Social Exclusion and Ethnic Segregation in Schools: The Role of Teacher's Ethnic Prejudice*

Sule Alan[1], Enes Duysak[2], Elif Kubilay[2], and Ipek Mumcu[3]

[1]European University Institute
[2]University of Essex
[3]University of Exeter

June 6, 2020

Abstract

Using uniquely detailed data on primary school children, we show that teachers who hold prejudicial attitudes towards an ethnic group create socially segregated classrooms. We identify this relationship by exploiting a natural experiment where newly arrived refugee children are randomly assigned to teachers. We elicit children's social networks to construct multiple measures of social exclusion and ethnic segregation in classrooms. We find that teachers' ethnic prejudice, measured by an Implicit Association Test, significantly lowers the prevalence of social ties between host and refugee children, increases homophily amongst host children, and puts refugee children at a higher risk of bullying victimization. Our results suggest that teachers' ethnic prejudice may be a significant barrier against building cohesive schools in ethnically diverse communities.

JEL Codes: I24, J15
Keywords: ethnic prejudice; integration; social exclusion; ethnic segregation

*We are grateful to IPA Peace and Recovery, J-PAL initiatives of Crime and Violence, for funding the data collection. We thank Mert Gumren, Ozge Seyrek, Melek Celik and Yusuf Agus for their invaluable assistance.

1 Introduction

Ethnic prejudice and stereotypes are the root causes of ethnic tensions and conflicts around the world (Arbatli et al. (2020)). These harmful beliefs easily breed in ethnically diverse environments and tend to take hold in difficult sociopolitical contexts. In a world with growing anti-immigrant sentiments reinforced by massive population movements, understanding how ethnic biases shape our social interactions with the members of out-groups is crucial to building empowered and cohesive communities. Public education is perhaps the most effective platform to engineer an integrated and cohesive society. Schools play a central role in cultivating tolerance toward ethnic diversity, shaping social preferences, and building social capital (Gradstein and Justman (2002)). However, ethnically diverse schools, especially in socioeconomically disadvantaged settings, are susceptible to ethnicity-based conflicts, acts of social exclusion, and inter-group violence.[1] Students belonging to the minority are generally at a higher risk of falling victim to these actions.

This paper examines how the school environment shapes young children's inter-ethnic relationships. Our focus is the role of teachers, in particular, teachers' implicit bias toward a minority group, in determining the nature of social ties formed among students. Teachers' implicit bias against a minority group may manifest itself as discriminatory behaviors, ranging from poor quality teacher-pupil interactions and neglect to outright unfair treatment of minority students, such as biased grading and disproportionate reprimanding (Lavy (2008), Burgess and Greaves (2013), Alesina et al. (2018), Alan et al. (2018)). The teacher's behaviors and attitudes toward ethnicity can implicitly guide students' socialization choices. Teachers who hold prejudicial attitudes toward an ethnic group may consciously or unconsciously prevent students from forming inter-ethnic friendships, creating segregated classrooms where minority students tend to be socially and spatially excluded. Research shows that socially excluded minority children are more likely to engage in self-defeating behavior, reinforcing further social exclusion and weaker inter-ethnic interactions (Buhs et al. (2006), Verkuyten and Thijs (2016), Glover et al. (2017)).

[1] Research on the effects of ethnic school composition on students' inter-ethnic relations produces mixed results. As a prominent approach, inter-group contact theory predicts considerable social and private benefits of integrated schools (see e.g. Allport (1954), Paluck et al. (2019)). From a broader perspective, diversity has been shown to contribute to greater creativity and production in a society (Ottaviano and Peri (2012)). On the other hand, the costs of ethnic diversity include deteriorating trust and social solidarity (Putnam (2007)), higher levels of conflict (O'Reilly III et al. (1998)), lower provision of public goods (Alesina and Ferrara (2005)) and less support for redistributive policies (Dahlberg et al. (2012)).

We leverage an ideal setting to identify the causal link between teachers' ethnic prejudice and students' inter-ethnic relationships. Our study site, Turkey, has received over 4 million refugees since the start of the Syrian Civil War in 2011. About 1 million of these refugees are school-aged children. As part of a multi-partnered EU initiative launched in 2016, most refugee children have been placed in Turkish state schools based on their registered addresses. The study site provides us with a setting where i) the ethnic composition in schools changed due to a massive refugee influx within a short period, ii) conditional on school, both host and refugee students were assigned to classrooms randomly, and iii) we are allowed to collect data from students and teachers in previously unavailable details. Our data set covers over 5000 primary school students and their teachers from 73 schools with a significant presence of Syrian refugee children. About 13% of the students in our sample are refugees, placed in schools between September 2017 and April 2018.

The literature on social networks offers powerful tools to study social interactions (Jackson (2008)). We utilize these tools to construct our outcome variables. Our primary outcomes are indicators of social exclusion of refugee students and ethnic segregation in classrooms. To construct these indicators, we elicit students' social networks within their classrooms. Specifically, students are asked to list the classmates i) they consider as their friends, ii) to whom they provide emotional and academic support, and iii) from whom they receive emotional and academic support, allowing the natural overlaps between these categories. Using the reported social ties, we construct student (node)-level measures of social exclusion and the degree of inter-ethnic interactions. We also construct classroom level measures (indices) of homophily.[2] In our case, the host (refugee) homophily in a classroom concerns the above-expected numbers of social ties amongst the host students (refugee students) and gives us a measure of ethnic segregation in the classroom.

We measure teachers' ethnic bias using an Implicit Association Test (IAT), developed by Greenwald et al. (1998). IAT aims to measure implicit attitudes toward a group using the difference in reaction times when individuals are asked to relate a group-specific concept to a negative or a positive attribute. Because it is hard to game, the test is considered to be free from social desirability bias inherent in surveys on socio-politically sensitive topics. In recent years, it has been used extensively in economics research in the context of gender and ethnic discrimination in schools (see e.g. Alesina et al. (2018), Carlana (2019)). Even

[2]Homophily, a term coined by Lazarsfeld and Merton (1954), refers to a tendency to interact exclusively with the members of own group.

though the test captures the degree of ethnic bias in teachers, it is still likely to be confounded with a variety of other teacher characteristics. To the extent that these characteristics are also predictive of the outcomes we are interested in, our estimated effects would be biased. The differential impacts of the measured implicit bias on hosts and refugee children help mitigate these concerns. Nevertheless, to control as many possible confounds as possible to isolate the effect of teacher's ethnic bias, we collect detailed data from teachers. These include previously unavailable characteristics, such as fluid cognitive ability, measured by Raven's Progressive Matrices (Raven et al. (2004)), and cognitive empathy, measured by Reading the Mind in the Eyes test (Baron-Cohen et al. (1997)). We also collect detailed information to construct measures of teaching styles, pedagogical practices, and motivation as well as standard demographic information and qualifications. Our IAT scores exhibit substantial variation across teachers, and teachers' cognitive ability emerges as the sole predictive characteristic. One standard deviation higher cognitive ability is associated with 0.29 standard deviation lower implicit ethnic bias in our data.

We first document that refugee students have significantly fewer social ties than host students. They have fewer friends and classmates who support them emotionally and academically. While on average, a host student receives 2.07 friendship nominations, a typical refugee student receives only 1.08 nominations. Considering that they arrived in their classrooms much later than their host classmates, this finding is not surprising. However, we find that teacher's ethnic bias significantly lowers the number of social ties enjoyed by refugee students. In classrooms where teachers have a stronger ethnic bias, refugee students are more likely to be socially excluded, i.e., they report to have fewer friends and fewer classmates who provide them emotional and academic support. For example, we find that a one standard deviation increase in teacher's ethnic bias score leads to 0.24 fewer friendship ties enjoyed by a refugee student, implying a 22% decline relative to the average.

We also estimate a significant effect of teacher bias on the degree of homophilic ties amongst host students. At the individual level, we find that as the teacher's ethnic bias increases, the number of friendship nominations extended from host students to refugee students decreases, whereas the number of friendship nominations extended from host students to host students increases significantly. This finding also emerges in our analysis of classroom level segregation (homophily). We find that teachers who have a stronger ethnic bias create more ethnically segregated classrooms, measured as excessive numbers of homophilic social ties. Consistently with Currarini et al. (2009), we observe this excess homophily only among host students. A one standard deviation increase in the teacher's ethnic bias increases

friendship homophily among host students by about 13%.

We show that refugee students who are exposed to teachers with a stronger ethnic bias face a higher risk of peer violence and bullying. We find that a one standard deviation increase in teacher's ethnic bias score leads to 5 percent increase in bullying reported by refugee students, with a precise null effect on host students. Exposure to the acts of social exclusion and victimization is likely to slow refugee students' progress in learning the host country's language. This, in turn, is likely to have a detrimental impact on their learning outcomes. Corroborating this conjecture, we estimate that teacher's ethnic bias has a significantly negative impact on refugee children's verbal ability in the host country's language. Specifically, a one standard deviation increase in teacher's ethnic bias score lowers refugee students' Turkish test scores by 0.15 standard deviation. The estimated effect on mathematics test scores is also negative and economically significant (0.07 standard deviation), but does not reach statistical significance. We estimate precise null effects on host students' achievement scores.

We explore various potential mechanisms that may explain our results. First, we rule out a possible reverse causality issue whereby the behavior of refugee children may influence the teacher's implicit bias. We rule out this mechanism by showing that teacher's ethnic bias bears no relation to their assessment of refugee students' behavioral conduct. We then show that an obvious mechanism is at play, that is, teachers transmit their ethnic attitudes to students. We also explore whether our results can be explained by teachers' discriminatory classroom practices and the internalization of such behaviors by host students. Research show that students who have good relationships with their teachers are more likely to be accepted by their classmates; see e.g. Hughes et al. (2001), Skinner and Belmont (1993), Birch and Ladd (1997). We find evidence that supports this mechanism. Students who are perceived to be the teacher's favorites enjoy more friendship ties in our data, and refugee students are much less likely to enjoy such status in classrooms with biased teachers. In these classrooms, refugee students are likely to be spatially segregated and often seated at the back corners of the classroom, away from the teacher's attention.

Our contribution in this paper is twofold. First, to the best of our knowledge, this is the first paper that shows how teachers' implicit attitudes toward an ethnic minority group shape students' social networks, particularly their inter-ethnic relationships. Our setting allows us to identify this relationship causally by exploiting a natural experiment generated by a massive refugee crisis and the way a school placement policy is implemented by the host

country. The second contribution pertains to the quality of our data. Because there is no secondary data available to answer the questions we pose in this paper, we collected detailed primary data from children and teachers using a diverse toolkit. Our data allow us to i) construct multiple measures of social exclusion and ethnic segregation by utilizing the tools of social network theory, ii) control for hard-to-measure and previously unavailable teacher characteristics that may confound our ethnic bias measure, and iii) explore mechanisms through which teachers' ethnic bias might influence students' inter-ethnic relationships.

Our paper contributes to several strands of literature. Several papers test the intergroup contact theory and show that prejudice toward out-groups may dissipate as intergroup relationships are allowed to take hold (see, e.g. Allport (1954), Putnam (2007), Paluck and Green (2008), Rao (2019)). Contrary to the predictions of intergroup contact theory, the literature on integration and diversity management show that negative sentiments against newly encountered out-group members may lead to intergroup conflict (Weiner (1978), Fearon and Laitin (2011), Bazzi et al. (2019)). By showing the factors that mediate the effects of inter-ethnic contact, our paper complements these studies. Our paper also fits in the literature that studies the effects of negative attitudes toward minorities on various outcomes. In the context of education, Van den Bergh et al. (2010) show that the achievement gap between ethnic groups in schools can be explained by the teacher's ethnic bias. Rosenthal and Jacobson (1968), Jussim and Harber (2016), Gershenson et al. (2018) document that teachers' biased expectations for minorities may become self-fulfilling. Alesina et al. (2018) show that teachers' bias against migrant students may be reversed by making teachers aware of their bias; see also Bertrand and Duflo (2017) for a comprehensive review of the broader literature on discrimination. Additionally, our paper contributes to the line of research that strives to understand the role of teachers in shaping children's socio-emotional skills.[3] Finally, by showing how social networks can be shaped and re-shaped via exogenous influences, we contribute to the empirical literature on social networks (Marmaros and Sacerdote (2006)).

Our results highlight the role of teachers in achieving ethnically integrated schools. They suggest that the type of teachers assigned to ethnically mixed schools may have a substantial impact on the nature of students' inter-ethnic relationships. The implication of our study for education policy is that selecting teachers carefully for ethnically diverse schools and offering them professional development opportunities for adopting inclusive classroom practices may

[3]There is now a sizeable literature that shows that interventions aiming at enhancing socio-emotional skills can be successful via intensive teacher training; see Alan et al. (2019), Alan and Ertac (2018).

go a long way in achieving cohesion in these schools and help to build social capital in ethnically diverse communities (Alan et al. (2020)).

The rest of the paper is organized as follows. Section 2 provides the background and the context for identification. Section 3 describes the data, the construction of our outcome measures, and the teacher's ethnic bias score. Our results are presented and discussed in Section 4. We conclude in Section 5.

2 Background and Context for Identification

Our study site, Turkey, has received over 4 million refugees since the start of the Syrian Civil War in 2011. About 1 million of these refugees are school-aged children. As part of a multi-partnered EU initiative launched in 2016, most refugee camps in Turkey are now closed, and along with them, the temporary education centers, where refugee children were receiving remedial education. The current Turkish Ministry of Education (MoE) policy is to place all school-aged refugee children in state schools based on their registered address.[4] School administrators are mandated to admit refugee children located in their catchment areas and, upon admission, distribute them to classrooms based only on their age.[5] All refugee children lacked essential Turkish language ability when they first arrived in their new schools. School administrators are mandated to distribute newcomers to classrooms as evenly as possible to avoid overwhelming teachers. The objective of the school placement policy is to achieve faster integration through total immersion; therefore, schools are not allowed to open separate classrooms for the refugee students.[6] In practice, the placement policy was activated in the 2017-2018 academic year, and most refugee students in our sample had joined their classrooms between 2017 and 2018.

In addition to the refugee placement program, our identification strategy is powered by the way Turkish teachers are appointed to state schools, and the way they are assigned to

[4]If a school is overwhelmed with the numbers of refugee pupils, transportation to nearby schools is provided, again, as part of the deal initiated in 2016 between the EU and the MoE.

[5]While mandated to place the children in grades based on their age, refugee children in a given grade are generally older than host children. This is because school officials tend to place them about 1 grade lower to facilitate better language learning.

[6]A new policy was announced in the summer of 2019 and allowed schools to open language preparation classes for refugee students who lacked proficiency in Turkish based on a nationwide language test. The opening of these classrooms is left to the discretion of the school administrators. Our sample period predates this development.

classrooms within the school. After completing their degree requirements, teachers are placed in a pool to be appointed to a public school in need. A new teacher typically has no say in which school he/she will be appointed to.[7] Currently, teachers may ask to be re-appointed before completing at least four years (over six years in actual practice) of service in their current school. Requests to be re-appointed are honored if i) there is a school in need in the preferred city/district, and ii) the teacher has higher service points than her competitors who stated the same location preference. The service points are accumulated based on years of service. As working in high-SES catchment areas is more desirable for teachers, there tends to be a high teacher turnover in low-SES district schools such as the ones in our sample. Even with a long tenure in the profession, it is extremely hard to be appointed to the generally desired (high-SES) districts. Note also that teachers can be re-appointed involuntarily by the Ministry if they are no longer needed in their current schools.

Our focus in this study is primary schools. Turkish primary schools provide an ideal setting to explore the effect of teacher's ethnic prejudice on students' outcomes because of two reasons: First, except for re-locations mentioned above, a primary school teacher teaches the students allocated to him/her from grade 1 to grade 4. School administrators are mandated to randomly allocate first graders to teachers through publicly held draws in the presence of parents. A primary school teacher spends considerable time with his/her pupils compared to a middle or high school branch teacher, so they are more likely to influence students' behaviors and attitudes. Second, the refugee school placement policy initially targeted primary schools. At the time of our data collection, very few refugee students were placed in middle or high schools.

Despite the institutional and contextual advantages provided by our study site, we face two threats to identification. First, our ethnic bias measure may be confounded by some teacher characteristics that may be relevant for predicting the student outcomes we consider. A notable example is teacher ability/quality. While not established in the literature, it is plausible that teacher quality may be relevant in shaping students' social skills and, therefore, peer interactions. In addition to ability, the ethnic bias measure may be confounded by the teacher's classroom practices (teaching styles) and teaching philosophy, which may also be relevant in determining the way social ties are formed amongst pupils. Concerned that these confounds may exist, we collected detailed data from teachers, including their fluid

[7] In practice, they have no say in which district and even which city they will eventually end up in. Teachers can refuse their placement, but this means giving up a public service job with social security, which is rare.

cognitive ability and cognitive empathy, teaching styles, pedagogical practices, and beliefs (teaching philosophy). All our empirical analyses control for these characteristics in addition to standard teacher demographics and paper qualifications.

The second threat to our identification is possible reverse causality. If teachers' ethnic bias is influenced by their observation of refugee children's social skills and behavioral conduct, it would be difficult to give our findings a causal interpretation. For example, if refugee children tend to act in a manner considered to be anti-social, they may have a hard time forming friendship ties with host children and, observing such anti-social behaviors, teachers may form an unfavorable opinion about refugees and their country of origin. Note also that what is considered anti-social can be culture-specific so that any unaccustomed act may easily be considered anti-social and reinforce ethnic prejudice. The opposite of this argument may also hold. A strongly biased teacher may update her views upon observing refugee students acting in ways contrary to her initial beliefs. In Section 3.4, we revisit these threats and show that our data allow us to address them.

3 Data and the Outcomes

Our data set is a sub-sample of combined baseline data collected as part of two independently run RCTs. These RCTs have been initiated in April 2018 and September 2018 in two Southeast provinces of Turkey, Sanliurfa, and Mersin. These two provinces have received massive numbers of refugees since the start of the Syrian Civil War, and they were part of the MoE's refugee placement program since 2017. Both RCTs include schools that are flagged as "socioeconomically disadvantaged" by the two provincial education authorities. The total number of schools for the RCT 1 and RCT 2 is 80 and 77, respectively, with no overlap. To answer the research question we pose in this paper, we exclude several schools from the original samples. First of all, as our paper focuses on primary school students and teachers, we exclude all middle schools from our sample (27 middle schools in RCT 2). We also exclude schools with no presence of refugee students and those with a single classroom.

Our final selection criterion is to help us achieve a clean identification of the effect of teacher's bias. While there is generally no room for parents to choose their children's teachers in the state system, we were informed that some parents might pressure the administrators for their children to be assigned to a particular teacher in practice. Parents usually prefer older and experienced teachers. These are also teachers who are more likely to be working

in schools that they prefer due to accumulated high service points. Parental involvement is rare in the socioeconomic segment our sample represents. Nevertheless, we collected detailed testimonies from teachers about how, in practice, host and refugee children were assigned to teachers by their school administrators. Using these testimonies, we applied a stringent selection criterion and chose the primary schools in which all students are reported to be allocated to teachers randomly, with no influence from parents.

Our sample selection criteria left us with the majority of the original (combined) baseline data. Our final sample covers over 5000 2nd, 3rd, and 4th-grade students and 192 teachers from 73 primary schools. About 13% of our sample consists of refugee students. We collected our data by physically visiting all schools and spending about three lecture hours in every classroom. In these three lecture hours, we collected data from children using a rich battery of tools with the help of field assistants. While we were collecting data from children, we asked teachers to fill up their own surveys in isolated rooms, so all student data collection took place in the absence of teachers.

Before giving detailed descriptive statistics, we discuss our data collection inventory and how we construct our outcome measures and variables of interest in detail.

3.1 Student Characteristics

We collected rich data on student characteristics using surveys and tests. To measure children's fluid IQ, we implemented "Raven's Progressive Matrices" (Raven et al. (2004)). To measure children's cognitive empathy, we implemented "Reading the Mind in the Eyes" test (Baron-Cohen et al. (1997)). Because there are no centrally administered objective tests at the grade level we are interested in, we also implemented math and Turkish language tests in classrooms. We prepared these tests separately for each grade level, based on the national curricula. Our student survey includes standard demographic information such as family structure, number of siblings, and other socioeconomic indicators, such as the availability of internet at home and the employment status of the father. We also asked students survey questions to elicit their ethnic attitudes, and their teacher's behavior and teaching practices. These questions are given in Appendix B.

To construct detailed measures of social exclusion and ethnic segregation, we elicited social ties between classmates by exercising great care. We now give a detailed account of how we elicit these ties and how we construct our outcomes of interests.

3.1.1 Social Networks in the Classroom

We are interested in whether the teacher's ethnic bias affects the formation of inter-ethnic ties in the classroom. In particular, whether it leads to the social exclusion of refugee children. To answer this question, we first construct social exclusion measures by eliciting social networks in classrooms. To do this, we provide children with a user-friendly paper template. The template asks nominations of classmates in three categories: friendship, emotional support, and academic support. The template also asks students to write down the names of classmates whom they perceive as the teacher's favorite. For each category, students are provided with three boxes to nominate (write down the names of) classmates. First, they are asked to nominate at most their three best friends in the classroom, and then three of the teacher's favorites. Then they move on to nominating at most three classmates from whom they get emotional support, then three classmates from whom they get academic support. Finally, they are asked to nominate at most three classmates whom they emotionally support and then three classmates whom they academically support.

Before the elicitation began, the children were instructed that they could also nominate friends who were absent then. Moreover, they were told that the provided template allows them to nominate up to three classmates for each category, which means that they are allowed to nominate no classmate, or only one, or only two classmates. However, they may not nominate more than three classmates for any category. They were told that they should choose the first three classmates if they have more than three nominations in mind. Finally, they were told that nominations across categories could overlap.[8] We piloted this elicitation method several times using templates that allow for 4 and 5 nominations. Our pilot sessions revealed that children have a hard time filling up larger templates; therefore, we decided to cap the number of nominations to three.

Using the elicited ties, we construct two sets of outcomes. The first set contains individual (node) level outcomes. The first of these outcomes is the standard in-degree centrality measure. This measure is simply the number of edges a node receives in a given category. Recall that our categories are friendship, emotional support, and academic support. A student's in-degree centrality for the friendship category gives the number of friendship

[8]The template was designed as boxes to facilitate children to write one classmate in each box neatly. We gave detailed examples before starting the elicitation to make sure they understood the procedures. Our examples were standard based on clearly written experimental instructions. The elicitation (the entire data collection for that matter) was performed with the utmost discretion to avoid social pressure. The template sample is presented in Appendix C.1.

nominations he/she receives from his/her classmates. The minimum number of nominations one can receive is zero, and the maximum is the number of students (minus the student himself/herself) in the classroom at the time of the elicitation. We construct the in-degree centrality for the other categories similarly. For example, in-degree centrality of a student for the "classmates who support me emotionally" category is the number of classmates who nominate him/her as a supportive classmate, i.e., someone who provides emotional support.

We also construct several individual-level outcomes using out-degree ties, i.e., nominations made by a student. Here, we are interested in inter-ethnic nominations. One of the outcomes we construct is the number of refugee nominations made by a student. For each category, the minimum number of refugee nominations is zero, and the maximum is 3, conditional on the number of refugee students in the classroom. We also use as an outcome variable the number of host nominations made by a refugee child and the number of refugee nominations made by a host child in all three categories. Note that none of these measures are independent of the class size or the proportion of refugee children in the classroom. Therefore, our analyses always control for these two classroom characteristics.

Our second set of measures are classroom-level indices of ethnic segregation. Specifically, following Coleman (1958), we construct a homophily index for refugees and hosts separately. Let R and H denote refugee and host, respectively. Denote the number of within-group ties of group i in classroom j as s_{ij}, and the total number of ties of group i in classroom j as t_{ij}, where $i \in \{H, R\}$. Then, $\frac{s_{ij}}{t_{ij}}$ gives us the share of within-group (homophilic) ties for group i.

Denoting w_{ij} as the population share of group i in classroom j, the excess homophily of group i is defined as $\frac{s_{ij}}{t_{ij}} - w_{ij}$. The intuition behind this definition is simple. Say the share of refugees in a classroom is 10%. Then we would expect refugee students to form around 10% of their friendships with other refugee classmates if the ties are formed at random. If the observed share of within-group ties, $\frac{s_{ij}}{t_{ij}}$, is greater than the expected share, w_{ij}, we conclude that the group exhibits "excess homophily". To make the measure invariant to group size, Coleman (1958) normalizes excess homophily by $1 - w_{ij}$, the maximum possible excess homophily that can be observed in group i.

Then, Coleman's Homophily Index for group $i \in \{R, H\}$ in classroom j:

$$C_{ij} = \frac{\frac{s_{ij}}{t_{ij}} - w_{ij}}{1 - w_{ij}}.$$

Note that since we have two groups (hosts and refugees), Coleman's measure gives us two normalized excess homophily scores, one for each group. We will explore the effect of teacher's ethnic bias on both these scores. Figure 1 presents actual friendship ties from two classrooms in our data for illustration purposes. Both classrooms have similar size (28 and 27 pupils) and a similar number of refugee students (6 and 8). It can be seen in visual clarity that classroom 1 is less ethnically segregated than classroom 2. Coleman Homophily Indices for the host (refugee) group are -0.016 (0.17) and 0.937 (0.635) for classroom 1 and 2 respectively. Note that the refugee excess homophily index is not defined in classrooms with only one refugee student.

3.2 Teacher Characteristics

We collected detailed information from teachers, including demographics, education, experience, tenure, as well as cognitive abilities, teaching styles, and motivation. We explain these after we explain the way we implemented the IAT and construct our ethnic bias measure.

3.2.1 Teacher's Ethnic Bias: An Implicit Association Test

Developed by Greenwald et al. (1998), IAT aims to elicit implicit attitudes toward a group. Because it is quite hard to game, the test is likely to be free from social desirability bias inherent in surveys on socially and politically sensitive topics. Traditionally, the IAT is implemented using computers or tablets. However, using the same protocol, the IAT can be administered using paper and pen in settings where using computers are impractical. Lemm et al. (2008) developed a data collection protocol for a paper-and-pen IAT as an alternative to the computer-based collection. In a paper-and-pen IAT, the respondent observes a word or an image (stimulus), presented in the middle of a line. The two attributes to associate the stimulus are located on the left and the right of the stimulus. The respondent is instructed to mark the attribute he/she associates with the stimulus. The implicit association is quantified by the number of correct responses within a given period. Note that what paper-and-pen IAT measure is an approximation of a respondent's reaction time.[9]

We designed our paper and pen IAT as four attribute categories; good, bad, Turkey, and Syria. We first allowed the teachers to familiarize themselves with the words and images

[9]For the applications of paper-and-pen IAT, see Teachman et al. (2003), Vargas et al. (2007), Lemm et al. (2008).

(stimuli) listed under each attribute category. The first part, "good vs. bad", and the second part, "Turkey vs. Syria", are both single attribute comparisons, and each involves 14 lines (14 stimuli). Teachers are given 14 seconds to complete each part. For example, in the first part (good vs. bad), the respondent sees the word "sad" in the middle of the line and expected to attribute it to the category "bad," which is located on the right of the stimulus. After completing the attribution of 14 stimuli in 14 seconds, the respondent moves to the second part. Here, the same exercise is done for categories Turkey and Syria. For example, the respondent sees the word "Aleppo" and expected to attribute it to the category "Syria," which is located on the right of the stimulus. After completing parts 1 and 2, the respondent moves to the last two parts of the test, which contain two attribute comparisons. In part 3, the categories are "Good/Turkey vs. Bad/Syria" and in part 4, "Good/Syria vs. Bad/Turkey". These parts have 28 lines each (28 stimuli), and teachers were given 30 seconds to complete each part. For example, in part 4, the respondent observes the word "enemy" and expected to attribute it to "Bad/Turkey," as it belongs to the "bad" attribute category. Each part is administered with a timer. We provide our IAT sheets in Appendix D. The first two parts of the task can be considered as warm-up sessions. We use the last two parts to construct our measure of ethnic bias. Following the protocol by Lemm et al. (2008), we construct our implicit bias score as follows:

Let A and B denote the number of correct answers in part 3 (Good/Turkey vs. Bad/Syria) and part 4 (Good/Syria vs. Bad/Turkey), respectively. Then, the ethnic bias score η is:

$$\eta = \frac{X}{Y} * \sqrt{X - Y},$$

where $X = max\{A, B\}$ and $Y = min\{A, B\}$. If $B > A$, then the resulting values are multiplied by -1 in order to retain the direction of the effect.[10]

Higher values of score η indicate a stronger implicit bias toward Syria and Syrians. After constructing the score in the way described above, we standardize it to give mean zero and standard deviation of 1. This transformation facilitates a more intuitive interpretation of our coefficient estimates. Figure 2 presents the distribution of the standardized ethnic bias score in our data, showing substantial variation across teachers.[11]

[10]Note that the formula does not allow for $Y = 0$. We do not have a case where either A or B is zero, so we let set minimum $B = 1$. Then as part 3 and part 4 contain 28 questions each, we have $-145.5 < \eta < 145.5$.

[11]We were not able to obtain accurate IAT scores for about 12% of the teachers. This is due either to the teacher's refusal to complete the task or the experimenter error of following the written protocol.

3.2.2 Teacher Ability, Teaching Styles and Motivation

In addition to standard information on demographics and qualifications, we were able to collect information from teachers in detail that is, to the best of our knowledge, previously unavailable in any dataset. We measure the teachers' cognitive ability and cognitive empathy using Raven's Progressive Matrices, and the Reading the Mind in the Eyes test, respectively. We also elicited teaching styles (pedagogy) through survey questions.[12] Using these questions, we construct five main teaching styles: Modern teaching, Growth mindset, Warmth, Extrinsic motivation, and Inquiry-based teaching.

It is also argued that conditional on ability, the teacher's effort and motivation are important indicators of his/her quality (Duckworth et al. (2009)). However, it is hard to observe and measure these characteristics. A commonly applied method to capture motivation is to measure professional satisfaction (Pool (1997)). For this, we asked several questions regarding job satisfaction and dedication to the teaching profession to construct a motivation measure, which we refer to as "professional satisfaction". Finally, we asked teachers their assessment of their own competency in teaching. Exemplary questions for each of the categories mentioned above are presented in Appendix B.

3.3 Descriptive Statistics

Table 1 presents a summary of student characteristics separately for host and refugee children. The last column gives the p-value from the test of equality in means across the two groups. It is clear that host and refugee student characteristics are significantly different. Refugee children are, on average, one year older than host children and their fathers are more likely to be unemployed. Apart from gender composition, only two characteristics are not different across host and refugee students, fluid cognitive ability (Raven score), and behavioral conduct grades assigned by teachers. However, refugee students perform significantly worse in cognitive empathy test. They scored 0.48 standard deviations lower in the Reading the Mind in the Eyes test. However, a caution is in order here as this test requires command of Turkish language with a relatively rich vocabulary.[13] As expected, refugee students

[12] Pedagogy has been shown to be important in determining student outcomes; see Bietenbeck (2014), Hidalgo-Cabrillana and Lopez-Mayan (2018).

[13] Even though we assigned an Arabic speaking assistant to help refugee students who do not have a good command of the Turkish language, this test is likely to be biased against non-native speakers.

perform significantly worse than host children in standardized math and verbal (Turkish) ability tests. Refugee children scored 0.40 and 0.52 standard deviations lower in math and Turkish, respectively. Under the network variables in Table 1 we summarize our individual (node)-level social exclusion measures. The first two provide friendship nominations received (in-degree ties), and friendship nominations made (out-degree ties) followed by nominations received for the status of teacher's favorite student. The next four present emotional and academic support received and provided. The last three provide average out-degree ties that involve refugee students. All these measures clearly show that refugee students have significantly fewer friends, and they provide and receive less support from their classmates than host children. Significant differences between host and refugee students in these measures indicate severe ethnic segregation in classrooms. Moreover, refugee students report a higher level of teacher neglect and, interestingly, they have higher ethnic bias than host students. Bullying is prevalent in our sample, for both hosts and refugees. Nevertheless, refugee students are significantly more likely to experience bullying from their peers compared to host students, and they report higher number of bullies.

Table 2 presents teacher demographics and classroom characteristics. About 38% of our teachers are male, with an average age of 35. The average teaching experience stands at 12 years with considerable variation, ranging from 1 year to 29 years. Panel 1 and 2 in Figure 3 present the distribution of teacher's (standardized) cognitive ability and cognitive empathy scores, respectively. In both cognitive measures, we observe significant variation across teachers. The proportion of refugee students in classrooms varies greatly, ranging from 0.02 to 0.47, with an average of 15%. The last six rows give the summary statistics of Coleman's Homophily Indices for hosts and refugees. Recall that this index varies between -1 and 1 with higher values indicating more homophilic ties. We observe high average host homophily, with substantial variation across classrooms. Variation in refugee homophily is large as well but it has much lower average.

3.4 Internal Validity Checks

As mentioned in Section 2, despite the MoE's mandate of randomly assigning pupils across classrooms, anecdotal evidence suggests that the rule may not always be adhered and school principals sometimes give in to pressure from some parents. This type of selection may lead to ability sorting and mask the actual teacher effect on student outcomes. Given the way we choose our sample, this selection is highly unlikely. Nevertheless, we perform several

checks to make sure that our results are internally valid. We first check whether there is any evidence of ability sorting, i.e., whether teacher characteristics predict students' cognitive ability. Table 3 presents the predictive power of teacher characteristics on pupil's cognitive ability scores for hosts and refugees separately. The first two columns control only for school and province fixed effects, class size, and grade level. Along with student age and gender, these covariates explain about 33% (25%) of the variation in the fluid cognitive ability of hosts (refugees). As we add teacher's demographic characteristics (Columns 3 and 4), then teacher's cognitive characteristics (Columns 5 and 6), and finally teaching styles (Columns 7 and 8), we do not gain extra explanatory power for hosts, but teaching styles (jointly) seem to matter for refugee students.

We conduct a similar internal validity check to see whether school administrators did follow the rule of fair allocation of refugee students across classrooms. Table 4 presents the predictive power of teacher characteristics on the proportion of refugee students in classrooms. As can be seen, province and school fixed effects, class size and grade level together explain 81% of the variation in the proportion of refugee students across classrooms. None of the teacher characteristics, including the teacher's ethnic bias score, have any predictive power on the proportion of refugees in the classroom. Almost all coefficient estimates are virtually zero. Confirming the declarations of school administrators and teachers in our sample, we find no evidence of ability sorting or non-random allocation of refugee students across teachers.

4 Results

Before estimating the effect of teacher's ethnic bias on student outcomes we describe above, we analyze what our ethnic bias score captures in terms of classroom and teacher characteristics. For this, we regress the standardized measure of ethnic bias on classroom and teacher characteristics. For the former, we use class size and the proportion of refugee students in the classroom. For the latter, we use teacher's demographic characteristics, cognitive characteristics, teaching styles, and motivation.

Table 5 presents our results controlling for school and province fixed effects. Our full covariate set (Column 4) collectively explains about 54% of the variation in teacher's ethnic bias score. Teacher's cognitive ability emerges as the only significant predictor of his/her ethnic bias score. A one standard deviation increase in teacher's Raven score is associated

with 0.29 standard deviations decline in ethnic bias score, and this association is significant at the 10% level. Note that demographics, experience, cognitive empathy, teaching styles, and motivation do not appear to predict a teacher's ethnic bias. Nevertheless, we will control for all these characteristics when estimating the effect of teacher's ethnic bias on our outcomes of interest.

4.1 Empirical Specification

Our empirical specification for individual-level outcomes is as follows:

$$y_{ics} = \alpha_0 + \alpha_1 \text{TBIAS}_{cs} + \text{SC}_{ics}^{'}\beta + \text{TC}_{cs}^{'}\gamma + \delta_s + \varepsilon_{ics}, \tag{1}$$

where y_{ics} is the outcome of interest for child i in classroom c in school s. TBIAS_{cs} is the standardized measure of teacher's ethnic bias, the variable of interest. Vector SC_{ics} is a vector of observables for student i in classroom c school s that are likely predictive of the outcome y. These include gender, age, standardized cognitive ability, and cognitive empathy score. Vector TC_{cs} contains teacher and classroom characteristics, and δ_s are school fixed effects. The coefficient of interest is $\hat{\alpha}_1$. We cluster standard errors at the school level. Since the variable of interest is continuous, when appropriate, we plot linear predictions from our estimates for easier interpretations.

Throughout the paper, we estimate the above empirical model for host and refugee students separately, and test the equality of coefficient estimates. The estimated differential effects of teacher's bias on the outcomes of hosts and refugees provide a strong support for the validity of our ethnic bias measure. Because we test multiple hypotheses for individual level outcomes, we also provide Romano-Wolf corrected p-values in Appendix A.

For classroom level outcomes (ethnic segregation/homophily measures) we also perform a semi-parametric estimation and provide visually the non-parametric relationship between teacher's ethnic bias and segregation in the classroom. Our semi-parametric estimation uses the following specification:

$$y_{cs} = \alpha_0 + \text{TC}_{cs}^{'}\gamma + f(\text{TBIAS}_{cs}) + \delta_d + \varepsilon_{cs}, \tag{2}$$

where y_{cs} is the outcome of interest for classroom c in school s. Vector TC_{cs} contains teacher and classroom characteristics, such as teaching styles, classroom size, and proportion of

refugees. And δ_s are school fixed effects. The objective here is to visualize the non-parametric relationship between TBIAS_{cs} and y_{cs}, controlling (parametrically) for the aforementioned characteristics. Again, we cluster standard errors at the school level.

4.2 Teacher's Ethnic Bias and Social Exclusion of Refugee Children

We now examine the impact of teacher's ethnic bias on the student outcomes as detailed above. While we use a broad set of controls in all our analyses, for the sake of space, we present only the estimates of the coefficient of interest, that is, the coefficient on teacher's ethnic bias score, $\hat{\alpha}_1$. Tables that present all coefficient estimates are available upon request.

We begin by discussing the results for our social exclusion measures. As explained in Section 3.1.1, these are our individual-level outcome measures, constructed using reported social ties. Table 6 presents the effect of teacher's ethnic bias on friendship ties for host and refugee students separately. The first two columns give the estimated effects of teacher's ethnic bias on the number of friendship nominations received (in-degree centrality), the last two columns give the same for friendship nominations made (out-degree centrality). We do not estimate a significant difference in the impact between refugee and host children with respect to the number of friends nominated (the last two columns). However, the estimated effect on the number of friendship nominations a refugee child receives is negative and statistically different from that of a host child; see the last row for the test for the equality of impact estimates. A one standard deviation increase in teacher's ethnic bias score leads to, on average, 0.24 fewer friendship nominations that a refugee child receives. To put this estimate into perspective, the average number of friendship nominations received by refugee children is 1.08 (see Table 1), which is about 50% less than the ties enjoyed by the host children. 0.24 fewer ties imply a 22% decline in friendship ties enjoyed by refugee children due to a one standard deviation increase in teacher's ethnic bias. Teacher's bias has an opposite, albeit small effect on host students, and this effect is statistically significant at 10%.

Table 7 presents the same estimates for other social ties categories. Columns 1 and 2 present the estimated effects of teacher's bias on the number of classmates a student claims to receive emotional support from, plus the number of such nominations she receives from her classmates. Columns 3 and 4 present the same estimates for the academic support category. Columns 5 and 6 present the effects on the total number of classmates a student reports to

provide emotional support to, plus the number of such nominations she obtains from her classmates. Columns 7 and 8 present the same estimates for the academic support category. A simple illustration (for a hypothetical class size of four) of how we construct these outcomes is given in Appendix C.2. This table shows the same pattern we observe in Table 6. That is, as the teacher's ethnic bias increases, the number of social ties a refugee student enjoys declines significantly. Take emotional support received for an example (Column 2). A one standard deviation increase in teacher's ethnic bias score leads to an average of 0.22 fewer number of emotional support ties a refugee student enjoys. The results for academic ties are similar. We estimate no significant impacts on host students' emotional and academic support ties.

The above analyses explore the effect of teacher's bias on the number of social ties for host and refugee students separately. Our next question pertains to the effect of teacher's bias on the formation of inter-ethnic social ties. For this, we first analyze the effect of teacher's bias on the prevalence of nominations extended to refugee students. To put our estimates in perspective, keep in mind that the average number of refugee classmates nominated as friends already stands low (0.19 nominations). See also in Table 1, the average number of refugee classmates nominated as friends by host children is only 0.12. Table 8 presents the estimated effects on the number of refugee nominations for all three categories. Column 1 presents the estimates for the number of refugee classmates a student nominates as his/her friends, Column 2 and Column 3 present the number of refugee classmates a student claims to extend emotional and academic support to, respectively. We obtain similar findings in these analyses. As the teacher's ethnic bias increases, the number of nominations that involve refugee students declines significantly. Specifically, a one standard deviation increase in teacher's ethnic bias score leads to 0.03 fewer refugee students nominated as friends, and 0.02 fewer refugee students nominated as classmates to provide emotional and academic support to, respectively. While seemingly small, these estimates are likely to be economically significant. Considering that the average number of refugee classmates nominated as friends is 0.19, the coefficient estimate of 0.03 implies an approximately 16% decline in the number of friendship nominations extended to refugee students. We now analyze these inter-ethnic nominations further.

Table 9 presents the estimated effects of teacher's ethnic bias on students' inter-ethnic friendship nominations. The first 2 columns present nominations made by host children, the last two columns refugee children. Panel 1 presents the results based on friendship ties, Panel 2 and Panel 3 emotional and academic support ties, respectively. Panel 1 confirms

our earlier findings that teacher's ethnic bias lowers the number of friendship nominations refugee students receive. A one standard deviation increase in teacher's ethnic bias leads to 0.02 fewer refugee classmate nominations, and 0.05 more host classmate nominations made by host children. What is particularly striking in this panel is that we not only estimate a sizable negative impact on the number of refugee nominations made by refugee students, but also a positive impact on the number of host nominations made by refugee students. For the latter, we find that a one standard deviation increase in teacher's ethnic bias leads to 0.10 more friendship nominations extended by refugee students to host students. However, the estimate does not reach significance. These findings are consistent with Currarini et al. (2009), which show that preferences toward socializing with the same ethnicity imply more popularity for the members of the dominant ethnic group than those of minority groups. Our results show how teacher's negative attitudes toward the minority can play a role in reinforcing this result. We obtain similar, albeit statistically weaker results on emotional and academic support categories. As in the case of friendship nominations, we estimate significantly fewer nominations extended from hosts to refugee students in both categories. We do not estimate significant effects on the host to host nominations in these two categories.

4.3 Teacher's Ethnic Bias and Ethnic Segregation in the Classroom

The above results imply that ethnically biased teachers are likely to create ethnically segregated classrooms. Such classrooms would be characterized by strong in-group (homophilic) ties with minimal inter-group socialization. To test this hypothesis, we use Coleman's Homophily Index described in Section 3.1.1 as our outcome and estimate the effect of teacher's ethnic bias on the level of excess homophily within the host and refugee students separately. Table 10 presents the results. Consistently with our individual-level analyses, we estimate significant effects of teacher's ethnic bias on host homophily. A one standard deviation increase in teacher's ethnic bias leads to 0.09 units (about 13%) increase in host homophily index concerning friendship ties. Both absolute and relative effect sizes concerning emotional and academic support ties are similar (about 10% and 13% increase respectively). We find no evidence of excess homophily within refugee students, which is consistent with our individual-level results. Figure 4 depicts the estimated non-parametric relationship between teacher's ethnic bias and excess host homophily using the semi-parametric specification in Equation 2. These non-parametric estimates present, in visual clarity, the strong relationship between teacher's ethnic bias and ethnic segregation in the classroom.

4.4 Teacher's Ethnic Bias and Peer Violence

Ethnic segregation likely leads to ethnic tensions and conflict in the school ground. In ethnically segregated schools, socially excluded minority students may be at a higher risk of falling victim to verbal and physical bullying. Recall that refugee students report higher exposure to bullying than host students do in our sample. Specifically, the 93% (89%) of refugee (host) students reported to be bullied regularly at school. We estimate a significant effect of teacher's ethnic bias on the probability of a refuge student reporting bullying victimization. Table 11 presents the results for hosts and refugees separately. As seen in the table, refugee students who are exposed to teachers with stronger ethnic bias face a higher risk of being bullied by their peers. A one standard deviation increase in teacher's ethnic bias leads to 5 percentage points increase in the probability of being bullied by peers. We estimate a precise null effect on host students. Results for the number of reported bullies are similar.

4.5 Teacher's Ethnic Bias and Achievement of Refugee Children

Taken together, these results strongly suggest that teachers have a significant role in the social exclusion of refugee students in their new schools. This exclusion is likely to slow refugee students' progress in learning the host country's language, and negatively impact their achievement outcomes. Table 12 presents results that corroborate this prediction. We estimate that teacher's ethnic bias has a significantly detrimental effect on refugee students' verbal ability in the Turkish language. A one standard deviation increase in teacher's ethnic bias score lowers refugee students' Turkish test scores by 0.15 standard deviation. The estimated effect on mathematics scores is negative and economically significant (0.07 standard deviation) but does not reach statistical significance. Note that the teacher's ethnic bias has no impact on host students' achievement scores. These precise null estimates confirm that our ethnic bias measure does not proxy an unobserved teacher quality after controlling for available teacher characteristics. Figure 5 provides linear predictions obtained from our estimates and show our results in visual clarity.

4.6 Potential Mechanisms

An important threat to our identification strategy is possible reverse causality. The reverse causality may be driving our results if the teacher's IAT score is- rather than representing an inherent bias against refugees- influenced by her observation of refugee students. For example, if refugee students have a higher tendency to behave in an anti-social manner than the host students, teachers may form an unfavorable opinion about them, and this may be partly reflected in their IAT scores. Anti-social students may find it hard to form social ties with their classmates. Such a mechanism would be consistent with our findings. If this mechanism is at play, we would observe that refugee students receive less favorable behavioral conduct grades from biased teachers. Recall that we find no evidence of differential grading of behavioral conduct (p-value=0.59, see Table 1). Importantly, we do not find any association between teacher's ethnic bias and behavior grades she assigns either to refugees or to hosts. This suggests that our results are unlikely to be driven by reverse causality. Figure 6 presents the coefficient estimates for hosts and refugees, neither of which are statistically different from zero.[14]

What are the likely mechanisms that explain our results? Why refugee children who are assigned to teachers with ethnic bias are more likely to be socially excluded and subject to more peer bullying? Moreover, why host students with such teachers tend to form more homophilic ties? An obvious channel to explore is a mechanism whereby the teacher's ethnic bias is transmitted to students, lowering their willingness to form social ties with different ethnic groups. We test this by estimating the effect of teacher's ethnic bias on students' ethnic bias. Another mechanism may be related directly to the teacher's behavior and the fact that he/she is in a powerful position to set behavioral norms in the classroom, i.e, he/she is a role model. A teacher with a strong prejudice against an ethnic group may adopt discriminatory practices in the classroom and treat minority students differently. These may include neglect or disproportionate reprimanding. Host students might internalize these behaviors as norms and act accordingly. To test this mechanism, we construct two measures of discriminatory teacher practices. The first measure is based on the students' report on teacher behavior, and the other one on the teacher's classroom seating arrangement. For the

[14]We collected behavioral conduct grades from teachers using a template that instructs the teacher to assign for each pupil a grade between 1 and 5, 1 representing no behavioral problem and 5 severe behavioral problems. These grades were collected as part of our endline data collection for RCT 1. The intervention has shown to have no effect on these grades. Nevertheless, we control for the treatment status of the school when analyzing this variable.

former, students answered survey questions asking whether their teacher cares about their problems, whether they make an effort to explain complex problems, and whether they set high expectations for their academic achievement. We construct student reported "teacher neglect" using these questions (see Appendix B). The latter measure is more subtle than the former. Classroom seating arrangements are considered to be essential management tools, and research shows that these arrangements impact learning outcomes (Wannarka and Ruhl (2008)). In our setting, host students, students who are popular (with higher in-degree centrality), and those who are reported to be teacher's favorites are less likely to be seated at the back, where it is harder to get teacher's attention.[15] A teacher with a strong prejudice against an ethnic group may choose to push the minority students to back seats and interact with them less frequently, which may influence the way students socialize with each other.

Figure 6 plots the coefficient estimates for behavior grades (for evidence against reverse causality), student ethnic bias, student-reported teacher neglect, the number of nominations for the status of teacher's favorite, and the probability of sitting at the back corner of the classroom. We plot these estimates separately for host and refugee students to highlight the differential impacts. It is evident in this picture that the transmission mechanism is an important channel. The striking finding here is that this transmission holds for both hosts and refugee students. In fact, the effect is particularly strong for refugee students. It appears that students who are exposed to teachers with ethnic prejudice are likely to develop such biases toward those whom they consider as out-group.

Contrary to our initial conjecture, however, the teacher's ethnic bias does not affect refugee students' perception of teacher neglect. Interestingly, it has a significant impact on host students' perception. Specifically, hosts students with ethnically biased teachers report significantly less teacher neglect. This result comes across counter-intuitive at first. However, it is consistent with our main results. The teacher faces a choice of allocating her effort between regular teaching and helping newcomers who have limited language ability. Teachers with less ethnic prejudice may allocate more time to helping refugee students. Their teacher's extra attention to newcomers may not go unnoticed by host students and possibly interpreted as "neglect" on their part. However, while host students in these classrooms report more teacher neglect, they do not reflect this as resentment toward their refugee classmates. On the contrary, inter-ethnic social ties are much more prevalent in such classrooms, and this is likely due to the teacher's inclusive practices and personal efforts to mix students.

[15] All related p-values are less than 0.01.

The mechanism of teacher's discriminatory practices is more directly evident in the results of seating arrangements. As shown in Figure 6, ethnically biased teachers tend to push refugee students to the back corners of the classroom. Refugee students are 8 percentage points more likely to sit at back desks in classrooms with biased teachers. The effect is virtually zero for host students. Consistently with these findings, refugee students are significantly less likely to enjoy the status of "teacher's favorite" in classrooms with biased teachers. Specifically, a one standard deviation increase in teacher bias is associated with 0.21 fewer nominations for refugee students for the status of teacher's favorite. All said, the transmission of ethnic prejudice from teachers to students and teachers' discriminatory classroom practices internalized by students are likely to drive our results.

5 Conclusion

We show that teachers' implicit bias against a minority group has a significant effect on students' socialization choices in the classroom, particularly their inter-ethnic relationships. For identification, we exploit a setting where the ethnic composition in schools changed due to massive refugee influx, and students, including refugees, are randomly assigned to classrooms. We elicit students' social networks to construct indicators of social exclusion of minority students and ethnic segregation in each classroom.

We find that teachers' ethnic bias, measured by an Implicit Association Test, significantly lowers the prevalence of social ties between host and refugee children and hinders refugee children's progress in learning the host country's language. We show that teachers who have a stronger ethnic bias toward the refugees' country of origin create socially segregated classrooms, where host students tend to form homophilic social ties. In these classrooms, refugee students are more likely to be spatially segregated and face higher risk of bullying victimization as well. The transmission of ethnic attitudes from teachers to students and teacher's discriminatory classroom practices internalized by students appear as likely mechanisms driving these results.

Given the importance of the childhood period for the development of socio-emotional and cognitive skills, our results imply that the type of teachers children are exposed to in ethnically diverse schools can have significant societal impacts. Our results suggest that ethnic prejudice can breed in schools, marginalize minority children, and deprive the majority children of realizing the benefits of ethnic diversity. To the extent that these harmful

effects persist into adolescence and adulthood, they may damage communities' cohesiveness by increasing the risk of ethnic tension and conflict. This study shows that ethnically diverse schools are obvious grounds to take preemptive measures against these adverse social outcomes.

References

Alan, S., Baysan, C., Gumren, M., and Kubilay, E. (2020). Building Inter-Ethnic Cohesion in Schools: An Intervention on Perspective-Taking. Technical Report 2020-009, Human Capital and Economic Opportunity Working Group.

Alan, S., Boneva, T., and Ertac, S. (2019). Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit. *The Quarterly Journal of Economics*, 134(3):1121–1162.

Alan, S. and Ertac, S. (2018). Fostering Patience in the Classroom: Results from Randomized Educational Intervention. *Journal of Political Economy*, 126(5):1865–1911.

Alan, S., Ertac, S., and Mumcu, I. (2018). Gender Stereotypes in the Classroom and Effects on Achievement. *The Review of Economics and Statistics*, 100(5):876–890.

Alesina, A., Carlana, M., Ferrara, E. L., and Pinotti, P. (2018). Revealing Stereotypes: Evidence from Immigrants in Schools. Working Paper 25333, National Bureau of Economic Research.

Alesina, A. and Ferrara, E. L. (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, 43(3):762–800.

Allport, G. W. (1954). *The Nature of Prejudice*. The Nature of Prejudice. Addison-Wesley, Oxford, England.

Arbatli, C. E., Ashraf, Q. H., Galor, O., and Klemp, M. (2020). Diversity and Conflict. *Econometrica*, 88(2):727–797.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(7):813–822.

Bazzi, S., Gaduh, A., Rothenberg, A. D., and Wong, M. (2019). Unity in Diversity? How Intergroup Contact Can Foster Nation Building. *American Economic Review*, 109(11):3978–4025.

Bertrand, M. and Duflo, E. (2017). Field Experiments on Discrimination. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, volume 1 of *Handbook of Field Experiments*, pages 309–393. North-Holland.

Bietenbeck, J. (2014). Teaching practices and cognitive skills. *Labour Economics*, 30(C):143–153.

Birch, S. H. and Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, 35(1):61–79.

Buhs, E. S., Ladd, G. W., and Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology*, 98(1):1–13.

Burgess, S. and Greaves, E. (2013). Test Scores, Subjective Assessment, and Stereotyping of Ethnic Minorities. *Journal of Labor Economics*, 31(3):535–576.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics*, 134(3):1163–1224.

Coleman, J. (1958). Relational Analysis: The Study of Social Organizations with Survey Methods. *Human Organization*, 17(4):28–36.

Currarini, S., Jackson, M. O., and Pin, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, 77(4):1003–1045.

Dahlberg, M., Edmark, K., and Lundqvist, H. (2012). Ethnic Diversity and Preferences for Redistribution. *Journal of Political Economy*, 120(1):41–76.

Duckworth, A. L., Quinn, P. D., and Seligman, M. E. P. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology*, 4(6):540–547.

Fearon, J. D. and Laitin, D. D. (2011). Sons of the Soil, Migrants, and Civil War. *World Development*, 39(2):199–211.

Gershenson, S., Hart, C. M. D., Hyman, J., Lindsay, C., and Papageorge, N. W. (2018). The Long-Run Impacts of Same-Race Teachers. Working Paper 25254, National Bureau of Economic Research.

Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores. *The Quarterly Journal of Economics*, 132(3):1219–1260.

Gradstein, M. and Justman, M. (2002). Education, Social Cohesion, and Economic Growth. *American Economic Review*, 92(4):1192–1204.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480.

Hidalgo-Cabrillana, A. and Lopez-Mayan, C. (2018). Teaching styles and achievement: Student and teacher perspectives. *Economics of Education Review*, 67:184–206.

Hughes, J. N., Cavell, T. A., and Willson, V. (2001). Further support for the developmental significance of the quality of the teacher–student relationship. *Journal of School Psychology*, 39(4):289–301.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.

Jussim, L. and Harber, K. D. (2016). Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies:. *Personality and Social Psychology Review*.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, 92(10):2083–2105.

Lazarsfeld, P. F. and Merton, R. K. (1954). Friendship as a Social process: A Substantive and Methodological Analysis. In *Freedom and Control in Modern Society*, volume 18, pages 18–66. New York: Octagon.

Lemm, K. M., Lane, K. A., Sattler, D. N., Khan, S. R., and Nosek, B. A. (2008). Assessing implicit cognitions with a paper-format Implicit Association Test. In *The Psychology of Modern Prejudice*, pages 123–146. Nova Science Publishers, Hauppauge, NY, US.

Marmaros, D. and Sacerdote, B. (2006). How Do Friendships Form? *The Quarterly Journal of Economics*, 121(1):79–119.

O'Reilly III, C. A., Williams, K. Y., and Barsade, S. (1998). Group demography and innovation: Does diversity help? In *Composition*, Research on Managing Groups and Teams, Vol. 1., pages 183–207. Elsevier Science/JAI Press, US.

Ottaviano, G. I. P. and Peri, G. (2012). Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10(1):152–197.

Paluck, E. L. and Green, D. P. (2008). Prejudice Reduction: What Works? A Review and Assessment of Research and Practice. *Annual Review of Psychology*, 60(1):339–367.

Paluck, E. L., Green, S. A., and Green, D. P. (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy*, 3(2):129–158.

Pool, S. W. (1997). The Relationship of Job Satisfaction With Substitutes of Leadership, Leadership Behavior, and Work Motivation. *The Journal of Psychology*, 131(3):271–283.

Putnam, R. D. (2007). E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, 30(2):137–174.

Rao, G. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, 109(3):774–809.

Raven, J., Raven, J. C., and Court, J. H. (2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Pearson, San Antonio, TX. OCLC: 697438611.

Rosenthal, R. and Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, 3(1):16–20.

Skinner, E. A. and Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4):571–581.

Teachman, B. A., Gapinski, K. D., Brownell, K. D., Rawlins, M., and Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology*, 22(1):68–78.

Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., and Holland, R. W. (2010). The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap. *American Educational Research Journal*, 47(2):497–527.

Vargas, P. T., Sekaquaptewa, D., and von Hippel, W. (2007). Armed Only with Paper and Pencil: "Low-Tech" Measures of Implicit Attitudes. In *Implicit Measures of Attitudes*, pages 103–124. The Guilford Press, New York, NY, US.

Verkuyten, M. and Thijs, J. (2016). Ethnic discrimination and global self-worth in early adolescents: The mediating role of ethnic self-esteem:. *International Journal of Behavioral Development*.

Wannarka, R. and Ruhl, K. (2008). Seating arrangements that promote positive academic and behavioural outcomes: A review of empirical research. *Support for Learning*, 23(2):89–93.

Weiner, M. (1978). *Sons of the Soil: Migration and Ethnic Conflict in India*. Princeton University Press.

6 Tables

Table 1: Refugee and Host Student Characteristics

	Host		Refugee		
	Mean	SD	Mean	SD	P-value
Student Demographics:					
Male	0.51	(0.50)	0.49	(0.50)	0.54
Age in Months	104.84	(10.27)	116.53	(15.77)	0.00
SES Indicators:					
Working Mother	0.28	(0.45)	0.35	(0.48)	0.00
Working Father	0.89	(0.31)	0.76	(0.43)	0.00
Computer at Home	0.44	(0.50)	0.34	(0.48)	0.00
Internet at Home	0.54	(0.50)	0.57	(0.50)	0.27
Cognitive & Behavioral Outcomes:					
Raven Score	-0.00	(1.00)	0.00	(1.01)	0.95
Cognitive Empathy Score	0.00	(1.00)	-0.48	(0.95)	0.00
Math Score	0.00	(1.00)	-0.40	(0.98)	0.00
Verbal Score	0.00	(1.00)	-0.52	(0.93)	0.00
Teacher's Behavioral Conduct Grade	0.00	(1.00)	-0.04	(1.01)	0.59
Network Variables:					
Friendship Ties (In-Degree)	2.07	(2.04)	1.08	(1.41)	0.00
Friendship Ties (Out-Degree)	2.32	(0.92)	1.90	(1.09)	0.00
Nominations for Teacher's Favorite	1.59	(2.78)	0.70	(1.59)	0.00
Emotional Support Ties (Received)	2.31	(1.89)	1.36	(1.59)	0.00
Academic Support Ties (Received)	1.94	(1.71)	1.23	(1.46)	0.00
Emotional Support Ties (Provided)	2.32	(1.89)	1.34	(1.57)	0.00
Academic Support Ties (Provided)	1.96	(1.85)	1.11	(1.47)	0.00
Refugee Friendship Nominations	0.12	(0.36)	0.65	(0.84)	0.00
Refugee E. Support Nominations	0.10	(0.34)	0.30	(0.58)	0.00
Refugee A. Support Nominations	0.10	(0.34)	0.30	(0.60)	0.00
Student Beliefs:					
Teacher Neglect	-0.00	(1.00)	0.24	(0.93)	0.00
Ethnic Bias	-0.00	(1.00)	0.15	(0.94)	0.00
Reported Bullying:					
Proportion Reports Bullying	0.89	(0.32)	0.93	(0.26)	0.00
Number of Bullies Reported	7.02	(4.84)	8.04	(4.64)	0.00

Cognitive & behavioral outcomes, and students beliefs are standardized to have mean zero for host students. Network variables give the number of edges (ties) for each category and represent our individual (node) level outcomes.

Table 2: Teacher and Classroom Characteristics

	Mean	SD	Min	Max
Teacher Demographics:				
Male	0.38	0.49	0.00	1.00
Age	35.28	7.19	23.00	49.00
Married	0.74	0.44	0.00	1.00
Number of Children	1.15	0.97	0.00	4.00
Tenured	0.95	0.21	0.00	1.00
Years of Experience	11.64	7.04	1.00	29.00
Number of Semesters in Current Class	4.31	1.81	1.00	8.00
Classroom Variables:				
Class Size	31.71	7.50	15.00	53.00
Proportion of Refugees	0.15	0.09	0.02	0.47
Coleman Host Homophily (Friendship)	0.67	0.33	-0.15	1.00
Coleman Host Homophily (Emotional S.)	0.61	0.36	-0.27	1.00
Coleman Host Homophily (Academic S.)	0.67	0.39	-0.52	1.00
Coleman Refugee Homophily (Friendship)	-0.16	0.67	-1.00	1.00
Coleman Refugee Homophily (Emotional S.)	-0.19	0.72	-1.00	1.00
Coleman Refugee Homophily (Academic S.)	-0.30	0.73	-1.00	1.00

Table 3: Student Ability Sorting Across Classrooms

	(1)		(2)		(3)		(4)	
	Host	Refugee	Host	Refugee	Host	Refugee	Host	Refugee
Teacher Ethnic Bias	-0.02	-0.04	-0.01	-0.02	-0.03	-0.03	-0.03	-0.03
	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)	(0.04)	(0.02)	(0.04)
Male	-0.06**	-0.11	-0.06**	-0.11	-0.06**	-0.11	-0.06**	-0.10
	(0.03)	(0.09)	(0.03)	(0.09)	(0.03)	(0.09)	(0.03)	(0.09)
Student Age in Months	0.00	0.02***	0.00	0.02***	0.00	0.02***	0.00	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Teacher Demographics:								
Male Teacher			-0.02	0.04	-0.01	0.06	0.00	0.11
			(0.05)	(0.11)	(0.05)	(0.11)	(0.05)	(0.12)
Teacher Age			-0.02*	0.01	-0.02*	0.01	-0.02*	0.00
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Married Teacher			0.08	0.06	0.09	0.08	0.09	-0.01
			(0.08)	(0.13)	(0.07)	(0.13)	(0.07)	(0.13)
Number of Children			-0.01	-0.05	-0.01	-0.06	-0.02	-0.07
			(0.04)	(0.07)	(0.04)	(0.07)	(0.04)	(0.07)
Tenured Teacher			-0.11	-0.10	-0.11	-0.13	-0.12	-0.01
			(0.15)	(0.23)	(0.15)	(0.24)	(0.15)	(0.27)
Years of Experience			0.02*	0.00	0.02*	0.00	0.02	0.01
			(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Teacher Cognition:								
Teacher Raven Score					-0.05**	-0.04	-0.03	0.02
					(0.02)	(0.04)	(0.02)	(0.06)
Teacher Cognitive Empathy Score					0.02	-0.01	0.02	-0.00
					(0.03)	(0.06)	(0.03)	(0.06)
Teaching Styles:								
Growth Mindset							0.04	-0.05
							(0.04)	(0.07)
Inquiry-based Pedagogy							-0.04	-0.12
							(0.04)	(0.10)
Modern Teaching							-0.03	0.04
							(0.03)	(0.07)
Extrinsic Motivation							0.00	0.06
							(0.03)	(0.08)
Warmth							-0.00	-0.04
							(0.03)	(0.07)
Teacher Motivation:								
Competence							0.01	0.10
							(0.04)	(0.08)
Professional Satisfaction							0.04	0.09
							(0.03)	(0.06)
N	4417	670	4401	663	4401	663	4373	659
P-Value of Joint Significance	0.443	0.318	0.278	0.914	0.186	0.833	0.390	0.038
R-Squared	0.326	0.252	0.326	0.258	0.328	0.259	0.331	0.269

Reported results are from OLS estimation. The dependent variable is the student's standardized cognitive ability (Raven) score. All regressions include classroom size, grade level, province and school fixed effects. P-values of the joint significance tests are obtained by setting the coefficient estimates on teacher variables to zero. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 4: Allocation of Refugee Students Across Teachers

	(1)	(2)	(3)	(4)
Teacher Ethnic Bias	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Teacher Demographics:				
Male Teacher		-0.01	-0.01	-0.01
		(0.01)	(0.01)	(0.01)
Teacher Age		0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)
Years of Experience		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
Married Teacher		-0.01	-0.01	-0.01
		(0.02)	(0.02)	(0.02)
Number of Children		-0.01	-0.01	-0.01
		(0.01)	(0.01)	(0.01)
Tenured Teacher		0.02	0.02	0.02
		(0.06)	(0.06)	(0.06)
Teacher Cognition:				
Teacher Raven Score			0.00	0.00
			(0.01)	(0.01)
Teacher Cognitive Empathy Score			-0.00	-0.00
			(0.01)	(0.01)
Teaching Styles:				
Growth Mindset				-0.00
				(0.01)
Inquiry-based Pedagogy				0.01
				(0.01)
Modern Teaching				-0.01
				(0.01)
Extrinsic Motivation				-0.00
				(0.01)
Warmth				0.01
				(0.01)
Teacher Motivation:				
Competence				0.00
				(0.01)
Professional Satisfaction				0.00
				(0.01)
N	192	191	191	190
P-Value of Joint Significance	0.832	0.916	0.924	0.980
R-Squared	0.810	0.812	0.813	0.819

Reported results are from OLS estimation. The dependent variable is the proportion of refugee students in the classroom. All four specifications include class size, grade level, province and school fixed effects. P-values of the joint significance tests are obtained by setting the presented coefficient estimates to zero. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 5: What Predicts Teacher's Ethnic Bias?

	(1)	(2)	(3)	(4)
Proportion of Refugees	0.39	0.60	0.89	0.11
	(1.78)	(1.90)	(1.69)	(1.69)
Class Size	-0.02	-0.03	-0.01	-0.03
	(0.05)	(0.05)	(0.05)	(0.05)
Teacher Demographics:				
Male Teacher		-0.06	0.03	0.18
		(0.20)	(0.20)	(0.23)
Teacher Age		0.05	0.05	0.04
		(0.05)	(0.04)	(0.04)
Years of Experience		-0.03	-0.03	-0.00
		(0.06)	(0.05)	(0.05)
Married Teacher		0.11	0.19	0.15
		(0.32)	(0.33)	(0.37)
Number of Children		0.15	0.15	0.21
		(0.18)	(0.15)	(0.15)
Tenured Teacher		0.44	0.33	0.17
		(0.67)	(0.64)	(0.69)
Teacher Cognition:				
Teacher Raven Score			-0.28**	-0.29*
			(0.13)	(0.15)
Teacher Cognitive Empathy Score			0.19	0.16
			(0.15)	(0.15)
Teaching Styles:				
Growth Mindset				0.05
				(0.12)
Inquiry-based Pedagogy				0.17
				(0.13)
Modern Teaching				0.02
				(0.16)
Extrinsic Motivation				0.14
				(0.17)
Warmth				-0.01
				(0.15)
Teacher Motivation:				
Competence				0.18
				(0.17)
Professional Satisfaction				0.02
				(0.15)
N	192	191	191	190
R-Squared	0.407	0.450	0.502	0.539

Reported results are from OLS estimation. The dependent variable is teacher's ethnic bias. All regressions include province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 6: Teacher's Ethnic Bias and Social Exclusion: Friendship Ties

	In-Degree		Out-Degree	
	Host	Refugee	Host	Refugee
Teacher Ethnic Bias	0.06*	-0.24***	0.03	0.03
	(0.04)	(0.08)	(0.02)	(0.07)
N	4373	659	4373	659
P-Value (Host=Refugee)	0.001		0.961	

Reported results are from OLS estimation. The dependent variable in Column 1 and 2 is the number of friendship nominations a student receives from his/her classmates, i.e, in-degree centrality. The dependent variable in Column 3 and 4 is the number of friendship nominations a student makes, i.e, out-degree centrality. Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 7: Teacher's Ethnic Bias and Social Exclusion: Support Ties

	E.S. Received		A.S. Received		E.S. Provided		A.S. Provided	
	Host	Refugee	Host	Refugee	Host	Refugee	Host	Refugee
Teacher Ethnic Bias	-0.02	-0.22***	-0.05	-0.15*	-0.02	-0.29***	-0.03	-0.23***
	(0.04)	(0.06)	(0.05)	(0.08)	(0.05)	(0.07)	(0.05)	(0.08)
N	4373	659	4373	659	4373	659	4373	659
P-Value (Host=Refugee)	0.004		0.194		0.001		0.008	

Reported results are from OLS estimation. Dependent variables are, Column 1 and 2: the total number of classmates a student claims to receive emotional support from, plus the number of such nominations she receives from her classmates. Column 2 and 3: same as column 1 and 2 for academic support. Column 5 and 6: the total number of classmates a student claims to provide emotional support to, plus the number of such nominations she obtains from her classmates. Column 7 and 8: same as column 5 and 6 for academic support. Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 8: Teacher's Ethnic Bias and Nominations Involving Refugee Classmates

	Friendship	Emotional Support	Academic Support
Teacher Ethnic Bias	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.00)
Observations	5032	5032	5032

Reported results are from OLS estimation. Dependent variables are the number of refugee classmates a student nominates as friends (Column 1), the number of refugee classmates to whom a student claims to provide emotional support (Column 2) and the number of refugee classmates to whom a student claims to provide academic support (Column 3). Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 9: Teacher's Ethnic Bias and Inter-Ethnic Ties

	Host Nominations		Refugee Nominations	
	Host → Refugee	Host → Host	Refugee → Refugee	Refugee → Host
Panel 1: Friendship Ties				
Teacher Ethnic Bias	-0.02**	0.05**	-0.07	0.10
	(0.01)	(0.03)	(0.06)	(0.07)
Panel 2: Emotional Support				
Ties Teacher Ethnic Bias	-0.01	-0.02	-0.04*	-0.10*
	(0.01)	(0.03)	(0.02)	(0.06)
Panel 3: Academic Support Ties				
Teacher Ethnic Bias	-0.01**	-0.01	-0.05**	0.01
	(0.01)	(0.03)	(0.02)	(0.05)
Observations	4373	4373	659	659

Reported results are from OLS estimation. Dependent variable is the number of friendship nominations (Panel 1), the number of classmates to whom a host student claims to provide emotional support and a refugee student claims to receive emotional support (Panel 2), the number of classmates to whom a host student claims to provide academic support and a refugee student claims to receive academic support (Panel 3). Column 1 (from host to host) and Column 2 (from host to refugee), Column 3 (refugee to refugee) and Column 4 (refugee to host). Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 10: Teacher's Ethnic Bias and Ethnic Segregation in the Classroom: Coleman's Homophily Index

	Frienship		Emotional Support		Academic Support	
	Host	Refugee	Host	Refugee	Host	Refugee
Teacher Ethnic Bias	0.09**	-0.01	0.06*	0.05	0.09***	-0.04
	(0.04)	(0.11)	(0.03)	(0.09)	(0.03)	(0.11)
N	190	171	187	159	190	151
P-Value (Host=Refugee)	0.217		0.851		0.057	

Reported results are from OLS estimation. Dependent variables are Coleman's Homophily Index constructed for three social tie categories: friendship, emotional support, academic support. The index is not defined for classrooms with only one refugee student at the time of the elicitation. Regression controls include the proportion of refugees, class size, teacher characteristics (demographics, qualifications, cognitive ability, cognitive empathy, styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table 11: Teacher's Ethnic Bias and Bullying

	Probability Bullied		Number of Bullies	
	Host	Refugee	Host	Refugee
Teacher Ethnic Bias	-0.00	0.05**	0.10	0.46**
	(0.01)	(0.03)	(0.12)	(0.22)
N	4051	322	4276	633
P-Value (Host=Refugee)	0.070		0.088	

The first two columns present the average marginal effects from probit regressions. The dependent variable is a binary indicator that takes the value 1 if the student reports to be bullied regularly by his/her peers and zero otherwise. The next two columns present OLS results where the dependent variable is the number of bullies a student reports. Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

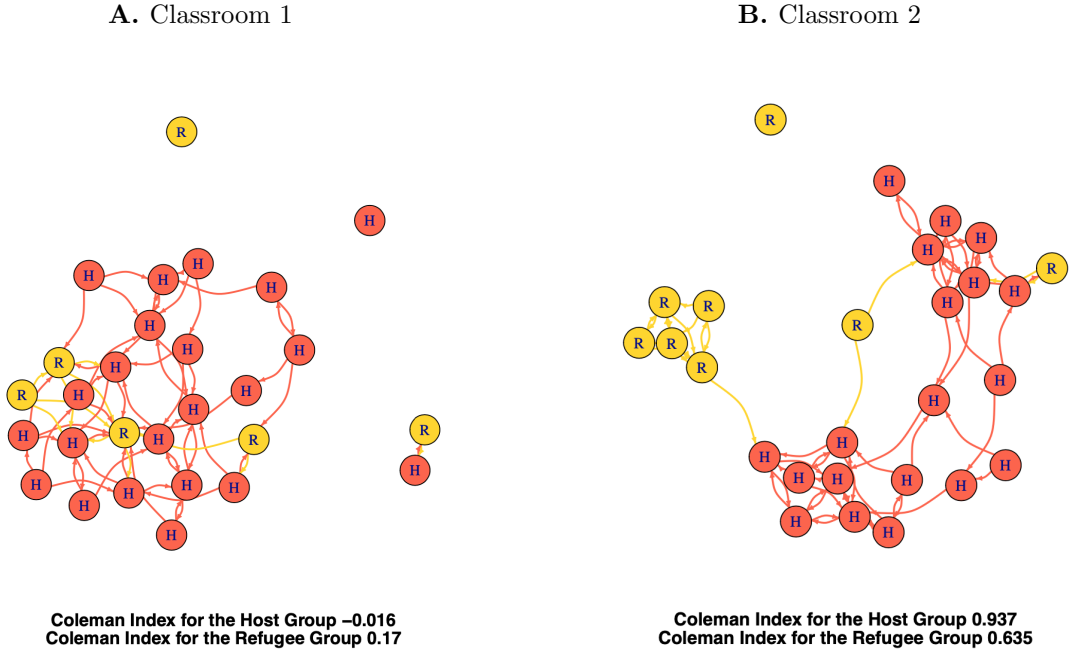
Table 12: Teacher's Ethnic Bias and Student Achievement

	Turkish		Mathematics	
	Host	Refugee	Host	Refugee
Teacher Ethnic Bias	-0.00	-0.15**	-0.01	-0.07
	(0.02)	(0.06)	(0.02)	(0.06)
N	4316	651	4323	650
P-Value (Host=Refugee)	0.013		0.272	

Reported results are from OLS estimation. Dependent variables are standardized scores obtained from the math and Turkish tests implemented in classrooms. Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. Standard errors clustered at school level. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

7 Figures

Figure 1: Coleman's Homophily Index for Two Classrooms



Each circle denotes a node (student). Letters H and R refer to host and refugee child, respectively. Nominations are shown with directional edges between nodes. Coleman Homophily Index is calculated separately for host and refugee group as described in 3.1.1, with higher numbers indicating higher segregation.

Figure 2: Distribution of Teacher's Ethnic Bias

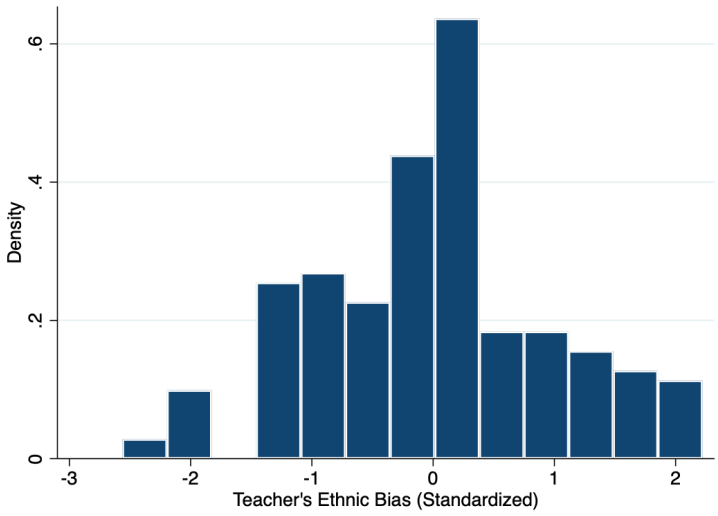


Figure 3: Distribution of Teacher's Cognitive Ability and Cognitive Empathy

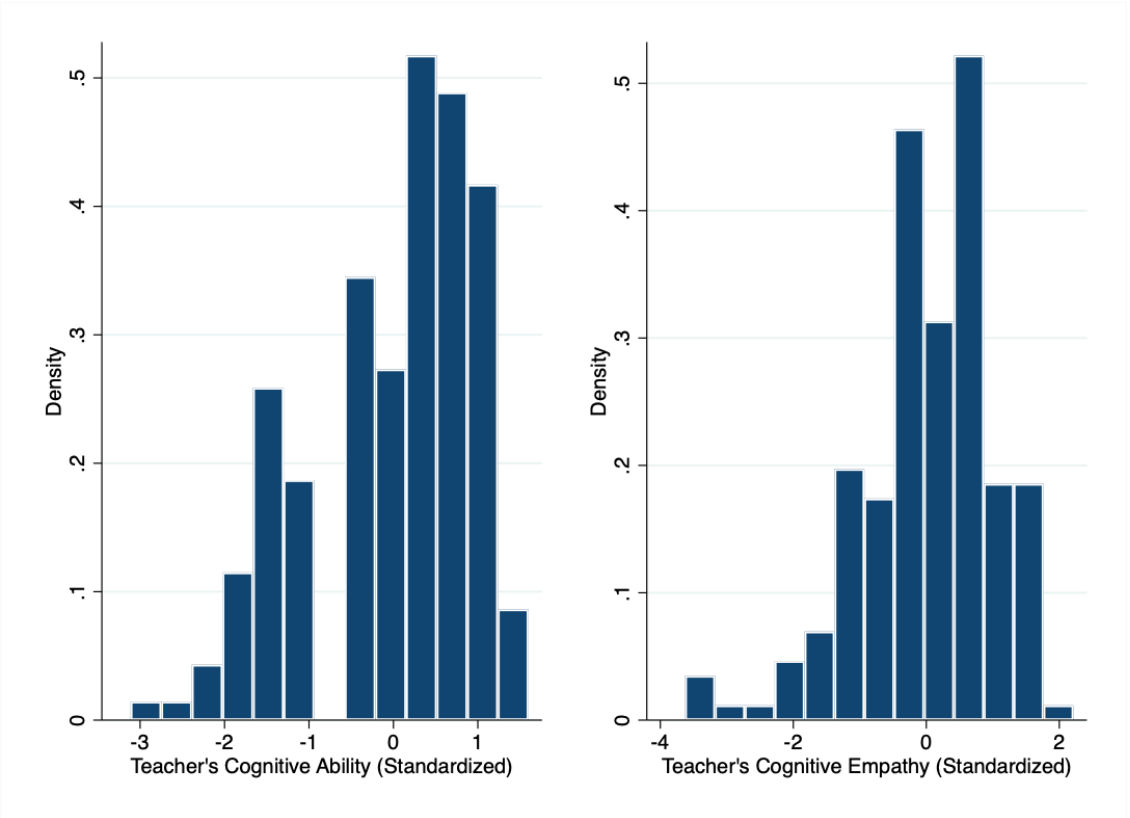
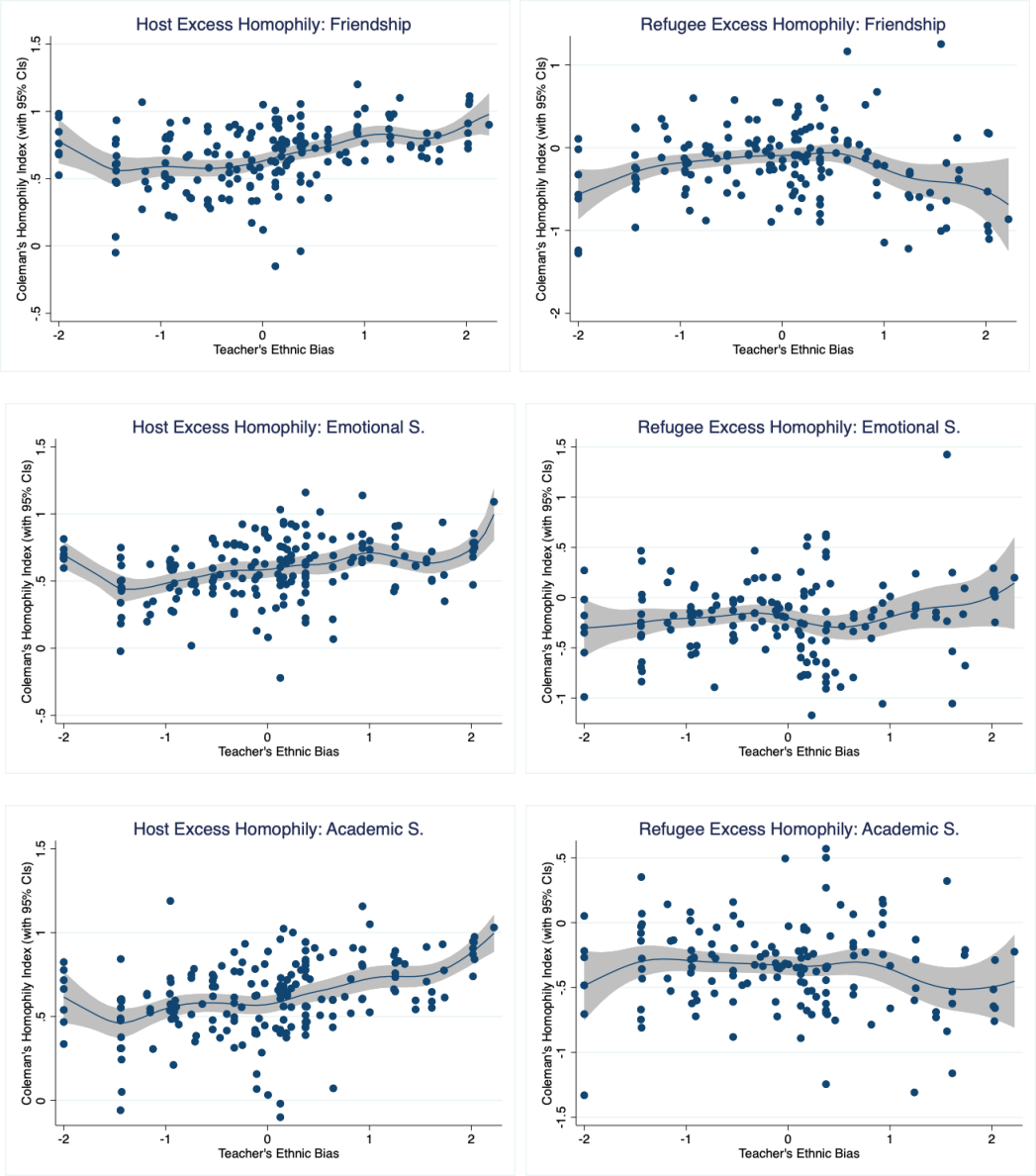
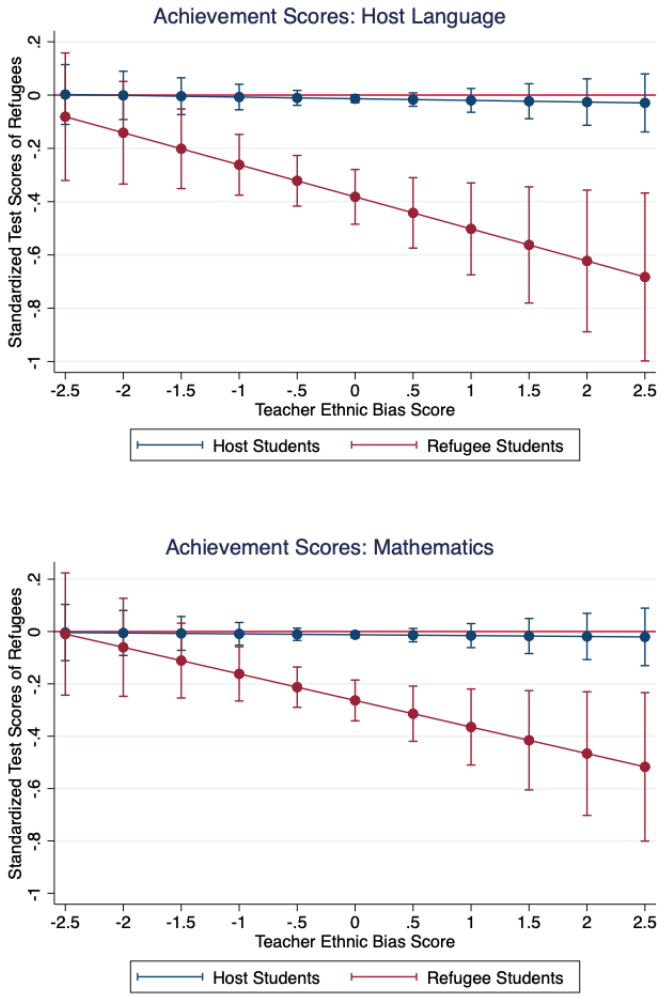


Figure 4: Teacher's Ethnic Bias and Excess Host Homophily



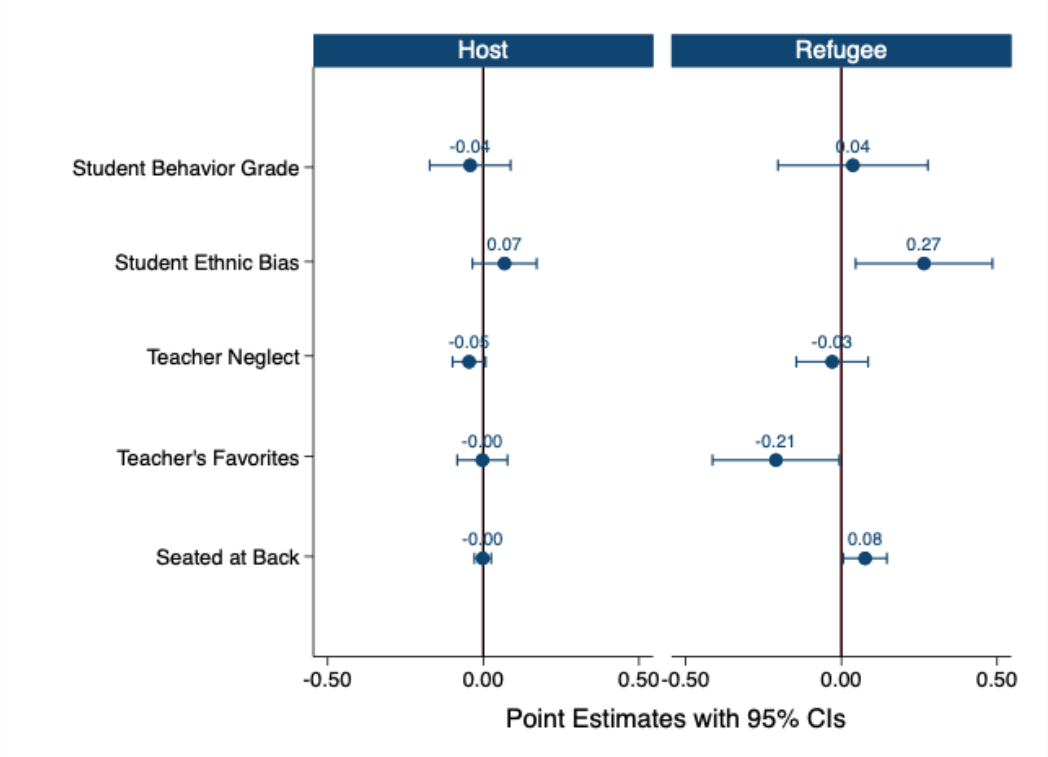
Figures plot the non-parametric estimates (and 95% confidence bands) of the effect of teacher's ethnic bias on excess homophily based on friendship ties (top panel), emotional support ties (mid panel) and academic support ties (bottom panel). All student characteristics (student gender, age, cognitive ability, cognitive empathy), classroom characteristics (class size and proportion of refugees), and teacher characteristics (demographics, qualifications, teaching styles, and motivation) enter the model linearly. Specifications include grade level, province and school fixed effects. Shaded areas contain 95% confidence intervals. Note: We dropped one outlier school to achieve convergence for refugee homophily estimations.

Figure 5: Teacher's Ethnic Bias and Achievement



Figures plot the linear predictions from the OLS regressions of standardized tests scores. Panel 1 provides predicted margins for Turkish tests scores, Panel 2 for math test scores. Both regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. 95% confidence intervals are based on standard errors clustered at the school level. The horizontal line indicates an effect of zero.

Figure 6: Potential Mechanisms



The figure depicts the estimated effects of teacher's ethnic bias on behavioral conduct grades, student ethnic bias, student reported teacher neglect (all standardized), the number of nominations received as teacher's favorite, and the probability of sitting at a back desk. Regressions control for student gender, age, cognitive ability, cognitive empathy, classroom characteristics (class size and proportion of refugees), teacher characteristics (demographics, qualifications, teaching styles, and motivation), grade level, province and school fixed effects. 95% confidence intervals are based on standard errors clustered at the school level. The vertical line indicates an effect of zero.

Appendix

A Correction of Multiple Hypotheses Testing: Romano-Wolf P-values

Table A1: Original and Romano-Wolf p-values (Host & Refugee Sample)

	Host		Refugee	
	Original	Romano Wolf	Original	Romano Wolf
Friendship Ties (In-Degree)	0.078	0.762	0.005	0.066
Friendship Ties (Out-Degree)	0.207	0.735	0.684	0.667
Emotional Support Ties (Received)	0.721	0.978	0.001	0.098
Academic Support Ties (Received)	0.289	0.784	0.064	0.335
Emotional Support Ties (Provided)	0.702	0.976	0.000	0.026
Academic Support Ties (Provided)	0.541	0.936	0.003	0.066
Verbal Score	0.854	0.978	0.023	0.066
Math Score	0.749	0.978	0.265	0.403
Number of Bullies Reported	0.401	0.934	0.035	0.351

The table provides p-values corrected for multiple hypotheses testing using Romano-Wolf algorithm. Number of replications is set to 500.

B Survey Inventories

Table A2: Student Survey Inventories

4-point likert scale: completely agree, agree, disagree, completely disagree	
Inventory	**Exemplary Items**
Teacher Neglect	My teacher does not care whether something is bothering me.
Ethnic Bias	I do not want to be friends with children who come from another country.
	Children from other countries are not as smart as us.
	I like children who come from other countries as much as I like my friends here.

Table A3: Teacher Survey Inventories

4-point likert scale: completely agree, agree, disagree, completely disagree	
Inventory	**Exemplary Items**
Teaching Styles	Punishment is necessary to create a disciplined class. (Extrinsic Motivation)
	Teachers should be serious and authoritative in their relationships with students. (Warmth)
	Your intelligence is something that you can't change very much. (Growth Mindset)
	It does not matter if there is noise in the classroom as long as the students are busy with something productive. (Modern Teaching)
	I encourage my students to do research on topics they are interested in and discuss these topics with me. (Inquiry-based Pedagogy)
Professional Satisfaction	I am very pleased to have chosen teaching as a profession.
Competence	It is difficult for me to communicate effectively with students.

C Social Networks

C.1 Network Elicitation Templates

Table A4

	1	2	3
My best friends in the class			
My teacher's favorite students			

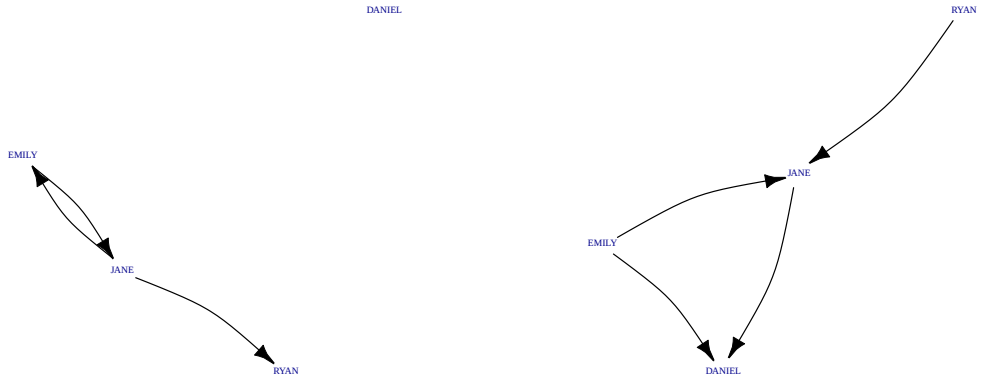
	1	2	3
Classmates whom I academically support			
Classmates whom I emotionally support			

	1	2	3
Classmates who support me emotionally			
Classmates who support me academically			

C.2 Network Measure Construction Guideline

Figure A1

A. I emotionally support ... **B.** I get emotional support from ...



Suppose that there is a classroom with four students, Daniel, Emily, Jane and Ryan. Their answers to i) classmates whom I emotionally support and ii) classmates who support me emotionally, are visualized in panels A and B of Figure A1, respectively. We now construct our network measures, "Emotional Support Provided" and "Emotional Support Received", for Emily to set an example and the rest is calculated in a similar manner for the other students.[16]

- **Emotional Support Provided:** Emily states that she supports Jane (see Panel A). However, none of her classmates indicates that Emily is supportive (see Panel B).

[16]The cases of "academic support provided and academic support received" are analogous to emotional support provided and received.

Therefore, our measure of emotional support provided for Emily is 1 as she appears to support Jane (based on her self-report).

- **Emotional Support Received:** Jane state that she supports Emily (see Panel A). Also, Emily states that she gets emotional support from Jane and Daniel (see Panel B). Then, emotional support received for Emily is 2 as she appears to get support from Daniel and Jane.

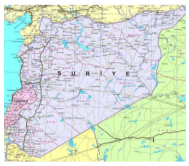
	Emotional Support Provided	Emotional Support Received
Daniel	1	0
Emily	1	2
Jane	2	2
Ryan	0	1

D Implicit Association Test

A. Which of the following statements are good and which are bad?

	GOOD		BAD
1	☐	SAD	☐
2	☐	PEACE	☐
3	☐	WAR	☐
4	☐	FRIEND	☐
5	☐	FREEDOM	☐
6	☐	LOVE	☐
7	☐	DISEASE	☐
8	☐	ENEMY	☐
9	☐	HEALTH	☐
10	☐	HAPPY	☐
11	☐	FIGHT	☐
12	☐	HATE	☐
13	☐	FRIEND	☐
14	☐	DISEASE	☐

B. Which of the statements below are relevant to Turkey and which are relevant to Syria?

	TURKEY		SYRIA
1	☐	ALEPPO	☐
2	☐		☐
3	☐	ANKARA	☐
4	☐	DABKE	☐
5	☐	TURKISH	☐
6	☐	ISTANBUL	☐
7	☐	ARABIC	☐

C. Which of the statements below belong to "Good or Turkey" group and which belong to "Bad or Syria" group?

	GOOD/TURKEY		BAD/SYRIA
1	☐	ALEPPO	☐
2	☐	HAPPY	☐
3	☐	DISEASE	☐
4	☐	FREEDOM	☐
5	☐		☐
6	☐	FIGHT	☐
7	☐		☐

D. Which of the statements below belong to "Bad or Turkey" group and which belong to "Good or Syria" group?

	GOOD/SYRIA		BAD/TURKEY
1	☐	FREEDOM	☐
2	☐	DABKE	☐
3	☐		☐
4	☐	ALEPPO	☐
5	☐	HAPPY	☐
6	☐	DISEASE	☐
7	☐	FRIEND	☐
8	☐		☐