I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14039

The Effect of Refugees on Native Adolescents' Test Scores: Quasi-Experimental Evidence from Pisa

Semih Tumen

JANUARY 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14039

The Effect of Refugees on Native Adolescents' Test Scores: Quasi-Experimental Evidence from Pisa

Semih Tumen
TED University, IZA and ERF

JANUARY 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effect of Refugees on Native Adolescents' Test Scores: Quasi-Experimental Evidence from Pisa*

Existing evidence suggests that low-skilled refugee influx may increase educational attainment among native adolescents due to reduced opportunities and returns in the lower segment of the labor market. In this paper, I test whether refugee influx can also increase the intensity of human capital accumulation among native adolescents who are enrolled in school. Using the PISA micro data and implementing a quasi-experimental empirical strategy designed to exploit (i) the time variation in regional refugee intensity and (ii) institutional setting in the Turkish public education system, I show that the Math, Science, and Reading scores of Turkish adolescents increased following the Syrian refugee influx. The increase in test scores mostly comes from the lower half of the test score distribution and from native adolescents with lower maternal education. The empirical design embeds a framework where the estimated refugee impact can solely be attributed to the labor market mechanism. In particular, I use the observation that refugee adolescents are enrolled more systematically into the Turkish education system after 2016, which gave me the opportunity to use 2015 and 2018 PISA waves in a way to isolate the the effect of the labor market mechanism from the potentially negating force coming from the education experience mechanism. I conclude that the labor market forces that emerged in the aftermath of the refugee crisis have led native adolescents, who would normally perform worse in school, to take their high school education more seriously.

JEL Classification: I21, I25, I26, J61

Keywords: Syrian refugees, test scores, PISA, labor markets

Corresponding author:
Semih Tumen
TED University
Department of Economics
Ziya Gokalp Cad., No.48
06420 Kolej
Ankara
Turkey
E-mail: semihtumen@gmail.com

* I thank Juan Dolado, Ibrahim Elbadawi, Andrew Foster, Paul Makdissi, Naci Mocan, Caglar Ozden, Pinar Ozel, Serdar Polat, Irene Selwaness, the participants of the UNHCR/WB Conference on Forced Migration in Copenhagen, ERF Workshop in Cairo, 12th Migration and Development Conference in Madrid, and the 5th Annual Istanbul Meeting on Human Capital for useful comments and suggestions. I also appreciate the support from the Economic Research Forum through the "Non-Monetary Dimensions of Inequality and Poverty among the Youth in the ERF Region" project call. The usual disclaimer holds.

1 Introduction

Large refugee waves generated by the Syrian crisis have posed severe challenges on the education systems in major host countries and the educational outcomes of native children/adolescents in those countries. In this paper, I investigate the impact of a large influx of Syrian refugees on the PISA scores—a widely-referenced source of standardized test scores compiled and published by the OECD—of native adolescents in Turkey. The PISA test is implemented on 15-year-old children and measures "... 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges."[1] Age 15 is an important milestone in human development. It is an age children start developing more complex working habits, having more concrete plans for future education and occupation, and taking the responsibility of their own choices. The human capital investment literature widely documents that non-cognitive skills quickly develop during adolescence, which substantially increases labor market awareness, and makes further human capital investment decisions more responsive to labor market developments (Borghans et al., 2008). The Syrian refugee influx substantially increased the supply of young and low-skilled labor in the Turkish labor markets. This labor supply shock may have altered human capital investment behavior of adolescents, especially the ones on the margin of dropping out of school, who are directly exposed to this shock. This paper asks the following questions: How do the PISA test scores of native adolescents in Turkey responded to the Syrian refugee influx? What are the relevant mechanisms? Which groups have been affected the most? What are the implications for short- and long-term policy making?

There is a large empirical literature studying the impact of immigration on the educational outcomes of native youth in host countries. The existing evidence suggests that immigration may affect native children's test scores through two main mechanisms that operate in opposite directions (Hunt, 2017). The first one is the *labor market mechanism* that improves the educational outcomes of natives and provides additional incentives to continue education due to increased competition for available jobs in the lower segment of the labor market. The studies testing the relevance of this mechanism generally document that immigrants tend to crowd native adoles-

[1]For more information, see https://www.oecd.org/pisa/.

cents into education. Papers studying this mechanism—such as Denisova (2003), Smith (2012), McHenry (2015), Jackson (2016), Hunt (2017), and Brunello et al. (2020)[2]—generally focus on the change in natives' school enrollment decisions due to increased competition in the low-skill labor market in response to increased immigration concentration and most of those papers document "crowding in" effects. In a companion work (Tumen, 2018), I show that the high school enrollment rates among Turkish adolescents have increased in response to increased refugee concentration. The main mechanism in that paper is that refugees displace low-skilled natives in the labor market and increased competition for jobs with low skill requirements generates a downward pressure on the returns to those jobs. The punchline is that reduced expected returns to staying low-skilled provides incentives for young individuals to spend more time in school, which generated a notable increase in high school enrollment rates—especially among males with lower parental education, who are more likely to leave school early and start working in "bad" jobs. In this paper, I focus on the relationship between refugee inflows and intensive margin of human capital accumulation. In particular, I ask whether the Syrian refugee influx changed the test scores of native adolescents; and, if yes, in what direction.

The second mechanism is the *educational experience mechanism* that negatively affects the educational outcomes of natives. The main insight behind this mechanism is that native children interact with immigrant children in school and/or classroom environments, and this interaction may have important implications for the quality of education they receive. In particular, immigrant concentration in a region, school, or classroom is shown to be negatively correlated with scholastic achievement of native children in host countries. According to this mechanism, increased refugee concentration may reduce the quality of instruction due to various factors such as lower-quality peer interactions, language barriers, and looser teaching standards having to be implemented by instructors. Studies including Betts (1998), Hoxby (1998), Betts and Lofstrom (2000), Borjas (2007), and Gould et al. (2009) document that immigrants either crowd natives out of education or reduce their test scores due to a combination of factors such as limited command of English, cultural diversities, and within-class negative externalities.[3]

[2] See also Eberhard (2012) and Llull (2017).

[3] See also Betts and Fairlie (2003), Peri and Sparber (2009), Jensen and Rasmussen (2011), Cascio and Lewis (2012), Foster (2012), Brunello and Rocco (2013), Orrenius and Zavodny (2015), Roed and Schone (2016), Frattini and Meschi (2019), Tumen (2019), Bossavie

The two mechanisms may be operating simultaneously and canceling out each other. Therefore, one of the main empirical challenges in this literature is to separately identify the respective roles of those two mechanisms. The empirical design in this paper embeds a framework where the estimated refugee impact can solely be attributed to the labor market mechanism, while the educational experience mechanism is shut down naturally.

Using the PISA micro-level data sets and implementing a quasi-experimental empirical strategy designed to exploit (*i*) the time variation in regional refugee intensity and (*ii*) institutional setting in the Turkish public education system, I find that the Math, Science, and Reading scores of native adolescents have increased notably in Turkey following the Syrian refugee influx. Importantly, the increase in test scores comes almost entirely from the lower half of the test score distribution and also from students with maternal education strictly less than high school. There is also suggestive evidence that the PISA scores of male adolescents increased more than those of females. I argue that the labor market forces that emerged in aftermath of the Syrian refugee crisis have led native adolescents, who would normally perform worse in school, to invest in their human capital more intensively.

The empirical analysis is subject to two main challenges: first, eliminating the potential endogeneity bias due to self-selection of refugees into regions based on their location preferences and, second, separately identifying the potentially concurrent effects coming from the labor market mechanism versus the school experience mechanism. To deal with the potential endogeneities due to the self-selection of refugees into locations, I use the diff-in-diff specification proposed by Ceritoglu et al. (2017) and the IV-diff-in-diff specification developed by Del Carpio and Wagner (2015), both of which are widely used in the literature. To address the second challenge, I exploit the educational setting, which almost entirely excluded Syrian refugees from secondary education until 2016. After 2016, the number of Syrian students enrolled in secondary education institutions in Turkey increased gradually. PISA is a triennial study implemented seven times from 2000 to 2018. Year 2015 represents a setting in which there are close to 3 million refugees in Turkey, but

(2019), and Ransom and Winters (2020). Neymotin (2009), Geay et al. (2013), Ohinata and van Ours (2013), Shih (2017), Assaad et al. (2018), and Figlio and Ozek (2019), on the other hand, report zero or positive impact of increased immigrant concentration within the class/school on natives' educational outcomes.

almost zero refugee students are enrolled in secondary education as of 2015. This means that estimates as of 2015 can be directly attributable to the labor market mechanism, since the school experience mechanism is naturally ruled out. I also show that including the 2018 wave into the analysis weakens the estimates, although the net effect is still positive.

To better understand why the labor market mechanism might be a relevant one, it is critical to analyze the impact of Syrian refugees on host country labor market outcomes—especially in Turkey. There is an emerging literature investigating this issue. The main finding is that refugees in Turkey have, on average, lower skill levels than natives. They do not have easy access to work permit; so, they enter the labor market through informal manual jobs and displace natives informally employed in those jobs—see, e.g., Del Carpio and Wagner (2015), Tumen (2016), Ceritoglu et al. (2017), Aksu et al. (2018), and Altindag et al. (2020).[4] Informally employed refugee workers provide important labor cost advantages to firms and, accordingly, wage increases are lower than expected in the lower segment of the labor market (Balkan and Tumen, 2016). Informal refugee workers employed in manual tasks are complementary to formal native workers employed in more complex tasks (Akgunduz et al., 2018; Akgunduz and Torun, 2018). These results suggest that competition between refugee and native workers for low-skill jobs imposes a downward pressure on employment probabilities and potential wages in the lower segment of the labor market. At the same time, increased availability of formal jobs with higher skill requirements encourages skill acquisition. As a result, the decline in the expected returns to staying low-skilled and the increase in the availability of jobs with high skill requirements may jointly increase the intensity of human capital accumulation among native youth.

This paper directly contributes to the literature investigating the impact of immigration on the educational outcomes of natives. The paper offers three novelties. First, it builds on a unique empirical setup which allows for observing how two counteracting mechanisms—the labor market mechanism and the education experience mechanism—operate against each other over different episodes of the Syrian refugee influx. Second, together with Assaad et al. (2018), it is one of the first papers presenting evidence in a forced migration setting, while most of the earlier papers focus

[4] See Tumen (2015) for a summary of the main insights in this literature.

on voluntary migration settings. Unlike Assaad et al. (2018), who find null effect of Syrian refugees on the educational outcomes of Jordanian adolescents, I find statistically significant effects. Fallah et al. (2019) find that there is no statistically significant impact of Syrian refugees on the labor market outcomes of natives in Jordan. Given that the labor market impact of refugees is ignorable in Jordan, a null education effect on Jordanian adolescents does not contradict with the results of the current paper. Finally, this is the first paper exploiting the sub-regional sampling feature of the PISA test to study the educational impact of Syrian refugees. The standardized nature of the PISA test scores allows for a healthy basis for econometric analysis.

The plan of the paper is as follows. Section 2 provides the details of the institutional setting. Section 3 describes the data set used in the empirical analysis and the empirical methods used throughout the paper. Section 4 presents the results, shows the heterogeneous effects among different groups, and discusses policy implications. Section 5 concludes.

2 Institutional details

After the time of first refugee entry into Turkey (January 2012), the number of Syrian refugees in the country has increased steadily over time. The official figures suggest that the number of registered Syrian refugees in Turkey is approximately 3.6 million as of the end of 2020—see Figure (1). Around 50% of all refugees in Turkey are of age 17 and below—i.e., they are at school age. This suggests that public provision of education services is key to facilitate refugee integration. The education services provided to refugee children in Turkey can be described over three main episodes: (1) from January 2012 to mid-2014, (2) from mid-2014 to the beginning of 2016, and (3) from 2016 to date.

The number of Syrian refugees was rather small—around 800,000—and manageable during the initial episode. Both the authorities and the refugees were expecting that political turmoil in Syria would be a temporary one and the crisis would end soon. The Turkish government built a number of refugee camps nearby the Syrian border. The main purpose of the refugee camps was to provide emergency humanitarian assistance—such as food, clothing, shelter, health, security, and counseling services—to refugees on a temporary basis. Only a small fraction of refugees were living out of camps, but they were also settled in the immediate vicinity of the camps to benefit from the services provided in the camps. Some very basic education services were provided by the local NGOs—at primary school level—in small groups within the camps under the supervision of the Ministry of National Education, but the overall scale of those efforts was rather small and project specific. The instruction was provided in Arabic by volunteers as refugee parents did not want their kids to receive school instruction in Turkish due to assimilation concerns. There was no systematic attempt to integrate refugee children into the education system in Turkey during this episode.

After mid-2014, the intensity of conflict and violence in Syria increased substantially. The number of refugees was more than tripled between mid-2014 and the beginning of 2016. It became clear that there would be no early resolution to the crisis. The camps, which were already overcrowded, could not accommodate the new refugee waves. Refugees spread across the country and camps became dysfunctional. An urgent need to provide education services to refugee children had emerged. As a response to this need, the Ministry of National Education established Temporary Education Centers (TECs), which served as a second-shift education in regular public school buildings. Instruction was provided by Syrian teachers/volunteers in Arabic and based on a curriculum consistent with the education system in Syria. The teaching staff received intensive training.

It should be highlighted at this stage that until the end of the second episode, there were no Syrian students of age 13 or above in Turkish public schools—apart from a few specific cases. This suggests that, at the time the PISA 2015 test was administered, there was no systematic intra-school contact between native and refugee students among adolescents; so, the education experience mechanism discussed in Section 1 is automatically ruled out. In other words, any impact of the Syrian refugee influx on the test scores of native adolescents—as it is measured by the PISA 2015 test—can be solely attributed to the labor market mechanism. The identification strategy I explain in the next section is motivated by this observation.

The education services provided in TECs were subject to criticism such as (i) problems about the recognition of degrees and certification, (*ii*) lack of sustainability of the dual system, (*iii*) issues about the employee rights and benefits of Syrian teachers (due to their residency status in Turkey), and (*iv*) physical capacity constraints. Aside from those issues, the major problem with the TECs was the lack of contact between the refugee and native communities. In other words, the TEC system were not conducive of integration of refugee children into the Turkish society.

By 2016, building on the deficiencies of the TECs, "full integration" became a policy priority. In line with this priority, the education policy shifted again and the entire setup is redesigned so as to fully integrate Syrian children to the Turkish public education system. The integration program contained three major elements: language training, integration into the Turkish education system, and social integration. All those three elements required substantial investment into educational resources. In line with the principle of international burden sharing, the EU Facility for Refugees in Turkey (FRIT) provided financial support in two tranches to help the required investments be undertaken. Other international organizations also provided project-specific financing. Various programs have been implemented to facilitate integration along several dimensions through a specific project—Project on Promoting Integration of Syrian Kids into the Turkish Education (PIKTES).[5]

Basic figures obtained from the Ministry of National Education suggest that the enrollment of Syrian children and adolescents into public education institutions in Turkey has increased after 2016. As of the 2019/20 academic year, the enrollment rates of refugees with respect to different age groups are as follows: 34 percent of 95,094 kids at age 5, 95.5 percent of 382,748 kids of age 6-9, 58 percent of 300,458 kids of age 10-13, and 27 percent of 269,239 adolescents of age 14-17. Based on this figure, there are approximately 1,050,000 Syrian refugees of school age (5-17) in Turkey as of the 2019/20 academic year and the average enrollment rate is around 62 percent. The age group relevant for this study is 14-17. Clearly, the enrollment rates are lower for that age group as they are the ones who have more incentives to enter the labor market (mostly for males) and engage in early marriages (mainly for females). Nevertheless, the PISA 2018 test features an environment in which both the labor market and education experience mechanisms jointly operate

[5]For more details, see https://piktes.gov.tr/Home/IndexENG.

and affect the test scores of native adolescents. The empirical setup described in the next section uses the PISA 2015 and PISA 2018 tests both separately and jointly to test the relevance of the two channels.

3 Empirical setting

3.1 Data description

I use the Programme for International Student Assessment (PISA) data compiled and published by the OECD.[6] PISA is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students. Micro-level data from the 2000, 2003, 2006, 2009, 2012, 2015, and 2018 waves are publicly available from the OECD website. Since the beginning of the project, the PISA tests have involved more than 90 countries and around 3,000,000 students—who have taken the internationally agreed two-hour test—globally. Students have been assessed in science, mathematics, reading, collaborative problem solving, and financial literacy.

The PISA data set includes regional variation in test scores for Turkish youth. It also includes information on gender, mother's and father's education, grade level, month of birth, and other characteristics. The standardized nature of the PISA test scores provides a strong basis for comparison across students exposed and non-exposed to refugees. I focus on four waves of the PISA micro data: 2009, 2012, 2015, and 2018. The years before 2009 does not include a regional classification comparable to that for the post-2009 tests. The waves 2009 and 2012 represent the years with no refugee intensity and they are set as the pre-influx period.[7] 2015 and 2018 are defined as the post-influx periods. The PISA data set uses NUTS1-level regional categorization for Turkey. To match the change in refugee intensity with this categorization, the Ministry of Interior data is used to construct the refugee-to-population ratios at NUTS1 level. Table (2) presents a set of basic summary statistics for the sample used in the empirical analysis.

[6] See http://www.oecd.org/pisa/aboutpisa/ for more detailed information on the PISA database.

[7] In fact, the refugees started to enter the country in 2012, but the number of refugees as of 2012 was very low to have an impact on the test scores of natives. Removing 2012 from the analysis does not change the results.

3.2 Econometric model and identification

The empirical setup in this paper relies on a diff-in-diff estimation. Following Tumen (2018), I use two versions of the diff-in-diff setup. The first one is a simple before-after comparison of the regions exposed and not exposed to the refugee influx, which is similar to Ceritoglu et al. (2017). The second is an IV-diff-in-diff approach exploiting the variation in refugee concentration over time/across regions and using the weighted distance from the source governorates in Syria to destination provinces in Turkey as an IV—similar to Del Carpio and Wagner (2015). The Syrian refugee inflows started in 2012 and accelerated over time—see Figure (1). Until late 2012, the number of Syrian refugees was almost zero in the entire country. From late 2012 to mid-2014, the refugees were mostly located close to the Turkey-Syria border. After mid-2014, part of the refugees moved toward the western regions of the country. Figure (3) shows the distribution of refugees across NUTS1 regions as of the end of 2015.[8]

The baseline diff-in-diff specification performs a basic before-after comparison across regions with high refugee concentration versus those with almost no refugees in the spirit of Card (1990) and Card and Krueger (1994). The baseline diff-in-diff analysis may suffer from a classical selection bias as refugees started to select into provinces after 2014/15. Building on this observation, the IV specification addresses the potential selection problem due to the endogenous sorting of Syrian refugees into locations in Turkey.

3.2.1 The diff-in-diff model

The first specification is based on the difference-in-differences strategy implemented by Balkan and Tumen (2016), Ceritoglu et al. (2017), and Tumen (2018). The post-influx period is defined by the dummy variable A_{iy} as:

$$A_{iy} = \begin{cases} 1 & \text{if year} > 2012; \\ 0 & \text{if year} \leq 2012, \end{cases}$$

where i and y indexes individuals and years, respectively. The pre-influx years are 2009 and 2012, while the post-influx period is 2015—and both 2015 and 2018, depending on the specification.

[8] Table (1) and Figure (2) provide a detailed description of NUTS1-level regional categorization in Turkey.

Similarly, two groups of regions are defined as treatment and control groups by the dummy variable T_{ir} as:

$$T_{ir} = \begin{cases} 1 & \text{for the treatment group;} \\ 0 & \text{for the control group,} \end{cases}$$

where r indexes regions. There are three main specifications for the treatment and control regions. Figure (4) visually characterizes those specifications. In the first specification, the treatment group consists of region 12, while the control group includes regions 10 and 11. The second specification extends the treatment and control regions as follows: the treatment group consists of regions 6 and 12, while the control group includes regions 8, 9, 10, and 11. Finally, the third specification uses the entire country and defines the treatment and control groups as follows: the treatment group consists of regions 1, 6, and 12, while the control group includes regions 2, 3, 4, 5, 7, 8, 9, 10, and 11.

The choice of regions is not arbitrary. The treatment and control regions in the first specification consist of neighboring regions with similar economic, social, ethnic, cultural, religious, and historical characteristics. They are the least developed regions in the country and they have immediate comparability. Refugee intensity is among the major distinguishing factors between those regions—as Tumen (2016) and Ceritoglu et al. (2017) argue. The difference in refugee intensity can easily be observed from Table (3). The first specification is the narrowest definition of treatment and control regions. The second specification slightly extends the first specification by including region 6—the region with the second highest refugee intensity—into the treatment regions, and regions 8 and 9—regions neighboring the narrowest control regions and also with almost zero refugee intensity—into the control regions. The third and final specification includes the Istanbul region (the region 1), which has the third highest refugee intensity, into the treatment regions, while the rest of the regions are placed into the control regions. As we move from the first to the third specification, the immediate comparability between the treatment and control regions becomes less and less obvious, and self selection starts becoming a more serious issue—which I address in Section 3.3.2 using an appropriately designed IV strategy.

The diff-in-diff regression model can be formally specified as follows:

$$S_{irt} = \beta_0 + \beta_1(T_{ir} \times A_{it}) + \boldsymbol{\beta_3'}\boldsymbol{X}_{irt} + f_r + f_t + \epsilon_{irt}, \tag{1}$$

where S_{irt} is variable characterizing the PISA test score of individual i of age 15 in region r and in year t, $\boldsymbol{X}_{irt}$ is a vector of individual-level characteristics, f_r and f_t are region and year fixed effects, respectively, and ϵ_{irt} is an error term. The coefficient (β_1) of the interaction between T_{ir} and A_{it} gives the causal effect of interest.

The vector of individual-level covariates, $\boldsymbol{X}_{irt}$, include gender, father's education, mother's education, grade fixed effects, regional per capita real GDP (in logs), regional real trade volumes between Syria and Turkey (in logs), and month-of-birth fixed effects. Parental education variables control for the intensity of parental investment in human capital and can also be used as a proxy for unobserved ability. The grade and month-of-birth fixed effects are included to control for the within-cohort maturity level and education level factors. Regional per capita real GDP and regional real trade volumes between Syria and Turkey (sum of exports and imports) are included to control for time-varying economic factors.

3.2.2 The IV-diff-in-diff model

To address the endogenous location choices of refugees, I use the IV specification developed by Del Carpio and Wagner (2015) and later extended by other papers in the literature. This specification exploits the time-region variation in refugee-to-population ratio across Turkey and uses data from the entire Turkey. So, this specification is mainly a diff-in-diff with continuous treatment, where the treatment variable is instrumented to remove potential endogeneities. The main estimating equation can be formulated as follows:

$$S_{irt} = \alpha_0 + \alpha_1 R_{rt} + \alpha_2 \ln(D_{rt}) + \boldsymbol{\alpha_3'}\boldsymbol{X}_{irt} + f_r + f_t + \epsilon_{irt}, \tag{2}$$

where R_{rt} is the region-level refugee-to-population ratio and D_{rt} is the year-specific shortest distance between the most populated province of the region and the nearest border-crossing. The variable characterizing the shortest distance between the most populated province of the region and the nearest border-crossing is defined such that $D_{rt} = 0$ before (and including) 2012 and $D_{rt} = D_r$ after 2012. Following Del Carpio and Wagner (2015), I put the distance variable into the estimating equation in natural logarithms. The motivation comes from the empirical gravity models in the international trade literature. The inclusion of the year-specific distance variable ensures that the estimates are not contaminated by the omission of variables correlated with distance to border and affecting the outcome variable of interest.

To address the potential endogeneity of the refugees' location decisions within Turkey, I follow Del Carpio and Wagner (2015) and Akgunduz et al. (2018) to construct an IV strategy as follows. The variable R_{rt} is potentially correlated with ϵ_{irt} in Equation (2), which can bias the estimates. The reason is that the refugee concentration may be disproportionately high in regions offering better labor-market options and other socio-economic opportunities. In other words, R_{rt} and S_{irt} may be indirectly correlated through an unobserved factor in ϵ_{irt}. To address this concern, the following IV is constructed:

$$\mathrm{IV}_{rt} = N_t \sum_j \pi_j \frac{1}{L_{jr}}, \tag{3}$$

where N_t is the total number of refugees in Turkey in year t, π_j is the fraction of Syrian population living in each Syrian governorate j in the pre-conflict period (I use 2010), and L_{jr} is the shortest travel distance between each Syrian governorate j and the most populated city of each region r in Turkey.[9] One possibility is that the outcomes may be correlated with distance to border as the Syrian crisis directly hits the border regions and its impact diminishes as distance to border goes up. However, I directly control for the distance to nearest border-crossing by including the log of year-specific distance to nearest border crossing, D_{rt}, into the estimating equation. Since there are multiple—exactly 6—border-crossings between Syria and Turkey, it is possible to separate the distance effect from the location choice decision using this IV strategy. There is a single instrument and I use the 2SLS estimator in instrumenting R_{rt} with the distance-based variable/metric IV_{rt} specified in Equation (3).

[9]Google maps is used to calculate the shortest travel distances. There are 14 Syrian governorates and 12 NUTS-level regions in Turkey, which means that the distance is calculated between 168 distinct routes.

4 Results and discussion

In this section, I report the results of the diff-in-diff and IV regressions using alternative specifications. There are three main outcome variables: Math, Science, and Reading test scores obtained from the PISA data base. Following the convention in the literature, the standardized values of the test scores are used in the regressions. The standard errors are clustered at region level in all regressions. Since the number of clusters is low, bootstrapped standard errors are also reported in the tables—which is a common practice in modern applied microeconomics. The main idea is that, when the number of clusters is low, clustering the standard errors may increase the likelihood of a type-2 error (MacKinnon et al., 2017)—i.e., it may reduce the standard errors and, therefore, may lead to non-rejection of a false hypothesis. Since the structure of the PISA data set forces me to use the NUTS1-level regional categorization, which divides Turkey into 12 broadly defined regions, the estimates I present may be subject to this criticism. To address this issue, I also calculate the standard errors using the wild cluster bootstrap approach developed by Roodman et al. (2019) and report the resulting p-values in the tables along with the clustered standard errors.[10] A heterogeneity analysis along with some robustness exercises and further extensions are also performed and reported at the end of this section.

4.1 Main estimates

Table (4) reports the results of the baseline diff-in-diff analysis. The test scores used in the regressions are standardized, so the coefficients are interpreted in terms of standard deviations. In Panel A, the post-influx period is 2015, while the post-influx period also includes 2018 in Panel B. So, the Panel A can be interpreted as the impact of refugees on PISA test scores purely through the labor market mechanism, while the coefficients reported in Panel B also include the negating effect coming from the education experience mechanism. The results from all three diff-in-diff specifications reveal similar patterns. In Panel A, we see that the Math, Science, and Reading scores increased by 0.11-0.19, 0.14-0.18, and 0.14-0.22 standard deviations, respectively, in response to increased refugee concentration. Other than the Science scores, the estimated increase comes almost entirely from males. The estimates tend to be larger for narrowly-defined treatment and

[10] The -boottest- Stata command is used in the calculations.

control groups, while they get smaller as extended treatment and control groups are introduced. In Panel B, again we see some positive and statistically significant coefficients; but, in general, the estimates become smaller and statistical significance is diminished. In particular, statistical significance remains for the males sample only. The diff-in-diff estimates suggest that the labor market channel generated a significant increase in PISA test scores for all three types of tests, and the estimates start to decline as the negating force coming from the education experience channel kicks in.

Table (5) documents the baseline IV estimates.[11] The results of the IV analysis also confirm that the refugee influx has generated an increase in the test scores of native youth in Turkey for all three test types in Panel A. The coefficients for Math, Science, and Reading scores—for all sample—are 0.048, 0.051, and 0.058 standard deviation, respectively. Unlike the diff-in-diff estimates, the magnitudes of the IV estimates have a more natural interpretation. As an example, think about the coefficient estimate for the Math score, which is 0.048. This estimate means that one percentage point increase in refugee-to-population ratio increase the Math scores by 0.048 standard deviation. Thinking that there are roughly 3.6 million Syrian refugees and the population is around 80 million, the total effect is approximately $0.048 \times (3.6/80) \times 100 \approx 0.22$ standard deviations. Table (2) suggests that 1 standard deviation is approximately 22 percent of the average Math score (for year 2015); therefore, the estimated coefficient translates into an approximately 4 percent increase in test scores for the Math test. It should be noted that, similar to the diff-in-diff estimates, the estimates become smaller and tend to lose statistical significance in Panel B.

Tables (6) and (7) present the results obtained from DID and IV regressions, respectively, where the mean of all three test scores are used as the dependent variable rather than subject-specific test scores. The results are almost unchanged. Overall, the baseline results suggest that the intensity of human capital accumulation—as it is measured by the Math, Science, and Reading scores of the PISA test—has increased among 15-year-old natives in response to the increased Syrian refugee

[11] To support IV validity and rule out the possibility that the key controls used in the regressions may be systematically associated with the instrument, I regress the instrument on the time-varying regional variables used in the regressions—namely, log per capita real GDP, log real trade volumes between Syria and Turkey (for exports and imports, separately, and also for total trade volume), and mean values for mother's and father's education—controlling for year and region fixed effects. Table (11) reports the results, which suggest that the instrument is not systematically related to those variables. Standard errors are clustered at region level, which is a less conservative approach than wild-bootstrapped standard errors, and still all coefficients are statistically and economically insignificant.

concentration in Turkey. The Panel A in both diff-in-diff and IV analysis suggest that the increase in test scores as of 2015 can be attributed to the labor market mechanism only. The influx of low-skilled Syrian refugees increased the competition for jobs with low skill requirements in Turkey. The increase in competition for those jobs has reduced the employment opportunities and also the starting wages. Tumen (2018) shows that the increase in low-skilled refugee concentration increased the high school enrollment rates among Turkish native youth. The findings of the current study complements Tumen (2018) as follows: the main finding is that the increase in low-skilled refugee concentration also increased the intensity of human capital accumulation in the intensive margin. Labor market is an alternative for young males in Turkey. Consistent with this observation, the majority of the regression specifications suggests that the increase in test scores mostly come from males. There are some specifications in which females' test scores have also been estimated to increase in a statistically significant way. Those specifications generally include the entire country. The western regions in Turkey also offer employment opportunities for young females; so, in the regressions for the entire country, it is not unexpected to see some increase in females' test scores.[12]

As a general empirical pattern, I also find that the increase in test scores in response to the refugee influx starts to be negated by the activation of the education experience mechanism after 2016. The estimates presented in Panel B of both the diff-in-diff and IV analyses capture this effect as the "post-influx" period also includes 2018 in addition to 2015. Still the coefficients are positive and the labor market mechanism seems to dominate the negative impact coming from the education experience mechanism—mostly for males. It should be noted that the enrollment rates for adolescent refugees to the secondary education institutions in Turkey is still below 30 percent, which suggests that, as the enrollment rates improve with continued investment by the Ministry of National Education, the education experience channel may erode further the test score gains that the labor market mechanism has been generating.

Testing common trends. The empirical approach used throughout the paper is based on versions of diff-in-diff analysis, which means that the underlying assumption is common trends across regions

[12]It should also be noted that the gender-specific results exhibit rather mixed patterns across specifications and, thus, the result that males' test scores increased more than those of females in response to refugee inflows should be interpreted as suggestive rather than conclusive evidence. See Aksu et al. (2018) for further discussion about gender differences in the impact of refugees on natives' outcomes.

in the pre-treatment period. If this assumption does not hold, then the estimates will not be valid. I formally test the common-trends assumption using a simple event-study approach adopted by Autor (2003) and Tumen (2018). To perform this task, I construct four dummy variables for the four PISA waves: 2009, 2012, 2015, and 2018—taking 1 for the corresponding wave and 0 for others. I also create a treatment dummy variable taking 1 for regions 1, 6, and 12, and 0 for the others—i.e., this is the third specification in the baseline diff-in-diff analysis. Those three regions are the ones with the highest refugee concentration—see Table (3).

I regress the standardized value of the test score (separately for Math, Science, and Reading scores) on the interactions between the wave dummies and the treatment dummy, year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, mother's education, father's education, and a gender dummy. Standard errors are clustered at region level. Figure (5) plots the coefficients of the interaction terms along with the 95 percent confidence bands. The analysis suggests that the trends in all three test scores were common prior to the influx and they differentiated after the influx—confirming the validity of the common trends assumption.

One disadvantage of the PISA data set is that there are only two observations in the pre-treatment period (due to lack of consistent regional categorization in earlier waves), which may raise some concerns about the reliability of standard common trends tests. To supplement this analysis, I use auxiliary data from the 2004–2011 waves (8 years) of the Turkish Household Labor Force Survey (LFS) to test for the existence of pre-trends in two related outcomes for native adolescents: high school enrollment and youth employment. The LFS data set—and other publicly available micro-level data sets—does not directly include variables on test scores. However, testing the existence of pre-trends in multiple youth-related key labor market and educational outcomes may provide valuable supportive evidence that there are no statistically significant differences in trends for youth outcomes prior to the refugee influx between treated and control regions. To perform this task, I use the same regions as above and I regress (in two separate regressions) the high-school enrollment and employment of natives of age 15-17 on the interactions between the wave dummies and the treatment dummy, year fixed effects, region fixed effects (NUTS2), regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, age fixed effects, parental education fixed effects, and a gender dummy. I plot the coefficients and 95 percent confidence bands—2004 is the ignored year category—in Table (6), which shows that there are no statistically significant pre-trends in both outcomes for a period covering 8 years prior to the Syrian refugee influx.

Heterogeneous effects. To answer the question "Which groups have been affected the most?," I perform two different heterogeneity analysis. The first one splits the sample based on mother's education. Accordingly, two different sub-samples for native adolescents are formed: one for students with mothers less than high school educated and the other one with mothers of high school education and above. Mother's education is a good indication of the social status of the family and can give an idea whether the student is coming from a stronger versus a weaker parental background. The second heterogeneity analysis splits the sample into two based on the distribution of test scores. Specifically, I focus on the upper and lower halves of the distribution, i.e., above and below the median score. Different from the baseline analysis and without loss of generality, I pool the three test scores in the regression and include test subject dummies to control for test-specific factors. For both heterogeneity analysis, I use the IV-2SLS specification, which is more general than the simple diff-in-diff specifications and has a more natural interpretation for the magnitudes of coefficient estimates.

The purpose of this exercise is to understand whether the refugee influx more heavily affects the test scores of disadvantaged students or not. Lower returns in the lower segment of the labor market should be more binding for the educational decisions of the disadvantaged group. In this part, I formally test whether this conjecture is true or not.

Tables (8) and (9) presents the IV-2SLS results for students with low- versus high-educated mothers, respectively. The findings suggest that, for the three test types, the main effect comes almost entirely from native adolescents with mothers of strictly less than high school education. For the group with higher maternal education, the coefficients are mostly statistically insignificant. Similar to the aggregate results, adding the 2018 wave into the analysis reduces the size and statistical significance of the estimates and this effect is relevant for the disadvantaged group. Again, the effects are more visible on males than females.

Table (10) presents the results for the heterogeneity analysis performed based on the test score distribution. Panel A and B reports the estimates for below-median and above-median samples, respectively. The results suggest that the increase in test scores following the refugee influx comes almost entirely from the bottom half of the test-score distribution, while the estimates for the upper half are statistically insignificant. This result holds for both males and females, while the difference is starker for males than females.

The heterogeneity in the estimates supports the validity of the proposed mechanism. The increase in the low-skilled labor supply following the refugee influx reduced the employment opportunities and wages for low-skilled natives. Tumen (2018) argues that the decline in the labor market opportunities for natives has increased high school enrollment rates among natives. This paper documents that the test scores of natives have also increased following the influx and the increase mostly comes from the lower portion of the skill distribution. The increase in the intensity of human capital investment among low-skilled natives suggests that the refugee influx has provided incentives for educational upgrading. The evidence presented in this part also suggests that the negating effect coming from the education experience mechanism is also more relevant for the disadvantaged groups.

The 2012 national education reform. A nation-wide compulsory education law, which increased mandatory education from eight to twelve years, became operational in Turkey as of September 2012. The law effectively made high school education compulsory and the timing of treatment coincides with the timing of refugee influx. Unlike the refugee influx, which exhibited substantial variation across sub-regions, the compulsory education law covered all regions equally since it was a national reform. However, the impact of the reform might vary across regions and such as a potential regional variation might contaminate the estimates. Using rich nationally representative micro data combined with administrative data on public education system in Turkey and various statistical techniques, Tumen (2018) documents that the compulsory education reform does not affect the causal impact of Syrian refugee inflows on high school enrollment behavior of natives in Turkey. Unfortunately, the structure of PISA data set is not rich enough to allow for similar formal tests of any potential impact of the compulsory education reform on PISA test scores. However, I argue that the compulsory education reform is less likely to plague the causal relationship between refugee influx and PISA scores for three main reasons. First, if it does not contaminate the causal impact of refugees on high school attainment, then there is no clear reason to expect that it may contaminate the causal link in the intensive margin—i.e., the test scores. Second, the increase in school enrollment caused by the reform comes from the ones who are less likely to register and/or finish high school absent the reform. This group is less likely to drive an economically significant jump in "overall average PISA test scores" in response to modest labor market incentives. Finally, I argue that the causal impact of refugees on test scores is likely more prominent for males than females. However, due to socio-cultural reasons, the school enrollment rates and other educational outcomes of adolescent females have been traditionally low in Turkey and, if the reform generates any impact on educational outcomes, it would be much more clearly observed for females rather than males.

Alternative mechanisms. The estimates presented in this paper are discussed within the framework of two main mechanisms that may explain how Syrian refugee influx can influence the test scores of native adolescents: the labor market mechanism and the education experience mechanisms. These mechanisms are clearly defined as the two main mechanisms through which interaction with immigrants may impact the educational outcomes of natives—see, for example, Hunt (2017). However, there can be alternative mechanisms that may be operating in the background. For example, families may have heightened safety concerns after the arrival of refugees especially in areas with high refugee concentration and they may be more strict in following up their kids' school enrollment/attendance. There can also be alternative sub-explanations related to the labor market mechanism. If, for example, students lost their after school informal part-time jobs, this may have also encouraged them to put more effort in school. Although the list of rather minor alternative mechanisms can be extended further, data limitations do not allow for testing each possible explanation in a convincing way. So, this paper follows the guidance provided by the literature and focuses on the two main mechanisms.

5 Concluding remarks

In this paper, I aim to come up with a set of estimates pertaining to the impact of refugees on standardized test scores of native adolescents in Turkey. Differences in test scores proxy differences in human capital development. Therefore, immigration may affect the educational dynamics in a society through its impact on the quantity and quality of early human capital acquisition. The type of immigration and the skill composition of immigrants are important determinants of the nature of this impact. This is among the first papers estimating the impact of Syrian refugees on the standardized test scores of natives in host countries. I also develop an empirical approach that helps me to separate the two mechanisms—the labor market mechanism and the education experience mechanism—that tend to negate each other. In particular, I use the observation that refugee adolescents are enrolled more systematically into the Turkish education system after 2016, which gave me the opportunity to use 2015 and 2018 waves in a way to isolate the the effect of the labor market mechanism from the negating force coming from the education experience mechanism.

I show that the Math, Science, and Reading scores of Turkish native adolescents have notably increased following the Syrian refugee influx—conditional on parental education, which is used as a proxy for unobserved ability. There is suggestive evidence that the increase in PISA scores is more pronounced for males than females. Most importantly, the increase in test scores mostly comes from the lower half of the test score distribution and from adolescents with lower maternal education. This suggests that refugee influx has reduced the test score inequality among natives. The results survive using a variety of alternative specifications and other robustness checks. I argue that the labor market forces that emerged in aftermath of the refugee crisis have led native adolescents, who would normally perform worse in school, to take their high school education more seriously.

I also provide evidence that, as the presence of refugee adolescents increase in the Turkish education system, the education experience mechanism starts eroding the test scores gains that are initially obtained through the labor market mechanism. Currently the enrollment rate of refugee adolescents—age 14-17—is below 30 percent and, as refugee integration proceeds, the enrollment rates will increase further. This suggests that the test score gains are less likely be observed in future cohorts. The Turkish government has intensively invested in steps to improve the quality of the educational integration process. Those efforts aims to minimize the negative effect coming from the education experience mechanism. With better integration of refugees, excess competition in the lower segment of the labor market will decline. In the long term, both the labor market mechanism and the education experience mechanism will likely lose power. Timely policy measures taken by the policy makers—in terms of both labor market and education policies—would improve the quality of refugee integration and smooth out this process.

References

Akgunduz, Y. E., W. Hassink, and M. Van den Berg (2018): "The Impact of the Syrian Refugee Crisis on Firm Entry and Performance in Turkey," *World Bank Economic Review*, 32, 19–40.

Akgunduz, Y. E. and H. Torun (2018): "Two and a Half Million Syrian Refugees, Skill Mix, and Capital Intensity," GLO Discussion Paper #186.

Aksu, E., R. Erzan, and M. Kirdar (2018): "The Impact of Mass Migration of Syrians on the Turkish Labor Market," IZA Discussion Paper #12050.

Altindag, O., O. Bakis, and S. V. Rozo (2020): "Blessing or Burden? The Impact of Refugees on Businesses and the Informal Economy," *Journal of Development Economics*, 146, 102490.

Assaad, R., T. Ginn, and M. Saleh (2018): "Impact of Syrian Refugees in Jordan on Education Outcomes for Jordanian Youth," ERF Working Paper #1214.

Autor, D. (2003): "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics*, 21, 1–42.

Balkan, B. and S. Tumen (2016): "Immigration and Prices: Quasi-Experimental Evidence from Syrian Refugees in Turkey," *Journal of Population Economics*, 29, 657–686.

Betts, J. R. (1998): "Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities?" in *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, ed. by D. Hamermesh and F. Bean, New York, NY: Russell Sage Foundation.

Betts, J. R. and R. W. Fairlie (2003): "Does Immigration Induce 'Native Flight' from Public Schools into Private Schools?" *Journal of Public Economics*, 87, 987–1012.

Betts, J. R. and M. Lofstrom (2000): "The Educational Attainment of Immigrants: Trends and Implications," in *Issues in the Economics of Immigration*, ed. by G. J. Borjas, Chicago, IL: University of Chicago Press, chap. 2, 51–116.

Borghans, L., L. A. Duckworth, J. J. Heckman, and B. ter Weel (2008): "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 43, 972–1059.

Borjas, G. J. (2007): "Do Foreign Students Crowd Out Native Students from Graduate Programs?" in *Science and the University*, ed. by P. Stephan and R. Ehrenberg, Madison, WI: University of Wisconsin Press.

Bossavie, L. (2019): "The Effect of Immigration on Natives' School Performance: Does Length of Stay in the Host Country Matter?" *Journal of Human Resources*, forthcoming.

Brunello, G., E. Lodigiani, and L. Rocco (2020): "Does Low Skilled Immigration Affect the Education of Natives? Evidence from Italian Provinces," *Labour Economics*, 63, 101794.

Brunello, G. and L. Rocco (2013): "The Effect of Immigration on the School Performance of Natives," *Economics of Education Review*, 32, 234–246.

Card, D. (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43, 245–257.

Card, D. and A. B. Krueger (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84, 772–793.

Cascio, E. U. and E. G. Lewis (2012): "Cracks in the Melting Pot: Immigration, School Choice, and Segregation," *American Economic Journal: Economic Policy*, 4, 91–117.

Ceritoglu, E., H. B. Gurcihan Yunculer, H. Torun, and S. Tumen (2017): "The Impact of Syrian Refugees on Natives' Labor Market Outcomes in Turkey: Evidence from a Quasi-Experimental Design," *IZA Journal of Labor Policy*, 6, 1–28.

Del Carpio, X. and M. Wagner (2015): "The Impact of Syrian Refugees on the Turkish Labor Market," Unpublished manuscript, World Bank.

Denisova, A. (2003): "Immigration and the Educational Choices of Native-Born Workers: The Role of Income," in *Essays on Immigration*, Washington, DC: Georgetown University, PhD Dissertation, chap. 1.

EBERHARD, J. (2012): "Immigration, Human Capital, and the Welfare of Natives," Unpublished manuscript, University of Southern California.

FALLAH, B., C. KRAFFT, AND J. WAHBA (2019): "The Impact of Refugees on Employment and Wages in Jordan," *Journal of Development Economics*, 139, 203–216.

FIGLIO, D. AND U. OZEK (2019): "Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students," *Journal of Labor Economics*, forthcoming.

FOSTER, G. (2012): "The Impact of International Students on Measured Learning and Standards in Australian Higher Education," *Economics of Education Review*, 31, 587–600.

FRATTINI, T. AND E. MESCHI (2019): "The effect of immigrant peers in vocational schools," *European Economic Review*, 113, 1–22.

GEAY, C., S. MCNALLY, AND S. TELHAJ (2013): "Non-native Speakers of English in the Classroom: What are the Effects on Pupil Performance?" *Economic Journal*, 123, F281–F307.

GOULD, E. D., V. LAVY, AND D. M. PASERMAN (2009): "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence," *Economic Journal*, 119, 1243–1269.

HOXBY, C. M. (1998): "Do Immigrants Crowd Disadvantaged American Natives Out of Higher Education?" in *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, ed. by D. Hamermesh and F. Bean, New York, NY: Russell Sage Foundation.

HUNT, J. (2017): "The Impact of Immigration on the Educational Attainment of Natives," *Journal of Human Resources*, 52, 1060–1118.

JACKSON, O. (2016): "Does Immigration Crowd Natives into or out of Higher Education?" Unpublished manuscript, Federal Reserve Bank of Boston.

JENSEN, P. AND A. W. RASMUSSEN (2011): "The Effect of Immigrant Concentration in Schools on Native and Immigrant Children's Reading and Math Skills," *Economics of Education Review*, 30, 1503–1515.

Llull, J. (2017): "Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model," *Review of Economic Studies*, 85, 1852–1896.

MacKinnon, J. G., M. O. Nielsen, and M. D. Webb (2017): "Bootstrap and Asymptotic Inference with Multiway Clustering," Unpublished manuscript, Queen's University.

McHenry, P. (2015): "Immigration and the Human Capital of Natives," *Journal of Human Resources*, 50, 34–71.

Mincer, J. (1974): *Schooling, Experience, and Earnings*, New York, NY: Columbia University Press.

Neymotin, F. (2009): "Immigration and Its Effects on the College-Going Outcomes of Natives," *Economics of Education Review*, 28, 538–550.

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

Orrenius, P. M. and M. Zavodny (2015): "Does Immigration Affect Whether US Natives Major in Science and Engineering?" *Journal of Labor Economics*, 33, S79–S108.

Peri, G. and C. Sparber (2009): "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, 1, 135–169.

Ransom, T. and J. V. Winters (2020): "Do Foreigners Crowd Natives out of STEM Degrees and Occupations? Evidence from the US Immigration Act of 1990," *Industrial and Labor Relations Review*, forthcoming.

Roed, M. and P. Schone (2016): "Impact of Immigration on Inhabitants' Educational Investments," *Scandinavian Journal of Economics*, 118, 433–462.

Roodman, D., J. MacKinnon, M. Nielsen, and M. Webb (2019): "Fast and Wild: Bootstrap inference in Stata using boottest," *Stata Journal*, 19, 4–60.

Shih, K. (2017): "Do International Students Crowd-Out or Cross-Subsidize Americans in Higher Education?" *Journal of Public Economics*, 156, 170–184.

SMITH, C. L. (2012): "The Impact of Low-Skilled Immigration on the Youth Labor Market," *Journal of Labor Economics*, 30, 55–89.

TUMEN, S. (2015): "The Use of Natural Experiments in Migration Research," *IZA World of Labor*.

——— (2016): "The Economic Impact of Syrian Refugees on Host Countries: Quasi-Experimental Evidence from Turkey," *American Economic Review*, 106, 456–460.

——— (2018): "The Impact of Low-Skill Refugees on Youth Education," IZA Discussion Paper #11869.

——— (2019): "Refugees and 'Native Flight' from Public to Private Schools," *Economics Letters*, 181, 154–159.

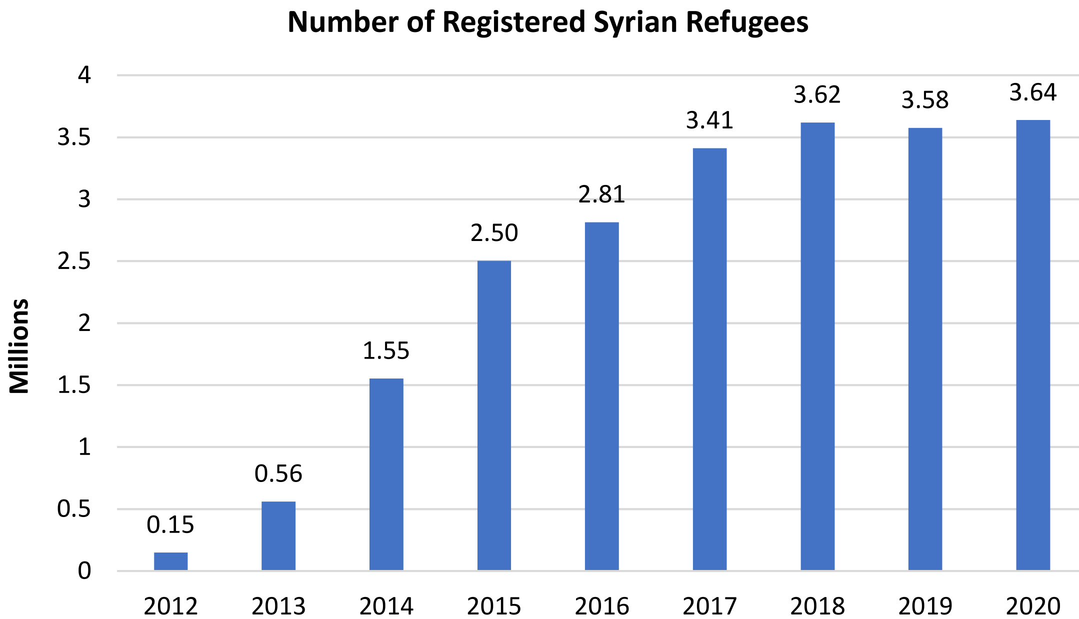


Figure 1: Number of registered Syrian refugees in Turkey. This figure plots the number of registered Syrian refugees in Turkey from 2012 to 2020—as of December 2020. The data sources are the UNHCR and the Government of Turkey. See: https://data2.unhcr.org/en/situations/syria/location/113.

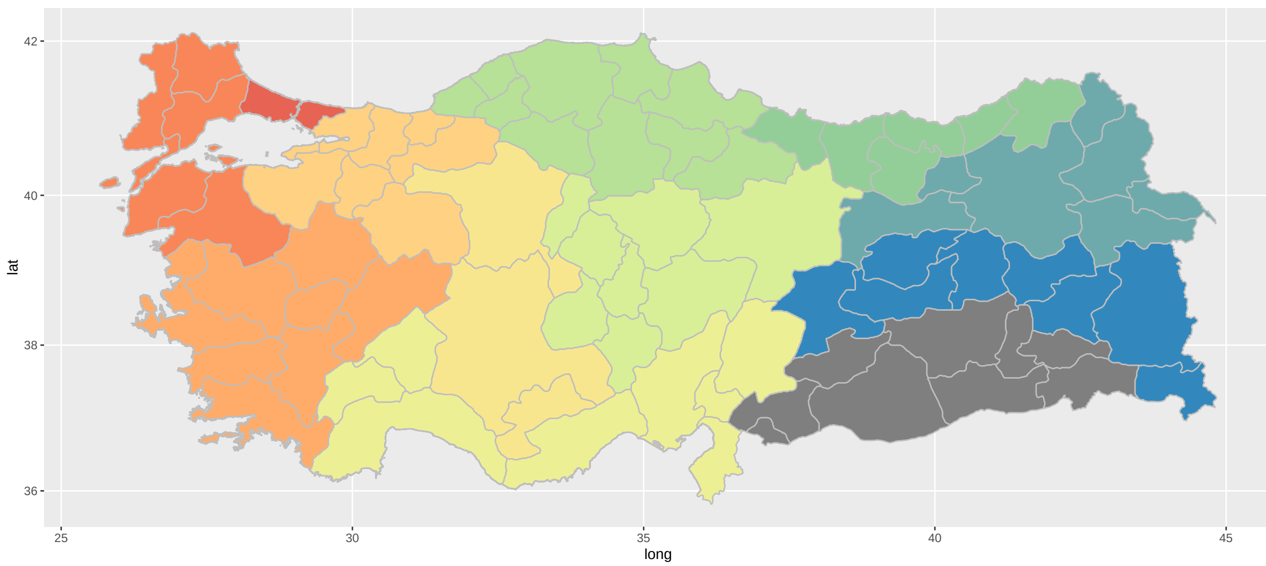


Figure 2: Regional map (Turkey). This figure displays the NUTS1-level regional classification for Turkey. Table (1) below lists the provinces included into each NUTS1 region.

Region #	Provinces
Region 1	Istanbul
Region 2	Tekirdag, Edirne, Kirklareli, Balikesir, Canakkale
Region 3	Izmir, Aydin, Denizli, Mugla, Manisa, Afyonkarahisar, Kutahya, Usak
Region 4	Bursa, Eskisehir, Bilecik, Kocaeli, Sakarya, Bolu, Duzce, Yalova
Region 5	Ankara, Konya, Karaman
Region 6	Antalya, Isparta, Burdur, Adana, Mersin, Hatay, Kahramanmaras, Osmaniye
Region 7	Kirikkale, Aksaray, Nigde, Nevsehir, Kirsehir, Kayseri, Sivas, Yozgat
Region 8	Zonguldak, Karabuk, Bartin, Kastamonu, Cankiri, Sinop, Samsun, Tokat, Corum, Amasya
Region 9	Trabzon, Ordu, Giresun, Rize, Artvin, Gumushane
Region 10	Erzurum, Erzincan, Bayburt, Agri, Kars, Igdir, Ardahan
Region 11	Malatya, Elazig, Bingol, Tunceli, Van, Mus, Bitlis, Hakkari
Region 12	Gaziantep, Adiyaman, Kilis, Sanliurfa, Diyarbakir, Mardin, Batman, Sirnak, Siirt

Table 1: Provinces in NUTS1 regions in Turkey. There are 81 provinces in 12 NUTS1-level regions in Turkey. This table shows the provinces included in each NUTS1 region.

Variable	**2009**	**2012**	**2015**	**2018**
Means				
Male	0.511	0.504	0.502	0.507
Math score	446.51	446.65	415.81	452.61
Reading score	465.71	474.37	425.42	464.59
Science score	455.36	462.29	421.91	467.60
Standard deviations				
Male	0.500	0.500	0.500	0.500
Math score	88.07	87.62	73.63	81.22
Reading score	77.56	81.86	74.54	84.36
Science score	75.09	75.32	73.20	78.03
# of observations	4,996	4,780	5,895	6,890

Table 2: Basic summary statistics. This table presents basic summary statistics for the PISA sample used in the empirical analysis.

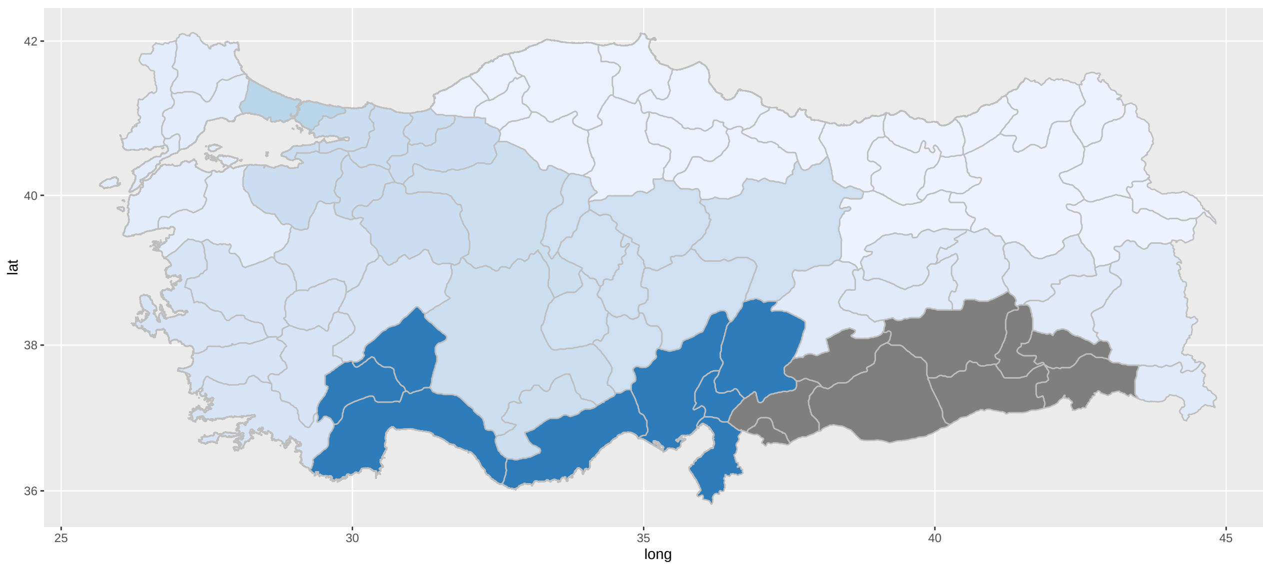


Figure 3: Refugee shares as of 2015. This figure displays the refugee shares in Turkey at NUTS1-level as of the end of 2015. Table (3) below documents the numerical refugee share values in each region—calculated as the ratio of the number of registered Syrian refugees to the native population in each NUTS1 region as of the end of 2015. Darker areas represent the regions with higher refugee-to-population ratios.

Region #	**Refugee share (%)**
Region 1	2.45
Region 2	0.51
Region 3	1.09
Region 4	1.59
Region 5	1.52
Region 6	7.37
Region 7	1.46
Region 8	0.14
Region 9	0.10
Region 10	0.07
Region 11	0.65
Region 12	11.32

Table 3: Refugee shares in NUTS1 regions in Turkey. This table shows the refugee shares in each of the NUTS1-level region as of the end of 2015 in Turkey. The regions with highest refugee concentration are Region #12 (11.32%) represented with black color in Figure (2), Region #6 (7.37%) represented in dark blue color in Figure (2), and Region #1 (2.45%), which is the Istanbul region.

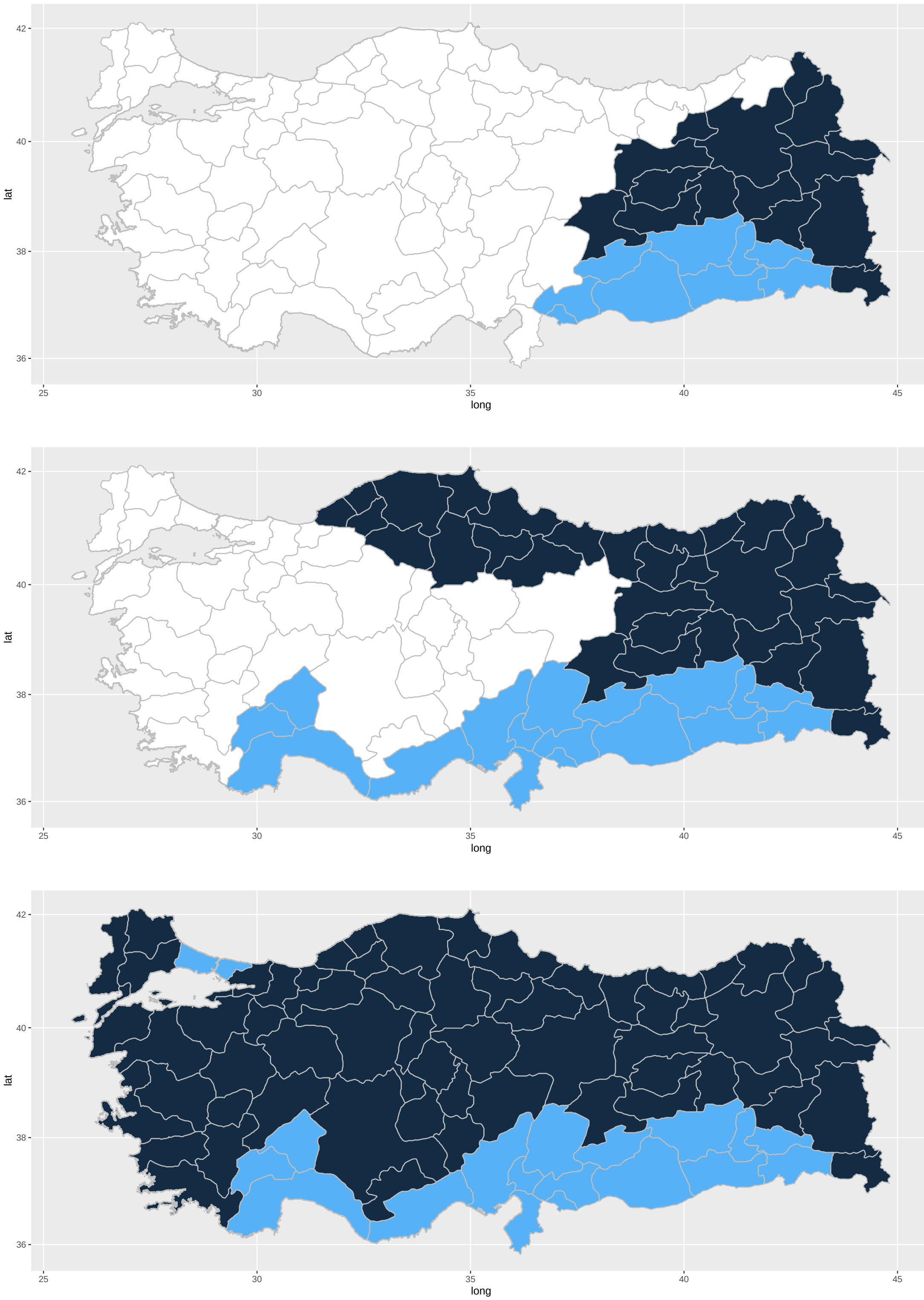


Figure 4: DID specifications. This figure displays the three different regional specifications used in the DID estimations. Dark blue color represents the control regions, while the light blue color represents the treatment regions.

DID ESTIMATION

Dependent variable: Standardized value of the corresponding test score

Panel A	**Period of observation: 2009–2015**								
	Math			**Science**			**Reading**		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
	Specification I								
Refugee effect	0.185*	0.196***	0.149	0.138*	0.042*	0.214**	0.173	0.101	0.208
(clustered s.e.)	(0.054)	(0.013)	(0.101)	(0.038)	(0.011)	(0.059)	(0.123)	(0.082)	(0.145)
(w. b. p-values)	(0.069)	(0.016)	(0.303)	(0.079)	(0.061)	(0.033)	(0.245)	(0.413)	(0.401)
# of obs.	2,580	1,317	1,263	2,580	1,317	1,263	2,580	1,317	1,263
	Specification II								
Refugee effect	0.156	0.203**	0.094	0.183*	0.159*	0.204*	0.222**	0.253**	0.171
(clustered s.e.)	(0.089)	(0.053)	(0.111)	(0.071)	(0.063)	(0.081)	(0.077)	(0.082)	(0.123)
(w. b. p-values)	(0.156)	(0.054)	(0.594)	(0.059)	(0.076)	(0.082)	(0.039)	(0.049)	(0.245)
# of obs.	6,079	3,123	2,956	6,079	3,123	2,956	6,079	3,123	2,956
	Specification III								
Refugee effect	0.111**	0.114*	0.102	0.159*	0.153**	0.168*	0.139**	0.158*	0.113
(clustered s.e.)	(0.042)	(0.056)	(0.082)	(0.069)	(0.064)	(0.083)	(0.059)	(0.069)	(0.082)
(w. b. p-values)	(0.034)	(0.060)	(0.254)	(0.078)	(0.038)	(0.089)	(0.043)	(0.079)	(0.379)
# of obs.	14,876	7,464	7,412	14,876	7,464	7,412	14,876	7,464	7,412
Panel B	**Period of observation: 2009–2018**								
	Specification I								
Refugee effect	0.120	0.129**	0.081	0.107	0.131**	0.071	0.101	0.129*	0.062
(clustered s.e.)	(0.067)	(0.043)	(0.103)	(0.076)	(0.044)	(0.085)	(0.069)	(0.055)	(0.091)
(w. b. p-values)	(0.203)	(0.045)	(0.348)	(0.195)	(0.031)	(0.386)	(0.171)	(0.057)	(0.311)
# of obs.	3,722	1,929	1,793	3,722	1,929	1,793	3,722	1,929	1,793
	Specification II								
Refugee effect	0.122	0.138*	0.083	0.111	0.123*	0.081	0.144*	0.159*	0.104
(clustered s.e.)	(0.086)	(0.061)	(0.113)	(0.077)	(0.057)	(0.093)	(0.066)	(0.074)	(0.085)
(w. b. p-values)	(0.186)	(0.065)	(0.234)	(0.195)	(0.075)	(0.222)	(0.085)	(0.073)	(0.230)
# of obs.	9,098	4,580	4,518	9,098	4,580	4,518	9,098	4,580	4,518
	Specification III								
Refugee effect	0.101	0.121**	0.084	0.105	0.126**	0.075	0.104	0.123*	0.072
(clustered s.e.)	(0.072)	(0.041)	(0.083)	(0.065)	(0.040)	(0.088)	(0.079)	(0.068)	(0.091)
(w. b. p-values)	(0.174)	(0.057)	(0.199)	(0.145)	(0.029)	(0.201)	(0.178)	(0.068)	(0.192)
# of obs.	21,766	10,958	10,808	21,766	10,958	10,808	21,766	10,958	10,808

Table 4: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, mother's education, and father's education are included as control variables into all regressions. A gender dummy is also included in regressions for all sample (i.e., columns 1, 4, and 7).

IV-2SLS ESTIMATION

Dependent variable: Standardized value of the corresponding test score

Panel A	**Period of observation: 2009–2015**								
	Math			Science			Reading		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
OLS	0.032***	0.036***	0.028	0.045***	0.039***	0.047***	0.044***	0.048***	0.034**
(clustered s.e.)	(0.008)	(0.008)	(0.018)	(0.009)	(0.010)	(0.011)	(0.007)	(0.009)	(0.016)
(w. b. p-values)	(0.016)	(0.011)	(0.136)	(0.014)	(0.018)	(0.009)	(0.014)	(0.019)	(0.044)
1st stage	2.171***	2.168***	2.162***	2.171***	2.168***	2.162***	2.171***	2.168***	2.162***
(clustered s.e.)	(0.309)	(0.316)	(0.304)	(0.309)	(0.316)	(0.304)	(0.309)	(0.316)	(0.304)
(w. b. p-values)	(0.024)	(0.026)	(0.029)	(0.024)	(0.026)	(0.029)	(0.024)	(0.026)	(0.029)
IV-2SLS	0.048***	0.041***	0.046**	0.051***	0.034**	0.064***	0.058***	0.046***	0.062***
(clustered s.e.)	(0.015)	(0.012)	(0.024)	(0.014)	(0.014)	(0.020)	(0.017)	(0.013)	(0.021)
(w. b. p-values)	(0.016)	(0.009)	(0.044)	(0.021)	(0.050)	(0.028)	(0.024)	(0.023)	(0.043)
F-stat	49.26	50.21	44.89	49.26	50.21	44.89	49.26	50.21	44.89
# of obs.	14,876	7,464	7,412	14,876	7,464	7,412	14,876	7,464	7,412
Panel B	**Period of observation: 2009–2018**								
OLS	0.027	0.031*	0.015	0.025	0.031*	0.019	0.027	0.035*	0.021
(clustered s.e.)	(0.019)	(0.017)	(0.019)	(0.018)	(0.014)	(0.018)	(0.019)	(0.015)	(0.020)
(w. b. p-values)	(0.234)	(0.099)	(0.222)	(0.198)	(0.101)	(0.234)	(0.194)	(0.103)	(0.276)
1st stage	2.223***	2.206***	2.193***	2.223***	2.206***	2.193***	2.223***	2.206***	2.193***
(clustered s.e.)	(0.301)	(0.305)	(0.345)	(0.301)	(0.305)	(0.345)	(0.301)	(0.305)	(0.345)
(w. b. p-values)	(0.009)	(0.008)	(0.014)	(0.009)	(0.008)	(0.014)	(0.009)	(0.008)	(0.014)
IV-2SLS	0.035	0.033**	0.027	0.030	0.039**	0.026	0.020	0.034*	0.010
(clustered s.e.)	(0.023)	(0.015)	(0.022)	(0.021)	(0.015)	(0.022)	(0.019)	(0.017)	(0.021)
(w. b. p-values)	(0.176)	(0.058)	(0.201)	(0.167)	(0.049)	(0.227)	(0.145)	(0.091)	(0.301)
F-stat	48.16	44.21	50.17	48.16	44.21	50.17	48.16	44.21	50.17
# of obs.	21,766	10,958	10,808	21,766	10,958	10,808	21,766	10,958	10,808

Table 5: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, natural logarithm of the year-specific distance to nearest border crossing, mother's education, and father's education are included as control variables into all regressions. A gender dummy is also included in regressions for all sample (i.e., columns 1, 4, and 7). The OLS model regresses the test score on refugee share and other controls, and the IV-2SLS model uses the distance-based variable as an IV for the refugee share to regress the test score on refugee share and other controls.

DID results for mean test scores

Dep. var.: Stand. value of mean test score

Panel A	**Period of obs.: 2009–2015**		
	All	**Male**	**Female**
	Specification I		
Refugee effect	0.166*	0.144*	0.181*
(clustered s.e.)	(0.056)	(0.047)	(0.071)
(w. b. p-values)	(0.078)	(0.089)	(0.123)
# of obs.	2,580	1,317	1,263
	Specification II		
Refugee effect	0.184*	0.226**	0.135
(clustered s.e.)	(0.069)	(0.055)	(0.086)
(w. b. p-values)	(0.099)	(0.051)	(0.231)
# of obs.	6,079	3,123	2,956
	Specification III		
Refugee effect	0.141*	0.149*	0.129
(clustered s.e.)	(0.048)	(0.049)	(0.078)
(w. b. p-values)	(0.101)	(0.096)	(0.201)
# of obs.	14,876	7,464	7,412
Panel B	**Period of obs.: 2009–2018**		
	Specification I		
Refugee effect	0.109	0.130**	0.076
(clustered s.e.)	(0.068)	(0.042)	(0.081)
(w. b. p-values)	(0.178)	(0.053)	(0.341)
# of obs.	3,722	1,929	1,793
	Specification II		
Refugee effect	0.128	0.141*	0.092
(clustered s.e.)	(0.078)	(0.054)	(0.111)
(w. b. p-values)	(0.201)	(0.096)	(0.368)
# of obs.	9,098	4,580	4,518
	Specification III		
Refugee effect	0.104	0.126*	0.080
(clustered s.e.)	(0.064)	(0.042)	(0.083)
(w. b. p-values)	(0.188)	(0.083)	(0.301)
# of obs.	21,766	10,958	10,808

Table 6: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, mother's education, and father's education are included as control variables into all regressions. A gender dummy is also included in the regressions for all sample (i.e., column 1).

IV-2SLS results for mean test scores

Dep. var.: Stand. value of mean test score

Panel A	**Period of obs.: 2009–2015**		
	All	**Male**	**Female**
OLS	0.041***	0.043***	0.038**
(clustered s.e.)	(0.008)	(0.009)	(0.016)
(w. b. p-values)	(0.013)	(0.009)	(0.043)
1st stage	2.171***	2.168***	2.162***
(clustered s.e.)	(0.309)	(0.316)	(0.304)
(w. b. p-values)	(0.024)	(0.026)	(0.029)
IV-2SLS	0.053***	0.044***	0.060**
(clustered s.e.)	(0.011)	(0.009)	(0.020)
(w. b. p-values)	(0.019)	(0.013)	(0.048)
F-stat	49.26	50.21	44.89
# of obs.	14,876	7,464	7,412
Panel B	**Period of obs.: 2009–2018**		
OLS	0.026	0.032*	0.020
(clustered s.e.)	(0.017)	(0.015)	(0.016)
(w. b. p-values)	(0.197)	(0.102)	(0.241)
1st stage	2.223***	2.206***	2.193***
(clustered s.e.)	(0.301)	(0.305)	(0.345)
(w. b. p-values)	(0.009)	(0.008)	(0.014)
IV-2SLS	0.029	0.035**	0.021
(clustered s.e.)	(0.019)	(0.013)	(0.019)
(w. b. p-values)	(0.184)	(0.056)	(0.245)
F-stat	48.16	44.21	50.17
# of obs.	21,766	10,958	10,808

Table 7: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, natural logarithm of the year-specific distance to nearest border crossing, mother's education, and father's education are included as control variables into all regressions. A gender dummy is also included in regressions for all sample (i.e., column 1). The OLS model regresses the test score on refugee share and other controls, and the IV-2SLS model uses the distance-based variable as an IV for the refugee share to regress the test score on refugee share and other controls.

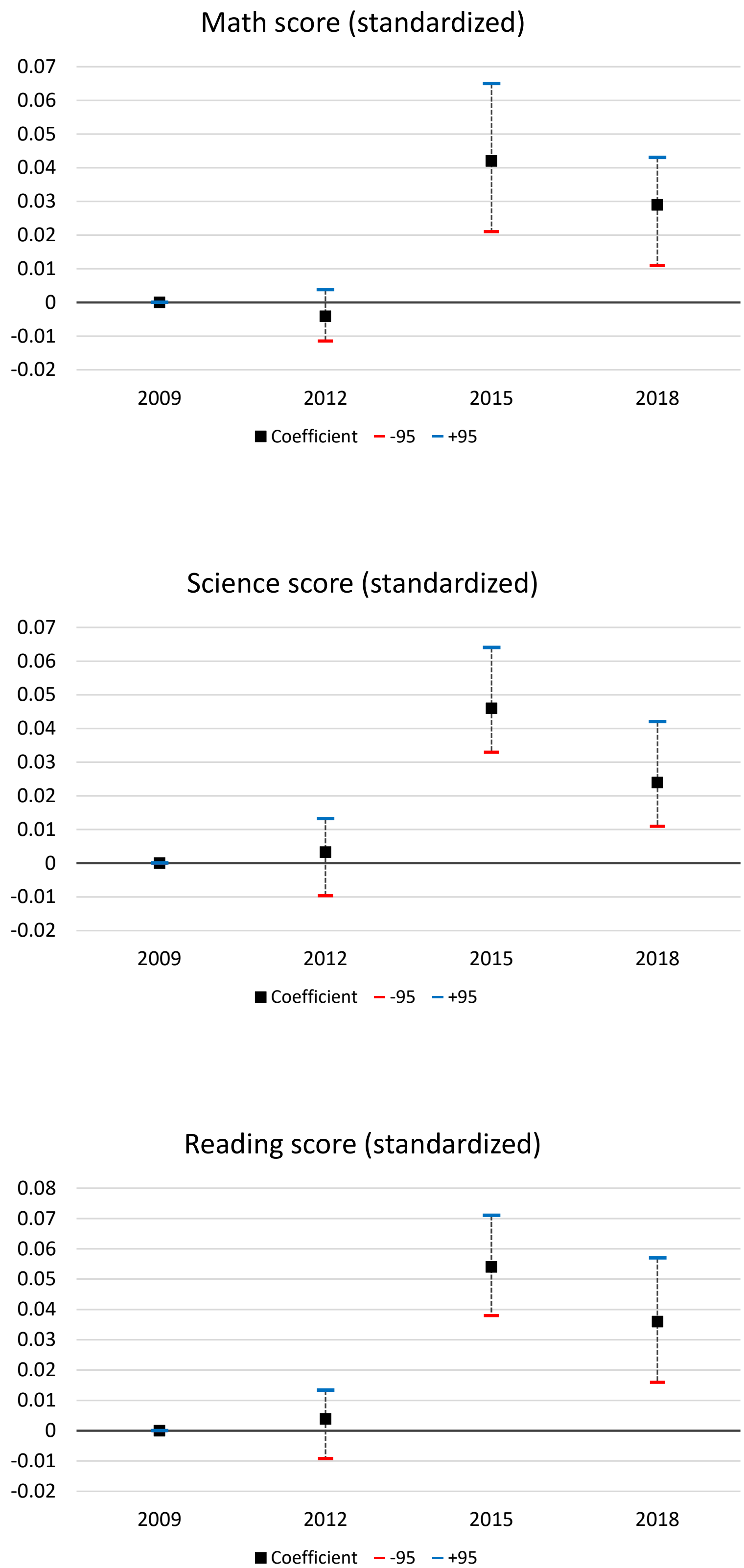
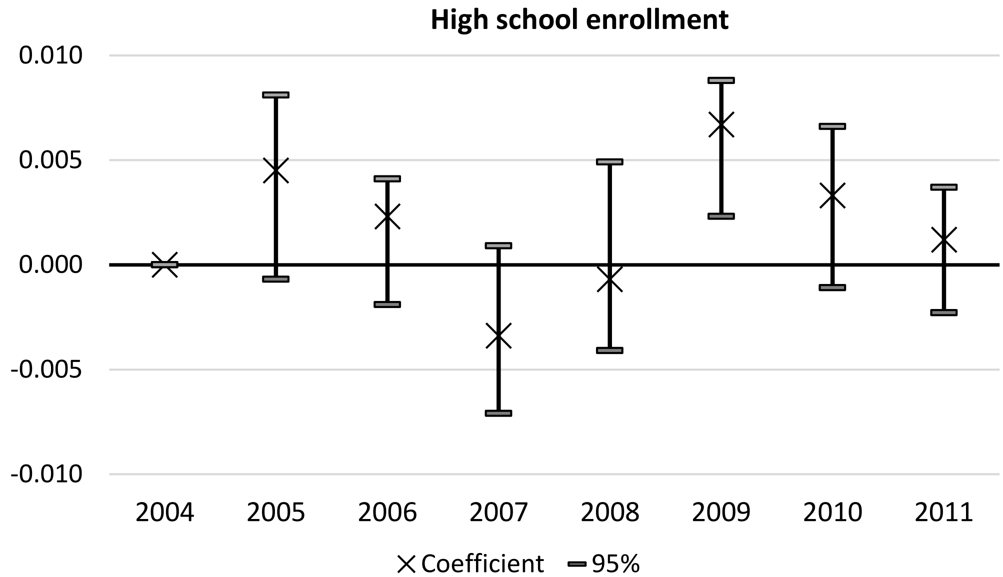


Figure 5: Testing common trends. Estimated coefficients of the interaction between treatment and wave dummies are plotted together with the 95% confidence intervals. Standard errors are clustered at region level. The results are robust to using the wild-bootstrap procedure. The estimation procedure is described in Section 4.



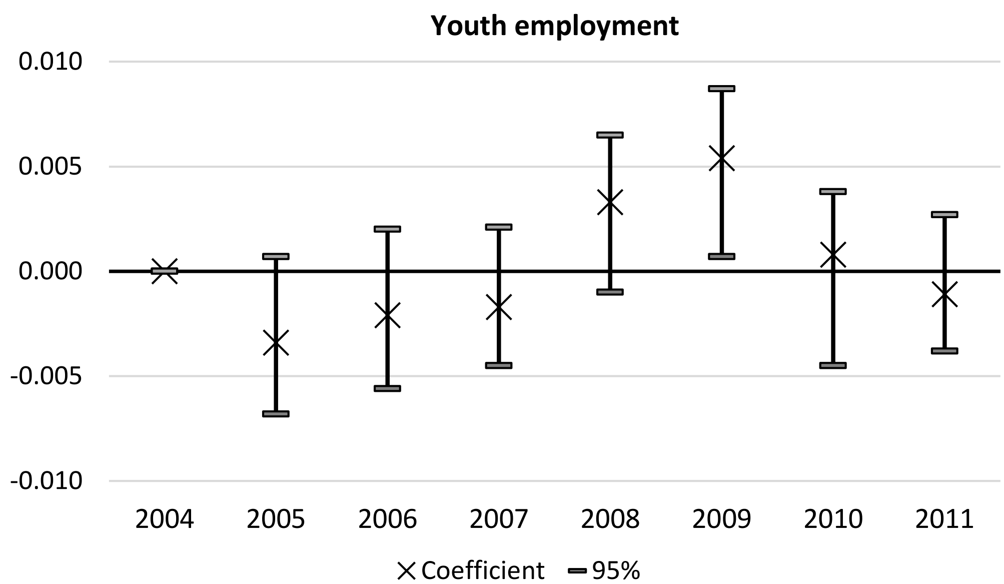


Figure 6: Testing pre-trends in alternative outcomes. Estimated coefficients of the interaction between treatment and wave dummies are plotted together with the 95% confidence intervals. Standard errors are clustered at region level. The results are robust to using the wild-bootstrap procedure. The estimation procedure is described in Section 4.

IV-2SLS ESTIMATION (Mother w/ less than high school education)

Dependent variable: Standardized value of the corresponding test score

Panel A	**Period of observation: 2009–2015**								
	Math			**Science**			**Reading**		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
OLS	0.041***	0.048***	0.028*	0.047***	0.058***	0.037**	0.046***	0.065***	0.028*
(clustered s.e.)	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)	(0.015)	(0.012)	(0.016)	(0.014)
(w. b. p-values)	(0.014)	(0.009)	(0.087)	(0.011)	(0.010)	(0.051)	(0.017)	(0.007)	(0.101)
1st stage	2.043***	2.091***	2.095***	2.043***	2.091***	2.095***	2.043***	2.091***	2.095***
(clustered s.e.)	(0.271)	(0.288)	(0.267)	(0.271)	(0.288)	(0.267)	(0.271)	(0.288)	(0.267)
(w. b. p-values)	(0.017)	(0.012)	(0.020)	(0.017)	(0.012)	(0.020)	(0.017)	(0.012)	(0.020)
IV-2SLS	0.045*	0.071***	0.023	0.042***	0.069***	0.021	0.039**	0.062***	0.022
(clustered s.e.)	(0.022)	(0.015)	(0.020)	(0.011)	(0.012)	(0.020)	(0.016)	(0.016)	(0.019)
(w. b. p-values)	(0.098)	(0.032)	(0.286)	(0.021)	(0.009)	(0.192)	(0.048)	(0.021)	(0.203)
F-stat	45.01	43.12	47.76	45.01	43.12	47.76	45.01	43.12	47.76
# of obs.	11,705	5,859	5,846	11,705	5,859	5,846	11,705	5,859	5,846
Panel B	**Period of observation: 2009–2018**								
OLS	0.027	0.031	0.021	0.027	0.032*	0.024	0.030	0.038**	0.024
(clustered s.e.)	(0.018)	(0.019)	(0.020)	(0.017)	(0.016)	(0.020)	(0.020)	(0.016)	(0.017)
(w. b. p-values)	(0.145)	(0.123)	(0.198)	(0.122)	(0.097)	(0.187)	(0.132)	(0.050)	(0.167)
1st stage	2.065***	2.004***	2.073***	2.065***	2.004***	2.073***	2.065***	2.004***	2.073***
(clustered s.e.)	(0.276)	(0.292)	(0.271)	(0.276)	(0.292)	(0.271)	(0.276)	(0.292)	(0.271)
(w. b. p-values)	(0.021)	(0.023)	(0.031)	(0.021)	(0.023)	(0.031)	(0.021)	(0.023)	(0.031)
IV-2SLS	0.031	0.036*	0.025	0.029	0.037*	0.022	0.025	0.036*	0.017
(clustered s.e.)	(0.021)	(0.018)	(0.020)	(0.021)	(0.019)	(0.018)	(0.018)	(0.017)	(0.014)
(w. b. p-values)	(0.207)	(0.083)	(0.267)	(0.187)	(0.104)	(0.255)	(0.176)	(0.105)	(0.298)
F-stat	49.11	45.89	52.09	49.11	45.89	52.09	49.11	45.89	52.09
# of obs.	15,864	7,948	7,916	15,864	7,948	7,916	15,864	7,948	7,916

Table 8: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, natural logarithm of the year-specific distance to nearest border crossing, mother's education, and father's education are included as control variables into all regressions. A gender dummy is also included in regressions for all sample (i.e., columns 1, 4, and 7). The OLS model regresses the test score on refugee share and other controls, and the IV-2SLS model uses the distance-based variable as an IV for the refugee share to regress the test score on refugee share and other controls. The sample consists of students with maternal education strictly less than high school.

IV-2SLS ESTIMATION (Mother w/ at least high school education)

Dependent variable: Standardized value of the corresponding test score

Panel A	**Period of observation: 2009–2015**								
	Math			**Science**			**Reading**		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
OLS	0.012	0.016	0.010	0.022*	0.020	0.024	0.022	0.025	0.019
(clustered s.e.)	(0.016)	(0.014)	(0.013)	(0.010)	(0.012)	(0.016)	(0.016)	(0.018)	(0.015)
(w. b. p-values)	(0.190)	(0.186)	(0.211)	(0.121)	(0.123)	(0.132)	(0.145)	(0.123)	(0.189)
1st stage	2.101***	1.999***	2.189***	2.101***	1.999***	2.189***	2.101***	1.999***	2.189***
(clustered s.e.)	(0.274)	(0.288)	(0.282)	(0.274)	(0.288)	(0.282)	(0.274)	(0.288)	(0.282)
(w. b. p-values)	(0.025)	(0.021)	(0.029)	(0.025)	(0.021)	(0.029)	(0.025)	(0.021)	(0.029)
IV-2SLS	0.028	0.041**	0.016	0.024	0.037***	0.014	0.019	0.042**	0.001
(clustered s.e.)	(0.022)	(0.015)	(0.022)	(0.019)	(0.018)	(0.017)	(0.020)	(0.015)	(0.016)
(w. b. p-values)	(0.248)	(0.082)	(0.301)	(0.198)	(0.048)	(0.259)	(0.187)	(0.034)	(0.283)
F-stat	47.01	48.94	45.92	47.01	48.94	45.92	47.01	48.94	45.92
# of obs.	3,171	1,605	1,566	3,171	1,605	1,566	3,171	1,605	1,566
Panel B	**Period of observation: 2009–2018**								
OLS	0.005	0.007	0.003	0.009	0.011	0.008	0.012	0.010	0.013
(clustered s.e.)	(0.020)	(0.018)	(0.015)	(0.015)	(0.015)	(0.014)	(0.019)	(0.020)	(0.018)
(w. b. p-values)	(0.198)	(0.185)	(0.211)	(0.201)	(0.186)	(0.218)	(0.187)	(0.182)	(0.198)
1st stage	2.001***	1.934***	2.056***	2.001***	1.934***	2.056***	2.001***	1.934***	2.056***
(clustered s.e.)	(0.209)	(0.192)	(0.204)	(0.209)	(0.192)	(0.204)	(0.209)	(0.192)	(0.204)
(w. b. p-values)	(0.034)	(0.031)	(0.037)	(0.034)	(0.031)	(0.037)	(0.034)	(0.031)	(0.037)
IV-2SLS	0.006	0.010	0.009	0.010	0.011	0.009	0.013	0.016	0.009
(clustered s.e.)	(0.021)	(0.022)	(0.024)	(0.017)	(0.019)	(0.018)	(0.019)	(0.018)	(0.021)
(w. b. p-values)	(0.243)	(0.211)	(0.266)	(0.265)	(0.204)	(0.287)	(0.226)	(0.191)	(0.250)
F-stat	48.04	46.11	50.04	48.04	46.11	50.04	48.04	46.11	50.04
# of obs.	5,902	3,010	2,892	5,902	3,010	2,892	5,902	3,010	2,892

Table 9: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, natural logarithm of the year-specific distance to nearest border crossing, mother's education, and father's education are included as control variables into all regressions. A gender dummy is also included in regressions for all sample (i.e., columns 1, 4, and 7). The OLS model regresses the test score on refugee share and other controls, and the IV-2SLS model uses the distance-based variable as an IV for the refugee share to regress the test score on refugee share and other controls. The sample consists of students with mothers of at least high school education.

IV-2SLS ESTIMATION: Heterogeneity

	All	**Male**	**Female**
Panel A		**Below median**	
1st stage	2.167***	2.134***	2.175***
(clustered s.e.)	(0.163)	(0.170)	(0.172)
(w. b. p-values)	(0.017)	(0.021)	(0.020)
IV-2SLS	0.009***	0.011***	0.006**
(clustered s.e.)	(0.002)	(0.003)	(0.002)
(w. b. p-values)	(0.032)	(0.017)	(0.056)
F-stat	139.31	130.90	142.12
# of obs.	21,800	11,618	10,182
Panel B		**Above median**	
1st stage	2.202***	2.254***	2.173***
(clustered s.e.)	(0.343)	(0.351)	(0.334)
(w. b. p-values)	(0.014)	(0.011)	(0.031)
IV-2SLS	0.005	0.008	0.003
(clustered s.e.)	(0.005)	(0.005)	(0.004)
(w. b. p-values)	(0.271)	(0.194)	(0.301)
F-stat	47.14	50.32	42.98
# of obs.	22,828	10,774	12,054

Table 10: ***, **, and * refer to 1%, 5%, and 10% significance levels, respectively. Standard errors are clustered at region level. In addition to standard clustering, a wild bootstrap exercise is also performed as described by Roodman et al. (2019). The corresponding p-values are reported in parentheses below the clustered standard errors. Year fixed effects, region fixed effects, regional per capita real GDP (in logs), regional real trade volumes (in logs) between Syria and Turkey, month-of-birth fixed effects, grade fixed effects, test subject fixed effects, natural logarithm of the year-specific distance to nearest border crossing, mother's education, father's education, and gender are included as control variables into all regressions. The upper (lower) panel restricts the sample to test scores below (above) the 50th percentile.

Dependent Variable	Estimate
Log per capita real GDP	0.0000
	(0.0067)
Log total real trade volume	0.0002
	(0.0073)
Log real export volume	0.0003
	(0.0071)
Log real import volume	-0.0002
	(0.0069)
Mother's education (average)	0.0007
	(0.0061)
Father's education (average)	0.0003
	(0.0060)
# of observations	48

Table 11: The relationship between the instrument and time-region-varying regressors. Each row reports the result from a separate regression, where the name of the dependent variable is indicated in each row and the independent variables are the instrument, year fixed effects, and region fixed effects. The reported estimate in each row corresponds to the coefficient of the instrument in that regression. The number of observations is 48 for each regression (since there are 12 regions in the analysis observed over four consecutive PISA waves from 2009 to 2018). The average years of education for mothers and fathers are calculated using the Mincerian approach for converting educational attainment dummies into years of schooling (Mincer, 1974). Standard errors are clustered at region level.