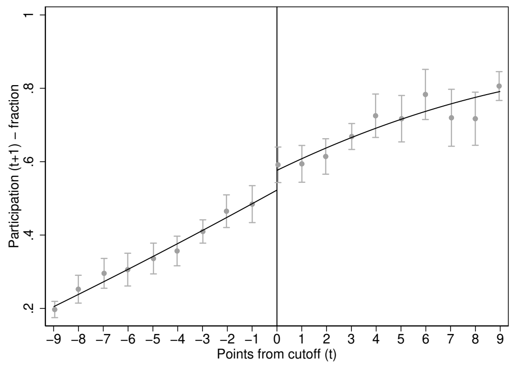
(a) Panel A: Z5 category



(b) Panel B: Z6 category



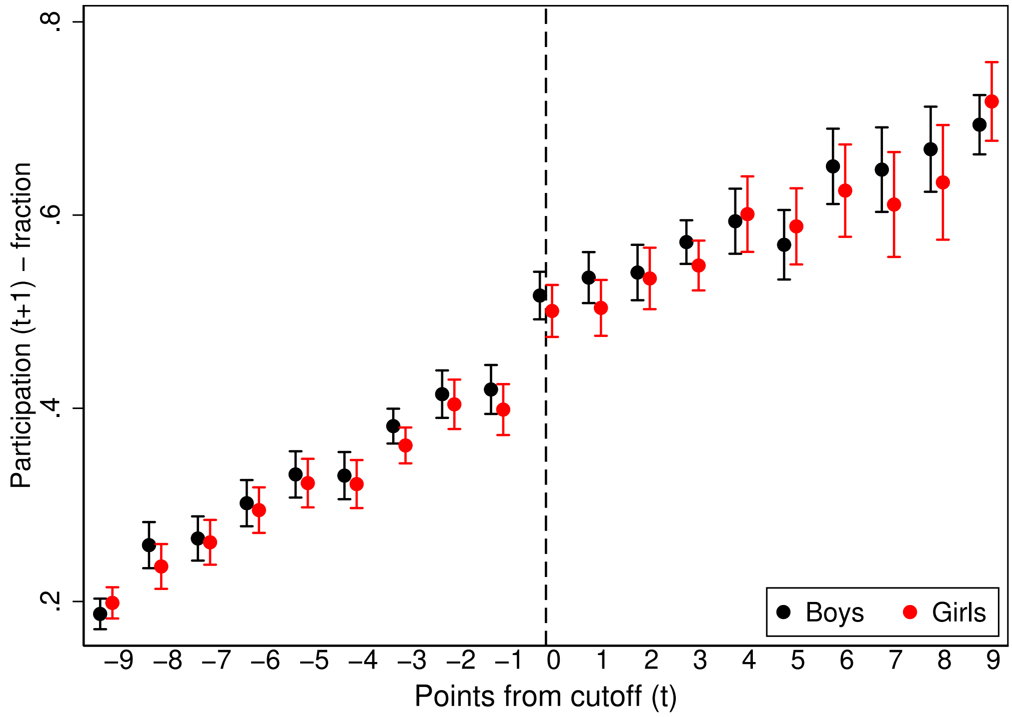
(c) Panel C: Z7 category



(d) Panel D: Z8 category

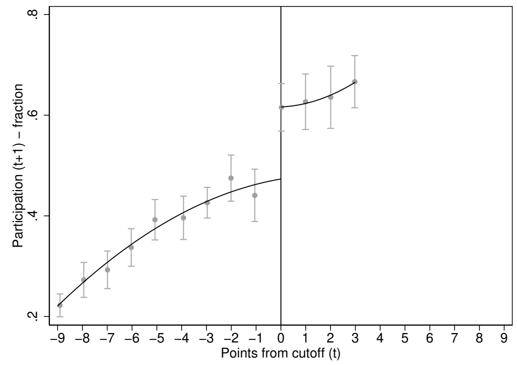
Notes: Vertical axis shows the fraction of students who participate in the Olympiads the following year by category. Horizontal axis shows normalised score S_i to the diploma cutoff, with $S_i \geq 0$ indicating the provision of diplomas and $S_i \leq 0$ indicating no diploma was awarded. Mean and 95% confidence intervals are shown for the outcome within each bin (score). Sample sizes for categories Z5, Z6, Z7, and Z8 are 16,334, 11,777, 10,291, and 8,566, respectively.

Figure 6: Diploma Effect on Future Participation by Gender

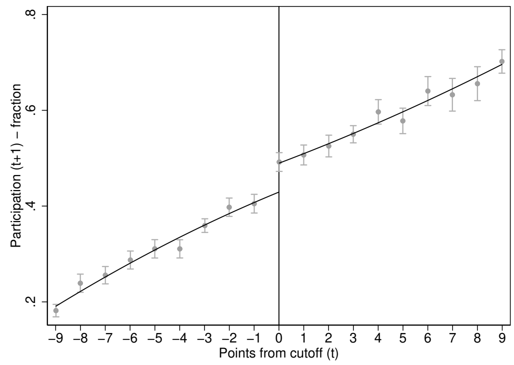


Notes: Vertical axis shows the fraction of students who participate in the Olympiads the following year by gender. Horizontal axis shows normalised score S_i to the diploma cutoff, with $S_i \geq 0$ indicating the provision of diplomas and $S_i \leq 0$ indicating no diploma was awarded. Mean and 95% confidence intervals are shown for the outcome within each bin (score). Sample is all student-test results in categories Z5 to Z8 between 2011-2017 (as in Table 1, column 2). Sample size for boys and girls are 25,032 and 21,936, respectively.

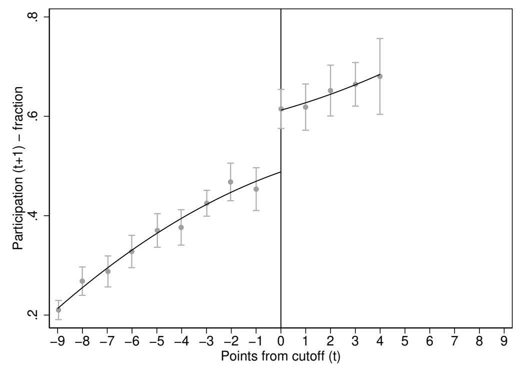
Figure 7: Diploma Effect on Future Participation by Maximum Score Achieved in Districts



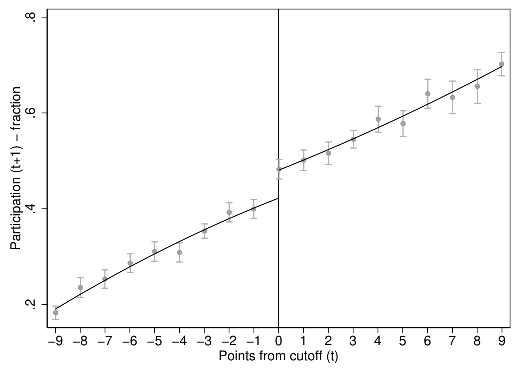
(a) Max district score< 13 (margin< 4)



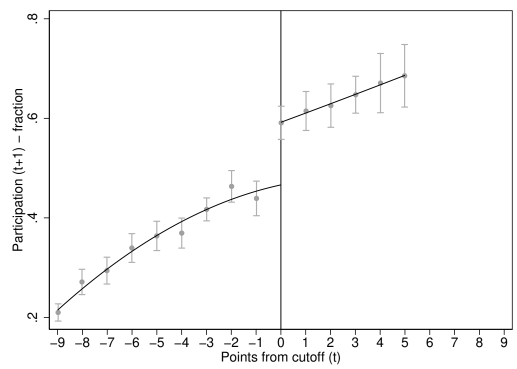
(b) Max district score≥ 13 (margin≥ 4)



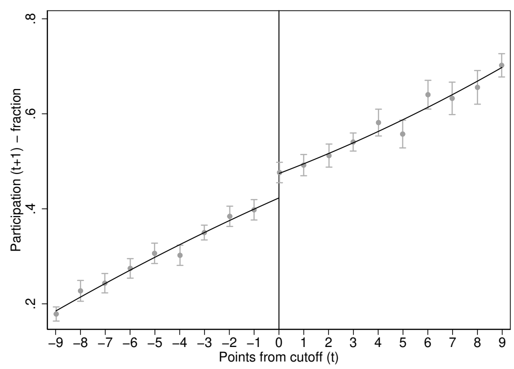
(c) Max district score< 14 (margin< 5)



(d) Max district score≥ 14 (margin≥ 5)



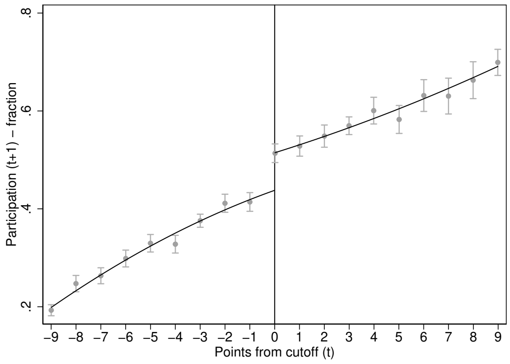
(e) Max district score< 15 (margin< 6)



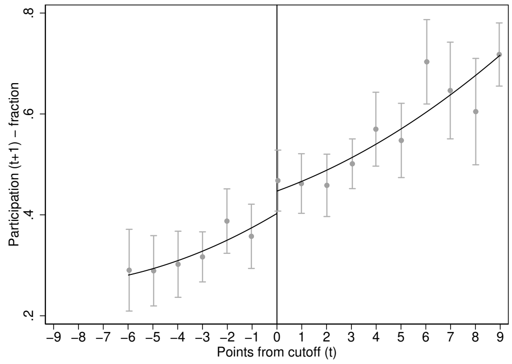
(f) Max district score≥ 15 (margin≥ 6)

Notes: Vertical axis shows the fraction of students who participate in the following category. Horizontal axis shows normalised score S_i to the diploma cutoff, with $S_i \geq 0$ indicating the provision of diplomas and $S_i \leq 0$ indicating no diploma was awarded. Mean and 95% confidence intervals are shown for the outcome within each bin (score).

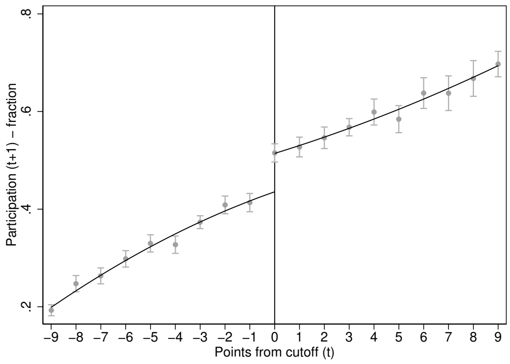
Figure 8: Diploma Effect on Future Participation by Minimum Score Achieved in Districts



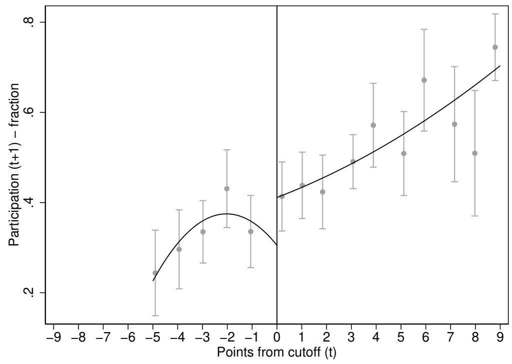
(a) Min district score< 3 (margin< -6)



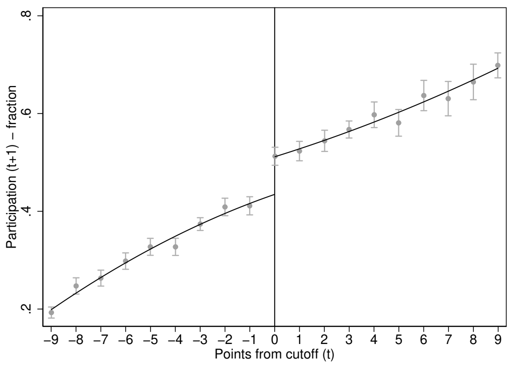
(b) Min district score≥ 3 (margin≥ -6)



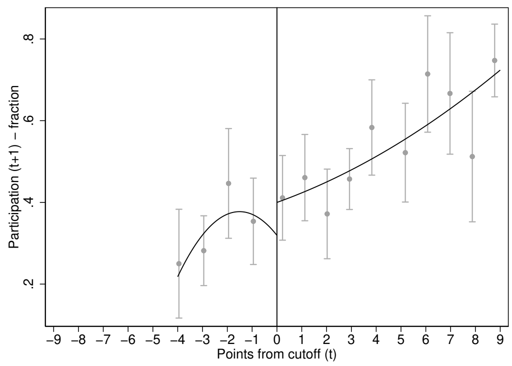
(c) Min district score< 4 (margin< -5)



(d) Min district score≥ 4 (margin≥ -5)



(e) Min district score< 5 (margin< -4)



(f) Min district score≥ 5 (margin≥ -4)

Notes: Vertical axis shows the fraction of students who participate in the Olympiads in the next category. Horizontal axis shows normalised score S_i to the diploma cutoff, with $S_i \geq 0$ indicating the provision of diplomas and $S_i \leq 0$ indicating no diploma was awarded. Mean and 95% confidence intervals are shown for the outcome within each bin (score).

Table 1: Sample Characteristics

		RD samples (points from cutoff)	
	All results (1)	+/-2 points (2)	+/-3 points (3)
Panel A. All transitions			
Past participation (percent)	37	37	38
Female (percent)	47	47	47
Hungarian (percent)	8	8	8
Grammar school (percent)	14	14	14
Observations	46,968	11,182	18,656
Panel B. 5th to 6th			
Past participation (percent)	0	0	0
Female (percent)	47	47	47
Hungarian (percent)	8	7	8
Grammar school (percent)	5	5	5
Observations	16,334	4,115	6,653
Panel C. 6th to 7th			
Past participation (percent)	46	48	47
Female (percent)	47	48	48
Hungarian (percent)	8	8	8
Grammar school (percent)	18	17	17
Observations	11,777	2,876	4,558
Panel D. 7th to 8th			
Past participation (percent)	61	63	63
Female (percent)	47	46	47
Hungarian (percent)	8	8	9
Grammar school (percent)	19	19	19
Observations	10,291	2,552	4,487
Panel E. 8th to 9th			
Past participation (percent)	66	71	71
Female (percent)	46	44	44
Hungarian (percent)	9	7	8
Grammar school (percent)	18	20	19
Observations	8,566	1,639	2,958

Notes: Sample in column 1 includes one observation per student-test (categories Z5 to Z8) between 2011 and 2017. Sub-samples in columns 2-4 include student-test observations whose scores fell within +/- 2 and 3 points from the cutoff, respectively.

Table 2: The Diploma Effect on Predetermined Characteristics of Competitors

	Gender		Past participation	
	RD sample		RD sample	
	+/-2pts	+/-3pts	+/-2 pts	+/-3pts
	(1)	(2)	(3)	(4)
$1(S \geq 9)$	-0.0188	-0.0215**	0.00349	0.01676
	(0.0104)	(0.00760)	(0.00304)	(0.00856)
S	0.00137	-0.00155	0.01881***	0.00855**
	(0.00146)	(0.00324)	(0.00155)	(0.00147)
$S \times 1(S \geq 9)$	-0.0111	-0.00121	-0.01285*	0.00105
	(0.0154)	(0.00253)	(0.00406)	(0.00303)
θ_{dtc}	Yes	Yes	Yes	Yes
θ_s	Yes	Yes	Yes	Yes
Observations	10,912	18,522	10,912	18,522
R2	0.255	0.188	0.533	0.503

Notes: This table contains regression discontinuity based estimates of the effect of receiving the diploma on predetermined characteristics for two different bandwidths. Gender is an indicator variable indicating whether the participant is a girl, and past participation is an indicator variable indicating whether the participant participated in a previous category. Parentheses contain standard errors clustered at the participant level.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level

Table 3: The Diploma Effect on Subsequent Participation

	RD samples			
	+/-2 points		+/-3 points	
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	0.0526***	0.0337**	0.0388**	0.0388***
	(0.00805)	(0.0104)	(0.0121)	(0.00755)
$Score$	0.0145***	0.0253	0.0276***	0.0285***
	(0.00108)	(0.0117)	(0.00326)	(0.00440)
$Score \times 1(Score \geq 9)$	0.00826*	-0.0109	-0.0152***	-0.0145***
	(0.00329)	(0.00937)	(0.00349)	(0.00342)
Constant	0.346***	0.316**	0.392***	0.380***
	(0.0265)	(0.0543)	(0.0266)	(0.0295)
θ_d	Yes	No	Yes	No
θ_t	Yes	No	Yes	No
θ_c	Yes	No	Yes	No
θ_s	Yes	Yes	Yes	Yes
θ_{dtc}	No	Yes	No	Yes
Observations	11,054	10,912	18,567	18,522
Peer groups		1,553		1,714
R2	0.255	0.478	0.216	0.406

Notes: This table contains the estimates of receiving the diploma on future participation based on regression discontinuity regression. Controls included in the regressions are gender of competitors, whether they participated before, and their rank within their district. Parentheses contain standard errors doubled clustered at the participant and points levels.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level

Table 4: The Diploma Effect on Subsequent Participation by Gender

	RD samples			
	+/-2 points		+/-3 points	
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	0.0539*	0.0167	0.0466**	0.0326***
	(0.0170)	(0.0102)	(0.0141)	(0.0079)
$Score$	0.0161*	0.0337**	0.0262***	0.0280***
	(0.0067)	(0.0091)	(0.0033)	(0.0039)
$Score \times 1(Score \geq 9)$	0.0052	-0.0329**	-0.0191***	-0.0195**
	(0.0074)	(0.0057)	(0.0033)	(0.0056)
$Girl$	-0.0209	-0.0491***	-0.0095*	-0.0214**
	(0.0216)	(0.0061)	(0.0046)	(0.0080)
$Girl \times 1(Score \geq 9)$	-0.0028	0.0355***	-0.0169*	0.0127
	(0.0204)	(0.0061)	(0.0081)	(0.0089)
$Girl \times Score$	-0.0034	-0.0167*	0.0029	0.0016
	(0.0139)	(0.0059)	(0.0022)	(0.0039)
$Girl \times S \times 1(S \geq 9)$	0.0066	0.0469**	0.0089*	0.0108
	(0.0204)	(0.0133)	(0.0038)	(0.0072)
Constant	0.347***	0.338***	0.389***	0.385***
	(0.0175)	(0.0479)	(0.0261)	(0.0281)
θ_d	Yes	No	Yes	No
θ_t	Yes	No	Yes	No
θ_c	Yes	No	Yes	No
θ_s	Yes	Yes	Yes	Yes
θ_{dtc}	No	Yes	No	Yes
Observations	11,054	10,912	18,567	18,522
Peer groups		1,553		1,714
R2	0.255	0.478	0.217	0.407

Notes: This table contains the estimates of the effect of receiving the diploma on future participation by gender of participants based on the regression discontinuity design. All regressions control for past participation and rank within the district. Parentheses contain standard errors doubled clustered at the participant and points levels.
*** Significant at the 1 percent level
** Significant at the 5 percent level
* Significant at the 10 percent level

Table 5: The Diploma Effect on Subsequent Participation and Diploma Scarcity

	RD samples			
	+/-2 points		+/-3 points	
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	0.114***	0.114***	0.102***	0.101***
	(0.00930)	(0.0167)	(0.00841)	(0.00906)
$Score$	-0.00898**	-0.00601	0.0186*	0.0196*
	(0.00279)	(0.00439)	(0.00775)	(0.00902)
$Score \times 1(Score \geq 9)$	0.0149**	0.00845	-0.0108**	-0.0143***
	(0.00395)	(0.00742)	(0.00390)	(0.00329)
$Fraction$	0.454***	0.483**	0.194	0.196
	(0.0641)	(0.0964)	(0.116)	(0.143)
$1(Score \geq 9) \times Fraction$	-0.189***	-0.197**	-0.167***	-0.161***
	(0.0313)	(0.0367)	(0.0276)	(0.0360)
Constant	-0.0501	-0.0766	0.237	0.236
	(0.0523)	(0.0878)	(0.118)	(0.145)
Separate FE	Yes	No	Yes	No
Aggregate FE	No	Yes	No	Yes
Observations	10,821	10,793	18,158	18,154
R2	0.257	0.323	0.218	0.268

Notes: This table contains the estimates of the effect of receiving the diploma on future participation and its interaction with the fraction of diploma recipients in the district based on the regression discontinuity design. All regressions control for past participation and rank within the district. Parentheses contain standard errors doubled clustered at the participant and points levels.
*** Significant at the 1 percent level
** Significant at the 5 percent level
* Significant at the 10 percent level

Table 6: Sample Characteristics - Upper Score Comparison

	Break point=13		Break point=14		Break point=15	
	< 13	≥ 13	< 14	≥ 14	< 15	≥ 15
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Past participation (percent)	46	35	43	35	41	35
Female (percent)	48	46	48	46	48	46
Hungarian (percent)	7	8	6	9	6	9
Grammar school (percent)	11	14	10	15	10	15
Observations	7,290	39,678	10,276	36,692	13,976	32,992

Notes: This table describes competitors' characteristics under three different classifications of districts depending on the maximum score achieved. Columns (1) and (2) correspond to the partition based on a break point of 13 points. For instance, column (1) groups all competitors coming from districts where the maximum score obtained by a competitor was 12 or less, while column (2) groups all competitors coming from districts where the maximum score obtained by a competitor was 13 or more. Columns (3) and (4) correspond to the partition based on a break point of 14 points, while columns (5) and (6) correspond to the partition based on a break point of 15 points.

Table 7: Sample Characteristics - Lower Score Comparison

	Break point=3		Break point=4		Break point=5	
	< 3	≥ 3	< 4	≥ 4	< 5	≥ 5
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Past participation (percent)	36	48	36	48	37	50
Female (percent)	47	47	47	46	47	46
Hungarian (percent)	8	4	8	2	8	3
Grammar school (percent)	14	15	14	15	14	16
Observations	43,640	3,328	45,027	1,941	45,883	1,085

Notes: This table describes competitors' characteristics under three different classifications of districts depending on the minimum score achieved. Columns (1) and (2) correspond to the partition based on a break point of 3 points. For instance, column (1) groups all competitors coming from districts where the minimum score obtained by a competitor was 2 or less, while column (2) groups all competitors coming from districts where the minimum score obtained by a competitor was 3 or more. Columns (3) and (4) correspond to the partition based on a break point of 4 points, while columns (5) and (6) correspond to the partition based on a break point of 5 points.

Table 8: Tests for Random Assignment of Maximum and Minimum Scores in District

	RD samples			
	+/-2 points		+/-3 points	
	Max	Min	Max	Min
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	-0.0652	0.0033	-0.0440	-0.0226
	(0.0389)	(0.0184)	(0.0406)	(0.0218)
$Score$	0.0927**	0.0283	0.0552**	0.0438***
	(0.0215)	(0.0121)	(0.0149)	(0.0060)
$Score \times 1(Score \geq 9)$	-0.0256	-0.0504*	0.0349	-0.0292*
	(0.0238)	(0.0176)	(0.0223)	(0.0120)
Observations	10,626	10,626	17,815	17,815
R2	0.587	0.566	0.564	0.522

Notes: This table contains the estimated effect of receiving the diploma on the maximum and minimum scores achieved in a district. All regressions include district-by-year, category and school fixed effects. Parentheses contain standard errors doubled clustered at the participant and points levels.
*** Significant at the 1 percent level
** Significant at the 5 percent level
* Significant at the 10 percent level

Table 9: Diploma Effect and the Ability Gap

	RD samples			
	+/-2 points		+/-3 points	
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	0.227**	0.178**	0.220***	0.161***
	(0.041)	(0.043)	(0.032)	(0.032)
$Score$	0.019***	0.010**	0.029***	0.021***
	(0.001)	(0.003)	(0.003)	(0.002)
$Score \times 1(Score \geq 9)$	0.011*	0.002	-0.011**	-0.012***
	(0.004)	(0.003)	(0.004)	(0.003)
$Score^{max}$	-0.008	-0.005	-0.011***	-0.008**
	(0.004)	(0.004)	(0.003)	(0.003)
$Score^{min}$	0.013***	0.001	0.010***	-0.005
	(0.002)	(0.002)	(0.002)	(0.004)
$Score^{max} \times 1(Score \geq 9)$	-0.011**	-0.008**	-0.011***	-0.008***
	(0.003)	(0.003)	(0.002)	(0.002)
$Score^{min} \times 1(Score \geq 9)$	-0.000	0.005	-0.006	-0.002
	(0.005)	(0.004)	(0.006)	(0.007)
Constant	0.497***	0.346**	0.591***	0.441***
	(0.083)	(0.075)	(0.059)	(0.062)
θ_d	Yes	No	Yes	No
θ_t	Yes	No	Yes	No
θ_c	Yes	No	Yes	No
θ_s	Yes	Yes	Yes	Yes
θ_{dc}	No	Yes	No	Yes
Observations	10,665	10,656	17,845	17,842
R2	0.258	0.381	0.220	0.336

Notes: This table contains the estimated effect of receiving the diploma on future participation depending on score comparisons. All regressions control for gender, past participation and rank within the district. Parentheses contain standard errors doubled clustered at the participant and points levels.
*** Significant at the 1 percent level
** Significant at the 5 percent level
* Significant at the 10 percent level

Table 10: Tests for Random Assignment of High- and Low-performing Peers (Definition 1)

	RD samples			
	+/-2 points		+/-3 points	
	High (1)	Low (2)	High (3)	Low (4)
$1(Score \geq 9)$	-0.0008	-0.0001	0.0004	0.0040*
	(0.0006)	(0.0015)	(0.0007)	(0.0016)
$Score$	0.0022***	-0.0021*	0.0011***	-0.0055***
	(0.0004)	(0.0008)	(0.0003)	(0.0007)
$Score \times 1(Score \geq 9)$	-0.0012*	0.0007	-0.0002	0.0037***
	(0.0005)	(0.0009)	(0.0003)	(0.0009)
Observations	10,626	10,626	17,815	17,815
R2	0.565	0.582	0.538	0.543

Notes: This table contains the estimated effect of receiving the diploma on the fraction of high- and low-performing peers in the district. High-performing peer is defined as the competitor who scores 18 points while a low-performing peer scores 0 points. All regressions include district-by-year, category and school fixed effects. Parentheses contain standard errors doubled clustered at the participant and points levels.
*** Significant at the 1 percent level
** Significant at the 5 percent level
* Significant at the 10 percent level

Table 11: Tests for Random Assignment of High- and Low-performing Peers (Definition 2)

	RD samples			
	+/-2 points		+/-3 points	
	High	Low	High	Low
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	-0.0000	0.0024	0.0009	0.0046**
	(0.0011)	(0.0015)	(0.0013)	(0.0012)
$Score$	0.0026**	-0.0055***	0.0011**	-0.0073***
	(0.0007)	(0.0007)	(0.0004)	(0.0005)
$Score \times 1(Score \geq 9)$	-0.0019*	0.0028	0.0012	0.0048***
	(0.0008)	(0.0012)	(0.0006)	(0.0008)
Observations	10,626	10,626	17,815	17,815
R2	0.583	0.578	0.559	0.544

Notes: This table contains the estimated effect of receiving the diploma on the fraction of high- and low-performing peers in the district. High-performing peer is defined as the competitor who scores at least 17 points while a low-performing peer scores at most 1 point. All regressions include district-by-year, category and school fixed effects. Parentheses contain standard errors doubled clustered at the participant and points levels.
*** Significant at the 1 percent level
** Significant at the 5 percent level
* Significant at the 10 percent level

Table 12: Tests for Random Assignment of High- and Low-performing Peers (Definition 3)

	RD samples			
	+/-2 points		+/-3 points	
	High	Low	High	Low
	(1)	(2)	(3)	(4)
$1(Score \geq 9)$	-0.0007	-0.0016	0.0015	0.0010
	(0.0010)	(0.0019)	(0.0014)	(0.0015)
$Score$	0.0034**	-0.0054***	0.0013*	-0.0072***
	(0.0006)	(0.0008)	(0.0005)	(0.0006)
$Score \times 1(Score \geq 9)$	-0.0005	0.0029	0.0013*	0.0034**
	(0.0008)	(0.0015)	(0.0006)	(0.0010)
Observations	10,626	10,626	17,815	17,815
R2	0.576	0.577	0.558	0.539

Notes: This table contains the estimated effect of receiving the diploma on the fraction of high- and low-performing peers in the district. High-performing peer is defined as the competitor who scores at least 16 points while a low-performing peer scores at most 2 points. All regressions include district-by-year, category and school fixed effects. Parentheses contain standard errors doubled clustered at the participant and points levels.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level

Table 13: The Diploma Effect and the Ability Gap (High- and Low-performing Peers) - district-by-year fixed effects

	I: $H \geq 18pts, L \leq 0pts$		II: $H \geq 17pts, L \leq 1pt$		III: $H \geq 16pts, L \leq 2pts$	
	+/-2	+/-3	+/-2	+/-3	+/-2	+/-3
	(1)	(2)	(3)	(4)	(5)	(6)
$1(score \geq 9)$	0.039**	0.038**	0.041**	0.038**	0.042**	0.038**
	(0.010)	(0.011)	(0.010)	(0.011)	(0.009)	(0.011)
$score$	0.020***	0.027***	0.019***	0.026***	0.018***	0.026***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)
$1(score \geq 9) \times score$	0.003	-0.016***	0.003	-0.015***	0.005	-0.015***
	(0.007)	(0.003)	(0.006)	(0.003)	(0.006)	(0.004)
$High$	-0.002	-0.006	-0.018	-0.020**	-0.022	-0.026**
	(0.013)	(0.007)	(0.014)	(0.008)	(0.011)	(0.009)
Low	-0.021**	-0.026***	-0.039*	-0.036***	-0.050**	-0.041***
	(0.007)	(0.005)	(0.014)	(0.007)	(0.011)	(0.009)
$1(score \geq 9) \times High$	-0.023*	-0.027***	-0.023*	-0.026***	-0.018**	-0.022**
	(0.009)	(0.006)	(0.009)	(0.006)	(0.005)	(0.006)
$1(score \geq 9) \times Low$	-0.022**	-0.008	-0.012	-0.007	-0.001	-0.002
	(0.005)	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)
Constant	0.332***	0.375***	0.326***	0.371***	0.309***	0.364***
	(0.038)	(0.027)	(0.042)	(0.029)	(0.033)	(0.034)
Observations	10,793	18,154	10,793	18,154	10,793	18,154
R2	0.322	0.269	0.324	0.271	0.324	0.271

Notes: This table contains the estimated effect of receiving the diploma on future participation depending on score comparisons. All regressions control for gender, past participation and rank within the district. Parentheses contain standard errors doubled clustered at the participant and points levels. ***, **, * Significant at the 1, 5 and 10 percent level, respectively.

Table 14: The Diploma Effect and the Ability Gap (High- and Low-performing Peers) - separate fixed effects

	I: $H \geq 18pts, L \leq 0pts$		II: $H \geq 17pts, L \leq 1pt$		III: $H \geq 16pts, L \leq 2pts$	
	+/-2	+/-3	+/-2	+/-3	+/-2	+/-3
	(1)	(2)	(3)	(4)	(5)	(6)
$1(score \geq 9)$	0.047***	0.038**	0.049***	0.038**	0.050***	0.039**
	(0.007)	(0.011)	(0.007)	(0.011)	(0.007)	(0.011)
$score$	0.012**	0.024***	0.012**	0.024***	0.011**	0.024***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
$1(score \geq 9) \times score$	0.009	-0.013**	0.010*	-0.012**	0.012**	-0.012**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$High$	0.012	0.008	-0.001	-0.001	-0.009	-0.009
	(0.012)	(0.007)	(0.012)	(0.007)	(0.013)	(0.008)
Low	-0.014	-0.021**	-0.025*	-0.024***	-0.037**	-0.031***
	(0.006)	(0.006)	(0.008)	(0.005)	(0.009)	(0.006)
$1(score \geq 9) \times High$	-0.023**	-0.026***	-0.023**	-0.026***	-0.020**	-0.023**
	(0.006)	(0.004)	(0.006)	(0.005)	(0.005)	(0.006)
$1(score \geq 9) \times Low$	-0.023**	-0.005	-0.014*	-0.004	-0.005	0.001
	(0.006)	(0.010)	(0.005)	(0.006)	(0.005)	(0.006)
Constant	0.321***	0.360***	0.319***	0.363***	0.309***	0.358***
	(0.044)	(0.026)	(0.044)	(0.027)	(0.042)	(0.027)
Observations	10,821	18,158	10,821	18,158	10,821	18,158
R2	0.257	0.218	0.257	0.218	0.258	0.219

Notes: This table contains the estimated effect of receiving the diploma on future participation depending on score comparisons. All regressions control for gender, past participation and rank within the district. Parentheses contain standard errors doubled clustered at the participant and points levels. ***, **, * Significant at the 1, 5 and 10 percent level, respectively.