Training Teachers for Diversity Awareness: Impact on School Outcomes of Refugee Children [*]

Semih Tumen[†]
TED University and IZA

Michael Vlassopoulos[‡]
University of Southampton and IZA

Jackline Wahba[§]
University of Southampton and IZA

October 12, 2022

Abstract

Despite efforts to integrate refugee children into host country education systems, their low school attachment, poor academic performance, and high drop-out rates remain major policy challenges. Teachers can have important impact on these student outcomes, yet classroom diversity poses difficulties for teachers who might not always be adequately prepared to address the needs of minority students. Using administrative data and a regression discontinuity approach, we evaluate whether a teacher training program—designed to raise awareness of primary and secondary school teachers in Turkey based on a cascade-training approach—is effective in improving school outcomes of refugee students. We find that the program almost halves the absenteeism gap between native and refugee students and its effect persists into the next academic year, albeit fading in size. We also find that the program improves the grades of refugee students in Turkish language and Math subjects, and that there is a positive association between improved attendance and grades. We argue that the most likely channel through which the effects of the program operate is a school-wide champion role assumed by trained teachers, which has a broad impact on raising diversity awareness within schools.

JEL codes: I21, I28, J15.
Keywords: Teacher training, refugees, absenteeism, academic achievement, diversity.

[*]We thank Matias Cattaneo, Jamal Haidar, Pinar Ozel, Serdar Polat, Metin Catar, Secil Balkir, and seminar participants at the American University of Cairo for useful comments.

[†]TED University, Department of Economics, Turkey. Email: semih.tumen@tedu.edu.tr.
[‡]University of Southampton, Department of Economics, UK. Email: m.vlassopoulos@soton.ac.uk.
[§]University of Southampton, Department of Economics, UK. Email: j.wahba@soton.ac.uk.

1 Introduction

In 2019, children constituted 40% of the 79.5 million internally displaced and refugee individuals worldwide (UNHCR, 2019). The large number of refugee children of school age has put considerable pressure on many host countries to integrate them in their educational systems. However, refugee children face various difficulties and barriers to attending school in host countries, including language barriers, ethnic and cultural differences, financial pressures, and displacement-related trauma. Consequently, they have a rather tenuous attachment to education with lower school engagement and higher drop out rates than native children, particularly in secondary education—see, for example, Dryden-Peterson (2015) and Sieverding et al. (2018).[1] This can have dramatic consequences for the academic development of refugee children, creating gaps in their skills and knowledge with significant lifetime implications for their labor market prospects.

Despite the importance of refugee educational integration, there is little evidence on the educational outcomes of refugee children. In fact, the vast majority of existing research, which we review further on, focuses on the impact of refugee children on the educational outcomes of native children. This is to a large extent due to the lack of suitable micro-level data capturing educational outcomes of refugee students, and the perceived temporary nature of refugees' stay in host countries. Furthermore, the role of teachers in refugee educational integration remains an understudied issue. Teachers are potentially key actors in facilitating the integration of refugee children into the education systems of host countries and fostering their academic advancement. Yet, they might not always be adequately prepared for instructing/managing diverse classrooms, fully aware of the circumstances of refugee children, and able to address their needs. This paper aims to fill these gaps by studying a large-scale teacher training program in Turkey designed to facilitate the educational integration of Syrian refugee children, using rich administrative micro-level data. Our main contribution is to provide novel causal evidence regarding the key role that training and professional development of teachers can play in addressing the low school attachment and academic achievement of refugee children.

The teacher training program that we study was implemented during the semester break of the school year 2017-18 by the Turkish Ministry of National Education (MoNE) and was funded by the EU. The main goals of the training program were to increase teachers' awareness toward the immediate needs of refugee students and to encourage/equip them to actively engage in educational integration policies. We evaluate whether the training led to an improvement in refugee children's school attendance and learning outcomes. We use a large administrative micro-level data set encompassing the universe of primary and secondary schools in two Turkish provinces (Gaziantep and Sanliurfa) that are the largest border provinces with Syria, hosting a total of 872,000 Syrian refugees—corresponding to, on average, a 21% refugee to population ratio.

[1] Only 61% of refugee children attend primary school compared to 91% globally, and at secondary level, only 23% of refugee youth are enrolled, compared to 84% globally (UNHCR, 2017).

Our main outcome of interest is school absenteeism of refugee students, as keeping refugee children in school is widely agreed to be the most important step toward fully integrating refugee children into host-country education systems (UNHCR, 2020). In the setting that we study, the absenteeism rate of refugee children is more than double that of native children. Refugee children face various obstacles in adapting to the host-country education systems.[2] In fact, the average refugee student is chronically absent from school (being absent from school for at least 10% of the school days), which is widely used as a critical early warning signal for academic risk and school dropout (Bruner et al., 2011). More generally, school absenteeism is an important concern, especially for disadvantaged groups, as it has been linked to adverse outcomes, such as low academic performance, substantial learning losses, and high drop-out rates (Aucejo and Romano, 2016; Gershenson et al., 2017). Furthermore, absenteeism is a proxy for non-cognitive skills (Jackson, 2018), and is negatively correlated with the major components of those skills (Lounsbury et al., 2004). School absenteeism is also strongly related to various risky behaviors, such as drug/alcohol abuse and smoking, and other life-course problems, such as antisocial behavior, adolescent pregnancy, and juvenile delinquency (Gubbels et al., 2019). We also leverage Turkish language and Math grades of Syrian students to investigate how the teacher training program might have affected academic performance. Finally, we investigate the mechanisms through which the program effects operate.

Our identification strategy exploits the rules that determine the recruitment of teachers into the training program. These rules create a set of discontinuities in the probability that a school will have trained teachers and allow for employing a regression discontinuity (RD) analysis. The running variable is the number of refugee students in each school, which is a discrete variable. We adopt a local randomization RD framework, which is the recommended approach to deal with the discrete running variable that characterizes our setting (Cattaneo et al., 2018). The teacher training program is designed to provide training to teachers in schools with at least 15 refugee students. The number of teachers to be trained in each school is a function of the running variable. Schools with 15 to 19 refugee students have one trained teacher. The number of trained teachers is incrementally increased up to 6 teachers per school as the running variable crosses higher thresholds (20, 40, 60, 80, and 100), in order to maintain a stable ratio of refugee children to trained teachers across schools. In other words, there is an extensive margin of the treatment around the first threshold and from there on the intensive margin of the treatment is kept relatively constant across higher thresholds. We use these discontinuities to identify the causal effect of the training program on refugee students' school outcomes.

[2] These include language barriers, cultural and socio-economic differences, inevitable breaks in school education during the process of seeking refuge, displacement-related trauma, mental distress, vulnerability, as well as adjustment to a new country. In particular, refugee families face financial pressures, which may push their children (especially boys) into local labor markets, so they may end up missing classes or dropping out of school. Also, teenage marriages may prevent young refugee girls from attending school. The complex nature of these issues suggests that any policy attempt aiming to increase school attachment of refugee children requires acute awareness about those underlying problems.

We find that the training program led to a significant reduction in refugee students' school absenteeism. In particular, it reduced refugee students' school absence by around 2.7 days per semester, almost halving the absenteeism gap between native and refugee students. We also find that the impact of the training program on students' absenteeism persisted into the first semester of the following academic year, albeit the effect was less pronounced—around 1.5 days. Importantly, we find that the reduction in absenteeism comes entirely from the first cutoff, which is the extensive margin of program eligibility for a school. As the intensity of treatment (ratio of refugee students per trained teacher) does not vary much across higher cutoffs, we do not find that training additional teachers leads to further reductions in absenteeism. Additional analysis indicates that the effects are present for both genders and for both primary and secondary students. Finally, we document that the training program reduced chronic absenteeism—e.g., being absent from school for at least 10% of school days in an academic year—and school drop-out rates.

A number of validation and falsification tests provides reassurance about the integrity of these results. In particular, we perform balancing checks of student and school characteristics around the cutoffs, which show that schools and students on each side of the cutoffs do not differ in terms of observable characteristics. Furthermore, we show that there is no significant RD effect around the cutoffs, when we examine absenteeism in the semester immediately before the training took place. We also find no impact of the training on school absenteeism of Turkish students, as one would expect, as the training was designed to address the needs of refugee students.

We provide suggestive evidence that the most likely channel through which the teacher training program reduced refugee students' school absenteeism was what we refer to as an "ambassador" or "champion" effect. The program raised teachers' awareness toward the needs of refugee students, encouraged teachers to act as school-wide mentors, and empowered them to foster a broader refugee-friendly school atmosphere rather than only targeting intra-classroom interactions between trained teachers and their refugee students. The fact that refugee children are spread out across various classrooms and that the impact of the training is widespread among refugee students in treated schools lends support to this interpretation.

We also investigate whether the training program had an impact on the academic performance of refugee students who on average tend to perform worse than their Turkish peers. For this purpose, we rely on the end-of-semester grades that students received in the semester following the program in two core subjects: Turkish language and Math. These grades are based on various forms of assessment, including written exams, oral exams, quizzes, and homework. Grades in these two subjects are particularly important as performance in Turkish language measures social and educational integration of refugee students, while performance in Math proxies their cognitive/analytical capacity. Moreover, having adequate Turkish language skills is also a prerequisite to understand the material in other subjects. In addition to measuring academic performance, school grades have also been shown to be good predictors of a variety of life outcomes (Borghans et al., 2016). We find that refugee students just above the first cutoff improved performance relative to those just below the first cutoff, both in absolute terms and relative to the Turkish children in their school and year. We also document that there is a high correlation between gains in school presence and gains in grades (i.e., those who decreased their absenteeism the most are the ones who also improved the most academically). The fact that we find a positive association between improvement in attendance and grades suggests that our results capture some underlying improvement in learning as a consequence of decreased absenteeism. This implies that the benefits of the mentoring role extends beyond the absenteeism outcomes and also includes improvement in refugee students' academic achievement.

Finally, we provide some back-of-the-envelope calculations illustrating the cost effectiveness of the training program, and discuss the implications of our results for the set of policies aiming to integrate refugee children into host-country education systems. In particular, we argue that closing the gaps in educational outcomes between native and refugee students through awareness-raising programs is quite cost effective considering the large budgets allocated to other types of integration programs, such as conditional cash transfer schemes.

Teachers are among the most powerful actors to effectively tackle integration challenges in educational settings. We document substantial returns to an intervention aiming to increase teachers' awareness toward the needs of vulnerable minority groups. Although our analysis focuses on refugee pupils, the takeaway lessons from our paper extend beyond refugee settings and can be applied to other ethnically mixed educational environments with disadvantaged minorities.

The rest of the paper is structured as follows. Section 2 reviews the related literature. Section 3 explains the institutional background and the design of the training program. Section 4 describes the data used in our analysis and the empirical strategy. Section 5 presents results on absenteeism and Section 6 on academic performance. Section 7 discusses mechanisms and Section 8 offers an evaluation of policy effectiveness. Section 9 concludes.

2 Related literature

Our paper is related to three main strands of the literature. The first strand focuses on the effects of refugee inflows on host-country education systems and, particularly, on the educational outcomes of native and refugee students. The focus of the great majority of the papers in this strand is on the effects of refugees on native students' outcomes—see, e.g., Figlio and Ozek (2019), Assaad et al. (2019), Green and Iversen (2020), and Tumen

(2019, 2021).[3] Surprisingly, there are very few papers directly studying the educational performance of refugee children in host countries and assessing the role of specific policies aimed at facilitating their educational integration.[4] For example, Sieverding et al. (2018) report that, despite substantial investments and policy efforts towards the education of Syrian refugee children in Jordan, a sustained increase in their enrollment rates has not been achieved. A few recent studies examine the social integration of refugee children in educational settings. Alan et al. (2021) study the impact of an educational program that aims to build social cohesion amongst refugee and native children, while Boucher et al. (2021) examine the impact of exogenous mixing of pre-school children on refugees' language acquisition and interethnic friendship formation. Our paper is unique in the sense that we leverage detailed administrative micro data on refugee students' outcomes within an original quasi-experimental RD design, which allows us to evaluate the impact of an integration program—in particular, an awareness-raising training program for teachers—on refugee children's school outcomes.

The second related strand of the literature highlights the role played by teachers as input in the production of education and in shaping students' educational outcomes (Hanushek and Rivkin, 2006; Jackson et al., 2014). There is a large literature on the effectiveness of teachers in improving student's test scores, but less so on the effects of teachers on students' non-cognitive skills, and particularly on school attendance and absences (Gershenson, 2016; Jackson, 2018; Liu and Loeb, 2019). For instance, Jackson (2018) proxies students' non-cognitive skills using behaviors that include absences, suspensions, course grades, and repetition in the ninth grade/year.[5] He finds that teachers have more prominent effects on high school completion rates and related long-run outcomes of students than direct effects on test scores. Another body of research indicates that minority students benefit academically when assigned to teachers of their own race/ethnicity (Dee, 2004, 2005; Egalite et al., 2015), while a few recent studies document the role of teachers' bias against students from minority groups in explaining the gaps between immigrants/minorities and natives (Hanna and Linden, 2012; Botelho et al., 2015; Alesina et al., 2018; Alan et al., 2021). Our study contributes to this literature in two related but distinct ways: first, by showing strong evidence of the instrumental role that

[3] There is also a more sizeable literature focusing on the impact of immigrant children on natives' educational performance—see, for example, Gould et al. (2009), Ohinata and van Ours (2013) Hunt (2017), Ballatore et al. (2018), Frattini and Meschi (2019), Bossavie (2020), and Figlio et al. (2021).

[4] A number of studies focus on the related issue of the educational outcomes of immigrant children, though this group is distinct from refugees whose families have been forced to flee their home countries. The evidence indicates that immigrant children under-perform compared to native children, and that the gap in educational outcomes between immigrants and native children may persist for both first and second generations—see, for example, Smith (2006), Schnepf (2007), Dustmann and Glitz (2011), and Bratsberg et al. (2012). A few studies investigate the role of policies to integrate immigrant children. Felfe et al. (2020) find that the introduction of birthright citizenship enhances the educational integration of immigrant children in Germany. Carlana et al. (2022) estimate the impact of a program that provided tutoring and career counseling to immigrant children in Italy.

[5] Throughout the paper, the term "grade" is used to describe two different concepts: (1) grade as the level of education (e.g., the ninth grade) and (2) grade as a measure of academic performance (e.g., the marks). To differentiate these two, we use the term "grade/year" to define the former, while "grade" is used to capture the latter.

teachers can play for the educational inclusion and school attachment of diverse student populations and, second, by providing evidence that increased school attachment also leads to substantial gains in academic performance.

Finally, our paper connects to the literature investigating the effects of teacher training interventions on student performance. The results presented in this literature are rather mixed. Angrist and Lavy (2001) examine the impact of in-service teacher training on achievement in Jerusalem elementary schools using a matched-comparison design, and find that the program improves test scores. Similarly, Bressoux et al. (2009) find that training teachers substantially improves students' test scores in Math, but not for low-achieving students in France, while Cilliers et al. (2020) find positive effects of teacher training in reading proficiency of primary school students in South Africa. On the other hand, Jacob and Lefgren (2004) use school reform efforts in Chicago to examine the impact of teacher training on Math and Reading performance of primary students using a quasi-experimental research design. They find that teacher training has no statistically or academically significant effect. Harris and Sass (2011) also find that professional development training generally does not improve the productivity of teachers using administrative data from Florida. Finally, Loyalka et al. (2019) examine a large-scale randomized national professional development program for teachers in China, but find that such intervention failed to improve teacher and student outcomes after one year due to the overly theoretical nature and lack of usefulness of the training. We contribute to this literature by providing evidence that teacher training programs can address the low school attachment and academic achievement levels of refugee children, which are major obstacles to their educational development.

3 Institutional background

3.1 Education of Syrian refugees in Turkey

The Syrian civil conflict has driven around 6.5 million of Syrians to flee their homes—5.6 million of whom are hosted in countries near Syria.[6] Turkey has received more than 3.6 million Syrian refugees (as of March 2021) the majority of whom live outside camps. The Syrian population in Turkey is younger, on average, than the native population. Moreover, total fertility rate of Syrian women is much higher than that of Turkish women. These two facts suggest that younger people are over-represented among Syrian refugees and social integration efforts should therefore focus on school-level policies. Based on recent educational statistics, the number of refugee children of school age (5-17) has reached around 1.1 million as of the 2019-20 academic year and around 700,000 of them are enrolled—i.e., the enrollment rate is approximately 63-64% (Tumen, 2018).

During the early stages of the refugee crisis, e.g., between 2011-2016, the main policy priorities were shaped around humanitarian assistance and provision of basic services.

[6] For detailed statistics, see https://www.unhcr.org/syria-emergency.html.

Education of refugee children were handled as a service designed and implemented outside of the Turkish public education system.[7] Since 2016, EU funded school integration programs have been implemented and full integration of refugees into the Turkish public education system has become an explicit policy priority. The EU Facility for Refugees in Turkey (FRIT—a 6-billion EUR fund) is designed to ensure that the needs of refugees and host communities in Turkey are addressed in a comprehensive and coordinated manner. The Facility focuses on humanitarian assistance, education, migration management, health, municipal infrastructure, and socio-economic support.

Various school integration programs have been implemented in Turkey through the PIKTES (Promoting Integration of Syrian Kids into the Turkish Education System) project, which is administered by the Ministry of National Education (MoNE) and financed by FRIT funds. The main ones are: back-up training, catch-up training, Turkish language training, and teacher training programs. The back-up and catch-up training programs aim at providing academic support to enrolled and out-of-school refugee students, respectively. The language training program aims at improving Turkish language skills of refugee students. This paper focuses on the teacher training program, which is described in detail in the next subsection.

3.2 The Teacher Training Program as an awareness-raising activity

Syrian families have experienced a violent and devastating civil war, which displaced them from their homes and generated large refugee waves. Most refugee children have been exposed to violence in various forms during the transition. Even in the absence of a direct exposure to violence, being forced out of their home country and having to live in an unfamiliar culture is itself a traumatic experience. In addition, Syrian children of school age had to stop going to school as they move from Syria to a host country. Hence, refugee children constitute a sensitive group and their needs have to be addressed with specific care and attention. They need support not only academically or on education-related issues, but they additionally need help along various other dimensions such as psychological counseling, developmental mentoring, relationships with their Turkish peers and teachers, adaptation into the new culture, etc.

Teachers are in direct contact with refugee students and increasing their capacity to address the needs of those students is of primary importance from a social integration perspective. As we discuss in Section 4.2, refugee children have low attachment to school—measured in terms of their days of absence from school—which is a major threat for their educational and socio-cultural integration. A teacher training program was designed and implemented to increase Turkish teachers' awareness of Syrian children's needs/vulnerabilities and increase their attachment to school, i.e., reduce their absenteeism. The main goals of this training program are: (i) to increase the teachers' awareness on issues related to educational integration of Syrian children, (ii) to increase

[7] See Boucher et al. (2021) for a more detailed chronology of those education services.

the capacity of teachers to address the needs of Syrian students, and (iii) to encourage teachers to actively contribute to refugee integration policies. In line with these objectives, teachers from schools where Syrian students are educated constitute the main target group.

The teacher training program was designed based on a "cascade training" philosophy, which is the standard staff training approach used by MoNE. The cascade training model is a way of cost-effectively training staff in large organizations. Typically, master trainers are trained on a specific topic with the expectation that they will transfer their knowledge to other staff in the same organization. This suggests that the teacher training program aimed at affecting the broader education environment in schools, through awareness spillovers among teachers.

The teacher training program was implemented during the semester break of the 2017-18 academic year, in January/February of 2018, over a 5-day period with a 30-hour program. The content of the program focused on three main areas: (i) language and communication, which introduces techniques for teaching Turkish as a second language, (ii) socio-economic integration and counseling, which includes topics such as, education in multicultural environment, inter-cultural sensitivities, educational guidance services, student recognition and orientation, students with special needs, and parental outreach, and (iii) legislation and context, which includes topics in temporary protection legislation, international law and children's rights, and recent research/reports on Syrian refugees.

Eligibility for the program was determined at school level, based on the number of Syrian students that are assigned to each school in the corresponding catchment area.[8] In particular, teachers from schools with more than 15 Syrian students (preferably those who actually instructed Syrian children) were eligible for the program. For eligible schools, the total number of teachers that would be eligible to participate in the training ranged from 1 to 6 according to the total number of Syrian students enrolled. Specifically, the thresholds were defined as follows: 1 teacher from schools with 15-19 enrolled Syrian students, 2 teachers from schools with 20-39 enrolled Syrian students, 3 teachers from schools with 40-59 enrolled Syrian students, 4 teachers from schools with 60-79 enrolled Syrian students, 5 teachers from schools with 80-99 enrolled Syrian students, and 6 teachers from schools with 100 enrolled Syrian students and above. Therefore, the teacher training thresholds were defined as 15, 20, 40, 60, 80, and 100. A total number of 8900 teachers from the 26 provinces with the highest refugee concentration were eligible for the program.[9] 8661 teachers completed the 30-hour program, which suggests a 97.31% completion rate. All participating teachers were Turkish nationals.

[8]Therefore, when we refer to a school with, say, 15 refugees students, we mean a school located in a catchment area with 15 potential refugee students rather than actual enrollment. There is a centralized online database called "e-school," which is the core system recording all administrative details (such as absenteeism, grades, special needs, and a limited set of other personal characteristics) for every student registered to a school in Turkey.

[9]These are the provinces in which the FRIT-funded PIKTES projects were carried out and the provinces are determined based on a protocol signed by the European Commission and MoNE.

The selection of teacher(s) to participate in the training program was carried out as follows. Using school-level registration records, provincial MoNE administrators determined the eligible schools and the number of teachers to be assigned from each eligible school according to the thresholds described above. Then, in each eligible school, school directors sorted the teachers based on the number of Syrian students in their classrooms. In a school with n slots for attendance to the teacher training program, where $n = \{1, \ldots, 6\}$, the top n teachers from the sorted list were invited to the program. When there were more than 1 suitable candidates for the n^{th} slot, then the teacher with the lowest number of past attendances to an in-service MoNE training program was chosen. If there were still multiple eligible candidates after this step, then the last attendee was randomly assigned. Program participation was on a voluntary basis. In cases where the assigned teacher was unable to attend, the administrators selected the next eligible teacher following the same steps.

Although there are other concurrent interventions aiming to improve school integration of refugee children, there is no other program operating around the cutoffs specified for this program. This suggests that our discontinuity-based approach is not confounded by other programs and allows us to identify the causal impact of the teacher training program.

Overall, the "teacher training policy" is a school-level intervention aiming to assign the most suitable teachers to the training program, where suitability is defined by the intensity of recent involvement in teaching Syrian students. From the viewpoint of Syrian students, attending a school with a trained teacher (or an additional trained teacher) is as-if randomized around the thresholds described above, assuming that the assignment of teachers to classes with Syrian children is similar for schools around the thresholds. Our main focus is to estimate the causal impact of the training program on student outcomes exploiting the policy discontinuities around the thresholds. The next section describes our data, variables, identification strategy, and econometric setup in much more detail.

4 Empirical strategy and data

4.1 Empirical strategy

A local randomization RD approach. Our empirical strategy exploits the discontinuity in schools' eligibility to participate in the training program to identify the causal effect of the program on student outcomes in a regression discontinuity framework. In particular, a school becomes eligible to have one teacher trained once the number of enrolled refugee students crosses 15, two teachers when the number of refugee students crosses 20 etc., according to the schedule presented above.

In this setting, there are three features that our identification, estimation, and inference need to take into account: the fact that the running variable (number of foreign students within a school) is discrete and has a few mass points; the fact that there are multiple cumulative cutoffs governing treatment (number of teachers within a school that are eligible to receive the training); and the nature and interpretation of the treatment effect. To account for the first feature, we adopt the local randomization approach to RD analysis as our main approach, instead of the standard continuity-based approach—as using the continuity-based methods when the running variable is discrete can lead to significant biases (Cattaneo et al., 2018). Unlike the standard continuity-based approach to RD analysis, which relies on the assumption of continuity and smoothness of the conditional expectation of the potential outcomes in the neighborhood of the cutoff (Imbens and Lemieux, 2008), the local randomization approach rests on a stronger exclusion restriction identification assumption. That is, the potential outcomes are unrelated to the running variable inside the window in which treatment is as-if randomly assigned, which implies that the conditional expectation functions are flat and the average treatment effect can be estimated as the difference between the average outcomes of observations just above and below the cutoff (Cattaneo et al., 2018). In other words, the data can be analyzed as-if treatment is randomly assigned near the cutoff. One main advantage of the local randomization approach is that it permits the use of finite sample inference methods (Fisherian inference framework) that are valid even if the number of observations around the cutoff is limited (Cattaneo et al., 2015). To account for the second feature (multiple cutoffs), we present cutoff-specific RD treatment effects. The third feature of our empirical setting is that we do not know which teachers participated in the training program and their characteristics (and, therefore, cannot identify which students were taught by a trained teacher). Consequently, our results should be interpreted as intention-to-treat (ITT) estimates.

Formally, denoting by $\{Y_{is}(1), Y_{is}(0)\}$ the potential outcomes of student i in school s, and by r_s the number of refugee children in school s, an application of the local randomization approach to RD requires that the assignment of schools inside a window around the cutoff is random and that the potential outcomes $\{Y_{is}(1), Y_{is}(0)\}$ around the cutoff are unrelated to r_s. Under these assumptions, the local randomization sharp RD effect is given by:

$$E[Y_i(1)|r_s = c] - E[Y_i(0)|r_s = c_{-}], \tag{1}$$

where c denotes the cutoff and c_{-} denotes the closest mass point below the cutoff.[10] A local-randomization approach can be used to base inference on comparison of students in schools with r_s equal to the cutoff to those with r_s just below the cutoff. In particular, one can employ the finite-sample Fisherian framework proposed by Cattaneo et al. (2015) to test the sharp null hypothesis that the treatment has no effect for any unit. As a robustness check, we also report a test of the Neyman null hypothesis that the average treatment effect is zero.

[10]Note that we consider this setting to be amenable to a sharp RD design as compliance of schools and teachers with the training program in this setting is to our understanding near perfect. However, as we have no way of ascertaining compliance in the data and use eligibility as an instrument in a fuzzy RD framework, our estimates can equally be interpreted as Intention-to-Treat (ITT) estimates.

Window choice. To assess sensitivity of our RD estimates to the choice of window around the cutoff, we present results for the smallest possible window around the cutoff (the cutoff mass point and the one just below, w=2), but also for symmetric windows of size 4 and 6 mass points around a cutoff. As the smallest windows (w=2 and w=4) might contain a small number of schools-in particular, at the higher cutoffs-we consider the window of w=6 as the most reliable one. Following the suggested best practice in Cattaneo et al. (2018), we also present covariate balance tests for each window that we consider separately.

Falsification tests. As is standard in the literature, we offer several tests of the integrity of the RD design: (i) we check whether the number of observations just below the cutoff is considerably different from the number of observations at the cutoff; (ii) we check whether treated units at the cutoff are similar to control units in terms of covariates; (iii) we check whether a treatment effect is detected at the same cutoff but one semester earlier before the training program was implemented; (iv) we carry out a placebo outcome test, in which we estimate the impact of the same training program on the outcomes of Turkish students.

4.2 Data

We use micro-level student administrative records from two Turkish provinces: Gaziantep and Sanliurfa. These are the largest provinces bordering Syria in the Southeast region of the country hosting a total of 872,000 Syrian refugees, which amounts to about 20% of the population.[11] The two provinces are among the main implementation hubs for programs aiming to integrate refugee students into the Turkish education system.

Our analysis draws on administrative data from all public schools (primary and secondary, years 1-12) in those two provinces.[12] That is, a total of 2081 schools hosting 64,582 refugee and 743,301 Turkish students in the 2017-18 academic year. The analysis focuses on three semesters: one semester prior to the training program (Fall 2017) and two semesters after the training took place (Spring 2018 and Fall 2018) allowing us to examine short-term and longer-term impacts of the program.

Absenteeism. The data contain information on the days of absence from school at semester level, which is our main outcome of interest. School absenteeism is an important student outcome that proxies various aspects of human capital development. First, the days of absence variable is directly used to measure learning losses due to reduced school presence. For example, absenteeism is used by many researchers and policy institutions to quantify learning losses that emerged during the Covid-19 pandemic.[13] Second, it is a

[11] The province-level refugee numbers are provided by the Ministry of Interior, Directorate General of Migration Management.

[12] Primary schools in Turkey cover years 1-4, middle school 5-8, and high school 9-12.

[13] See, for example, IMF (2021).

proxy for non-cognitive skills. Several papers in the literature document that absenteeism is negatively correlated with the big 5 personality traits that constitute the core of non-cognitive skills (Lounsbury et al., 2004). Finally, school absenteeism is strongly related to various antisocial and/or risky behaviors that emerge later in life.[14]

Student achievement. To measure student achievement, we use the student grades that are contained in the administrative records. We focus on grades in two subjects, Turkish language and Math. Note that grades are reported on a three-category scale (i.e., from 1 to 3) during the first three years of primary education, and a 0-100 grading scale is used from year 4 to year 12. Moreover, the data contain information about the country of origin of the student allowing us to differentiate native students from refugee students. In our analysis of grades, we focus on grades from year 4 onwards. The observed grades in the data set correspond to end-of-semester average of scores obtained from a variety of assessment methods—such as written exams, oral exams, quizzes, projects, homework, etc. To obtain a comparison basis for econometric analysis, we standardize grades by grade/year around zero with unit variance. We also construct a measure of performance of refugee students relative to that of Turkish children in the same school and year. In doing so, we perform the standardization after subtracting the Turkish students' average grades from each Syrian student's grade in the same school and year—see Section 6 for more details.

Summary statistics. Table 1 provides summary statistics for two samples: the first sample encompasses all students in the range of schools that are eligible (or close to the eligibility cutoff of 15), that is, they have between 12 and 102 refugee students in enrollment. This sample includes 6,536 refugee students and 269,940 Turkish students across 415 schools. The second sample, which we refer to as the discontinuity sample, is a smaller sample of students that attend schools in the neighborhood of the cutoffs that we consider in the RD analysis: they fall within the maximum window around the cutoffs (3 mass points on each side). This gives us a total of 2,630 refugee students and 113,284 Turkish students across 190 schools.

Refugee students in the discontinuity sample are on average 9.8 years old, 68% are in primary schools (on average in grade/year 3), and 61% are located in Gaziantep. Turkish students are slightly older (average age 11), with the majority attending higher levels of education (on average in grade/year 5). The sample is balanced in terms of gender. Overall, the two samples are very comparable along these characteristics.

With respect to our main outcome of interest, we see that refugee children recorded on average 8.9 days of absence in the first semester of the 2017-18 academic year, before the training took place. This amounts to them being absent about 10% of the school days, thereby incurring a considerable loss in learning. Instead, Turkish children were

[14]See Gubbels et al. (2019) for a comprehensive meta-analytic review of the related literature.

absent on average 3.7 days in the same semester, suggesting a substantial gap in school attachment between the two groups, which the training program aims to close. There are also substantial academic achievement gaps between native and refugee students for both Turkish language and Math subjects.

Table 1: Summary statistics—student characteristics

	Analysis range (12-102)		Discontinuity sample ($w:6$)	
Variable name	Syrian	Turkish	Syrian	Turkish
Days of absence (pre-treatment)	8.90	3.68	8.87	3.72
	(6.25)	(3.67)	(8.86)	(3.69)
Turkish language grade (pre-treatment)	42.67	71.68	43.41	73.16
	(18.27)	(23.22)	(20.13)	(24.54)
Math grade (pre-treatment)	38.99	69.96	40.08	71.44
	(21.02)	(25.19)	(23.88)	(26.63)
Age	9.60	11.11	9.77	10.98
	(2.42)	(2.61)	(2.41)	(2.50)
Grade/year	2.99	5.15	3.11	4.99
	(2.04)	(2.44)	(2.11)	(2.29)
Number of siblings	3.11	2.53	3.11	2.51
	(2.40)	(1.95)	(2.35)	(2.04)
$\mathbb{P}$(Male = 1)	0.48	0.51	0.50	0.51
$\mathbb{P}$(Primary school = 1)	0.70	0.39	0.68	0.39
$\mathbb{P}$(Gaziantep = 1)	0.65	0.59	0.61	0.54
Number of observations	6,563	269,940	2,630	113,284

Notes: The first two columns provide student-level summary statistics for schools with number of foreign students between 12 and 102. This is our broadly defined analysis range. The last two columns include schools that are up to 3 mass points around the cutoffs, which we call the discontinuity sample. Standard deviations are reported in brackets.

Table 2 offers some summary statistics of the schools in the two samples. In the discontinuity sample, 51.6% of the schools are located in Gaziantep, and they are almost equally split between primary and secondary schools. The average school size is about 615 students, with on average about 30 of those being refugees. The average share of refugee students in the sample is 7.8%, with primary schools having a slightly higher share (9.3%) than secondary schools (5.9%). Schools are larger on average in the larger sample, however, the average share of refugees is very similar, about 8%.

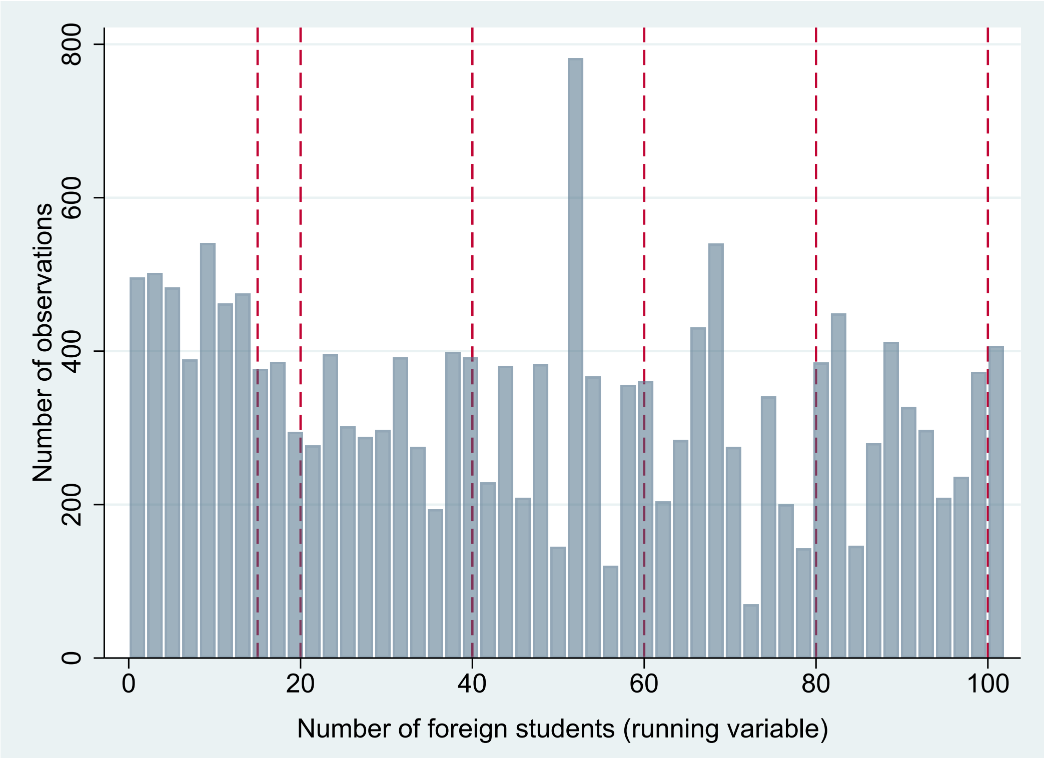
Figure 1 plots a histogram of the running variable (number of refugee students in the school). The figure does not indicate any systematic sorting above the cutoffs, an issue that we will test formally in the next section. In any case, manipulation of the running variable does not seem to be a plausible feature in our setting.

Table 2: Summary statistics—school characteristics

	Analysis range (12-102)	Discontinuity sample ($w:6$)
Total number of schools	415	190
Gaziantep	237	98
Sanliurfa	188	92
Primary	207	96
Secondary	208	94
Average school size	747.39	615.31
Gaziantep	772.24	577.43
Sanliurfa	710.18	658.69
Primary	630.90	544.71
Secondary	863.32	690.86
Average number of refugee students	39.23	29.81
Gaziantep	42.06	33.00
Sanliurfa	35.84	26.15
Primary	39.76	30.46
Secondary	38.71	29.11
Average refugee share	0.082	0.077
Gaziantep	0.085	0.088
Sanliurfa	0.077	0.064
Primary	0.099	0.093
Secondary	0.064	0.059

Notes: The first column presents school-level summary statistics for schools with number of foreign students between 12 and 102. The second column includes schools that are up to 3 mass points around the cutoffs.

Figure 1: Distribution of the running variable



Notes: This figure plots the distribution of schools in our analysis range by the number of refugee students they have, which is our running variable. Vertical dashed lines indicate the cutoffs.

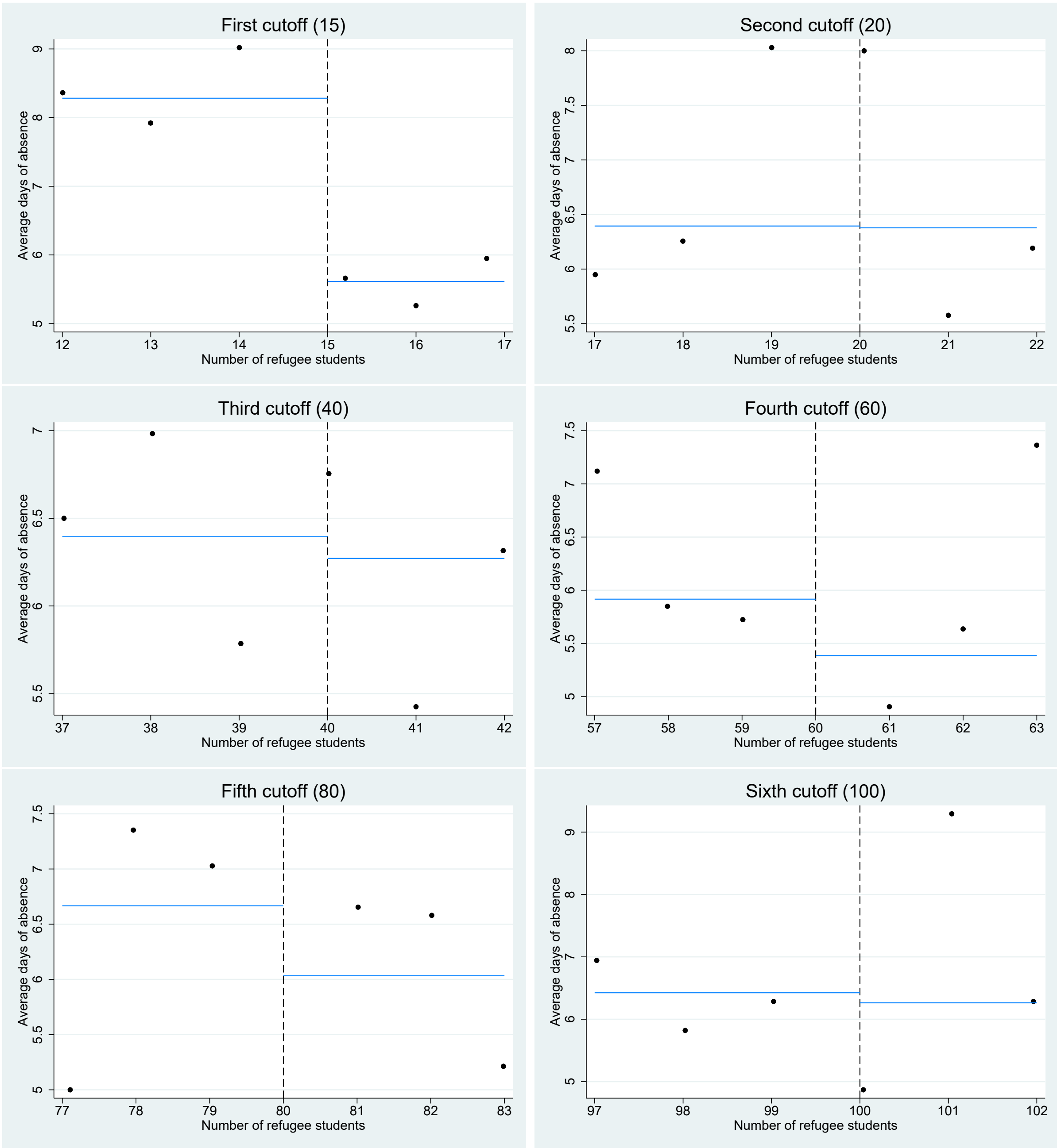
5 Effects on absenteeism

In this section, we first present our baseline RD results: the impact of a school becoming eligible to have an (additional) teacher receive the training on the school absenteeism of refugee students in the current semester. This is followed by tests of the validity of the regression discontinuity approach. We then present evidence on the longer-run impact of the training on the absenteeism of refugee students in the next academic year, assess the presence of heterogeneous treatment effects by gender and school level, and consider the impact of the training on other outcomes, such as, chronic absenteeism.

5.1 Baseline RD results

Figure 2 provides a first impression of the impact of training on refugee students' absenteeism. The figure plots the number of refugee students on the horizontal axis and the average days of absence of refugee students within the semester immediately after the program on the vertical axis. We do so separately for each of the six cutoffs that are associated with an additional teacher within a school that crosses a threshold receiving the training. Each dot in the figures represents the average days of absence corresponding to one of the distinct values that the running variable (number of refugee students in the school) takes. The horizontal lines represent the average days of absence of the three mass points on each side of the cutoff, so the distance between the lines gives the RD treatment effect, which visually appears to be most sizeable around the first cutoff (15).

Figure 2: Baseline visual evidence—local randomization



Notes: This figure provides visual evidence for our baseline local randomization RD analysis for 3 mass points around each of our cutoffs.

In Table 3, we report the number of observations in the two (one on each side), four (two on each side), and six (three on each side) mass points around the cutoff. The last column also reports the p-value of a binomial test, which assesses the density of the running variable around the cutoff, that is, whether the number of observations in the mass point(s) just above the cutoff are similar to those just below it (Cattaneo et al., 2017). For example, considering the first cutoff at 15, there are 52 observations just below the cutoff and 68 observations at the cutoff. The p-value of the Binomial test is

0.171 suggesting that there is no evidence of sorting of the running variable around this cutoff.

Table 3: Baseline RD estimates—local randomization

		Estimation & inference			Levels		Binomial test		
	Cutoff	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff	# of obs. below c	# of obs. above c	Bin. test p-value
A.	15 (w:2)	-3.358**	0.009	0.012	9.019	5.662	52	68	0.171
	15 (w:4)	-2.799***	0.000	0.000	8.220	5.421	191	171	0.318
	15 (w:6)	-2.669***	0.000	0.000	8.283	5.613	343	269	0.003
B.	20 (w:2)	-0.029	0.986	0.986	8.029	8.000	34	40	0.561
	20 (w:4)	-0.203	0.823	0.810	6.758	6.555	120	99	0.176
	20 (w:6)	-0.016	0.981	0.979	6.395	6.378	218	193	0.236
C.	40 (w:2)	0.969	0.342	0.277	5.786	6.755	70	147	0.000
	40 (w:4)	-0.078	0.922	0.910	6.339	6.261	130	234	0.000
	40 (w:6)	-0.124	0.818	0.835	6.395	6.272	200	291	0.000
D.	60 (w:2)	-0.818	0.332	0.320	5.724	4.905	105	74	0.025
	60 (w:4)	-0.580	0.408	0.333	5.797	5.217	251	129	0.000
	60 (w:6)	-0.531	0.412	0.363	5.917	5.386	276	140	0.000
E.	80 (w:2)	-0.373	0.812	0.798	7.028	6.655	36	55	0.059
	80 (w:4)	-0.639	0.503	0.512	7.258	6.619	124	105	0.234
	80 (w:6)	-0.633	0.395	0.391	6.667	6.033	168	180	0.555
F.	100 (w:2)	-1.417	0.294	0.257	6.286	4.868	91	38	0.000
	100 (w:4)	0.194	0.836	0.868	6.042	6.236	191	55	0.000
	100 (w:6)	-0.162	0.847	0.820	6.425	6.263	332	118	0.000

Notes: Panels A-F report the estimates around each of the cutoffs. All results are reported for three different window sizes: 2, 4, and 6 mass points around the cutoffs. Fisherian and Neyman p-values allow for finite-sample and large-sample statistical inference, respectively. The average values of the outcome variable before and after each cutoff are reported in columns 4 and 5. The binomial test checks whether the distribution of observations is balanced around the cutoffs or not. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Figure 2 suggests that a significant reduction in the average days of absence occurs around the first cutoff of 15 refugee students, whereas for the other cutoffs no such drop is visually obvious. Turning to the point estimates, Table 3 shows our baseline estimates of the treatment effect of training on days of absence. In Panel A, we present the RD estimate around the first cutoff, capturing the extensive margin of the training program. Then in Panels B to F, we present cutoff-specific RD estimates for each of the higher thresholds separately. Around the first cutoff (15), we find a negative and statistically significant RD effect of teacher training on days of absence that ranges from -3.4 for the window of 2, and decreases (in absolute value) to -2.8 days for the window of 4, and -2.7 days for the window of 6 mass points. Note also that in the case of this cutoff, the binomial test indicates no evidence of sorting of the running variable for w=2 and w=4, while the test rejects randomization for w=3. However, at w=3 the number of observations below the cutoff is smaller than above against what we would expect if sorting of school was taking place. In the rest of the cutoffs reported in Panels B-F, we do not find a statistically significant RD treatment effect. This is perhaps not surprising as the intensity of treatment (number of refugee children per trained teacher) is rather similar in those schools relative to schools that have only one trained teacher. We discuss further what this result implies for how the training program achieves its impact in section 7.1.

The treatment effect we estimate around the 15 cutoff is sizeable. Given that average days of absence during the semester following the training for the control group of this comparison (3 mass points to the left of the cutoff) is 8.3 days, our estimate of a reduction of 2.7 days implies a reduction of 32.5%. Note also that given that in the same window Turkish children have on average 3.8 days of absence, the treatment effect we estimate implies a closing of the gap between the absenteeism of Turkish and refugee children by 61% (initial gap is 4.4 days).

To summarize, the treatment effect seems to be concentrated around the 15 cutoff in which comparison is between schools with one trained teacher and control schools with no trained teachers—the extensive margin of the treatment. Therefore, to conserve space from here onwards, we present results only for the 15 cutoff.

5.2 Validation checks

In this section, we report various validation checks of the RD empirical framework:

Predetermined covariates. As a first validity check, we investigate whether the predetermined characteristics of students and schools are balanced around the cutoff. To this end, Table 4 presents RD effects on predetermined covariates for the cutoff at 15. Of all the tests reported in the table, there is only one instance in which there is a statistically significant difference between treatment and control at the 15 cutoff: male for w=2. After investigating more carefully the distribution of schools around w=2, we notice that, at r=14 (the LHS of the smallest window), there is a religious secondary school consisting of female students only (i.e., the type of the school is "Kiz Imam Hatip Ortaokulu"). Dropping that specific school from our sample removes the imbalance for gender around w=2. These tests are reported in italics (with †) in Table 4. Note that our baseline estimates reported in Table 3 are also not sensitive to removing that specific school from our sample.[15] Overall, we thus conclude that the covariates are balanced around the cutoff.

Pretreatment effect. We next investigate whether an RD treatment effect can be detected in the pretreatment period, that is, in the first semester of the academic year 2017-18 before the training had taken place. Reassuringly, the results illustrated in Figure 3 and reported in Table 5 show convincingly that no treatment effect is observed

[15] Specifically, the estimates for the first cutoff are 3.296, 2.781, and 2.659 for w=2, w=4, and w=6, respectively—with the same levels of statistical significance.

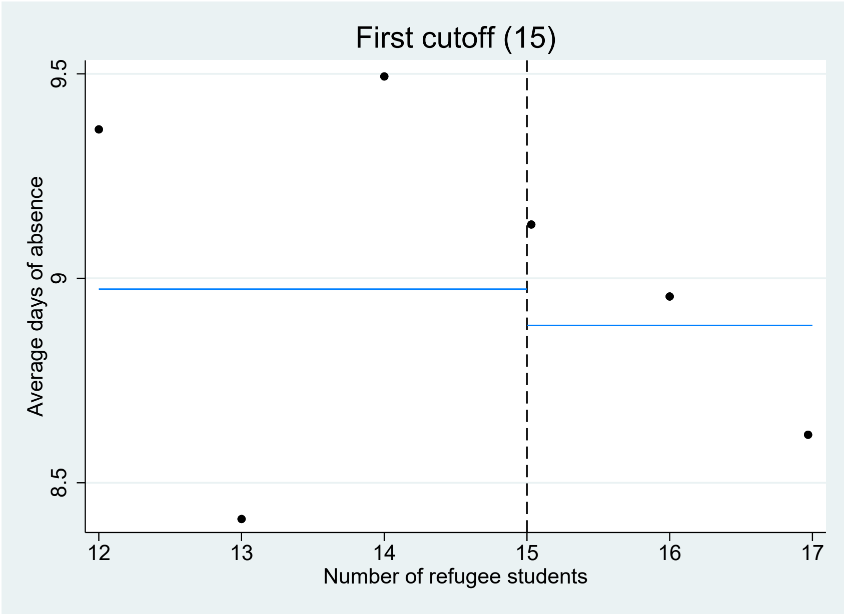
Table 4: RD effects on predetermined covariates

	Estimation & inference			**Levels**	
Cutoff: 15	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
Male (w:2)	0.208**	0.028	0.022	0.365	0.574
Male (w:4)	-0.068	0.205	0.190	0.466	0.398
Male (w:6)	-0.059	0.161	0.149	0.490	0.431
†*Male* (w:2)	0.113	0.143	0.123	0.461	0.574
†*Male* (w:4)	-0.101	0.171	0.159	0.499	0.398
†*Male* (w:6)	-0.080	0.139	0.121	0.511	0.431
Age (w:2)	-1.115	0.151	0.112	11.865	10.750
Age (w:4)	-0.613	0.353	0.341	11.105	10.491
Age (w:6)	-0.182	0.335	0.353	10.475	10.294
Grade/year≥mean (w:2)	-0.082	0.153	0.166	0.816	0.734
Grade/year≥mean (w:4)	-0.026	0.411	0.402	0.754	0.728
Grade/year≥mean (w:6)	-0.012	0.587	0.566	0.711	0.699
School size (w:2)	242.686	0.120	0.105	291.100	533.786
School size (w:4)	122.972	0.286	0.302	449.148	572.120
School size (w:6)	148.857	0.107	0.118	450.380	599.237

Notes: Estimates denoted by † refer to a sample in which we exclude a small all-female religious secondary school that has 14 refugee students. "Grade/year≥mean" refers to a dummy variable taking 1 if the grade/year of the refugee student is greater than the mean grade/year in the corresponding window and 0 otherwise. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

before the intervention.

Figure 3: Visual evidence—pretreatment semester



Notes: This figure provides visual evidence on refugee students' days of absence in the pre-treatment period (Fall 2017) as a falsification test.

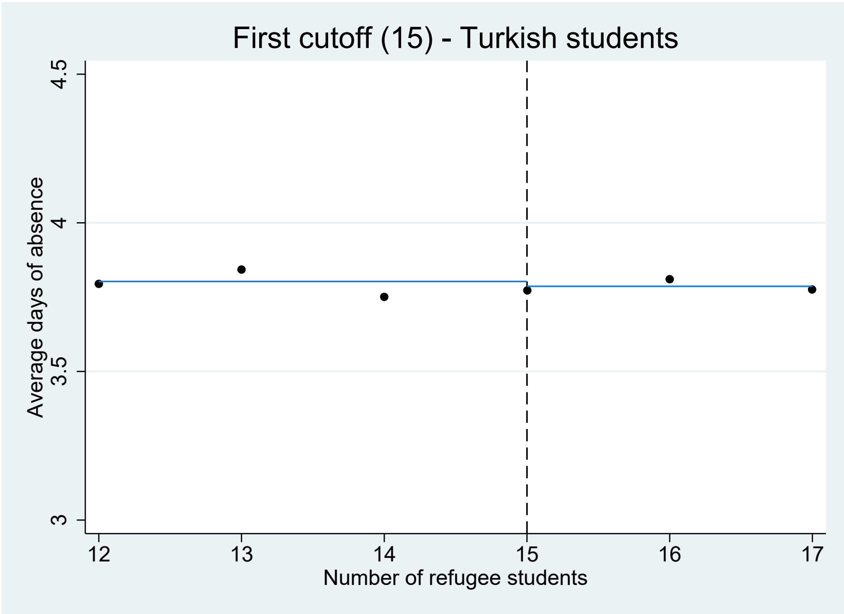
Table 5: RD effects for pretreatment semester

	Estimation & inference			**Levels**	
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
15 (w:2)	0.304	0.865	0.853	9.115	9.419
15 (w:4)	0.825	0.366	0.370	8.000	8.825
15 (w:6)	-0.013	0.989	0.986	8.898	8.885

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Placebo test: Effect on Turkish students. We next examine whether the training has an impact on the absenteeism of Turkish students. Recall that the training program was specifically designed to address the needs of refugee students, so we do not expect to find any effect on the outcomes of Turkish students, so this analysis constitutes a plausible placebo test. Figure 4 and estimates reported in Table 6 suggest that indeed there is no effect of the training on the absenteeism of Turkish students around the 15 cutoff. This test provides further reassurance as to the credibility of the main results on refugee students reported above.

Figure 4: Visual evidence—Turkish students



Notes: This figure provides visual evidence on Turkish students' days of absence.

Table 6: RD effects for Turkish students

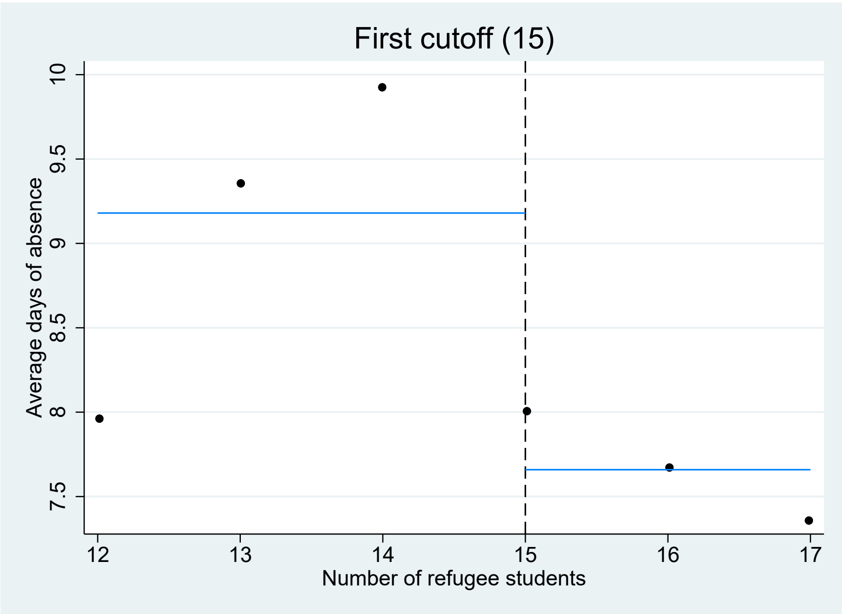
	Estimation & inference			**Levels**		**Binomial test**		
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff	# of obs. below c	# of obs. above c	Bin. test p-value
15 (w:2)	0.022	0.761	0.763	3.751	3.773	4,557	6,620	0.000
15 (w:4)	-0.018	0.693	0.695	3.811	3.792	12,983	13,890	0.000
15 (w:6)	-0.016	0.653	0.644	3.803	3.787	25,285	21,488	0.000

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

5.3 Longer-term effects

We next ask whether the effect of the teacher training program had persistent effects by examining refugee students' absenteeism in the following academic year. Figure 5 and Table 7 present this analysis. At the 15 cutoff, the estimate starts at -1.9 for w=2 and decreases (in absolute value) to -1.8 and -1.5 as the window widens. The effect becomes statistical significant as the sample becomes larger (p-values of 0.032 and 0.017, respectively). In terms of size of this effect, given the average days of absence for w=3 in the control group is 9.2 days, our estimate of -1.5 implies a 16.3% reduction. Compared to the 32.5% immediate reduction over the control, this indicates that the effect of training dissipates over time, which is perhaps not too surprising given that the training was one-off and also some of the teachers who have received the training might have moved to schools outside our analysis sample.

Figure 5: Visual evidence—longer term effects



Notes: This figure provides visual evidence on refugee students' days of absence in the longer-term (Fall 2018) to see the persistence of the training program's effects.

Table 7: Longer-term RD effects

	Estimation & inference			**Levels**	
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
15 (w:2)	-1.919	0.135	0.129	9.925	8.006
15 (w:4)	-1.762**	0.032	0.018	9.570	7.808
15 (w:6)	-1.521**	0.017	0.014	9.180	7.659

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

5.4 Heterogeneous treatment effects by gender and school level

We next examine whether there are any underlying heterogeneous treatment effects by gender and school level. Table 8 presents RD estimates separately for boys and girls in our sample. The results show no evidence of gender differences, as the levels on each side of the 15 cutoff are remarkably similar across the two genders.

Table 8: Estimates by gender

	Males				
	Estimation & inference			**Levels**	
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
15 (w:2)	-3.005	0.191	0.218	8.211	5.205
15 (w:4)	-2.010*	0.101	0.075	7.539	5.529
15 (w:6)	-2.814**	0.022	0.010	8.167	5.983
	Females				
	Estimation & inference			**Levels**	
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
15 (w:2)	-3.209**	0.038	0.028	9.485	6.276
15 (w:4)	-3.464***	0.000	0.001	8.814	5.350
15 (w:6)	-3.061***	0.000	0.000	8.394	5.333

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 9 presents RD estimates separately for primary schools, and upper-level schools (mid-schools and high-schools). These results indicate that the negative impact of the training on the absenteeism of refugee students at the 15 cutoff is present at both the primary and secondary level. So, the overall treatment effect estimated above does not seem to mask any significant heterogeneity along the school level dimension.

5.5 Chronic absenteeism and drop-out

Our main outcome variable throughout the paper is the days of absence from school per semester. However, absenteeism is also studied in the literature by using alternative definitions. In this subsection, we use some of these alternative definitions to study and understand the impact of the teacher training program on different aspects of school absenteeism.

We first focus on "chronic absenteeism," which is typically defined as missing at least 10% of the available school days in a semester. It is often used as an early warning signal for academic risk and school dropout (Bruner et al., 2011). The number of available school days in a typical semester is 90 days in Turkey. Panels A-B and C-D in Table 10 present results for two outcomes: missing at least 10 days and 20 days (i.e., chronic absenteeism in

Table 9: Estimates by school level

Primary education (years 1-4)					
	Estimation & inference			**Levels**	
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
15 (w:2)	-4.700**	0.028	0.029	9.700	5.000
15 (w:4)	-1.836*	0.056	0.055	7.036	5.200
15 (w:6)	-2.454***	0.001	0.001	8.021	5.567
Secondary education (years 5-12)					
	Estimation & inference			**Levels**	
Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
15 (w:2)	-2.732	0.118	0.106	8.857	6.125
15 (w:4)	-3.536***	0.000	0.001	9.130	5.594
15 (w:6)	-2.961***	0.002	0.001	8.620	5.659

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

varying degrees), respectively, in the semester following the training program. Specifically, in panel A-B (C-D), the dependent variable is a dummy variable taking 1 if the student misses at least 10 days (20 days) during the Spring 2018 semester and 0 otherwise. Our estimates around the first cutoff suggest a decline in chronic absenteeism in the interval of 7.6-16.4 percentage points depending on the specification. These results are similar in nature to our baseline estimates: the training program reduces chronic absenteeism and the reduction in chronic absenteeism is mainly driven by the decline at the first cutoff.

We also analyze the impact of the teacher training program on the probability of dropping out of school. In the Turkish education system, students automatically fail (and, therefore, have to repeat the corresponding grade/year) if they miss more than 20 school days per academic year without a valid excuse. The days of absence limit increases to 30 school days per academic year if the student presents a legitimate health report. We do not have access to information on medical reports. Based on these definitions, panels E-F and G-H report the estimates for which the dependent variable describes the probability of dropping out of school in two different ways: missing 30 days in the semester following the training program (Spring 2018) and missing 30 days in the academic year encompassing the training program (2017-18 academic year), respectively. The estimates around the first cutoff point to a 2.6-5.2 percentage points decline in the probability of dropping out. Overall, these results suggest that part of the decline in days of school absence following the teacher training program comes from the improvements in the tendency to be chronically absent from school and the tendency to dropout.

Table 10: RD effects for alternative outcomes

		Estimation & inference			Levels	
	Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Before cutoff	After cutoff
	Absenteeism ≥10 days in the post-treatment semester					
A.	15 (w:2)	-0.164*	0.072	0.053	0.385	0.221
	15 (w:4)	-0.131***	0.008	0.004	0.330	0.199
	15 (w:6)	-0.091**	0.017	0.011	0.306	0.216
	Absenteeism ≥20 days in the post-treatment semester					
B.	15 (w:2)	-0.101**	0.043	0.033	0.115	0.015
	15 (w:4)	-0.076***	0.003	0.002	0.100	0.023
	15 (w:6)	-0.080***	0.000	0.000	0.103	0.022
	Absenteeism ≥30 days in the post-treatment semester					
C.	15 (w:2)	-0.024	0.584	0.439	0.039	0.015
	15 (w:4)	-0.026*	0.129	0.067	0.031	0.006
	15 (w:6)	-0.028***	0.008	0.006	0.032	0.004
	Absenteeism ≥30 days in the academic year					
D.	15 (w:2)	-0.070	0.278	0.278	0.173	0.103
	15 (w:4)	-0.037	0.372	0.278	0.136	0.099
	15 (w:6)	-0.052**	0.062	0.049	0.149	0.097

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

5.6 Continuity-based RD estimates

For completeness, we also present estimates obtained by using the conventional RD continuity-based approach. In this approach we estimate an equation that takes the form:

$$y_{is} = \alpha + \beta 1\{r_s - r_s^* \geq 0\} + f(r_s) + \gamma X_i' + region_s + \epsilon_i, \tag{2}$$

where y_{is} denotes days of absence of student i in school s. r_s is the number of refugee children in school s, r_s^* is a cutoff number of refugee children (15, 20, 40, 60, 80, 100) above which a school has an additional teacher receiving training, and $f(r_s)$ is a flexible polynomial function of r_s. The vector X contains student characteristics and the term $region_s$ captures region/province fixed effects. The parameter of interest is β, which is the intention-to-treat causal effect of being in a school just above a cutoff. Standard errors are clustered at the school level.

Baseline results using this approach are presented in Figure A1 and Table A1 in the Appendix, while long-term effects in Figure A2 and Table A2, using two alternative methods of modeling $f(r_s)$ (linear, quadratic) and two alternative ways of choosing optimal bandwidth. In line with estimates presented above, these results confirm that there is a negative and statistically significant effect of the training program on the days of absence of Syrian students around the first cutoff (15). These estimates are generally larger in absolute size than the ones obtained using the local randomization approach, suggesting that our preferred approach provides more conservative estimates.

6 Effects on academic performance

The results presented in Section 5 suggest that the teacher training program improved meaningfully the school attachment of refugee children. If the program increased school attendance of refugee students, then it might have also improved their academic performance. Therefore, a natural follow-up question is whether the program also affected the grades of treated students.

Our administrative data set—which is quite rich and contains information spanning several academic years, two provinces, many schools, different grade/year levels, and various courses—includes end-of-semester grades for different subjects. These grades represent the average score obtained in exams and other in-class activities during the semester. School grades are not only relevant proxies of cognitive skills but have been shown to capture different personality characteristics than standardized test scores and to be good predictors of a variety of life outcomes (Borghans et al., 2016). We implement our RDD analysis using these grades to assess whether the program affected the academic achievement of refugee students.

We focus on grades in Turkish language and Math subjects, which are available for every student across all schools and grade/year levels. Turkish language grades proxy language and communication skills, while Math grades can be linked to cognitive/analytical abilities. To improve comparability, we standardize the subject-specific grades for Syrian students—i.e., center them around zero with a unit standard deviation—for each grade/year level, which means that the coefficient estimates are presented in terms of standard deviations. It should be noted that when a student entirely misses exams in a given subject, then the end-of-semester average grade is recorded as missing, which appears as "zero" at the end-of-academic-year grade card. For this reason, we treat missing grades as zeros throughout our analysis.

Parallel to our analysis in Section 5, we mainly focus on Syrian students around the first cutoff—i.e., schools just below and above the 15 cutoff—as it is the only margin that the program affects school attachment of refugee students.[16] We use three outcome variables measuring different aspects of achievement. The first outcome variable measures the grades of Syrian students in terms of levels, which we refer to as "absolute performance." Second, we construct a "relative performance" variable measuring the grades of refugee students in a given grade/year relative to the grades of Turkish students in the same school and grade/year. In doing so, first we calculate the mean grade for Turkish students; then, we subtract this mean from the grade of each Syrian student; and, finally, standardization is performed over this relative measure. And, third, we construct a dummy variable indicating whether a Syrian student has a missing grade, which measures academic participation. This variable can be interpreted as a measure of willingness to engage in academic activities, striving for success, or getting involved in academic competition.

Similar to our analysis of the absenteeism outcomes, we estimate the impact of the teacher training program on grades both in the short-term and longer-term. Short-term estimates refer to the effect observed during the semester immediately after the program implementation, while the longer-term estimates explore whether the program had any persistent effects in the following academic year. Panel A in Table 11 presents the estimates for absolute and relative performance for Turkish grades in the short-term. The results suggest that absolute performance improved by 0.228-0.241 standard deviation, while the improvement in relative performance is in the order of 0.398-0.466 standard deviation. Panel B shows that the longer-term effects are somewhat smaller and less significant in statistical terms. Panel C presents the short-term estimates for Math. The absolute performance improved by 0.147-0.164 standard deviation and the relative performance improved by 0.191-0.224 standard deviation. Panel D suggests that, similar to Turkish grades, the impact of the program on Math grades tends to shrink over time. The estimates for all three window sizes are statistically significant for the short-term estimates, while longer-term estimates become statistically insignificant as window size increases. These results suggest that the teacher training program improved the academic performance—measured in terms of language and cognitive skills—of refugee students.

[16] Note that we present visual evidence for all cutoffs in the Appendix. See Figures A3 and A4.

Table 11: Grades

		Turkish					
		Absolute performance			**Relative performance**		
	Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Difference in means	Fisherian p-value	Neyman p-value
A. Short-term	15 (w:2)	0.241**	0.043	0.038	0.466***	0.002	0.005
	15 (w:4)	0.228**	0.028	0.024	0.398***	0.009	0.010
	15 (w:6)	0.238**	0.022	0.019	0.405**	0.016	0.034
B. Longer-term	15 (w:2)	0.188**	0.047	0.044	0.317**	0.014	0.021
	15 (w:4)	0.181*	0.065	0.076	0.277**	0.023	0.031
	15 (w:6)	0.167	0.106	0.111	0.265**	0.032	0.033
		Math					
		Absolute performance			**Relative performance**		
	Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Difference in means	Fisherian p-value	Neyman p-value
C. Short-term	15 (w:2)	0.164**	0.049	0.042	0.224**	0.038	0.041
	15 (w:4)	0.151*	0.055	0.052	0.191*	0.056	0.061
	15 (w:6)	0.147*	0.068	0.076	0.206**	0.049	0.048
D. Longer-term	15 (w:2)	0.122*	0.091	0.089	0.181**	0.049	0.050
	15 (w:4)	0.101	0.155	0.161	0.159*	0.078	0.084
	15 (w:6)	0.093	0.187	0.193	0.148*	0.091	0.094

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Estimates are reported in terms of standard deviations.

The improvement was recorded in both absolute terms and relative to Turkish students. The relative improvement is particularly interesting and can be interpreted as reducing ethnic gap/inequality in academic achievement. Note that we find no gender differences in the estimates.

Table 12 presents the short-term and longer-term estimates for the outcome variable that measures the probability of having missing grades for refugee students. The results suggest that, in the short-term, the program reduced the probability by 0.121-0.131 percentage points for Turkish grades and 0.108-0.115 percentage point for Math grades. In other words, the teacher training program improved the academic participation of refugee students by 0.11 to 0.13 percentage points. Similar to the absolute and relative performance results, the estimates do not exhibit any statistically significant gender differences. The estimates also suggest that the program effects tend to dissipate over time.

Table 12: Probability of having missing grades

		Turkish			**Math**		
	Cutoff (window size)	Difference in means	Fisherian p-value	Neyman p-value	Difference in means	Fisherian p-value	Neyman p-value
A. Short-term	15 (w:2)	-0.121***	0.009	0.006	-0.109**	0.046	0.050
	15 (w:4)	-0.124**	0.026	0.034	-0.115**	0.034	0.033
	15 (w:6)	-0.131**	0.019	0.016	-0.108*	0.057	0.071
B. Longer-term	15 (w:2)	-0.079**	0.048	0.044	-0.091*	0.087	0.089
	15 (w:4)	-0.061*	0.089	0.093	-0.093*	0.085	0.091
	15 (w:6)	-0.058	0.105	0.121	-0.079	0.109	0.107

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Estimates are reported in terms of percentage points.

7 Exploring the mechanisms

7.1 Absenteeism

The design of the teacher training program has three main properties: (i) the training is not provided to teachers in schools with less than 15 refugee students; (ii) the number of trained teachers is not increased further above the last cutoff (100); and (iii) only up to 6 teachers are trained out of a possible 30 teachers per school, on average.[17]

This design implies that the teacher training program has a "partial" nature in the sense that it aims to train a small fraction of teachers in schools that are densely populated by refugee students. Recall that the training program keeps the number of refugee students per trained teacher in the 10-20 range for schools with 15-100 refugee students, and then lets this ratio increase after having 6 trained teachers per school. Therefore, the program

[17] School-level summary statistics given in Table 2 suggest that the average school size in the range of intervention is approximately 750. The average class size is typically 25 in Turkey. This means that the average number of teachers per school is roughly 30.

"injects" trained teachers into schools in a targeted way rather than training all possible teachers who directly interact with refugee students, and aims to maintain a certain ratio of trained teachers to refugee children in the school.

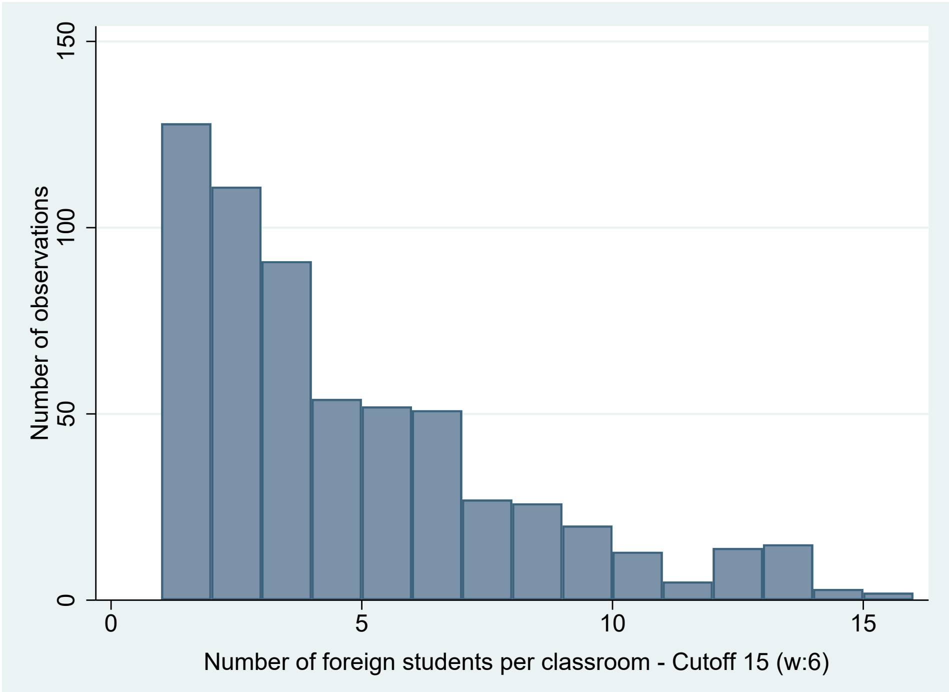
Table 13: Refugees per classroom around the 15 cutoff (w:6)

1st percentile	1
5th percentile	1
10th percentile	1
25th percentile	2
50th percentile	3
75th percentile	6
90th percentile	9
95th percentile	12
Mean	4.253
Std. deviation	3.250
# of observations	612

We argue that the main mechanism through which the teacher training program affects school attendance outcomes of refugee students can be interpreted as an "ambassador" or "champion" effect. The program aims to raise the general level of awareness toward the needs of refugee students and build a refugee-friendly school atmosphere rather than only trying to affect student outcomes through direct interactions between trained teachers and the refugee students in their classrooms. The cascade-training strategy used by MoNE in implementing the program supports this argument. Table 13 provides evidence in support of this interpretation. In particular, it shows the distribution of refugees per classroom around the 15 cutoff—the margin mainly driving the program effects. The average number of refugees per classroom is 4.3 and the median is 3. Figure 6 also shows the histogram of this distribution. Refugee students are spread across many classrooms rather than assigned to a single classroom as a large cluster. This suggests that refugee students are not assigned to only a few teachers in the treated schools; therefore, the program effects are not primarily driven by direct interactions between the trained teachers and the refugee students in their own classrooms. Instead, it is likely that the training program changed the overall school environment through potential spillovers to other teachers and school staff—and perhaps to native students, too. The trained teachers served as mentors to raise the school-wide awareness about the needs of refugee students. The program ingredients explained in Section 3.2 are also in line with this proposed mechanism.

Further evidence in support of this interpretation is provided in Figure 7, which plots the cumulative and probability distribution functions of the days of absence before and after the training program for students in schools just above the first cutoff. What is evident in these figures is that the reduction in days of absence in the post training semester is widespread among the students in treated schools. In particular, the pre-treatment cdf

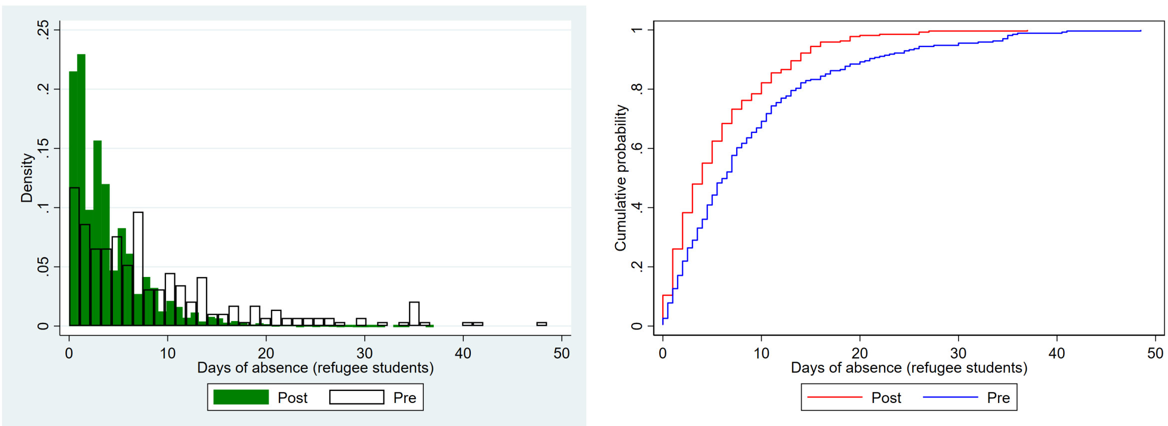
Figure 6: Refugee students per classroom



Notes: This figure plots the histogram of the number of refugee students per classroom for 3 mass points around the first cutoff.

first-order stochastically dominates that of the post-treatment—a Kolmogorov-Smirnov test rejects equality of the two cdfs ($p < 0.001$). This is consistent with the notion that indeed the training must have had a wider impact in the school beyond the possible changes in the teaching practices of the participating teacher.

Figure 7: Distribution of days of absence for students above the first cutoff



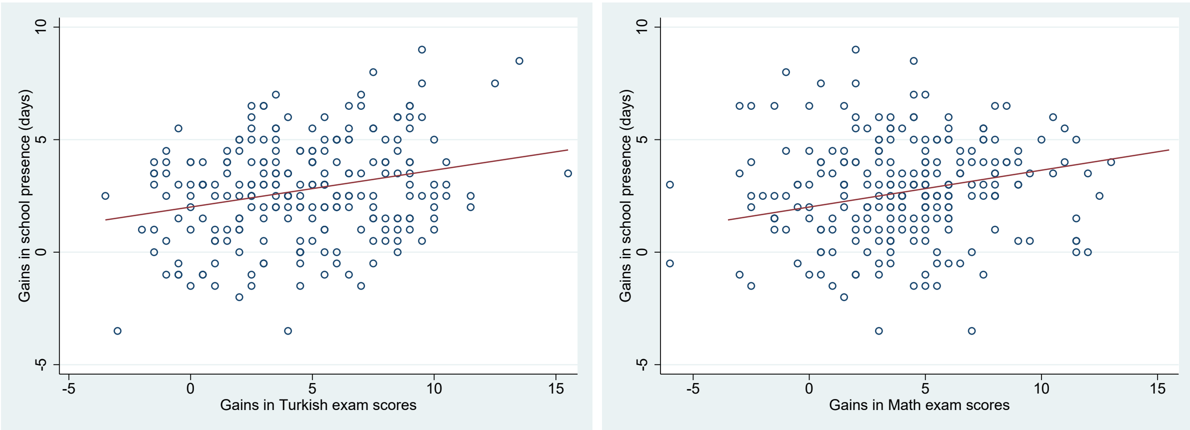
Notes: This figure displays the pdf (left) and cdf (right) of days of absence for schools with number of foreign students between 15 and 17, which constitute the treated group for the first cutoff. The figure plots the distribution of days of absence for the same students before and after the treatment.

It should also be noted that, although there may be other concurrent programs aiming to integrate refugee students into the Turkish education system, there is no other educational integration program operating around the 15 cutoff, which suggests that our estimates are not contaminated by any other program with a similar assignment mechanism.

7.2 Grades

Next we examine the potential drivers of the improvements in grades we found in section 6. One key channel that we explore is whether the decline in absenteeism generated by the teacher training program led to an improvement in refugee students' grades, i.e., whether higher school attachment has improved academic achievement.

Figure 8: Gains in absenteeism versus gains in grades



Notes: The figures plot the student-level gains in absenteeism and gains in (Turkish language and Math) grades for Syrian students in schools just above the first cutoff (N=269).

With the data set at hand, it is not easy to establish a causal relationship going from absenteeism to grades. But observing the joint distribution of grades and days of absence for each student in our sample allows us to examine the correlation between gains in school presence and gains in grades. Figure 8 illustrates scatter plots of these results (Turkish language on the left, Math on the right) for Syrian students. In this exercise, we focus on the Syrian students in treated schools just above the first cutoff—i.e., schools with 15, 16, and 17 refugee students. The sample size is 269. We calculate the decline in absenteeism and increase in grades from Fall (pre-treatment) to Spring (treatment) semester of 2017-18 academic year. The correlations reported in Figure 8 are statistically significant and approximately 0.2 for both grades.[18]

We interpret the finding that the gains in grades for Syrian students are positively correlated with gains in school presence as suggestive evidence of improved learning in response to increased school attachment among Syrian students. If the estimates were driven by changes in teachers' grading standards (e.g., teachers becoming more lenient

[18]It should be noted that there are students who did not participate in the exams in Fall, but took the exams in Spring and vice versa. For these students, participating in exams or not (i.e., having non-zero scores or not) is highly positively correlated with large changes in school presence. The number of these outlier observations is around 35 and including them largely improves the positive correlations reported in Figure 8. We prefer to exclude those outliers, which suggests that the reported correlations are conservative.

toward refugee students), then there would be no reason to see a meaningful positive correlation between changes in school attachment and changes in academic performance at student level. Instead, gains in grades would be more homogeneously distributed over the gains in absenteeism distribution. Moreover, only a small number of teachers are trained and the grade effect is driven by the first-cutoff schools with only one trained teacher, which makes it unlikely that the estimated improvement in grades comes from a change in school-wide grading behavior. Overall, we argue that the improvement in grades is a consequence of increased school attachment of Syrian children, which is likely caused by the broad school-wide mentoring impact of the teacher training program.

8 Evaluating policy effectiveness

The teacher training program was funded by the EU Facility for Refugees in Turkey (FRIT)—see Section 3.1 for the details—and the total budget allocated to the program was EUR 4.8 million. The main expenditure items were (i) technical preparation of the program curriculum, (ii) training for trainers, (iii) provision of training to the teachers, (iv) hard copies of program material, (v) travel and accommodation, (vi) program monitoring and post-program evaluation, (vii) administrative and operational expenses, and (viii) other incidentals.

Considering that the total number of teachers who completed the program and received certification is 8,661 (out of 8,900 registered teachers), the program cost per teacher is approximately EUR 554. Our findings suggest that the teacher training program non-negligibly improved school attendance of Syrian students and the effects persist into the following semesters. In particular, we find that the training program reduced the days of absence per semester by around 2.7 days in the short term and 1.5 days in longer term. The program was implemented in 26 PIKTES provinces in schools with 15 refugee students and above. This suggests that the program covered roughly 320,000 refugee students. Therefore, the cost of the program for each reduced day of school absence by refugee students is in the range of EUR 5.6-10. Consequently, closing the school attendance gap between Turkish and Syrian students—which is approximately 5.5 days—through teacher training programs costs around EUR 30.8-55 per refugee student per semester, which is quite cost effective considering the extremely large budgets allocated to refugee integration programs, especially the conditional cash transfer programs with huge budgets.[19] Note that the above cost-effectiveness calculations do not take into account the additional benefits of the program associated with gains in academic performance accompanying the improvement in school attendance and other potential long-term benefits, which can significantly add to the cost effectiveness of the program.

[19] See https://ec.europa.eu/commission/presscorner/detail/en/IP_20_1324 for the cash transfer programs that are implemented in Turkey.

9 Conclusions

The integration of refugee children into host countries' education systems is crucial for their academic development, social integration, and future labor market advancement. Refugee children face various obstacles in integrating into host countries' education systems, and their enrollment rates and academic achievement levels remain low. Despite the importance of the issue, research on policies to support the educational integration of refugee children, and more specifically on the role that teachers can play in this regard remains rather limited. We investigate whether a training program aimed to raise awareness of primary and secondary teachers about the needs of refugee students in Turkey is effective in improving educational outcomes of those students. We use school administrative records, and employ a regression discontinuity design that exploits discontinuities in the rule that determines eligibility to the training program.

Our findings show that teachers' training leads to a substantial reduction in absenteeism of refugee students, effectively closing by half the gap in absenteeism rate between native and refugee students. Specifically, the estimates suggest that refugee students' absenteeism decreased by around 2.7 days in the semester that followed the training, implying a 32.5% reduction relative to the previous semester. Examining the longer term effects of the program, we find lasting effects—around 1.5 days reduction in the first semester of the following academic year. We also find positive impacts on the educational attainment of refugee students, as measured by grades in Turkish language and Math subjects, and suggestive evidence that there is a positive correlation between improved attendance and grades. These findings suggest that this awareness training program prepared trained teachers to act as "ambassadors" or "mentors" for refugee children and resulted in a better overall educational environment where refugee children became less likely to miss school. The fact that the effect is fading over time suggests that such type of a training program would be more effective if it becomes more regular, perhaps annual, to ensure maximum effectiveness.

Our cost-effectiveness calculations suggest that host countries are potentially underinvesting in programs aiming to equip teachers with the necessary skills to address the needs of refugee students. There are several chronic problems in many host countries related to teachers' capacity to address refugee students' needs—e.g., language and communication skills, and other complementary skills related to teachers' professional development, such as teaching minorities, basic counselling, relationship with parents, intra-class conflict resolution, and classroom management in diverse environments (Cerna, 2019). Our findings indicate that better preparing teachers to face the multidimensional challenges in diverse educational settings could substantially improve the effectiveness of refugee integration policies.

References

Alan, S., C. Baysan, M. Gumren, and E. Kubilay (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *Quarterly Journal of Economics 136*(4), 2147–2194.

Alan, S., E. Duysak, E. Kubilay, I. Mumcu, et al. (2021). Social exclusion and ethnic segregation in schools: The role of teacher's ethnic prejudice. *Review of Economics and Statistics, forthcoming*.

Alesina, A., M. Carlana, E. L. Ferrara, and P. Pinotti (2018). Revealing stereotypes: Evidence from immigrants in schools. National Bureau of Economic Research, WP#25333.

Angrist, J. D. and V. Lavy (2001). Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools. *Journal of Labor Economics 19*(2), 343–369.

Assaad, R., T. Ginn, and M. Saleh (2019). Impact of Syrian refugees on education outcomes in Jordan. *CEPR Discussion Paper #14056*.

Aucejo, E. M. and T. F. Romano (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review 55*, 70–87.

Ballatore, R. M., M. Fort, and A. Ichino (2018). Tower of Babel in the classroom: Immigrants and natives in Italian schools. *Journal of Labor Economics 36*(4), 885–921.

Borghans, L., B. H. Golsteyn, J. J. Heckman, and J. E. Humphries (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences 113*(47), 13354–13359.

Bossavie, L. (2020). The effect of immigration on natives' school performance: Does length of stay in the host country matter? *Journal of Human Resources 55*, 733–766.

Botelho, F., R. A. Madeira, and M. A. Rangel (2015). Racial discrimination in grading: Evidence from Brazil. *American Economic Journal: Applied Economics 7*(4), 37–52.

Boucher, V., S. Tumen, M. Vlassopoulos, J. Wahba, and Y. Zenou (2021). Ethnic mixing in early childhood: Evidence from a randomized field experiment and a structural model. CEPR Discussion Paper #15528.

Bratsberg, B., O. Raaum, and K. Røed (2012). Educating children of immigrants: Closing the gap in Norwegian schools. *Nordic Economic Policy Review 3*(1), 211–251.

Bressoux, P., F. Kramarz, and C. Prost (2009). Teachers' training, class size and students' outcomes: Learning from administrative forecasting mistakes. *The Economic Journal 119*(536), 540–561.

Bruner, C., A. Discher, and H. Chang (2011). Chronic elementary absenteeism: A problem hidden in plain sight. Research Brief: Attendance Works and Child & Family Policy Center.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). rdrobust: Software for regression discontinuity designs. *Stata Journal 17*, 372–404.

Carlana, M., E. La Ferrara, and P. Pinotti (2022). Goals and gaps: Educational careers of immigrant children. *Econometrica 90*(1), 1–29.

Cattaneo, M. D., B. R. Frandsen, and R. Titiunik (2015). Randomization inference in the regression discontinuity design: An application to party advantages in the US Senate. *Journal of Causal Inference 3*(1), 1–24.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2018). A practical introduction to regression discontinuity designs: Volume II. *Cambridge Elements: Quantitative and Computational Methods for Social Science, II 113*.

Cattaneo, M. D., R. Titiunik, and G. Vazquez-Bare (2017). Comparing inference approaches for RD designs: A reexamination of the effect of Head Start on child mortality. *Journal of Policy Analysis and Management 36*(3), 643–681.

Cerna, L. (2019). Refugee education: Integration models and practices in OECD countries. OECD Education Working Paper #203.

Cilliers, J., B. Fleisch, C. Prinsloo, and S. Taylor (2020). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources 55*(3), 926–962.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics 86*(1), 195–210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review 95*(2), 158–165.

Dryden-Peterson, S. (2015). *The educational experiences of refugee children in countries of first asylum*. British Columbia Teachers' Federation.

Dustmann, C. and A. Glitz (2011). Migration and education. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 4, Chapter 4, pp. 327–439. Elsevier.

Egalite, A. J., B. Kisida, and M. A. Winters (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review 45*, 44–52.

Felfe, C., H. Rainer, and J. Saurer (2020). Why birthright citizenship matters for immigrant children: Short-and long-run impacts on educational integration. *Journal of Labor Economics 38*(1), 143–182.

Figlio, D. and U. Ozek (2019). Unwelcome guests? The effects of refugees on the educational outcomes of incumbent students. *Journal of Labor Economics 37*(4), 1061–1096.

Figlio, D. N., P. Giuliano, R. Marchingiglio, U. Ozek, and P. Sapienza (2021). Diversity in schools: Immigrants and the educational performance of US born students. *NBER Working Paper #28596*.

Frattini, T. and E. Meschi (2019). The effect of immigrant peers in vocational schools. *European Economic Review 113*, 1–22.

Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy 11*(2), 125–149.

Gershenson, S., A. Jacknowitz, and A. Brannegan (2017). Are student absences worth the worry in US primary schools? *Education Finance and Policy 12*(2), 137–165.

Gould, E. D., V. Lavy, and M. Daniele Paserman (2009). Does immigration affect the long term educational outcomes of natives? Quasi-experimental evidence. *Economic Journal 119*(540), 1243–1269.

Green, C. and J. M. V. Iversen (2020). Refugees and the educational attainment of natives. IZA Discussion Paper #13433.

Gubbels, J., C. E. van der Put, and M. Assink (2019). Risk factors for school absenteeism and dropout: a meta-analytic review. *Journal of Youth and Adolescence 48*, 1637–1667.

Hanna, R. N. and L. L. Linden (2012). Discrimination in grading. *American Economic Journal: Economic Policy 4*(4), 146–68.

Hanushek, E. A. and S. G. Rivkin (2006). Teacher quality. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 2, pp. 1051–1078. Elsevier.

Harris, D. N. and T. R. Sass (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics 95*(7-8), 798–812.

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources 52*(4), 1060–1118.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies 79*, 933–959.

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615–635.

IMF (2021). World Economic Outlook: Managing divergent recoveries. International Monetary Fund, April.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non-test score outcomes. *Journal of Political Economy 126*(5), 2072–2107.

Jackson, C. K., J. E. Rockoff, and D. O. Staiger (2014). Teacher effects and teacher-related policies. *Annual Review of Economics 6*(1), 801–825.

Jacob, B. A. and L. Lefgren (2004). The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in Chicago. *Journal of Human Resources 39*(1), 50–79.

Liu, J. and S. Loeb (2019). Engaging teachers: Measuring the impact of teachers on student attendance in secondary school. *Journal of Human Resources 56*, 343–379.

Lounsbury, J. W., R. P. Steel, J. M. Loveland, and L. W. Gibson (2004). An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence 33*, 457–466.

Loyalka, P., A. Popova, G. Li, and Z. Shi (2019). Does teacher training actually work? Evidence from a large-scale randomized evaluation of a national teacher training program. *American Economic Journal: Applied Economics 11*(3), 128–154.

Ohinata, A. and J. C. van Ours (2013). How immigrant children affect the academic achievement of native Dutch children. *Economic Journal 123*(570), F308–F331.

Schnepf, S. V. (2007). Immigrants' educational disadvantage: An examination across ten countries and three surveys. *Journal of Population Economics 20*, 527—545.

Sieverding, M., C. Krafft, N. Berri, C. Keo, and M. Sharpless (2018). Education interrupted: Enrollment, attainment, and dropout of Syrian refugees in Jordan. Economic Research Forum, WP#1261.

Smith, J. P. (2006). Immigrants and their schooling. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, pp. 155–187. Elsevier.

Tumen, S. (2018). The impact of low-skill refugees on youth education. *IZA Discussion Paper #11869*.

Tumen, S. (2019). Refugees and 'native flight' from public to private schools. *Economics Letters 181*, 154–159.

Tumen, S. (2021). The effect of refugees on native adolescents' test scores: Quasi-experimental evidence from PISA. *Journal of Development Economics*, 102633.

UNHCR (2017). Left behind: Refugee education in crisis. Technical report.

UNHCR (2019). Global trends: Forced displacement 2019. Technical report.

UNHCR (2020). Coming together for refugee education. Technical report.

Training Teachers for Diversity Awareness: Impact on School Outcomes of Refugee Children

Appendix

By Semih Tumen, Michael Vlassopoulos, and Jackline Wahba

A Additional Figures and Tables

Figure A1: Visual evidence—continuity-based RD analysis

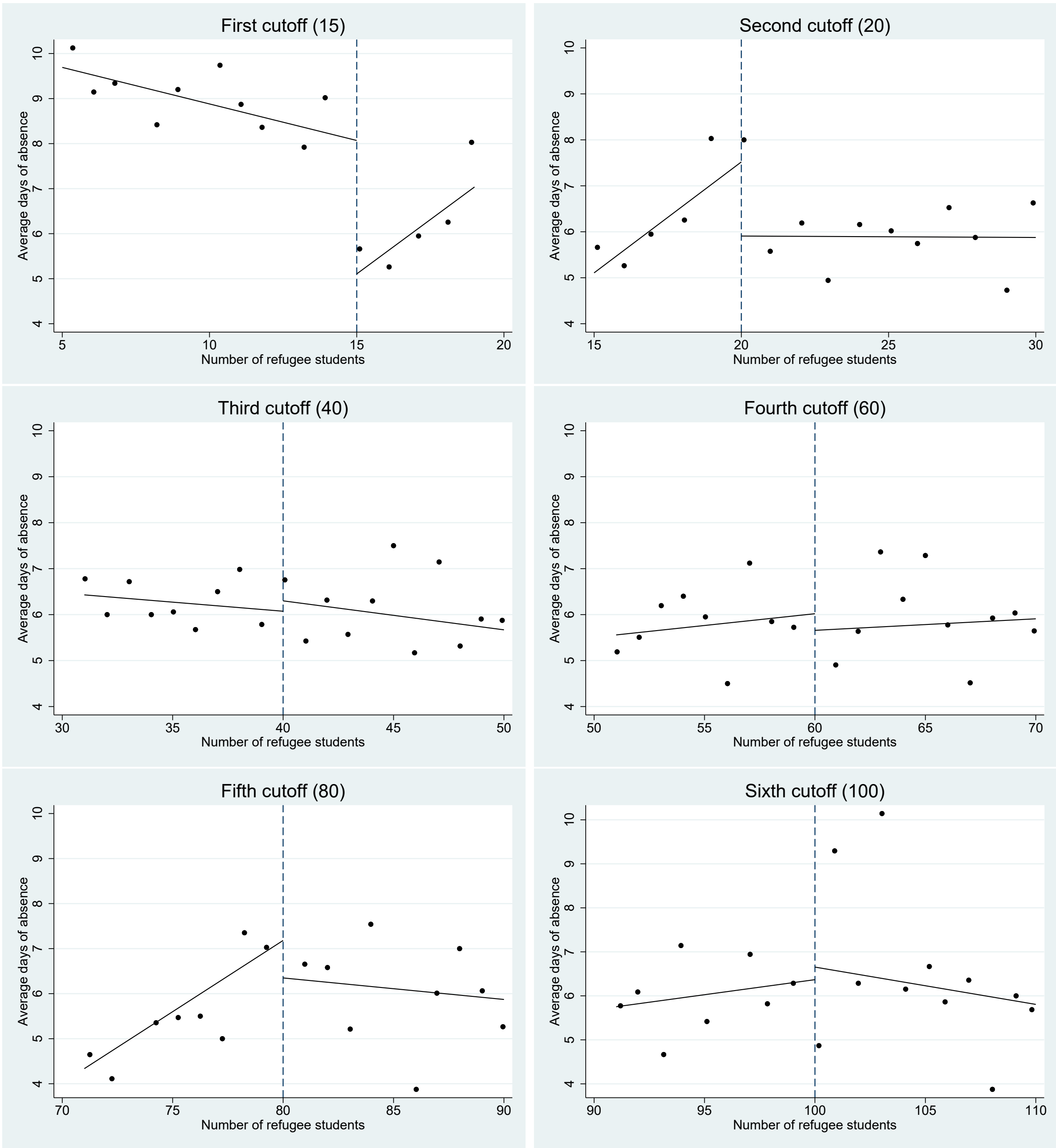


Table A1: Baseline estimates—continuity-based RD analysis

	Linear (CCT)	Quadratic (CCT)	Linear (IK)
First cutoff (15)	-4.4314**	-4.8025**	-4.5665**
s.e.	(2.2421)	(2.6026)	(2.2548)
Optimal bandwidth	3.27	3.76	2.39
# of observations	1,499	1,499	1,499
Second cutoff (20)	-2.1655	-3.5122	-2.1486
s.e.	(2.3643)	(3.8949)	(2.0862)
Optimal bandwidth	3.40	4.89	2.33
# of observations	1,244	1,244	1,244
Third cutoff (40)	3.2372	3.8095	2.0599
s.e.	(2.7594)	(3.8191)	(1.9459)
Optimal bandwidth	2.23	3.27	2.78
# of observations	1,498	1,498	1,498
Fourth cutoff (60)	-1.1794	-0.9125	-1.4239
s.e.	(2.5330)	(4.4377)	(1.4926)
Optimal bandwidth	2.25	3.48	2.05
# of observations	1,595	1,595	1,595
Fifth cutoff (80)	0.7722	0.5300	-0.6436
s.e.	(1.7624)	(2.0662)	(1.3280)
Optimal bandwidth	2.40	2.59	3.74
# of observations	1,070	1,070	1,070
Sixth cutoff (100)	-0.6475	-1.0981	-0.7226
s.e.	(0.6501)	(1.5278)	(0.8002)
Optimal bandwidth	3.17	2.64	3.38
# of observations	1,313	1,313	1,313

Notes:*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Heteroskedasticity and cluster-robust standard errors are in parentheses. The unit of clustering is school. CCT and IK refer to the optimal bandwidth implementation procedures developed by Calonico et al. (2017) and Imbens and Kalyanaraman (2012), respectively.

Figure A2: Visual evidence—continuity-based longer-term (Fall 2018) RD effects

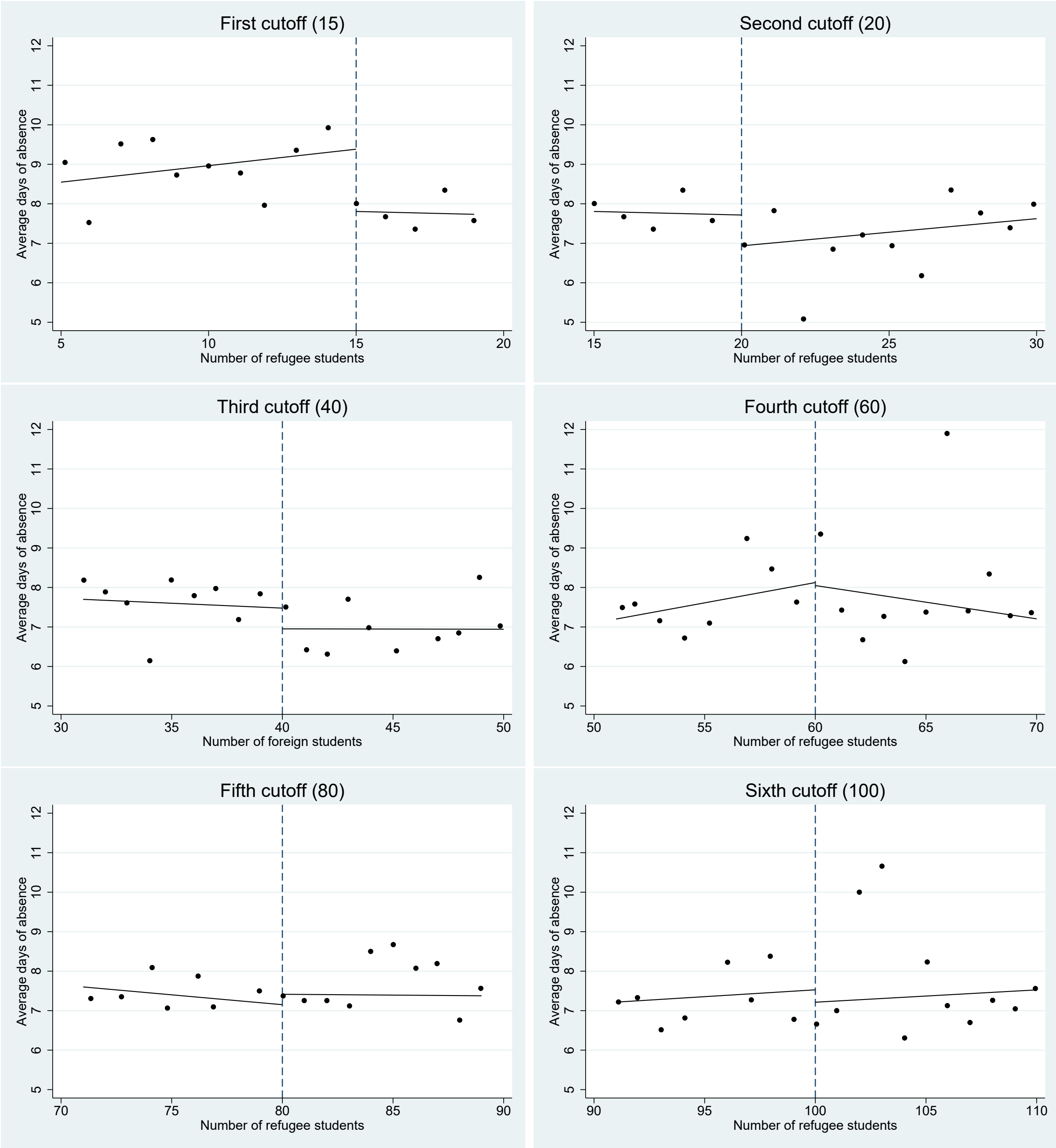
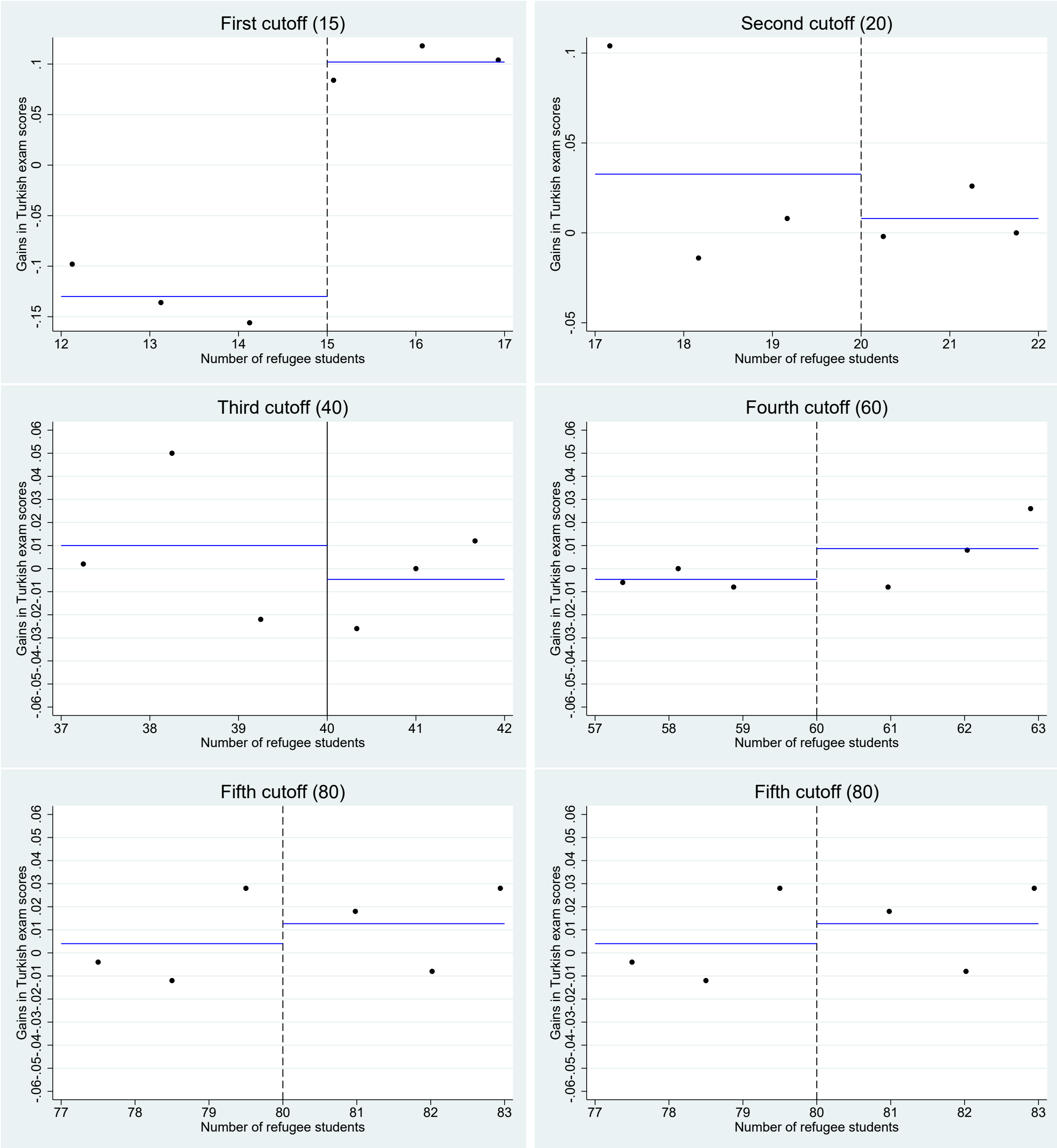


Table A2: Longer-term effects—continuity-based RD analysis

	Linear (CCT)	Quadratic (CCT)	Linear (IK)
First cutoff (15)	-3.3635*	-2.7641	-2.4901
s.e.	(2.0189)	(3.0812)	(1.8186)
Optimal bandwidth	3.33	4.29	2.28
# of observations	1,774	1,774	1,774
Second cutoff (20)	0.5479	3.9671	0.2952
s.e.	(1.8201)	(4.8521)	(1.3144)
Optimal bandwidth	3.19	3.80	2.79
# of observations	1,374	1,374	1,374
Third cutoff (40)	-0.9041	-3.4598	-0.4878
s.e.	(1.7270)	(4.3243)	(1.2429)
Optimal bandwidth	3.18	3.15	3.27
# of observations	1,500	1,500	1,500
Fourth cutoff (60)	1.6891	2.1957	1.5241
s.e.	(1.0398)	(1.6741)	(1.2869)
Optimal bandwidth	1.87	2.81	2.84
# of observations	2,098	2,098	2,098
Fifth cutoff (80)	-0.4625	-0.3712	-0.1358
s.e.	(1.9307)	(1.3226)	(0.9327)
Optimal bandwidth	3.45	3.73	2.90
# of observations	1,529	1,529	1,529
Sixth cutoff (100)	-0.6240	-0.6462	-0.3869
s.e.	(0.6058)	(0.7819)	(0.5893)
Optimal bandwidth	1.71	2.45	3.00
# of observations	1,484	1,484	1,484

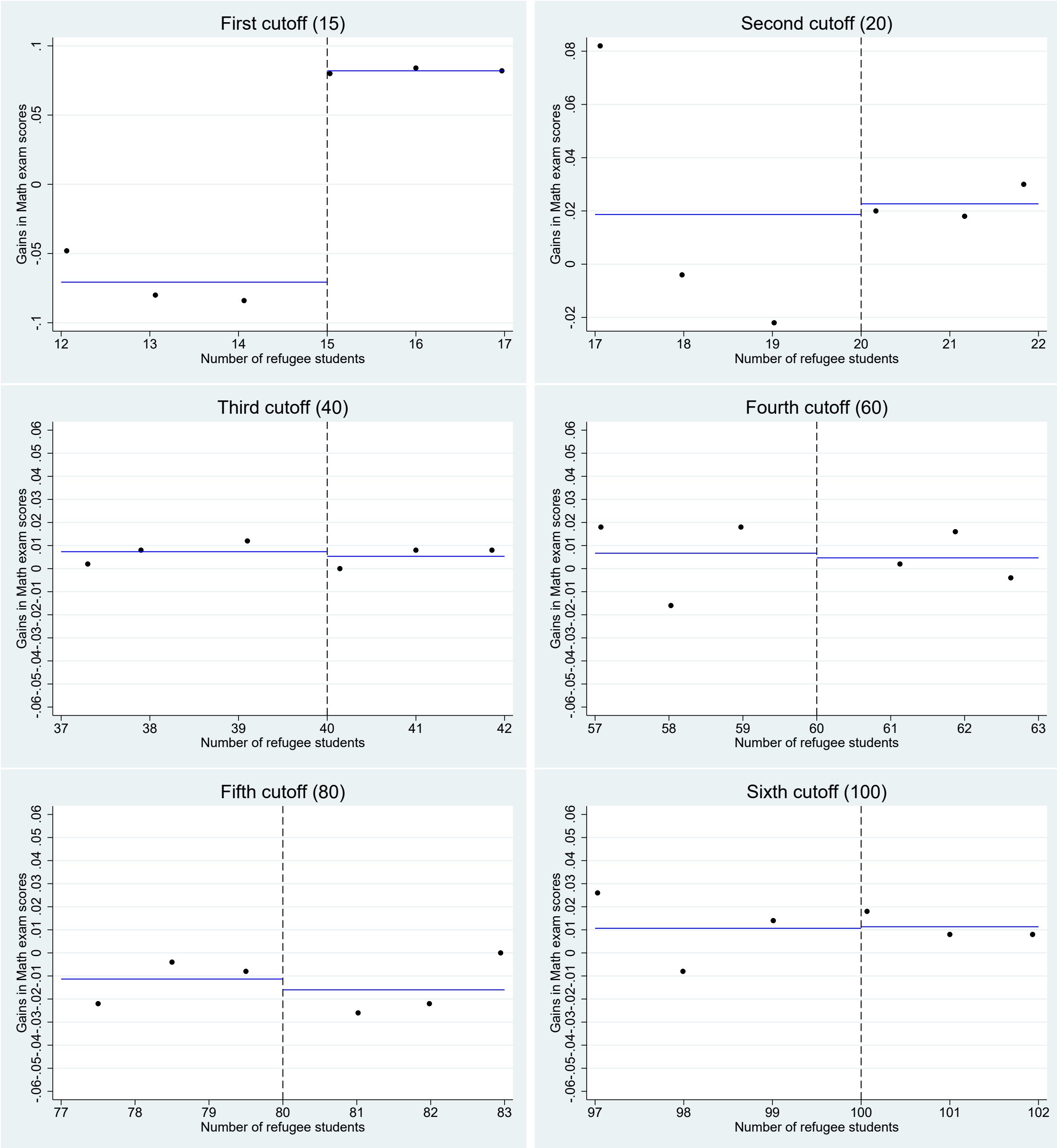
Notes:*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Heteroskedasticity and cluster-robust standard errors are in parentheses. The unit of clustering is school. CCT and IK refer to the optimal bandwidth implementation procedures developed by Calonico et al. (2017) and Imbens and Kalyanaraman (2012), respectively.

Figure A3: Visual evidence—local randomization (Turkish grades)



Notes: This figure provides visual evidence for our local randomization RD analysis based on Turkish language grades for 3 mass points around each of our cutoffs.

Figure A4: Visual evidence—local randomization (Math grades)



Notes: This figure provides visual evidence for our local randomization RD analysis based on Math grades for 3 mass points around each of our cutoffs.