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

Semih Tumen, Michael Vlassopoulos, Jackline Wahba

Journal of Human Resources, Volume 61, Number 1, January 2026, pp. 84-122 (Article)

Published by University of Wisconsin Press

Training Teachers for Diversity Awareness

Impact on School Outcomes of Refugee Children

Semih Tumen
Michael Vlassopoulos
Jackline Wahba

ABSTRACT

Using administrative data and a regression discontinuity approach, we evaluate the effectiveness of a teacher training program aiming to improve school outcomes of refugee students. The program halves the absenteeism gap between native and refugee students, and its effect persists into the next academic year. It also improves the grades of refugee students in Turkish language and math, and there is a positive association between improved attendance and grades. One possible channel through which the program's effects may operate is a school-wide champion role assumed by trained teachers, which has a broad impact on raising diversity awareness within schools.

I. Introduction

In 2019, children constituted 40 percent 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

Semih Tumen is at TED University, Department of Economics, Türkiye and IZA (corresponding author, semih.tumen@tedu.edu.tr). Michael Vlassopoulos is at University of Southampton, Department of Economics and IZA (m.vlassopoulos@soton.ac.uk). Jackline Wahba is at University of Southampton University of Southampton, Department of Economics, UK and IZA (j.wahba@soton.ac.uk). *The authors thank Matias Cattaneo, Jamal Haidar, Pinar Ozel, Serdar Polat, Metin Catar, Secil Balkir, Elif Kubilay, three anonymous referees, the participants of the 3rd LSE Workshop on Political Economy of Türkiye, the 3rd EBRD & King's College London Workshop on the Economics and Politics of Migration, the 6th Istanbul Human Capital Conference at Bahcesehir University, and seminar participants at the American University of Cairo, Pamukkale University, and TED University for useful comments. This paper uses confidential data from the Ministry of National Education of the Republic of Türkiye administrative student records. Models/codes are available upon request from the corresponding author.*
[Submitted June 2022; accepted May 2023]; doi:10.3368/jhr.0622-12378R2
JEL Classification: I21, I28, and J15
ISSN 0022-166X E-ISSN 1548-8004
Color version of this article is available online at: https://jhr.uwpress.org.
Supplementary materials are available online at: https://jhr.uwpress.org.
Semih Tumen https://orcid.org/0000-0002-8841-2290

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 Krafft et al. (2022).[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 below, 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 and managing diverse classrooms, fully aware of the circumstances of refugee children, and able to address their needs. This work aims to fill these gaps by studying a large-scale teacher training program in Türkiye 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 2017–2018 by the Ministry of National Education of the Republic of Türkiye (MoNE) and was funded by the EU. The main goals of the training program were to increase teachers' awareness of the immediate needs of refugee students and to encourage and 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, which corresponds to, on average, a 21 percent refugee to population ratio.

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

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

education systems.[2] In fact, the average refugee student is chronically absent from school (being absent from school for at least 10 percent of the school days), which is widely used as a critical early warning signal for academic risk and school dropout (Bruner, Discher, and Chang 2011). More generally, school absenteeism is an important concern, especially for disadvantaged groups, because 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, Jacknowitz, and Brannegan 2017). Furthermore, absenteeism is a proxy for noncognitive 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, van der Put, and Assink 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 explore 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, Idrobo, and Titiunik 2023). 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 15–19 refugee students have one trained teacher. The number of trained teachers is incrementally increased up to six 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.

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

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, and 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.

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—being absent from school for at least 10 percent of school days in an academic year—and school drop-out rates.

A number of validation and falsification tests provide 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 a 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 of 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 intraclassroom 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. The grades reflect academic performance based on various forms of assessment, including written exams, oral exams, quizzes, and homework. Grades in these two subjects are particularly important because performance in Turkish language measures social and educational integration of refugee students, while performance in math proxies their cognitive and 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. We also document that there is a strong positive correlation between gains in school presence and gains in grades (that is, those who decreased their absenteeism the most are the ones who also improved the most academically). The positive association between improvements in attendance and grades is consistent with the interpretation that the training program may have led to an underlying improvement in learning as a consequence of decreased absenteeism. This implies that the benefits of the training program extend beyond the absenteeism outcomes and also include improvements in refugee students' academic achievement.

Finally, we assess 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 absenteeism between native and refugee students through awareness-raising programs is quite cost-effective, considering the evidence on large lifetime costs of school absenteeism.

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 of 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.

In the following, Section II reviews the related literature. Section III explains the institutional background and the design of the training program. Section IV describes the data used in our analysis and the empirical strategy. Section V presents results on absenteeism and Section VI on academic performance. Section VII discusses mechanisms, and Section VIII offers an evaluation of policy effectiveness. Section IX concludes.

II. 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, for example, Figlio and Ozek (2019); Assaad, Ginn, and Saleh (2023); Green and Iversen (2022); 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, Krafft et al. (2022) 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

3. There is also a more sizeable literature focusing on the impact of immigrant children on natives' educational performance—see, for example, Gould, Lavy, and Paserman (2009); Ohinata and van Ours (2013); Hunt (2017); Ballatore, Fort, and Ichino (2018); Frattini and Meschi (2019); Bossavie (2020); and Figlio et al. (2024).

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 underperform 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, Raaum, and Røed (2012). A few studies investigate the role of policies to integrate immigrant children. Felfe, Rainer, and Saurer (2020) find that the introduction of birthright citizenship enhances the educational integration of immigrant children in Germany. Carlana, Ferrara, and Pinotti (2022) estimate the impact of a program that provided tutoring and career counseling to immigrant children in Italy.

educational settings. Alan et al. (2021) study the impact of an educational program that aims to build social cohesion among refugee and native children, while Boucher et al. (2021) examine the impact of exogenous mixing of preschool 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 improving students' educational outcomes and school climate (Hanushek and Rivkin 2006; Jackson, Rockoff, and Staiger 2014; Jackson et al. 2020, 2023).[5] There is a large literature on the effectiveness of teachers in improving students' test scores, but less so on the effects of teachers on students' noncognitive skills, and in particular, on school attendance and absences (Gershenson 2016; Jackson 2018; Liu and Loeb 2021). For instance, Jackson (2018) proxies students' noncognitive skills using behaviors that include absences, suspensions, course grades, and repetition in the ninth grade/year.[6] 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, Kisida, and Winters 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, Madeira, and Rangel 2015; Alesina et al. 2018; Alan et al. 2023). Our study contributes to this literature in two related but distinct ways: (i) by showing strong evidence of the instrumental role that teachers can play in the educational inclusion and school attachment of diverse student populations and (ii) by providing evidence that increased school attachment also leads to substantial gains in academic performance.

Finally, our work 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, Kramarz, and Prost (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

5. Collaborative teachers who actively engage in improving school climate significantly contribute to school effectiveness. See, for example, https://consortium.uchicago.edu/publications/investing-in-adolescents (accessed July 2, 2025).

6. Throughout the paper, the term "grade" is used to describe two different concepts: (i) grade as the level of education (for example, the ninth grade) and (ii) grade as a measure of academic performance (that is, the marks). To differentiate these two, we use the term "grade/year" to define the former, while "grade" is used to capture the latter.

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. Popova et al. (2022) find an association between certain features of teacher professional development programs (for example, the provision of incentives such as promotion or salary increases and a specific subject-oriented focus) and improved student test scores. Finally, Loyalka et al. (2019) examine a large-scale randomized national professional development program for teachers in China but find that the 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.

III. Institutional Background

A. Education of Syrian Refugees in Türkiye

The Syrian conflict has led around 6.6 million Syrians to flee their homes—5.5 million of whom are hosted by neighboring countries.[7] As of January 2023, Türkiye hosts around 3.5 million Syrian refugees, and the vast majority of them live outside of camps. The Syrian population in Türkiye is, on average, younger than the native population. Based on the recent figures, the number of school-age refugee children (5–17) is around 1.1 million as of the 2019–2020 academic year, and their enrollment rate is roughly 64 percent (Tumen 2018). Therefore, focusing on education policies could improve the long-term effectiveness of the overall refugee integration policy framework.

During the early stages of the crisis, for example, 2011–2016, the policy priorities were shaped around the humanitarian needs. Education of refugees was handled as a service separately provided outside of the Turkish education system.[8] After 2016, the EU Facility for Refugees in Türkiye (FRIT, a 6 billion euro fund) was launched, and a sizable portion of the FRIT funds were used to prioritize the full integration of refugees into the Turkish public education system. Accordingly, various school integration programs were implemented through the PIKTES (Promoting Integration of Syrian Kids into the Turkish Education System) project, which was administered by MoNE and financed by FRIT funds. The main ones were the back-up training, catch-up training, Turkish language training, and teacher training programs. The back-up and catch-up training programs aimed to provide academic support to enrolled and out-of-school refugee students, respectively. The language training program aimed to improve Turkish language skills of refugee students. We focus on the teacher training program, which is described in the next section.

7. For detailed statistics, see https://www.unhcr.org/syria-emergency.html (accessed July 2, 2025).
8. See Boucher et al. (2021) for a more detailed chronology of these education services.

B. The Teacher Training Program as an Awareness-Raising Activity

Most refugee children were exposed to violence during the transition from Syria to Türkiye. Even without a direct exposure to violence, being forced out of their home and having to live in an unfamiliar culture are traumatic experiences for refugee children. In addition, school-age Syrian children had to stop their education as they moved to a new country. Hence, refugee children constitute a sensitive group, and they need support not only on education-related issues, but on other issues, such as counseling, developmental mentoring, relationship with Turkish peers and teachers, and adaptation to a new culture.

Teachers are in direct contact with refugee students, and increasing teachers' capacity to address the needs of those students is important from an integration-policy perspective. As we discuss in Section IV.B, refugee children have low school attachment. A training program was designed and implemented to increase teachers' ability to address Syrian children's vulnerabilities and increase their attachment to school. The main goals of the program were: (i) to increase teachers' awareness on educational integration issues, (ii) to improve their capacity to identify and address the needs of Syrian students, and (iii) to encourage them to actively contribute to the integration policies. Teachers from schools with a high concentration of Syrian students were the target group.

The program was implemented during the semester break of the 2017–2018 academic year over five days, with a 30-hour curriculum. The focus was on three main areas: (i) language and communication; (ii) social integration, counseling, and guidance; and (iii) legislation and context (see Online Appendix Table A1 for a detailed breakdown of the topics in the syllabus and the time allocated to each topic). The content of the program suggests that the trained teachers would have high awareness of the circumstances and experiences of refugee students and their families; better knowledge about refugees' rights, obligations, and legal status; better communication skills that enable them to more effectively interact with refugee students and parents; awareness that Turkish is a foreign language for refugee students and that the Turkish classes designed for native students may not effectively improve the language skills of refugee students; and enough knowledge about the context, best practices, and long-term strategy of MoNE and the educational integration of refugees. Following this program, trained teachers were expected to more effectively address the educational needs of vulnerable refugee students and to help other teachers in the school implement the best practices.

The program was designed based on an implicit "cascade training" philosophy. In a standard cascade training model, master trainers are trained on a specific topic and are then asked to formally train other staff. Unlike this standard approach, the teachers who attended the program were not explicitly asked to train other teachers in their school. Instead, the program highlighted an "expectation" that the trained teachers would pass on knowledge and good practices to the other teachers in their school. This expectation is reflected in the structure of the program's syllabus presented in Online Appendix Table A1. Specifically, the program did not provide any pedagogical skills or tools related to formal cascade training. In the closing session of the program, there was a two-hour seminar explaining MoNE's integration strategy and the expectation that the trained teachers would serve as

"champions" in their schools. This suggests that the teacher training program aimed to affect the broader education environment in schools through awareness spillovers among teachers.

Each public school has a unique catchment area, and children residing within the boundaries of a particular catchment area are assigned to the same school. Note that there are no location restrictions within provinces for refugee families. The participating schools were selected based on the number of Syrian students assigned to them.[9] In particular, teachers from schools with more than 15 Syrian students were eligible for the program. For schools above this threshold, the number of teachers eligible to participate in the program ranged from one to six, depending on the number of Syrian students. Specifically, the thresholds were defined as follows: one teacher from schools with 15–19 Syrian students, two teachers from schools with 20–39 Syrian students, three teachers from schools with 40–59 Syrian students, four teachers from schools with 60–79 Syrian students, five teachers from schools with 80–99 Syrian students, and six teachers from schools with 100 or more Syrian students. Therefore, the thresholds were defined as 15 20, 40, 60, 80, and 100. A total of 8,900 teachers from the 26 provinces with the highest refugee concentration were eligible for the program.[10] Of these, 8,661 teachers completed the 30-hour program, which suggests a 97.31 percent completion rate. All participating teachers were Turkish nationals.

The process of selecting teachers for the program was carried out as follows. Using school-level administrative records, the provincial MoNE staff determined the eligible schools and the number of teachers to be assigned from each eligible school, according to the thresholds outlined above. 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 the teacher training program, where $n = \{1, \ldots, 6\}$, the top n teachers from the sorted list were invited to the program. In cases where there were multiple suitable candidates for the nth slot, 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. If an assigned teacher was unable to attend, the administrators selected the next eligible teacher following the same steps.

Although there were other concurrent interventions aiming to improve refugees' school integration, there was 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.

9. Therefore, when we refer to a school with, for example, 15 refugee 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 at a school in Türkiye.

10. These were the provinces in which the FRIT-funded PIKTES projects were carried out, and they were determined based on a protocol signed by the European Commission and MoNE.

Overall, the teacher training program is a school-level intervention that assigned the teachers to the program based on the rules described above. From the viewpoint of Syrian students, attending a school with a trained teacher (or an additional trained teacher) is as-if randomized around the specified thresholds, 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.

IV. Empirical Strategy and Data

A. Empirical Strategy

1. 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: (i) the fact that the running variable (number of foreign students within a school) is discrete and has a few mass points, (ii) that there are multiple cumulative cutoffs governing treatment (number of teachers within a school that are eligible to receive the training), and (iii) 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 because using the continuity-based methods when the running variable is discrete can lead to significant biases (Cattaneo, Idrobo, and Titiunik 2023). 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, Idrobo, and Titiunik 2023). 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, Frandsen, and Titiunik 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:

$$(1) \quad E[Y_i(1)|r_s = c] \; - E[Y_i(0)|r_s = c_-],$$

where c denotes the cutoff, and c_- denotes the closest mass point below the cutoff.[11] 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, Frandsen, and Titiunik (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.

2. 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 four and six 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, Idrobo, and Titiunik (2023), we also present covariate balance tests for each window that we consider separately.

3. Falsification tests

As is standard in the literature, we offer several tests to assess 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; and (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.

11. 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, because 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.

B. 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 percent of the population.[12] 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.[13] That is, a total of 2,081 schools hosting 64,582 refugee and 743,301 Turkish students in the 2017–2018 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 the short-term and longer-term impacts of the program.

1. Absenteeism

The data contain information on the days of absence from school at the 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 attendance. For example, absenteeism is used by many researchers and policy institutions to quantify learning losses that emerged during the COVID-19 pandemic.[14] Second, it is a proxy for noncognitive skills. Several papers in the literature document that absenteeism is negatively correlated with the "Big Five" personality traits that constitute the core of noncognitive skills (Lounsbury et al. 2004). Finally, school absenteeism is strongly related to various antisocial and/or risky behaviors that emerge later in life.[15]

2. Student achievement

To measure student achievement, we use the student grades that are contained in the administrative records. Specifically, we focus on grades in two core subjects: Turkish language and math. Note that grades are reported on a three-category scale (that is, 1–3) during the first three years of primary education and a zero to 100 grading scale is used for Years 4–12. Moreover, the data include information on the country of origin of the student, allowing us to differentiate between native Turkish students and refugee students. In our analysis of grades, we focus on grades from Year 4 onwards. The

12. The province-level refugee numbers are provided by the Turkish Ministry of Interior, Presidency of Migration Management.

13. In Türkiye, primary school covers Years 1–4, middle school covers Years 5–8, and high school covers Years 9–12.

14. See, for example, IMF (2021).

15. See Gubbels, van der Put, and Assink (2019) for a comprehensive meta-analytic review of the related literature.

Table 1
Summary Statistics—Student Characteristics

Variable Name	Analysis Range (12–102)		Discontinuity Sample ($w = 6$)	
	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)
P(Male = 1)	0.48	0.51	0.50	0.51
P(Primary school = 1)	0.70	0.39	0.68	0.39
P(Gaziantep = 1)	0.65	0.59	0.61	0.54
P(Preschool attendance = 1)	0.09	0.54	0.09	0.58
P(Origin is Aleppo = 1)	0.66	0	0.61	0
P(Absenteeism ≥ 10 days) (pre-treatment)	0.39	0.09	0.41	0.09
P(Absenteeism ≥ 20 days) (pre-treatment)	0.17	0.05	0.18	0.04
Observations	6,563	269,940	2,630	113,284

Notes: The first two columns provide student-level summary statistics for schools with a 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 three mass points around the cutoffs, which we refer to as the discontinuity sample. The grading scale for the Turkish language and math assessments is zero to 100. Standard deviations are reported in brackets.

observed grades in the data set correspond to end-of-semester averages obtained from a variety of assessment methods, such as written exams, oral exams, quizzes, projects, homework, etc. To obtain a basis of comparison for our analysis, we standardize the grades of Syrian students by school and grade/year around the mean grades of Turkish students in the same school and grade/year; see Section VI for more details.

3. 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. This sample includes

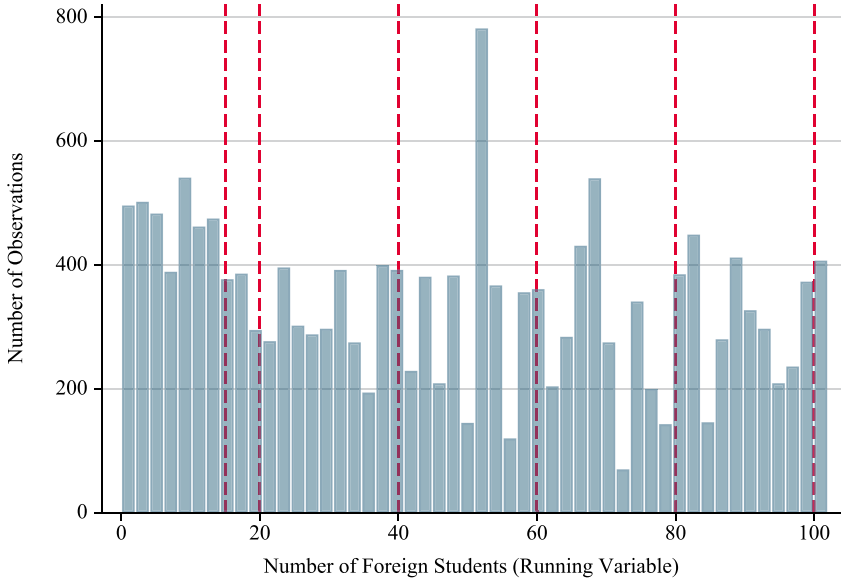

Figure 1

Distribution of Refugee Students

Notes: This figure plots the distribution of refugee students in our analysis range by the number of refugee students per school, which is our running variable. Vertical dashed lines indicate the cutoffs. The *y*-axis indicates the number of refugee students in the catchment area in each bin.

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 who attend schools in the neighborhood of the cutoffs that we consider in the RD analysis; they fall within the maximum window around the cutoffs (three 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 percent are in primary schools (on average in Grade/Year 3), 9 percent attended preschool, 61 percent originate from Aleppo, and 61 percent 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 in terms of 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 2017–2018 academic year, before the training took place. This amounts to them being absent about 10 percent of the school days, thereby incurring a considerable loss in learning. In contrast, Turkish children were 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 in both Turkish language and math subjects.

Online Appendix Table A2 offers some summary statistics of the schools in the two samples. In the discontinuity sample, 51.6 percent of the schools are located in Gaziantep, and they are almost equally split between primary and secondary schools. The average school size is approximately 615 students, with about 30 of those being refugees, on average. The average share of refugee students in the sample is 7.8 percent, with primary schools having a slightly higher share (9.3 percent) than secondary schools (5.9 percent). Schools are larger on average in the larger sample, but the average share of refugees is very similar, at about 8 percent.

Figure 1 shows a histogram of the number of refugee students in the catchment area plotted against the number of refugee students per school. The figure does not indicate any systematic sorting above the cutoffs, an issue that we will test formally in the next section. Overall, manipulation of the running variable does not seem to be a plausible feature in our setting.

V. 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.

A. Baseline RD Results

Figure 2 provides a first impression of the impact of training on refugee students' absenteeism. The figure displays 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 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 for 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).

In Table 2, 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, Titiunik, and Vazquez-Bare 2017). For example, considering the first cutoff at 15—there are 52 observations just below the cutoff and 68 observations at the

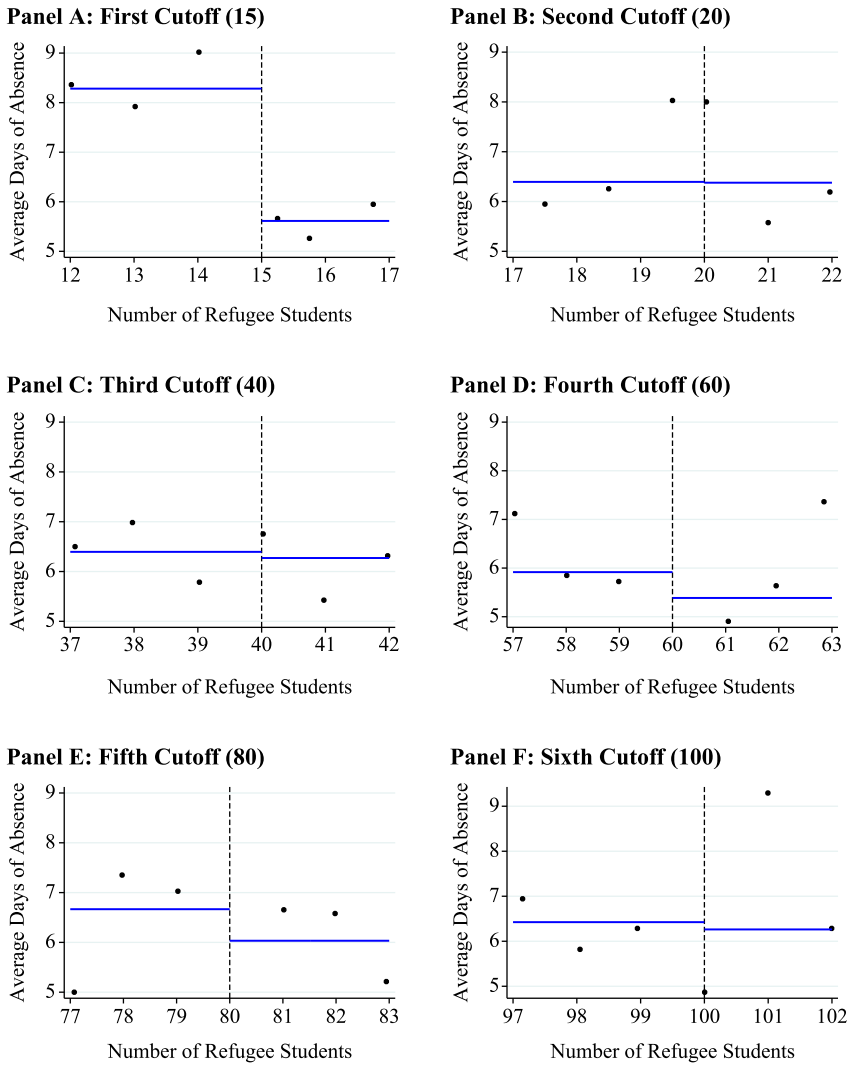

Figure 2

Baseline Visual Evidence—Local Randomization

Notes: This figure provides visual evidence for our baseline local randomization RD analysis for three mass points around each of the cutoffs. The outcome variable is the average days of absence per refugee student in the semester right after the training program.

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.

Figure 2 indicates 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

Table 2
Baseline RD Estimates—Local Randomization

	Estimation & Inference			Levels		Binomial Test		
Cutoff	Difference in Means	Fisherian p-Value	Neyman p-Value	Before Cutoff	After Cutoff	Obs. Below c	Obs. Above c	Bin. Test p-Value
Panel A: 15 Cutoff								
15 (w:2)	−3.357**	0.002	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
Panel B: 20 Cutoff								
20 (w:2)	−0.029	0.960	0.986	8.029	8.000	34	40	0.561
20 (w:4)	−0.203	0.866	0.810	6.758	6.556	120	99	0.176
20 (w:6)	−0.016	0.982	0.979	6.394	6.378	218	193	0.236
Panel C: 40 Cutoff								
40 (w:2)	0.969	0.348	0.277	5.786	6.755	70	147	0.000
40 (w:4)	−0.078	0.924	0.910	6.338	6.261	130	234	0.000
40 (w:6)	−0.124	0.834	0.835	6.395	6.271	200	291	0.000
Panel D: 60 Cutoff								
60 (w:2)	−0.818	0.342	0.320	5.724	4.905	105	74	0.025
60 (w:4)	−0.580	0.404	0.333	5.797	5.217	251	129	0.000
60 (w:6)	−0.531	0.408	0.363	5.917	5.386	276	140	0.000
Panel E: 80 Cutoff								
80 (w:2)	−0.373	0.822	0.798	7.028	6.655	36	55	0.059
80 (w:4)	−0.639	0.540	0.512	7.258	6.619	124	105	0.234
80 (w:6)	−0.633	0.396	0.391	6.667	6.033	168	180	0.555
Panel F: 100 Cutoff								
100 (w:2)	−1.417	0.262	0.257	6.286	4.868	91	38	0.000
100 (w:4)	0.194	0.856	0.868	6.042	6.236	191	55	0.000
100 (w:6)	−0.162	0.834	0.820	6.425	6.263	332	118	0.000
Panel G: Multiple Cutoff								
MC (w:2)	−0.442	0.340	0.348	6.631	6.190	388	422	0.246
MC (w:4)	−0.675**	0.032	0.032	6.667	5.992	1,007	793	0.000
MC (w:6)	−0.823***	0.004	0.002	6.822	5.999	1,439	1,191	0.000

Notes: The outcome variable is the days of absence per refugee student in the semester right after the training program. Panels A–F report the estimates around each of the cutoffs. Panel G reports the multiple-cutoff RD estimates (MC) obtained by pooling all 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. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 3
Regression Discontinuity 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**	0.032	0.022	0.365	0.574
Male (*w*:4)	−0.068	0.240	0.190	0.466	0.398
Male (*w*:6)	−0.059	0.198	0.149	0.490	0.431
Male (*w*:2)†	0.113	0.162	0.123	0.461	0.574
Male (*w*:4)†	−0.101	0.182	0.159	0.499	0.398
Male (*w*:6)†	−0.080	0.167	0.121	0.511	0.431
Age (*w*:2)	−1.115	0.155	0.112	11.865	10.750
Age (*w*:4)	−0.613	0.361	0.341	11.105	10.491
Age (*w*:6)	−0.182	0.380	0.353	10.475	10.294
# of siblings (*w*:2)	0.069	0.566	0.570	3.018	3.087
# of siblings (*w*:4)	−0.091	0.497	0.503	3.024	2.933
# of siblings (*w*:6)	−0.077	0.766	0.702	3.315	3.238
Preschool attendance (*w*:2)	0.006	1.000	0.878	0.038	0.044
Preschool attendance (*w*:4)	0.018	0.638	0.486	0.058	0.076
Preschool attendance (*w*:6)	0.018	0.496	0.439	0.079	0.097
Origin is Aleppo (*w*:2)	0.132	0.244	0.148	0.500	0.632
Origin is Aleppo (*w*:4)	0.022	0.804	0.674	0.586	0.608
Origin is Aleppo (*w*:6)	0.021	0.666	0.603	0.563	0.584
Grade/year ≥ mean (*w*:2)	−0.082	0.163	0.166	0.816	0.734
Grade/year ≥ mean (*w*:4)	−0.026	0.419	0.402	0.754	0.728
Grade/year ≥ mean (*w*:6)	−0.012	0.597	0.566	0.711	0.699
School size (*w*:2)	242.686	0.132	0.105	291.100	533.786
School size (*w*:4)	122.972	0.289	0.302	449.148	572.120
School size (*w*:6)	148.857	0.131	0.118	450.380	599.237
School type (*w*:2)	−0.041	0.361	0.366	0.453	0.412
School type (*w*:4)	0.028	0.509	0.528	0.465	0.493
School type (*w*:6)	0.011	0.611	0.634	0.471	0.482
# of teachers per school (*w*:2)	0.678	0.201	0.192	26.661	27.339
# of teachers per school (*w*:4)	0.826	0.165	0.171	25.958	26.784
# of teachers per school (*w*:6)	0.224	0.442	0.414	26.544	26.768

Notes: Estimates denoted by † refer to a sample in which we exclude a small all-female religious secondary school that has 14 refugee students. School type is a dummy variable indicating whether it is a vocational/religious school or a general school. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

such drop is visually obvious. Turning to the point estimates, Table 2 shows our baseline estimates of the treatment effect of training on days of absence. In Panel A, we report the RD estimate around the first cutoff, capturing the extensive margin of the training program. Then, in Panels B–F, we present cutoff-specific RD estimates

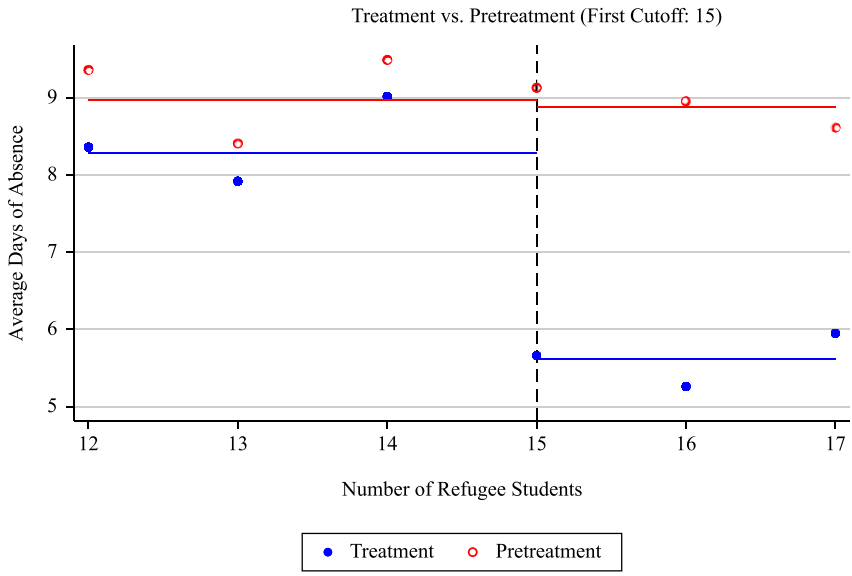

Figure 3

Visual Evidence—Pretreatment Semester

Notes: This figure provides visual evidence on refugee students' average days of absence in the pre-treatment period (fall 2017) as a falsification test. The results for the pretreatment and treatment semesters are indicated by gray (red online) and black (blue online) colors, respectively.

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 two and decreases (in absolute value) to −2.8 days for the window of four, and −2.7 days for the window of six mass points. Note that the binomial test indicates no evidence of sorting of the running variable for $w = 2$ and $w = 4$, while the test rejects randomization for $w = 6$. However, at $w = 6$, the number of observations below the cutoff is smaller than above against what we would expect if sorting into schools with more trained teachers was taking place.[16] In the rest of the cutoffs reported in Panels B–F, we do not find a

16. The right of the cutoff would be more desirable from the viewpoint of a refugee student than the left, as the environment would be more favorable when there are more trained teachers. Therefore, if Syrian children were sorting into schools on the basis of the treatment, we would expect to consistently observe a higher number of Syrian students on the right-hand side of the cutoffs than the left-hand side. Such a systematic imbalance does not exist in our analysis. Moreover, as Cattaneo, Idrobo, and Titiunik (2023) explain, having such an imbalance alone does not imply that the local randomization assumptions are violated, as imbalances are sometimes unavoidable around narrow windows. They suggest that when there are such imbalances, balancing tests on observed characteristics should be performed to establish whether treated and control students are similar around the cutoff. Our covariate balance analysis strongly establishes that there is no systematic imbalance in the covariates around any of our cutoffs (see Table 3 and Online Appendix Tables A5–A9). Note that the number of schools is also balanced around all of our cutoffs and for all window sizes (Online Appendix Table A3).

Table 4
Regression Discontinuity Effects for the Pretreatment Semester

	Estimation & Inference			Levels	
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Before Cutoff	After Cutoff
15 (*w*:2)	0.304	0.874	0.853	9.115	9.419
15 (*w*:4)	0.825	0.362	0.370	8.000	8.825
15 (*w*:6)	−0.013	1.000	0.986	8.898	8.885

Notes: The outcome variable is the days of absence per refugee student in the pretreatment semester. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

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 VII.A.

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 (three 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 percent. 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 percent (initial gap of 4.4 days).[17]

Note that our setting can be viewed as a multiple-cutoff RD design. In contrast to the canonical RD setup, where all units face the same cutoff value, the teacher training program exposes different units to different cutoffs. Panel G in Table 2 reports the multiple-cutoff RD estimates (MC) obtained by following the estimation procedure described by Cattaneo, Idrobo, and Titiunik (2023). We find a statistically significant average impact for $w = 4$ and $w = 6$ that ranges between −0.68 and −0.82 days per trained teacher. However, this impact is much smaller than our estimate for the first cutoff, which suggests that the statistically significant impact in the multiple-cutoff RD analysis is driven by the strong impact from the first cutoff.

To summarize, our results suggest that the treatment effect is concentrated around the first cutoff in which comparison is between schools with one trained teacher and

17. As a robustness check, we extended our window size to $w = 10$ and repeated our analysis. Online Appendix Table A4 and Figure A1 present the results. The qualitative and quantitative nature of our results are unchanged. Note that the first (15) and second (20) cutoffs are closer to each other than the other cutoffs. This means that if we use a window size of 12 (6 on either side of the cutoff), then the right side of the first cutoff will overlap with the second treatment and the left side of the second cutoff will start capturing the untreated category. This limitation does not apply to other cutoffs, but it also means that the largest feasible window size that can be uniformly applied to all cutoffs is ten.

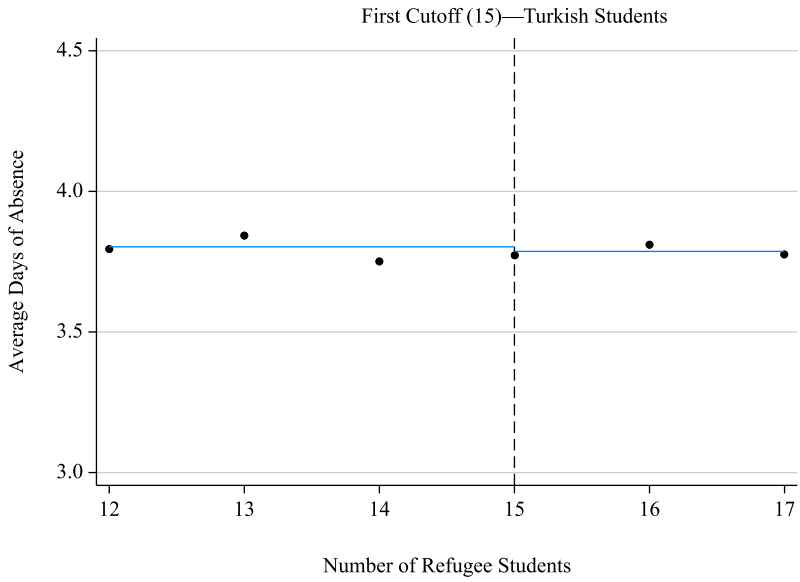

Figure 4

Visual Evidence—Turkish Students

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

control schools with no trained teachers—the extensive margin of the treatment. Therefore, to conserve space from here onward, we present and discuss results only for the first cutoff.

B. Validation Checks

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

1. 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 3 presents RD effects on predetermined covariates for the cutoff at 15. The individual-level predetermined covariates for the refugee students in our sample include age, gender, number of siblings, a dummy variable indicating whether the student attended preschool or not, a dummy variable indicating whether the district of origin is Aleppo or not,[18] and a dummy variable taking the value one if the grade/year of the refugee student is greater than the mean grade/year in the corresponding window and zero

18. Note that the region of origin for the majority (around 66 percent) of the refugee students is Aleppo; hence, whether a refugee student is from Aleppo or not might be a proxy for a variety of factors related to their (and their parents') access to refugee social networks.

Table 5
Regression Discontinuity Effects for Turkish Students

	Estimation & Inference			Levels	
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Before Cutoff	After Cutoff
15 (*w*:2)	0.024	0.706	0.740	3.751	3.775
15 (*w*:4)	−0.017	0.660	0.710	3.811	3.794
15 (*w*:6)	−0.022	0.506	0.529	3.810	3.787

Notes: The outcome variable is the days of absence per native student in the semester right after the training program. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

otherwise. The school-level predetermined covariates include school size, number of teachers, and a dummy variable taking the value one if the school type is vocational/religious and zero otherwise (that is, general school). Of all the tests reported in the table, there is only one instance in which there is a statistically significant difference between the 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 (the type of 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 marked with † in Table 3). Note that our baseline estimates reported in Table 2 are also not sensitive to removing that specific school from our sample.[19] We repeat the same exercise for the rest of the cutoffs; Online Appendix Tables A5–A9 present the results. This analysis indicates that there is also no evidence of any systematic imbalance in the predetermined covariates in those cutoffs. Overall, we thus conclude that the covariates are balanced around all the cutoffs.

2. 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 2017–2018 before the training had taken place. Reassuringly, the results illustrated in Figure 3 and reported in Table 4 show convincingly that no treatment effect is observed before the intervention.

3. 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

19. 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 6
Longer-Term RD Effects

	Estimation & Inference			Levels	
Cutoff	Difference in Means	Fisherian p-Value	Neyman p-Value	Before Cutoff	After Cutoff
15 (w:2)	−1.919	0.142	0.129	9.925	8.006
15 (w:4)	−1.762**	0.020	0.018	9.570	7.808
15 (w:6)	−1.521**	0.014	0.014	9.180	7.659

Notes: The outcome variable is the refugee students' days of absence in the academic year following the treatment semester. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

needs of refugee students, so we do not expect to find any effect on the outcomes of Turkish students, and thus, this analysis constitutes a plausible placebo test. Figure 4 and estimates reported in Table 5 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.

C. 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. Table 6 and Figure 5 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 statistically 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 $w = 6$ in the control group is 9.2 days, our estimate of −1.5 implies a 16.3 percent reduction. Compared to the 32.5 percent 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.

D. 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 7 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 presents RD estimates separately for primary schools and upper-level schools (middle 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

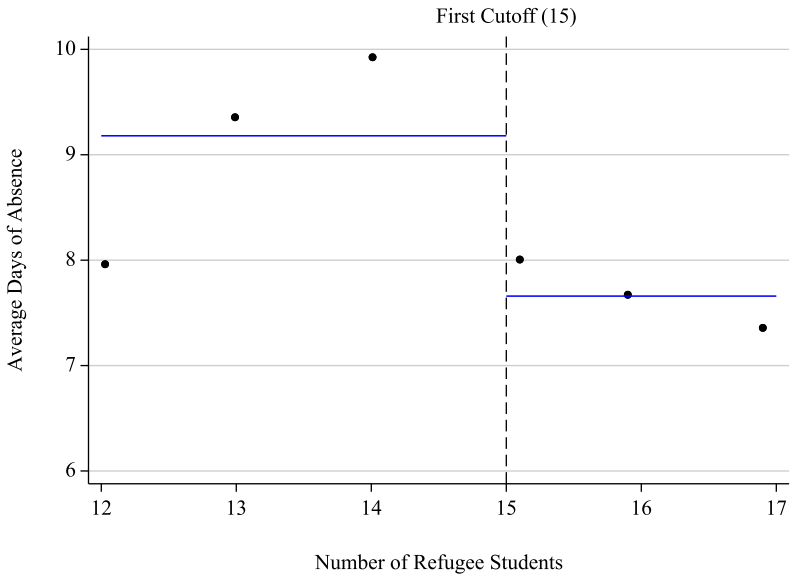

Figure 5

Visual Evidence—Longer Term Effects

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

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.

E. Chronic Absenteeism and Dropout

Our main outcome variable 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 percent of the available school days in a semester. It is often used as an early warning signal for academic risk and school dropout (Bruner, Discher, and Chang 2011). The number of available school days in a typical semester is 90 days in Türkiye. Panels A and B in Table 9 present results for two outcomes: missing at least ten days and 20 days (that is, chronic absenteeism in varying degrees), respectively, in the semester following the training program. Specifically, in Panel A (Panel B), the dependent variable is a dummy variable taking the value one if the student misses at least ten days (20 days) during the spring 2018 semester and zero 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 primary estimates presented in Table 2—the training program

Table 7
Estimates by Gender

	Estimation & Inference			Levels	
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Before Cutoff	After Cutoff
Males					
15 (*w*:2)	−3.005	0.174	0.218	8.211	5.205
15 (*w*:4)	−2.010*	0.088	0.075	7.539	5.529
15 (*w*:6)	−2.814**	0.018	0.010	8.167	5.983
Females					
15 (*w*:2)	−3.209**	0.030	0.038	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: The outcome variable is the days of absence per refugee student in the semester right after the training program. The upper and lower panels report the results for males and females, respectively. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

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

Table 8
Estimates by School Level

	Estimation & Inference			Levels	
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Before Cutoff	After Cutoff
Primary Education (Years 1–4)					
15 (*w*:2)	−4.700**	0.036	0.029	9.700	5.000
15 (*w*:4)	−1.836*	0.050	0.055	7.036	5.200
15 (*w*:6)	−2.454***	0.002	0.001	8.021	5.567
Secondary Education (Years 5–12)					
15 (*w*:2)	−2.732	0.112	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: The outcome variable is the days of absence per refugee student in the semester right after the training program. The upper panel reports the results for refugee students in primary education, while the lower panel reports the results for those in secondary education. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 9
Regression Discontinuity Effects for Alternative Outcomes

	Estimation & Inference			Levels	
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Before Cutoff	After Cutoff
Panel A: Absenteeism ≥10 Days in the Post-Treatment Semester					
15 (*w*:2)	−0.164*	0.070	0.053	0.385	0.221
15 (*w*:4)	−0.131***	0.012	0.004	0.330	0.199
15 (*w*:6)	−0.091**	0.018	0.011	0.306	0.216
Panel B: Absenteeism ≥20 Days in the Post-Treatment Semester					
15 (*w*:2)	−0.101**	0.040	0.033	0.115	0.015
15 (*w*:4)	−0.076***	0.010	0.002	0.099	0.023
15 (*w*:6)	−0.080***	0.000	0.000	0.102	0.022
Panel C: Absenteeism ≥30 Days in the Post-Treatment Semester					
15 (*w*:2)	−0.024	0.764	0.439	0.038	0.015
15 (*w*:4)	−0.026*	0.176	0.067	0.031	0.006
15 (*w*:6)	−0.028***	0.012	0.006	0.032	0.004
Panel D: Absenteeism ≥30 Days in the Academic Year					
15 (*w*:2)	−0.070	0.380	0.278	0.173	0.103
15 (*w*:4)	−0.037	0.368	0.278	0.136	0.099
15 (*w*:6)	−0.052**	0.078	0.049	0.149	0.097

Notes: The outcome is a dummy variable indicating whether the refugee student is (chronically) absent from school for more than ten days in the post-treatment semester, 20 days in the post-treatment semester, 30 days in the post-treatment semester, and 30 days during the 2017–2018 academic year in Panels A, B, C, and D, respectively. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

(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 C and D 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 (2017–2018 academic year), respectively. The estimates around the first cutoff point to a 2.6–5.2 percentage point 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 drop out.

F. Continuity-Based RD Estimates

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

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

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^* 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 $\mathbf{X}$ 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.

Online Appendix Figure A2 and Table A10 presents baseline results using this approach, while Online Appendix Figure A3 and Table A11 presents long-term effects, using two alternative methods of modeling $f(r_s)$ (linear, quadratic) and two alternative ways of choosing optimal bandwidth. In line with the 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. Note that the longer-term estimates tend to lack statistical significance in the continuity-based RD analysis.

VI. Effects on Academic Performance

The results presented in Section V suggest that the teacher training program meaningfully improved 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 contains rich information spanning several academic years, two provinces, many schools, and different grade/year levels—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 they 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 RD 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 and analytical abilities. To improve comparability, we standardize the subject-specific grades for Syrian students in each school and grade/year level relative to the

Table 10
Probability of Exam Participation

	Turkish			Math		
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value
Panel A: Short Term						
15 (*w*:2)	0.121***	0.007	0.006	0.109**	0.047	0.050
15 (*w*:4)	0.124**	0.028	0.034	0.115**	0.033	0.033
15 (*w*:6)	0.131**	0.018	0.016	0.108*	0.061	0.071
Panel B: Longer Term						
15 (*w*:2)	0.079**	0.047	0.044	0.091*	0.088	0.089
15 (*w*:4)	0.061*	0.088	0.093	0.093*	0.087	0.091
15 (*w*:6)	0.058	0.107	0.121	0.079	0.108	0.107

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

grades of Turkish students in the same school and grade/year. In doing so, first we calculate the mean grade for Turkish students, then subtract this mean from the grade of each Syrian student, and, finally, standardization is performed over this relative measure. This means that the coefficient estimates are presented in terms of standard deviations.

Parallel to our analysis in Section V, we mainly focus on Syrian students around the first cutoff—that is, schools just below and above the 15 cutoff—as it is the only margin for which the program affects school attachment of refugee students.[20] We use two outcome variables measuring different aspects of achievement. First, we construct a dummy variable indicating whether a Syrian student has a missing grade, which measures "exam 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. Second, we use the standardized grades of Syrian students for Turkish language and math subjects relative to the grades of Turkish students as described above.

Table 10 presents short-term and longer-term estimates for the outcome variable that measures the probability of exam participation for refugee students. The 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. The results suggest that, in the short- term, the program reduced the probability by 12.1–13.1 percentage points for the Turkish language subject and 10.8–11.5 percentage points for the math subject. In other words, the teacher training program improved the exam participation of refugee students by 11–13 percentage points. These effects do not differ by gender. The results also indicate that the program effects tend to dissipate over time.

20. Note that we present visual evidence for all cutoffs in Online Appendix Figures A4 and A5.

Table 11
Standardized Grades

	Turkish			Math		
Cutoff	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value	Difference in Means	Fisherian *p*-Value	Neyman *p*-Value
All Sample						
Panel A: Short Term						
15 (*w*:2)	0.257**	0.036	0.034	0.212**	0.043	0.041
15 (*w*:4)	0.264***	0.000	0.000	0.198**	0.049	0.049
15 (*w*:6)	0.267***	0.000	0.000	0.160*	0.060	0.051
Panel B: Longer Term						
15 (*w*:2)	0.211**	0.039	0.038	0.171*	0.084	0.088
15 (*w*:4)	0.209**	0.026	0.025	0.096	0.114	0.109
15 (*w*:6)	0.199**	0.018	0.016	0.122*	0.098	0.099
Conditioning on Exam Participation in the Pretreatment Semester						
Panel C: Short Term						
15 (*w*:2)	0.286**	0.024	0.031	0.221**	0.047	0.046
15 (*w*:4)	0.278***	0.009	0.007	0.187*	0.059	0.064
15 (*w*:6)	0.271***	0.000	0.000	0.191**	0.041	0.043
Panel D: Longer Term						
15 (*w*:2)	0.246**	0.031	0.033	0.178**	0.048	0.050
15 (*w*:4)	0.256**	0.026	0.022	0.170*	0.066	0.069
15 (*w*:6)	0.231**	0.011	0.013	0.169*	0.062	0.058

Notes: Estimates are reported in terms of standard deviations. Standardization of the refugee students' grades is performed relative to the mean scores of Turkish students in the same school and grade/year. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 11 presents our estimates for the impact of the teacher training program on grades, both in the short-term and longer-term. Panel A shows the estimates for Turkish and math grades in the short-term. The results indicate that the Turkish language grades improved by 0.257 standard deviation, while the improvement in the math grades is 0.212 standard deviation. Panel B shows that the longer-term effects are somewhat smaller and less statistically significant, especially in math. As we show in Table 10, the teacher training program improved exam participation among refugee students, which suggests that the student composition might have changed after the program. To correct for these potential compositional effects, we present another set of estimates by conditioning on exam participation—that is, by restricting our sample to the students who participated in exams in the pretreatment semester. Panel C presents the short-term estimates for Turkish and math grades conditional on exam participation in the previous semester. The Turkish language grades improved by 0.286

standard deviations, and the math grades improved by 0.221 standard deviations. Panel D suggests that, similarly to the estimates for the entire sample, the impact of the program tends to diminish over time.

These results suggest that the teacher training program improved the academic performance—measured in terms of language and cognitive skills—of refugee students.[21] These improvements can be interpreted as a reduction in the ethnic gap/inequality in academic achievement. We find no gender differences in these effects. Moreover, as a placebo test, we find that there is no impact on Turkish language and math grades of the refugee students in the pretreatment semester (Online Appendix Table A12 and Figure A6).[22]

VII. Exploring the Mechanisms

A. 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 beyond the last cutoff (100), and (iii) only up to six teachers are trained out of a possible 30 teachers per school on average.[23]

The design of 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. Back-of-the-envelope calculations suggest 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 as the number of trained teachers per school surpasses six. Therefore, the program "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.

One possible mechanism through which the teacher training program may affect 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 of the

21. It should be noted that, similar to the absenteeism outcomes, the program does not have any impact on exam participation and the Turkish language and math grades of the Turkish students.

22. As an alternative to the local randomization RD approach, we also estimate the following ANCOVA specification: $y_{i,s,2} = \alpha + \gamma y_{i,s,1} + \beta T_s + \boldsymbol{\theta}' \mathbf{X}_{i,s} + f_s + \epsilon_{i,s,2}$, where i and s index students and schools, respectively; $y_{i,s,2}$ is the post-treatment outcome (i.e., the outcome at the end of spring 2017–2018, the semester immediately following the training program); $y_{i,s,1}$ is the outcome in the pre-treatment semester (that is, fall 2017–2018); T_s is the school-level binary treatment variable taking the value 1 for schools with 15 or more refugee students and 0 otherwise; $\mathbf{X}_{i,s}$ is a vector of student characteristics (gender, age, grade/year, number of siblings, region of origin); f_s captures school fixed effects; and $\epsilon_{i,s,2}$ is an error term. Standard errors are clustered at the school level. The pre-treatment outcome ($y_{i,s,1}$) acts as an individual-level fixed effect as it controls for the level of the outcome realized in the pre-treatment semester. The main parameter of interest is β. The qualitative and quantitative nature of the estimates obtained from this ANCOVA model is very similar to that of our baseline local randomization RD estimates.

23. School-level summary statistics given in Online Appendix Table A2 suggest that the average school size in the range of intervention is approximately 750. The average class size is typically 25 in Türkiye. This means that the average number of teachers per school is roughly 30.

Table 12
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
SD	3.250
Observations	612

Notes: The table reports the percentiles of the distribution of the number of refugee students per classroom around the 15 cutoff for *w*:6.

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 implicit cascade-training strategy employed by MoNE in implementing the program supports this argument. Table 12 provides evidence consistent with 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 three. Figure 6 shows the histogram of this distribution, indicating that refugee students are spread across many classrooms rather than being 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 plausible that the training program changed the overall school environment through possible spillovers to other teachers and school staff—and potentially to native students as well. The trained teachers acted as mentors to raise the school-wide awareness about the needs of refugee students. The program ingredients detailed in Section III.B 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 the treated schools. In particular, the pre-treatment cumulative distribution function first-order stochastically dominates that of the post-treatment—a Kolmogorov–Smirnov test rejects equality of the two cumulative distribution functions ($p < 0.001$). 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.

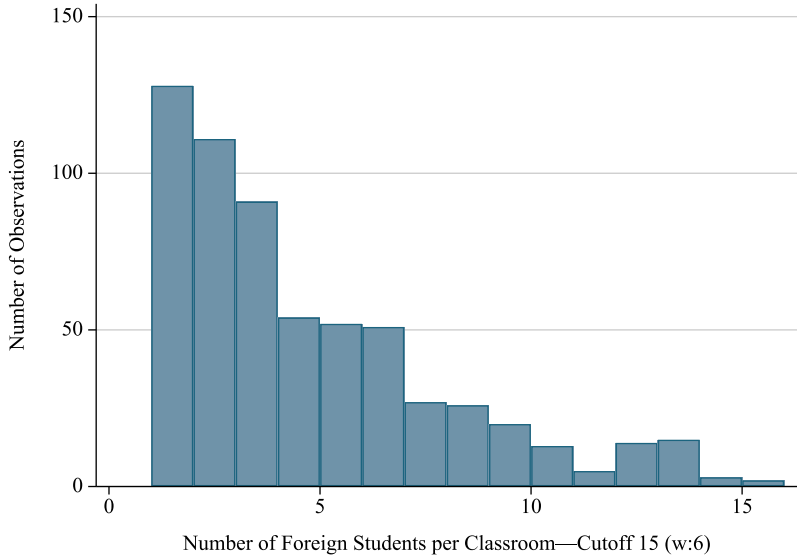

Figure 6

Refugee Students per Classroom

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

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.

B. Grades

Next, we investigate the potential factors contributing to the improvements in grades we found in Section VI. One possible channel that we consider is whether the decline in absenteeism resulting from the teacher training program played a role in improving the grades of refugee students, that is, whether a stronger school attachment led to improved academic performance.

With the data at hand, it is not easy to establish a causal relationship going from student absenteeism to grades. It is possible that the training program, by equipping trained teachers with the skills on how to effectively teach the Turkish language to foreigners, may have led to improvements in the refugee children's Turkish skills, which in turn could have helped improve their attendance and performance in math. Therefore, it is difficult to determine the exact causal chain linking the program to these student outcomes.

But examining the joint distribution of grades and days of absence for each student in our sample allows us to explore the correlation between gains in school presence and gains in grades. Figure 8 displays scatter plots of these results for Syrian students. In this analysis, we focus on the Syrian students in treated schools just above the first cutoff—that is, schools with 15, 16, and 17 refugee students. The sample size is 269. We

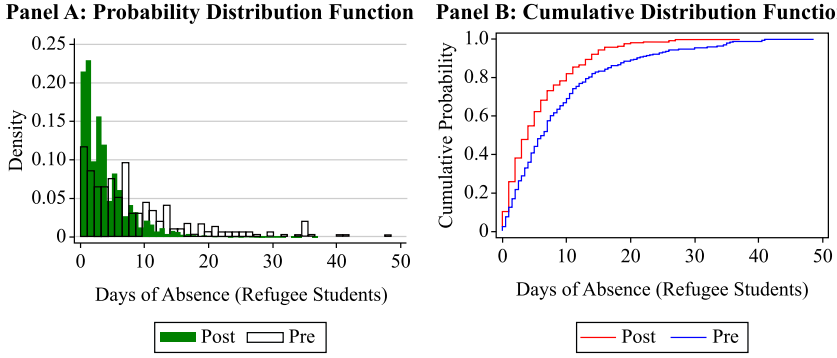

Figure 7

Distribution of Days of Absence for Students Above the First Cutoff

Notes: This figure displays the probability distribution function and cumulative distribution function of days of absence for schools with a number of foreign students between 15 and 17, which constitute the treated group for the first cutoff. The visuals plot the distribution of days of absence for the same students before and after the treatment.

calculate the decline in absenteeism and increase in grades from fall (pre-treatment) to spring (treatment) semester of the 2017–2018 academic year. The correlations reported in Figure 8 are statistically significant and approximately 0.2 for both grades.[24] The fact that the gains in grades for Syrian students are positively correlated with gains in school attendance is consistent with the idea that increased school attachment led to improved learning outcomes. However, as highlighted earlier, it is also possible that the training program had a direct impact on both outcomes (absenteeism and learning).

It is important to note that if the impacts on grades were driven by changes in teachers' grading standards (for example, teachers becoming more lenient toward refugee students), then we would not expect to see a meaningful positive correlation between changes in school attachment and changes in academic performance at the student level. Instead, gains in grades would be more evenly distributed across the gains in absenteeism distribution. Moreover, since only a small number of teachers were trained, and the grade effect is driven by the first-cutoff schools with only one trained teacher, it is unlikely that the estimated improvement in grades was caused by a change in school-wide grading behavior.

C. Reporting Bias

As grades are reported by teachers, one might be concerned that teachers' reporting might respond to the program. While we cannot entirely rule out some level of

24. It should be noted that there are students who did not participate in the exams in the Fall, but took the exams in the spring, and vice versa. For these students, participation in exams (that is, having nonzero scores) 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.

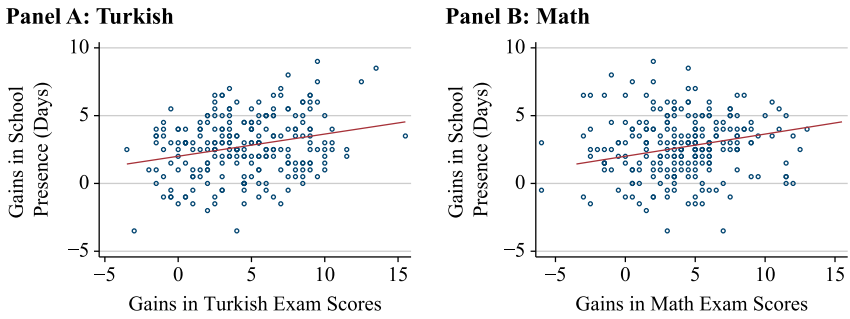

Figure 8

Gains in Absenteeism Versus Gains in Grades—Treated Students

Notes: The figures show the student-level gains in school presence and gains in (Turkish language and math) grades for Syrian students in schools just above the first cutoff ($N = 269$).

subjectivity or reporting bias in the grade results, we discuss several reasons why we do not think that reporting bias is likely to be the sole driver of the observed improvement in refugee students' outcomes.

First, it is important to note that the trained teachers have no performance targets and that the outcomes of students in treated schools are not monitored by MoNE. This absence of external incentives and pressures reduces the likelihood of inflated results for refugee students (Sandefur and Glassman 2015). In fact, the politically sensitive nature of the issue of the treatment of refugees in Türkiye suggests that there may be an opposite bias towards underrepresenting the positive outcomes of refugee students.

Second, the significant increase in exam participation of Syrian students in treated schools (Table 10) is a concrete outcome that cannot be influenced by reporting bias. Third, regarding the grades, we find that the program had a stronger impact on Turkish than math grades (Table 11). This asymmetry suggests that the results reflect actual differences in learning, rather than being driven by reporting bias, which would be expected to produce effects of similar magnitudes for both subjects.

Finally, as noted earlier, there is a positive correlation between absenteeism gains and grade gains. Given that it is virtually impossible to manipulate the absenteeism data because schools use a standardized system of monitoring and recording absences that minimizes any potential for manipulation of administrative records, this correlation provides additional evidence that the results on grades are also capturing actual learning outcomes.[25]

25. According to the MoNE Preschool and Primary Education Regulation (Official Gazette #29072) and the MoNE Secondary Education Regulation (Official Gazette #28758), teachers have the responsibility of accurately taking the official absenteeism records of their students. Meanwhile school administrators are responsible for entering these absenteeism records into the online database of the MoNE and regularly monitoring the database. This means that more than one party is involved in generating the absenteeism records. Moreover, if a student is absent from school, but their absence is not truthfully reported/recorded by the teachers and school administrators, then those teachers and administrators might be subject to disciplinary action (due to neglecting their official duties) if the student experiences a health issue or engages in criminal activity outside of school.

VIII. Evaluating Policy Effectiveness

The teacher training program was funded by the EU Facility for Refugees in Türkiye (FRIT) (see Section III.A 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 nonnegligibly 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. Note that these 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.

The literature investigating the impact of school absences on the short- and longer-term outcomes of students reports large estimates. For example, Liu, Lee, and Gershenson (2021) find that ten total absences across all subjects in the ninth grade/year reduce the probability of on-time graduation from high school and college enrollment by 2 percent. Goodman (2014) argues that four additional absences from math classes every year explain roughly one-fourth of the math achievement gap between poor and nonpoor students. Aucejo and Romano (2016) show that a reduction in student absences by ten days improves math and reading scores by 5.5 percent and 2.9 percent, respectively, and the improvement is more pronounced among low-performing students. Finally, Cattan et al. (2023) document that ten days of absence reduces academic performance in elementary school by 4.5 standard deviations and leads to a reduction in lifetime income measures by 1–2 percent. Given these large lifetime costs of school absenteeism, the teacher training program suggests a cost-effective way of improving the school attachment of disadvantaged students.

IX. 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 Türkiye 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 percent 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. One possible mechanism is 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 faded 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 discussion of cost-effectiveness suggests 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—for example, language and communication skills, and other complementary skills related to teachers' professional development, such as teaching minorities, basic counseling, relationship with parents, intraclass 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* 136:2147–94.

Alan, S., E. Duysak, E. Kubilay, and I. Mumcu. 2023. "Social Exclusion and Ethnic Segregation in Schools: The Role of Teacher's Ethnic Prejudice." *Review of Economics and Statistics* 105(5):1039–54.

Alesina, A., M. Carlana, E. La Ferrara, and P. Pinotti. 2018. "Revealing Stereotypes: Evidence from Immigrants in Schools." NBER Working Paper 25333. Cambridge, MA: NBER.

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* 19:343–69.

Assaad, R., T. Ginn, and M. Saleh. 2023. "Refugees and the Education of Host Populations: Evidence from the Syrian Inflow to Jordan." *Journal of Development Economics* 164:103131.
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* 36:885–921.
Borghans, L., B.H.H. Golsteyn, J.J. Heckman, and J.E. Humphries. 2016. "What Grades and Achievement Tests Measure." *PNAS* 113:13354–59.
Bossavie, L. 2020. "The Effect of Immigration on Natives' School Performance: Does Length of Stay in the Host Country Matter?" *Journal of Human Resources* 55:733–66.
Botelho, F., R.A. Madeira, and M.A. Rangel. 2015. "Racial Discrimination in Grading: Evidence from Brazil." *American Economic Journal: Applied Economics* 7:37–52.
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. Washington, DC: CEPR.
Bratsberg, B., O. Raaum, and K. Røed. 2012. "Educating Children of Immigrants: Closing the Gap in Norwegian Schools." *Nordic Economic Policy Review* 3:211–51.
Bressoux, P., F. Kramarz, and C. Prost. 2009. "Teachers' Training, Class Size and Students' Outcomes: Learning from Administrative Forecasting Mistakes." *Economic Journal* 119:540–61.
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* 90:1–29.
Cattan, S., D.A. Kamhofer, A. Karlsson, and T. Nilsson. 2023. "The Long-Term Effects of Student Absence: Evidence from Sweden." *Economic Journal* 133:888–903.
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* 3:1–24.
Cattaneo, M.D., N. Idrobo, and R. Titiunik. 2023. "A Practical Introduction to Regression Discontinuity Designs: Extensions." In *Cambridge Elements: Quantitative and Computational Methods for Social Science*, II, ed. R.M. Alvarez and N. Beck, Cambridge, UK: Cambridge University Press.
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* 36:643–81.
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* 55:926–62.
Dee, T.S. 2004. "Teachers, Race, and Student Achievement in a Randomized Experiment." *Review of Economics and Statistics* 86:195–210.
——. 2005. "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review* 95:158–65.
Dryden-Peterson, S. 2015. *The Educational Experiences of Refugee Children in Countries of First Asylum.* Washington, DC: Migration Policy Institute.
Dustmann, C., and A. Glitz. 2011. "Migration and Education." In *Handbook of the Economics of Education*, Volume 4, ed. Eric A. Hanushek, Stephen J. Machin, and Ludger Woessmann, 327–439. New York: 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* 38:143–82.
Figlio, D., P. Giuliano, R. Marchingiglio, U. Ozek, and P. Sapienza. 2024. "Diversity in Schools: Immigrants and the Educational Performance of US-Born Students." *Review of Economic Studies* 91(2):972–1006.
Figlio, D., and U. Ozek. 2019. "Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students." *Journal of Labor Economics* 37:1061–96.
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* 11:125–49.
Gershenson, S., A. Jacknowitz, and A. Brannegan. 2017. "Are Student Absences Worth the Worry in US Primary Schools?" *Education Finance and Policy* 12:137–65.
Goodman, J. 2014. "Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time." NBER Working Paper 20221. Cambridge, MA: NBER.
Gould, E.D., V. Lavy, and D.M. Paserman. 2009. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *Economic Journal* 119:1243–69.
Green, C., and J.M.V. Iversen. 2022. "Refugees and the Educational Attainment of Natives: Evidence from Norway." *Economics of Education Review* 88:102258.
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* 48:1637–67.
Hanna, R.N., and L.L. Linden. 2012. "Discrimination in Grading." *American Economic Journal: Economic Policy* 4:146–68.
Hanushek, E.A., and S.G. Rivkin. 2006. "Teacher Quality." In *Handbook of the Economics of Education*, Volume 2, ed. E.A. Hanushek and F. Welch, 1051–78. New York: Elsevier.
Harris, D.N., and T.R. Sass. 2011. "Teacher Training, Teacher Quality and Student Achievement." *Journal of Public Economics* 95:798–812.
Hunt, J. 2017. "The Impact of Immigration on the Educational Attainment of Natives." *Journal of Human Resources* 52:1060–118.
Imbens, G.W., and K. Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79:933–59.
Imbens, G.W., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142:615–35.
IMF. 2021. "World Economic Outlook: Managing Divergent Recoveries." Technical Report. Washington, DC: International Monetary Fund.
Jackson, C.K. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes." *Journal of Political Economy* 126:2072–107.
Jackson, C.K., S. Kiguel, S.C. Porter, and J.Q. Easton. 2024. "Who Benefits from Attending Effective High Schools?" *Journal of Labor Economics* 42(3):717–51.
Jackson, C.K., S.C. Porter, J.Q. Easton, A. Blanchard, and S. Kiguel. 2020. "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment." *American Economic Review: Insights* 2:491–508.
Jackson, C.K., J.E. Rockoff, and D.O. Staiger. 2014. "Teacher Effects and Teacher-Related Policies." *Annual Review of Economics* 6:801–25.
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* 39:50–79.

Krafft, C., M. Sieverding, N. Berri, C. Keo, and M. Sharpless. 2022. "Education Interrupted: Enrollment, Attainment, and Dropout of Syrian Refugees in Jordan." *Journal of Development Studies* 58:1874–92.

Liu, J., M. Lee, and S. Gershenson. 2021. "The Short- and Long-Run Impacts of Secondary School Absences." *Journal of Public Economics* 199:104441.

Liu, J., and S. Loeb. 2021. "Engaging Teachers: Measuring the Impact of Teachers on Student Attendance in Secondary School." *Journal of Human Resources* 56:343–79.

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* 33:457–66.

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* 11:128–54.

Ohinata, A., and J.C. van Ours. 2013. "How Immigrant Children Affect the Academic Achievement of Native Dutch Children." *Economic Journal* 123:F308–F331.

Popova, A., D.K. Evans, M.E. Breeding, and V. Arancibia. 2022. "Teacher Professional Development Around the World: The Gap Between Evidence and Practice." *World Bank Research Observer* 37:107–36.

Sandefur, J., and A. Glassman. 2015. "The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics." *Journal of Development Studies* 51:116–32.

Schnepf, S.V. 2007. "Immigrants' Educational Disadvantage: An Examination Across Ten Countries and Three Surveys." *Journal of Population Economics* 20:527–45.

Smith, J.P. 2006. "Immigrants and Their Schooling." In *Handbook of the Economics of Education*, Volume 1, ed. by E.A. Hanushek and F. Welch, 155–87. New York: Elsevier.

Tumen, S. 2018. "The Impact of Low-Skill Refugees on Youth Education." IZA Discussion Paper 11869. Bonn, Germany: IZA.

——. 2019. "Refugees and 'Native Flight' from Public to Private Schools." *Economics Letters* 181:154–59.

——. 2021. "The Effect of Refugees on Native Adolescents' Test Scores: Quasi-Experimental Evidence from PISA." *Journal of Development Economics* 150:102633.

UNHCR. 2017. "Left Behind: Refugee Education in Crisis." Technical Report. United Nations High Commissioner for Refugees.

——. 2019. "Global Trends: Forced Displacement 2019." Technical Report. United Nations High Commissioner for Refugees.

——. 2020. "Coming Together for Refugee Education." Technical Report. United Nations High Commissioner for Refugees.