H i C N Households in Conflict Network
The Institute of Development Studies - at the University of Sussex - Falmer - Brighton - BN1 9RE
www.hicn.org

The impact of low-skill refugees on youth education

Semih Tumen[1]

HiCN Working Paper 283

November 2018

Abstract: This paper examines the impact of Syrian refugees on high school enrollment rates of native youth in Turkey. Syrian refugees are, on average, less skilled and more willing to work in low-pay informal jobs than Turkish natives. Refugees can influence native youth's school enrollment likelihood negatively through educational ex- perience. But, at the same time, they can affect enrollment rates positively as they escalate competition for jobs with low-skill requirements. Using micro data from 2006 to 2016 and employing quasi-experimental methods, I find that high-school enrollment rates increased 2.7-3.6 percentage points among native youth in refugee-receiving regions. Furthermore, a one-percentage point increase in the refugee-to-population ratio in a region generates around 0.4 percentage point increase in native's high school enrollment rates. Most of the increase in high school enrollment comes from young males with lower parental backgrounds, which is consistent with the hypothesis that the main mechanism operates through the low-skill labor market. The regressions control for (i) variables proxying parental investment in human capital such as parental education, being in an intact family, and household size, (ii) regional economic activity, and (iii) regional availability of high schools and high school teachers.

Keywords: Low-skill Syrian refugees; youth education; high school enrollment.

JEL Codes: I25; J61.

Acknowledgements: I thank Onur Altindag, Yusuf Kenan Bagir, George Borjas, Jean-Francois Maystadt, Ahmet Ozturk, Sandra Rozo, Insan Tunali, Mathis Wagner, Vasco Yasenov, the participants of the "Conference on the Impacts of Refugees on Hosting Economies" held at the University of Southern California, and seminar participants at TED University and Gebze Technical University for useful comments and suggestions. The usual disclaimer holds.

[1] TED University, IZA, and ERF. Contact: semih.tumen@tedu.edu.tr.
TED University, Department of Economics. Ziya Gokalp Cad., No.48, 06420 Kolej, Ankara, Turkey.

1 Introduction

The violent civil conflict in Syria has generated a huge wave of forced displacement. Countries in the region—in particular, Jordan, Lebanon, and Turkey—have been hosting more than 5 million registered Syrian refugees as of 2018.[1] The incumbent regime in Syria has gained power against rebels recently, which is expected to intensify clashes and trigger additional refugee waves toward these countries. In addition, the huge physical destruction in the Syrian economy makes the return of refugees a remote possibility in the short term (Ceylan and Tumen, 2018). This suggests that hosting countries in the region will be facing high refugee concentration for a prolonged period of time, which will likely have long-term consequences on the socio-economic outcomes of natives in these countries.

There is an emerging literature investigating the impact of Syrian refugee inflows on host country labor market outcomes. The main findings in this literature can be summarized as follows. Refugees are, on average, less skilled than natives. In Turkey, the major host country, refugees do not have easy access to work permit; so, they penetrate 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), Bagir (2017), and Ceritoglu et al. (2017)].[2] Informally employed refugee workers provide important labor cost advantages and, accordingly, wages decline in the low-skill 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).[3]

These results suggest that competition between refugees and natives for low-skill jobs imposes a downward pressure on potential wages in the low-skill labor market. As a result, the decline in the expected returns to staying low-skilled and low-educated may increase school enrollment among native youth. In this paper, I investigate the impact of Syrian refugees on youth education in Turkey. I focus on the change in native high school enrollment rates as a response to Syrian

[1] For more detailed statistics, see https://data2.unhcr.org/en/situations/syria.

[2] Akgunduz et al. (2015), Fakih and Ibrahim (2016), Cengiz and Tekguc (2018), and Fallah et al. (2018) document smaller or negligible effects. Aksu et al. (2018) report a decline in informal employment among natives, but they argue that the rest of the results may change depending on the identification assumptions.

[3] See Peri and Sparber (2009) for similar findings documented using the US data.

refugee influx. To deal with the potential endogeneity issues, I use both 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). The Turkish Household Labor Force micro-level data sets are used in the empirical analysis.

I find that high school enrollment rates have increased substantially among native youth as a response to refugee inflows. The increase in high school enrollment attributed to the refugee influx comes almost entirely from the increase in male enrollment, while there is no statistically significant increase in female enrollment in most specifications. Both the diff-in-diff and IV estimates confirm these findings. Variables proxying parental investment in children's human capital, changes in regional economic activity, and changes in regional availability of educational resources are also controlled in the regressions. In terms of magnitudes, the basic diff-in-diff estimates suggest that high school enrollment rates are 2.7-3.6 percentage points higher in regions with high refugee concentration relative to comparable regions with almost no refugee presence. The IV-diff-in-diff estimates suggest that a one percentage-point increase in the refugee-to-population ratio increases the high school enrollment rate by 0.4 percentage point. I test the validity of the common-trends assumption in two steps. First I perform a formal test following Autor (2003). Then, I relax the common trends assumption in the spirit of Stephens and Yang (2014). The results suggest that the main identifying assumptions are valid.

There are two main channels through which refugees can affect high school enrollment among native youth (Hunt, 2017). First, increased competition for educational resources may discourage native youth and, as a result, high school enrollment may decline. Second, the decline in potential wages due to increased competition for low-skill jobs may encourage high school enrollment among native youth. The results of this study suggest that the second channel dominates the first one for the case of Syrian refugees in Turkey.

The literature investigating the impact of immigration on natives' educational outcomes is vast and has several sub-branches.[4] Denisova (2003), Smith (2012), McHenry (2015), Jackson (2016), and Hunt (2017) focus on the change in natives' school enrollment decisions as a response to in-

[4]See Dustmann and Glitz (2011) (especially Section 4) for an excellent review of the literature.

creased competition in the low-skill labor market due to immigration.[5] These papers document that—despite the forces operating in the opposite direction—immigrants tend to crowd natives into higher education, since they drive the pay down in the low-skill market. 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 and within-class negative externalities.[6] The negative effect is more pronounced for disadvantaged natives. Neymotin (2009), Geay et al. (2013), Ohinata and van Ours (2013), and Shih (2017), on the other hand, report zero or positive impact of increased immigrant concentration within the class/school on natives' educational outcomes.[7] Betts and Fairlie (2003), Murray (2016), and Farre et al. (2018) show that natives switch from public to private schools as a response to increased immigrant concentration in public schools.[8] Finally, Orrenius and Zavodny (2015) document that increased immigrant concentration reduces the likelihood that US women major in a science or engineering field.

This paper can be placed into the literature focusing on the mechanism operating through labor market opportunities. The paper most closely related to my paper is Hunt (2017). Similar to her paper, I also find that the improvement in schooling outcomes is a response to increased competition in the low-skill labor market due to immigration. There are two main differences. First, she focuses on long-term immigration in the US; and, second, the main outcome variable is high school graduation. In contrast, I focus on shorter-term response in educational outcomes in Turkey following the sudden Syrian refugee influx and the outcome variable in my paper is high school enrollment. The estimates I report are slightly higher than Hunt's estimates and this difference can be attributed to the difference in outcome variables between the two papers. Assaad et al. (2018) use micro data from Jordan and show that Syrian refugees do not affect the education outcomes of Jordanian native youth. The difference between the results of Assaad et al. (2018) and my paper also reflects the difference between the conclusions reached by Fallah et al. (2018) versus

[5] See also Eberhard (2012) and Llull (2017) for in-depth analyses of the main theoretical channels in a structural setting.

[6] See also Jensen and Rasmussen (2011), Foster (2012), Brunello and Rocco (2013), and Roed and Schone (2016) for similar results. Berker (2009) documents that internal migration negatively affects natives' educational attainment in Turkey.

[7] Schneeweis (2015) reports that high immigrant concentration does not affect natives' educational outcomes, but it negatively affects the educational outcomes of immigrants in Austria. Assaad, Ginn, and Saleh (2018) find that Syrian refugees do not affect the education outcomes of Jordanian natives.

[8] Cascio and Lewis (2012) find that natives move to other districts as a response to increased Hispanic settlements in California. This contributes to immigrants' residential isolation.

Del Carpio and Wagner (2015) and Ceritoglu et al. (2017). The studies for Jordan do not report any effect of refugees on the labor market outcomes of Jordanian natives, while some negative effects are documented for Turkey especially in the low-skill market. So, in this sense, there is not a fundamental contradiction between the findings of the current paper and Assaad et al. (2018).

The plan of the paper is as follows. Section 2 describes the data. Section 3 explains the empirical framework and discusses the baseline findings. Section 4 presents extensions of the baseline results. Section 5 concludes.

2 Data and institutional details

The main data set used in this paper is the Turkish Household Labor Force Survey (HLFS), which is compiled and published by the Turkish Statistical Institute (TurkStat). HLFS micro-level modules are published annually on a cross-sectional basis. Official labor market and employment statistics in Turkey are produced using the HLFS. It samples Turkish citizens—non-institutional population—based on an addressed-based procedure. Eleven waves (2006-2016) of the HLFS are used in the empirical analysis.

There is no publicly available micro-level information on Syrian refugees in Turkey. Instead, I exploit the regional variation in Syrian refugee intensity and the exogeneity of the timing of influx to estimate the impact of refugees on natives educational outcomes. Table (1) compares the basic characteristics of natives versus refugees as of year 2016—based on Ministry of Interior data. Clearly, refugees are younger and much less-skilled than natives, on average. Note that there is a significant fraction of refugees for whom the education level is "unknown." It is quite likely that these refugees also have very low levels of education. Figure (1) displays the trends in Syrian refugee inflows in Turkey. Note that the figure includes only the registered refugees. Figure (2) shows the time-region variation in Syrian refugee intensity in Turkey. The figure suggests that the refugees were clustered around the Turkey-Syria border until 2014 and, then, they moved toward other regions.[9]

[9]See Tumen (2016) and Bagir (2017) for a detailed description of this two-stage process.

The main outcome variable is high school enrollment. The HLFS asks whether the individual is enrolled in school at the time of the survey. I construct a binary outcome variable taking 1 if the individual is enrolled in high school and 0 otherwise—excluding the ones enrolled in college. I mainly focus on individuals of age between 15 and 18. I re-organize the school enrollment variable such that if the individual is enrolled in school, then there are three possible enrollment categories: less than high school, high school, and college & above. The change in the definition of the enrollment variable in 2014 necessitates this re-organization. The high school category includes vocational high school. College & above category includes 2-year college, 4-year college, and post-graduate education. Until 2014, there was a separate distant education ("Acikogretim") category, which included all levels of distant education in one variable.[10] After 2014, TurkStat removed this enrollment category and placed the individuals taking distant education under the corresponding school enrollment category.[11] To reconcile these differences, I assigned the individuals taking distant education into one of the "less than high school", "high school," or "college & above" categories based on the "highest level of education completed" variable.[12]

It would also be interesting to use other outcome variables such as high school graduation and college enrollment rates. The targeted individuals will mostly be in the age range of 18 to 22. However, there are some peculiarities of the HLFS data set making those exercises infeasible. First of all, the HLFS covers the non-institutional population in Turkey, i.e., (*i*) young people who left their parents' region of residence for the purpose of taking higher education in another province and residing in school dormitories and (*ii*) the ones who go to army to complete their compulsory military service will not be observed in the data set. Moreover, part of the observed individuals have already left their home; thus, it would be impossible to map their characteristics to parental characteristics—since we can match parents with kids only if they live in the same house. For these reasons, we limit our attention to the high school enrollment variable only.

Parental investment is a key determinant of children's human capital (Becker, 1993a). I use three

[10] In this setting, we knew that the individual was continuing distant education, but we didn't explicitly know the level of distant education.

[11] Suppose that the individual is enrolled in a distant-learning high school program. Before the 2014 wave, s/he was classified under the "distant education" enrollment category. But, after 2014, s/he has been classified under the "high school" category, without any reference to distant education.

[12] For example, if the individual is enrolled in distant education and if his/her highest level of completed education is high school, then I assign him/her to the "college & above" enrollment category.

variables to control for various dimensions of parental investment. The first variable is the final educational degree obtained by the household head. I focus on the household head, because the educational attainment of parents may not be observed jointly if the family is dissolved. Second, I use a dummy variable taking 1 if the parents are intact (married and living together) and 0 otherwise, as it is well-documented that family structure affects success in school (Conti and Heckman, 2014). Finally, I use household size to control for other dimensions of human capital investment such as the quantity-quality tradeoff and intensity of resource allocation (Becker, 1993b).

Including variables proxying regional economic activity is also critical as school enrollment decisions may also respond to changes in economic conditions. I use two different variables following the conventions in the literature. First, I use the region-level international trade volume (exports plus imports) in real terms. This variable is available for all time periods and is included into all regressions. Second, I use the real per capita GDP at the region level. NUTS2-level real GDP per capita is manually constructed from province-level figures through simple aggregation. Note that the regional GDP information is available until 2014, so it is only used in some of the specifications. Both variables are converted into real terms using the regional CPI figures. All data are publicly available from TurkStat.

It may also be case that the number of educational resources exhibit region-year variation due to a combination of policy responses and demand factors. To address this possibility, I include the number of high schools and the number of high school teachers in each region over time. The numbers are calculated as the sum of general and vocational high schools. The data source is the National Education Statistics Yearbooks published by the Ministry of National Education.

There are 26 NUTS2-level geographical regions in Turkey and the HLFS include this regional information, which will be useful to characterize the time-region variation in Syrian refugee intensity in Turkey. The empirical analysis also includes region dummies, survey-year dummies, a gender dummy, and age dummies. Tables (2) and (3) report the summary statistics that are relevant for the diff-in-diff and IV estimations, respectively.

The province-level refugee-to-population data come from various sources. The data for 2013 and 2014 are taken from AFAD (2013) and Erdogan (2014), respectively. The data for 2015 and 2016 are obtained from the official statistics published by the Ministry of Interior, Directorate General of Migration Management.

Note that after September 2012, high school enrollment became compulsory nationwide in Turkey. The law passed on April 2012 briefly says that children who finish the 8th grade in June 2012 have to enroll in high school as of September 2012. Although this is a "compulsory schooling" reform, there are severe enforcement problems. As Tables (2) and (3) suggest, the enrollment rates were around 70 percent—60 percent in the eastern and southeastern regions—several years after the reform. This was a nation-level reform, so any effect of the reform (homogeneous across regions) should be differenced out in diff-in-diff settings. To address the possibility of region-specific effects, several additional exercises are performed to show that the reform does not contaminate the estimates. Section 4.3 discusses this issue in much more detail.

3 Econometric analysis

3.1 Motivation and strategy

The empirical setup is based on a diff-in-diff analysis. I use two versions of diff-in-diff. The first one is a simple before-after comparison of the regions exposed and not exposed to refugee influx, which is similar to Ceritoglu et al. (2017). The second is an IV-diff-in-diff model exploiting the variation in refugee concentration over time/across regions and using distance from the source governorates in Syria to destination provinces in Turkey as an IV—similar to Del Carpio and Wagner (2015). To check the relevance of the common-trends assumption: (i) I formally test whether the common-trends assumption holds and (ii) I relax the common-trends assumption to see whether the results are sensitive to introducing region-specific trends in the outcome variable.

To motivate the respective roles of the basic diff-in-diff model and the IV-diff-in-diff model, it is key to understand the dynamics of Syrian refugee movements in Turkey. The Syrian refugee inflows started in January 2012 and accelerated over time [see Figure (1)]. Before this date, the number of Syrians was almost zero in the entire country. From 2012 to 2014, the refugees were mostly located close to the Turkey-Syria border for two reasons: (1) they were hoping to go back home once the crisis was resolved and (2) the Turkish government built large refugee camps along the border regions to provide basic services (such as health, security, food, education, etc). So, between 2012 and 2014, the Syrians who were forced out of their home country were in some sense forced to cluster around the border regions in Turkey. After the end of 2014, however, it became clear that the crisis would not end soon and the refugees started to actively seek permanent homes. Some of them preferred to stay close to the border regions, while others chose to move out of the southeastern regions toward the western regions of the country. Figure (2) roughly displays these location-choice patterns.

The patterns suggest that the two episodes should be treated separately in empirical analysis. What happened until mid- to end-2014 can be treated as a plausibly exogenous demographic shock to the hosting regions, while refugees self-selected into locations after the second half of 2014. This difference will be shaping the nature of empirical analysis throughout this section.

3.2 Basic diff-in-diff specification

The first specification is based on the difference-in-differences strategy implemented by Balkan and Tumen (2016) and Ceritoglu et al. (2017).[13] The post-influx period is defined by the dummy variable A_{iy} as:

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

where i and y indexes individuals and years, respectively. The pre-influx years are 2010 and 2011, while the post-influx years are 2012, 2013, and 2014. Similarly, two groups of regions are defined as treatment and control groups by the dummy variable T_{ir} as:

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

[13]The setting is also similar to the famous diff-in-diff example by Card and Krueger (1994).

where r indexes regions. The treatment group consists of Gaziantep, Hatay, Sanliurfa, Mardin, and Adana NUTS2 regions, while the control group includes Erzurum, Malatya, Agri, and Van NUTS2 regions. Figure (3) shows the treatment and control groups at a province level with dark and light blue colors, respectively. Table (2) compares the basic summary statistics in these two regions for the 15-18 age group suggesting that the two regions are similar with respect to individual-level characteristics. Moreover, both the treatment and control groups are in the eastern and southeastern parts of the country, which share similar cultural, ethnic, socio-economic, and economic characteristics. Most importantly, both regions have low levels of economic development (i.e., they are poorer regions relative to the rest of the country) and low education/school enrollment levels.

The main outcome variable is high school enrollment for age 15-18. It will be useful to see the high school enrollment trends in these two regions over time. Figure (4) plots the high school enrollment rates for age 15-18 using the HLFS data. In case the labor market survey mis-represents the school enrollment figures, I also use the National Education Statistics Yearbooks and plot the entire enrollment numbers in Figure (5). The figures suggest that both the levels and trends of high school enrollment are quite similar in the treatment and control groups in the pre-influx period, while there is some differentiation after the influx. So, the diff-in-diff setting is defensible based on a simple eyeball test.

At the end, the diff-in-diff model can be formally specified as follows:

$$E_{irt} = \beta_0 + \beta_1(T_{ir} \times A_{it}) + \boldsymbol{\beta_3'} \boldsymbol{X}_{irt} + \boldsymbol{\beta_4'} \boldsymbol{P}_{irt} + \boldsymbol{\beta_5'} \boldsymbol{Y}_{rt} + \boldsymbol{\beta_6'} \boldsymbol{S}_{rt} + f_r + f_t + \epsilon_{irt}, \tag{1}$$

where E_{irt} is a dummy variable taking 1 if individual i of age 15-18 in region r and in year t is enrolled in high school and 0 otherwise, $\boldsymbol{X}_{irt}$ is a vector of individual-level characteristics, $\boldsymbol{P}_{irt}$ is a vector of parental controls proxying the intensity of parental investment in children's human capital, $\boldsymbol{Y}_{rt}$ is a vector of controls for regional economic activity, $\boldsymbol{S}_{rt}$ is a vector of variables capturing changes in educational resources across regions over time, 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.

Baseline estimates. Table (4) presents the baseline estimates for the whole sample, males, and females, separately. Gender (for the whole sample), age, year, and region dummies are included. The baseline analysis does not include parental, regional, and educational resource controls. In the first three columns, the pre-influx and post-influx periods are specified in the baseline regressions as 2010-2011 and 2012-2014, respectively. The last three columns defines pre and post periods as 2009-2011 and 2012-2015, respectively. Note that 2009 was a period of crisis and in 2015 refugees started self-select into locations; so, the first three columns are our preferred specification. I include the results for the larger time window to verify that the results are not sensitive to slight changes in the baseline timing setup. The sample consists of young individuals of age 15-18. The standard errors are clustered at region-year level consistent with the source of identifying variation in the diff-in-diff setting.

The estimates suggest that there is a statistically significant effect only for males. Specifically, the high school enrollment rates for males increased about 3.3 percentage points after the refugee influx. There is no impact on females. The mean high school enrollment rate among males is approximately 55 percent in the treatment region before the influx. So, the increase in the high school enrollment rate among males is roughly about 6 percent after accounting for the respective increase in the control region.

Including parental controls. Although the quasi-experimental setup can eliminate certain type of endogeneities (such as the location choice), there may be additional concerns related to human capital investment. Schooling decisions of young individuals may depend on the correlation between the unobserved ability levels of children and their parental characteristics. Moreover, parents may have responded to the change in labor market conditions due to refugee influx by affecting the schooling decisions of their children. To address these concerns, I control for parental education, being in an intact family or not, and the size of the household. These variables capture various dimensions of parental investment in children's human capital and intra-family resource allocation. Note that "parent" means "household head" due to the structure of the HLFS data set.

The first three columns in Table (5) controls the variables representing parental/family characteristics. All parental characteristics yield results with theoretically correct sign—i.e., high parental education and being in an intact family affect high school enrollment positively, while a large household size has a negative effect—and high statistical significance. The results are similar in nature to the baseline results with two nuances. First, the coefficient for the entire sample [column 1] turns statistically significant after controlling for parental variables. And second, the coefficients become larger and their statistical significance improves once the parental controls are included. The estimates suggests a 2.7 percentage point increase in the overall high school enrollment rates, and for males the magnitude of the increase is 4.6 percentage points. For females, the coefficients are still statistically indistinguishable from zero. Clearly, controlling for parental characteristics increases the magnitude of the coefficients and improves the precision of the estimates.

Controlling for changes in regional economic activity. Different regional economic performances can generate differential incentives for high school enrollment among youth. To address this concern, I include the regional real GDP per capita and real international trade volume as controls. They enter the regressions in natural logarithms. Columns 4-6 in Table (5) show the estimates. The inclusion of this variables on top of the parental controls do not generate a major change in the results both in terms of magnitude and statistical significance, which suggests that the estimates can be attributable to refugee influx rather than the impact of Syrian conflict on differences in regional economic activity.

Including controls for regional availability of educational resources. Another potential concern is the differences in the change in the availability of educational resources across regions, which may be driving the estimates. To account for this possibility, I include the number of high schools and number of high school teachers available in each region as control variables. The last three columns in Table (5) document the estimates. The nature of the main results is unchanged after controlling for educational resources on top of parental and region-economy controls.

Separate regressions conditioned on parental background. The results up to this point suggest that young males (of age 15-18) respond to refugee influx by increasing their high school enrollment rates. Since they still live with their parents, it will be important to understand the parental characteristics influencing the enrollment decisions of the children. To address this question, I run separate regressions conditioning on parental education and employment status. Table (6) presents the results of the regressions conditioned on parental education levels. The estimates suggest that the entire increase in high school enrollment comes from male children with parents of less than high school education. Table (7) reports results conditioned on parental employment status for male children only. The estimates communicate the result that the increase in high school enrollment is observed for young males whose parents are out of formal employment. This story is consistent with the proposed mechanism that low-skill refugees increase competition for low-skill jobs and push native youth with lower backgrounds from the low-skill labor market toward high school education.

Placebo diff-in-diff regressions. I run several placebo regressions based on counterfactual pre- and post-influx periods. Six different combinations of alternative pre- and post-influx periods covering a wide range of counterfactual dates are used in the placebo regressions. The results are reported in Table (8). The placebo regressions do not yield any statistically significant coefficients for the main variable of interest. This result does not change when I include or exclude the parental, region-economy, and educational resource controls.

3.3 Testing the common-trends assumption

As Figures (4) and (5) suggest, an eyeball test reveals that the common-trends assumption likely holds. Unlike the labor market variables, changes in school enrollment rates are expected to be more homogeneous across the country and over years. Still, it will be useful to test formally whether the common-trends assumption holds as the validity of the diff-in-diff estimates rests on the relevance of this assumption.

To perform this validity check, I follow the strategy used by Autor (2003). Specifically, I focus on the period 2010-2014 and I construct five dummy variables each equal to 1 only in the relevant year and 0 in other years. Then I perform a regression of school enrollment on the interactions between these year dummies and the treatment dummy T_{ir}—gender dummy, age dummies, year dummies, region dummies, parental controls, regional controls, and educational-resource controls are also included similar to the saturated diff-in-diff estimations. Figure (6) plots the coefficients of the interaction terms along with the 95 percent confidence bands. Year 2010 is the omitted category. The results for the entire sample and males suggest that the coefficients are statistically indistinguishable from zero before 2012, while they become positive and turn statistically significant after 2012. Again, the results for females are statistically insignificant.

3.4 IV specification

The second specification uses the IV setting developed by Del Carpio and Wagner (2015). The IV specification again lies within the diff-in-diff universe as it exploits the time-region variation in refugee-to-population ratio across Turkey. Different from the first specification, it uses data from entire Turkey and covers the period in which refugees started to self-select into the locations that they prefer to live. The main estimating equation can be formulated as follows:

$$E_{irt} = \alpha_0 + \alpha_1 R_{rt} + \alpha_2 \ln(D_{rt}) + \boldsymbol{\alpha_3'} \boldsymbol{X}_{irt} + \boldsymbol{\alpha_4'} \boldsymbol{P}_{irt} + \boldsymbol{\alpha_5'} \boldsymbol{Y}_{rt} + \boldsymbol{\beta_6'} \boldsymbol{S}_{rt} + f_r + f_t + \epsilon_{irt}, \qquad (2)$$

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 2012 and $D_{rt} = D_t$ from 2012 onwards. 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. Moreover, the impact of policy changes on high school enrollment of natives may also be correlated with distance to border; so, the inclusion of the distance variable also addresses those concerns. Consistent with the motivation provided at the beginning of this section, this setup can address the selectivity issues arising in the post-2014 period via a suitably designed IV strategy.

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 E_{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.[14] 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, 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 IV_{rt}. The summary statistics for the sample used in the IV estimations are presented in Table (3).

Baseline IV estimates. The IV analysis uses the 2010, 2011, 2014, 2015, and 2016 waves of the HLFS. The main purpose is to capture the patterns of location choices made by refugees in the post-2014 period. Unlike the basic diff-in-diff setup, the sample covers the entire country. Table (9) presents the baseline estimates. The first three columns do not include parental, regional, and educational resource controls, while they are controlled in the last three columns.[15] The main

[14]Google maps is used to calculate the shortest travel distances. There are 14 Syrian governorates and 26 NUTS-level regions in Turkey, which means that the distance is calculated between 364 distinct routes.

[15]Note that the regional GDP data is available only until 2014 and, therefore, cannot be used in the IV analysis. Volume of

results are quite similar to the results of the basic diff-in-diff regressions. The OLS estimates are mostly downward biased due to the selection of refugees into locations. The IV estimates suggest that the refugee influx increases high school enrollment among natives, and the effect comes almost entirely from the increase in male enrollment rates. Controlling for parental, regional, and educational resource variables does not change the results in a meaningful way. The effect is nil for females. In terms of magnitudes, a one percentage point increase in the refugee-to-population ratio increases high school enrollment rates around 0.4 percentage point on aggregate and around 0.8 percentage point for males. There is a single instrument and the F-statistic is around 15 in the preferred specifications.

Performing separate IV regressions for youth with lower versus higher parental backgrounds yields somewhat different results than the diff-in-diff analysis [see Table (10)]. Similar to the diff-in-diff analysis, increase in the enrollment rates is driven by males whose parents have less than high school education. There is no statistically significant impact for females. The results differentiate when I condition parental education to having high school education and above. In this case, the enrollment rates of both males and females go up following the refugee influx.

The main difference between the results of the basic diff-in-diff and IV analyses is that the basic diff-in-diff analysis performs a comparison across the less developed regions in the country. The IV analysis, on the other hand, covers the entire country, which may be the force generating different results for females. The findings suggest that, unlike the eastern and southeastern parts of the country in which the level of development is low, parents with high school education and above in the western regions respond to the increase Syrian-refugee concentration by increasing the high school enrollment rate for their female children in addition to their male children. This result makes sense, because the Syrian border regions in Turkey are known to have different cultural attitudes toward women than the rest of the country. Female labor force participation rates are low and teenage marriage rates are high mostly due to paternalistic social norms. Still, female high-school enrollment for parents with high education is a rather small fraction in the sample; so, the main story can still be constructed on the increase in high school enrollment among males.

international trade by regions is used instead.

4 Extensions

4.1 Relaxing the common-trends assumption

In Section 3.3, I show that a formal test of the common-trends assumption in the basic diff-in-diff analysis yields the result that the assumption holds and the diff-in-diff estimates are likely valid. In this section, I relax the common-trends assumption following the strategy proposed by Stephens and Yang (2014) and Aksu et al. (2018)—i.e., I include year-region interaction terms to capture the possibility of differential trends in high school enrollment rates across regions. There are 81 provinces, 26 NUTS2 regions, and 12 NUTS1 regions in Turkey. The region variable in the HLFS data is provided at NUTS2 level. So, it would be ideal to include the full set of year-NUTS2 region interactions into both the basic diff-in-diff and IV regressions. However, the treatment is defined as the change across NUTS2 regions over time. Therefore, there will be a collinearity problem. Instead, I go one level up and include year-NUTS1 region interactions into the regressions, which can still capture any potential region-specific shocks to demand for higher education.

The year-region (NUTS1) interactions are included into both the basic diff-in-diff and IV regressions along with all the controls. Table (11) summarizes the results. For the IV regressions, the results are very similar to the baseline estimates albeit slightly smaller—0.68 for males as opposed to 0. The basic diff-in-diff estimates are also very similar in nature to the baseline estimates except two nuances. First, the estimates are larger; and, second, the coefficient for females turns statistically significant. But, the main result is still there: the increase in high school enrollment is strongly driven by males. It should also be noted, however, that there are two NUTS1 regions in each of the treatment and control regions, while the IV analysis uses variation coming from all of the 12 NUTS1 regions and, therefore, may be more reliable.

4.2 Testing the potential impact of compulsory military service enrollment

Military service is compulsory in Turkey. All men—with no disabilities and with normal health and BMI values—become eligible for military service after they turn 19. The ones enrolled in school can defer their service, but unskilled males typically complete their military service (which is 12 months) after finishing high school.[16] The HLFS gets the age information by directly asking to the respondent and there may be slight inconsistencies between the actual age and the reported age. Moreover, there may be a tendency to under-report the age if the individual is unofficially escaping from military obligations.

One can argue that the refugee influx may have triggered military service enrollment among unskilled young males whose labor market opportunities are restricted. In other words, increased refugee concentration may have pushed eligible young males toward military service. The HLFS samples only non-institutional population; so, the ones enrolled in military service are out of the sample. Under the reasonable assumption that the rate of high school enrollment would be highest for age 15 and lowest for age 18 in the age interval 15-18, ignoring this military enrollment scenario may bias the estimates upward. To address this concern, I repeat the diff-in-diff and IV regressions by dropping age 18 and the results are reported in Table (12). Both the diff-in-diff and IV estimates are very similar to the baseline findings, which means that the military service enrollment concern does not bias the estimates.

4.3 The 2012 national education reform: Does it contaminate the estimates?

In April 2012, a new compulsory schooling law became effective in Turkey, which increased compulsory education from 8 to 12 years beginning in September 2012.[17] This is a nation-wide educational reform and the timing of the reform more or less coincides with the timing of the refugee influx. As a result, it is necessary to assess the potential impact of this reform on the estimates reported in this paper. The econometric strategy is designed with this concern in mind, although I haven't explicitly discussed it until now. In this section, how I address this concern will be made explicit and discussed thoroughly.

The **first** question is whether there is a jump right after the reform both in terms of national and regional enrollment figures. The upper left panel in Figure (5) plots the entire high school enrollment numbers from 2009 to 2015 for everyone, males, and females using the National Education Statistics Yearbooks. The figure suggests that there is a secular increase (no jump) and the rate

[16] See Torun and Tumen (2016) for more information about institutional details of compulsory military service in Turkey.

[17] See OECD (2013) for a compact explanation of the 2012 education reform along with a summary of the education system in Turkey.

of increase is quite similar for males and females. Although high school enrollment became compulsory after this reform, it is well known that the enforcement mechanism is quite weak and the enforcement problem is currently an issue of public debate. As Table (3) displays, the high school enrollment rate for age 15-18 is still below 75 percent as of 2016. A related question is whether a jump in high school enrollment is observed in the border regions, which are the main areas affected from the refugee influx. Based on Figures (4) and (5), the trends are smooth for both males and females in the treatment and control regions—as they are defined in the basic diff-in-diff analysis. The enrollment rates are much lower in these regions (around 60 percent) and enforcement is also substantially weaker relative to rest of the country.[18] So, if a young individual wants to opt out of high school, s/he can do it without much effort—even more easily in the eastern and southeastern regions. Since this is a national reform and the main empirical strategy used in this paper is a diff-in-diff, any homogeneous impact of the reform would be eliminated through differencing.

The **second** question is whether the reform differentially affected the border regions relative to the rest of the country. In particular, the level of development might be an important factor determining the pace of increase in enrollment—i.e., regions with low enrollment rates may respond faster than the ones with high enrollment rates. My basic diff-in-diff setting compares regions with quite low and similar high school enrollment rates as Figure (4) clearly shows. In addition, I directly control for several variables proxying regional differences in development levels such as real per capita GDP, real international trade volume, and region dummies. Relaxation of the common-trends assumption introduces region-specific trends in enrollment rates and also serves for this purpose. Moreover, in the IV analysis, I also include year-specific distance from each region to the nearest border crossing. As Del Carpio and Wagner (2015) and Akgunduz et al. (2018) argue, this can also account for any distance-to-border-related differential impact of the national education reform on high school enrollment rates. After employing all these additional analyses, I conclude that the results do not change and the existing estimates are not contaminated by the 2012 national education reform.

Finally, one can also assert that the government have invested in educational resources in regions

[18]In these regions, the size of informal economy, rate of informal employment, and rate of unauthorized electricity consumption are the highest in Turkey.

with high refugee concentration and this has increased educational opportunities; as a result, demand for high school education may have increased at a faster rate in these regions than the rest of the country. To address this concern, I collect information from the Ministry of Education database on the number of high schools and the number of high school teachers from 2006 to 2016 (for both general and vocational schools) and use this information in all of the regressions, which means that the results already account for the change in the regional availability of educational resources.

Note that if there is any differential impact of the national education reform on high school enrollment rates in the refugee-receiving regions (i.e., the border regions), then we should be able to observe it on females rather than males. Due to a combination of cultural, religious, social, and economic factors, female school enrollment and labor force participation rates are quite low in these regions relative to the rest of the country, which are already quite low in Turkey relative to the OECD countries. So, any mechanism forcing young people to high school education would do so for females in a much higher magnitude. In contrast, my estimates suggest that the increase come from young males coming from lower backgrounds, which supports the validity of the labor market channel I propose in this paper.

5 Concluding remarks

In this paper, I show that high school enrollment increased among native youth in Turkey as a response to Syrian refugee influx. The effect comes almost entirely from males, while the impact on female enrollment rates is zero in most specifications. I also find that the increase in male enrollment rates is more pronounced for lower parental backgrounds—i.e., parental education strictly less than high school and parents out of formal employment.

These findings are consistent with the hypothesis that the Syrian refugee influx increased competition for low-skill jobs, which pushed youth with lower parental backgrounds toward school education. Most Syrian refugees do not have work permit and they are more willing to work in low-pay informal jobs than natives. Oversupply of informal Syrian workers increases competition in the low-skill labor market and pushes potential wages down especially for manual informal jobs with low-skill requirements. These labor market forces increase demand for high school education among young males with lower parental backgrounds, who would otherwise enter the low-skill labor market. Overall, the findings are in line with Hunt (2017).

Taken at face value, these results point to educational upgrading, which is expected to positively affect the average well-being of Turkish natives in the long-term. From the viewpoint of refugees, the prospects look worse. In the short-term, informal labor market seems useful since it generates labor income for refugees. However, if this trend continues, the interdependence between refugees and informal jobs may increase in the long run, which implies reduced access to social security and other public benefits. Extreme reliance of refugees to informal labor market may reduce the pace of integration and may deepen inequality concerns for the second-generation refugees.

References

AFAD (2013): "Syrian Refugees in Turkey, 2013: Field Survey Results," Turkish Disaster and Emergency Management Presidency.

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.

Akgunduz, Y. E., M. van den Berg, and W. Hassink (2015): "The Impact of Refugee Crises on Host Labor Markets: The Case of the Syrian Refugee Crisis in Turkey," IZA Discussion Paper No: 8841.

Aksu, E., R. Erzan, and M. Kirdar (2018): "The Impact of Mass Migration of Syrians on the Turkish Labor Market," Unpublished manuscript, Bogazici University.

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.

Bagir, Y. K. (2017): "Impact of the Syrian Refugee Influx on Turkish Native Workers: An Ethnic Enclave Approach," MPRA Working Paper #80803.

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

Becker, G. S. (1993a): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Chicago, IL: University of Chicago Press.

——— (1993b): *A Treatise on the Family*, Cambridge, MA: Harvard University Press.

Berker, A. (2009): "The Impact of Internal Migration on Educational Outcomes: Evidence from Turkey," *Economics of Education Review*, 28, 739–749.

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.

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.

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. 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.

Cengiz, D. and H. Tekguc (2018): "Is It Merely a Labor Supply Shock? Impacts of Syrian Migrants on Local Economies in Turkey," UMass PERI Working Paper #454.

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.

Ceylan, E. S. and S. Tumen (2018): "Measuring Economic Destruction in Syria from Outer Space," The Forum: ERF Policy Portal, Blog article, 19 June 2018, https://theforum.erf.org.eg/2018/06/19/measuring-economic-destruction-syria-outer-space/.

Conti, G. and J. J. Heckman (2014): "Economics of Child Well-Being," in *Handbook of Child Well-Being*, ed. by A. Ben-Arieh, F. Casas, I. Frones, and J. Korbin, Dordrecht, NL: Springer, 363–401.

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.

Dustmann, C. and A. Glitz (2011): "Migration and Education," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, New York, NY: Elsevier, vol. 4, chap. 4, 327–439.

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

Erdogan, M. M. (2014): "Syrians in Turkey: Social Acceptance and Integration," Hacettepe University Migration and Politics Research Center.

Fakih, A. and M. Ibrahim (2016): "The Impact of Syrian Refugees on the Labor Market in Neighboring Countries: Empirical Evidence from Jordan," *Defence and Peace Economics*, 27, 64–86.

Fallah, B., C. Krafft, and J. Wahba (2018): "The Impact of Refugees on Employment and Wages in Jordan," ERF Working Paper #1189.

Farre, L., F. Ortega, and R. Tanaka (2018): "Immigration and the Public-Private School Choice," *Labour Economics*, 51, 184–201.

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

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.

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

Murray, T. J. (2016): "Public or Private? The Influence of Immigration on Native Schooling Choices in the United States," *Economics of Education Review*, 53, 268–283.

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

OECD (2013): "Education Policy Outlook: Turkey," http://www.oecd.org/education/EDUCATION%20POLICY%20OUTLOOK%20TURKEY_EN.pdf.

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.

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

Schneeweis, N. (2015): "Immigrant Concentration in Schools: Consequences for Native and Migrant Students," *Labour Economics*, 35, 63–76.

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.

Stephens, M. and D. Yang (2014): "Compulsory Education and the Benefits of Schooling," *American Economic Review*, 104, 1777–1792.

Torun, H. and S. Tumen (2016): "The Impact of Compulsory Military Service Exemption on Education and Labor Market Outcomes: Evidence from a Natural Experiment," *Economics of Education Review*, 54, 16–35.

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

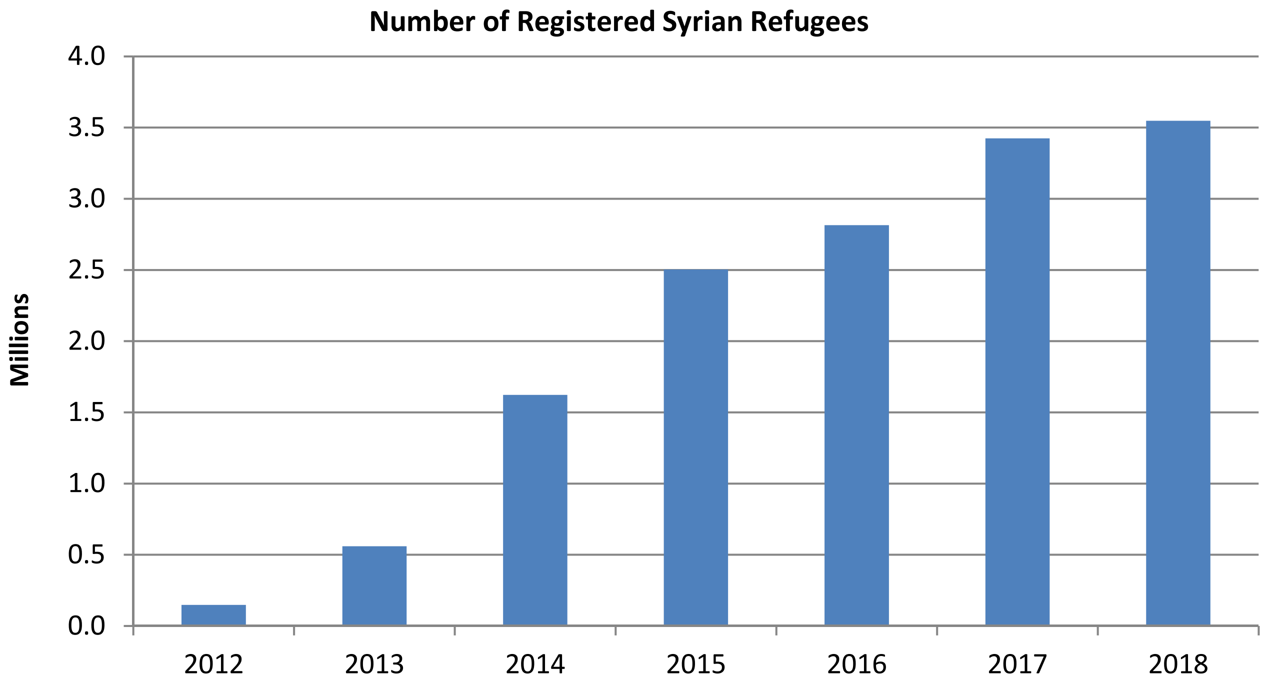


Figure 1: Number of registered Syrian refugees in Turkey. This figure plots the number of registered Syrian refugees in Turkey from 2012 to 2018—as of August 2018. The data sources are the UNHCR and the Government of Turkey. See: https://data2.unhcr.org/en/situations/syria/location/113.

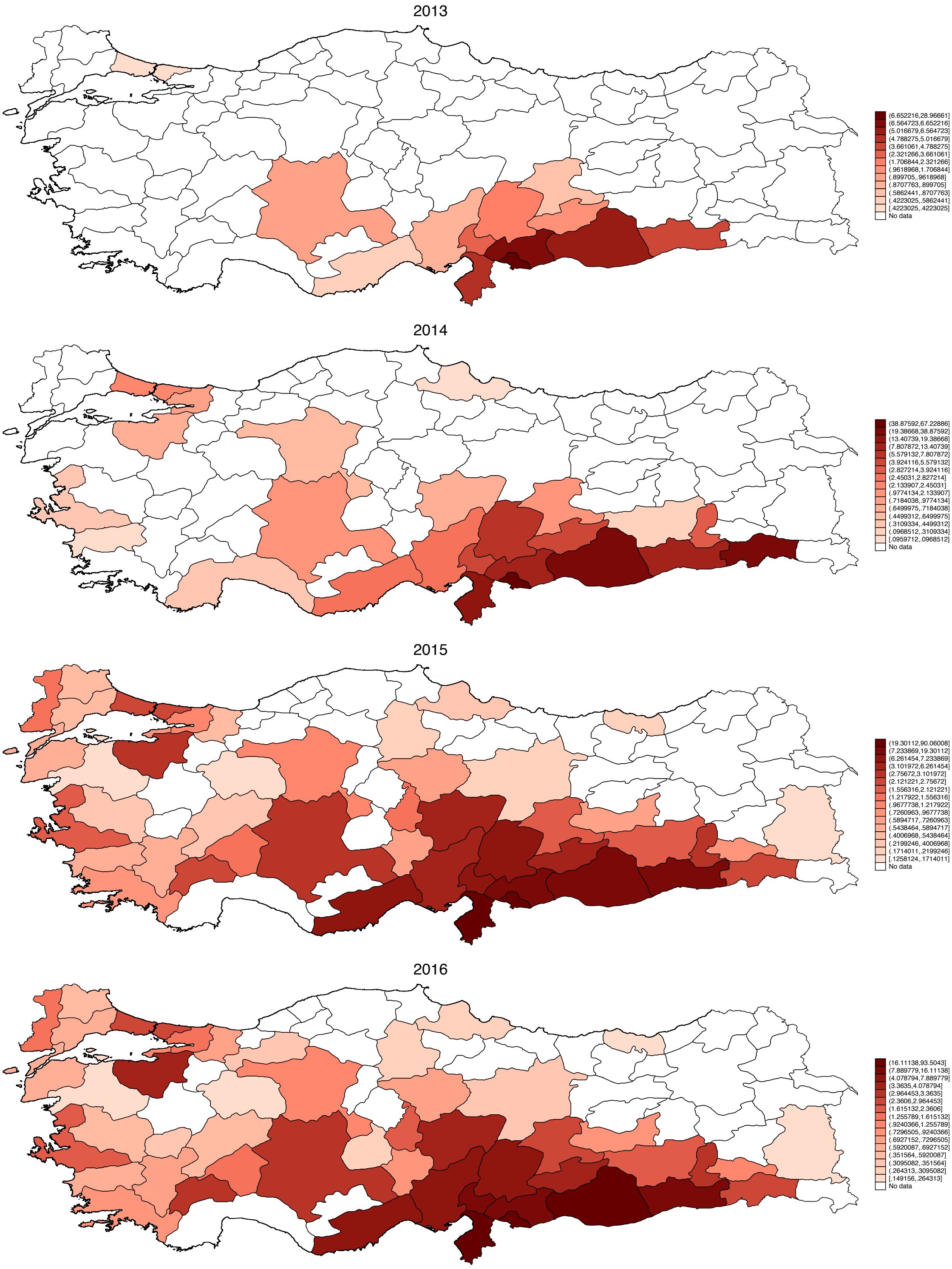


Figure 2: Regional refugee concentration. The figures display the refugee-to-population ratios in Turkish provinces from 2013 to 2016. To increase the visual accuracy of the figures, provinces with less than 1,000 refugees are assumed to have zero refugee-to-population ratio, although the estimations are performed based on full data. Zero refugee-to-population ratio is indicated with white color. The legends show the refugee-to-population ratios multiplied by 100. See Section 2 for a detailed explanation of the data sources.

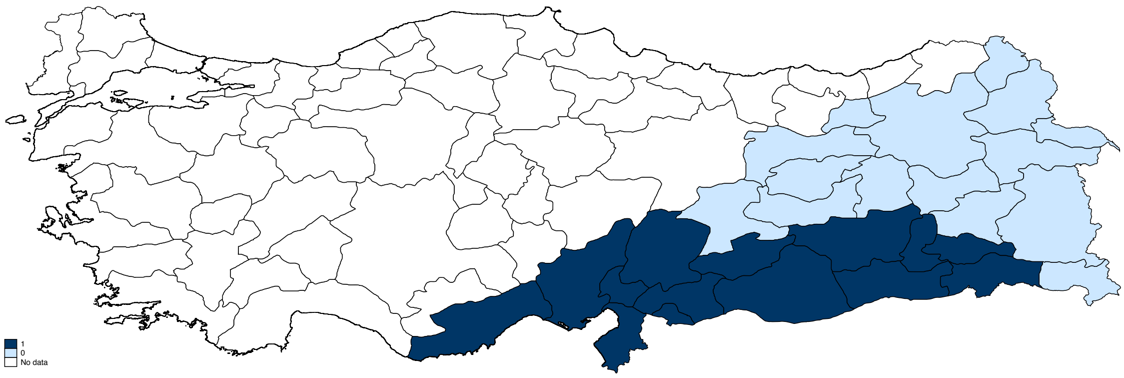
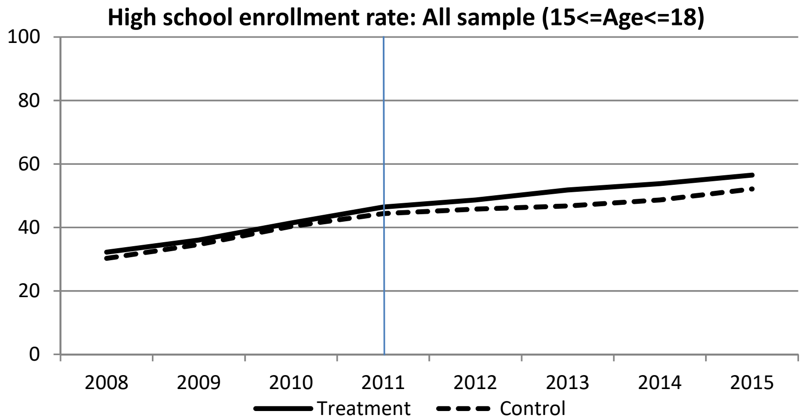
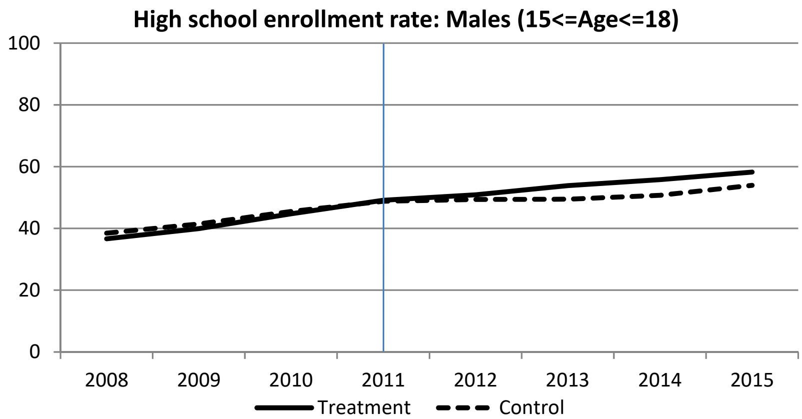


Figure 3: Treatment and control regions in the diff-in-diff specification. This figure shows the treatment and control regions in the diff-in-diff specification. The dark and light blue regions correspond to treatment ($T = 1$) and control ($T = 0$) regions, respectively. Based on the NUTS2 regional classification in Turkey, regions TR62 (Adana, Mersin), TR63 (Hatay, Kahramanmaras, Osmaniye), TRC1 (Gaziantep, Adiyaman, Kilis), TRC2 (Sanliurfa, Diyarbakir), and TRC3 (Mardin, Batman, Siirt, Sirnak) are the treatment regions; while regions TRA1 (Erzurum, Erzincan, Bayburt), TRA2 (Agri, Kars, Igdir, Ardahan), TRB1 (Malatya, Elazig, Bingol, Tunceli), and TRB2 (Van, Mus, Bitlis, Hakkari) are the control regions.





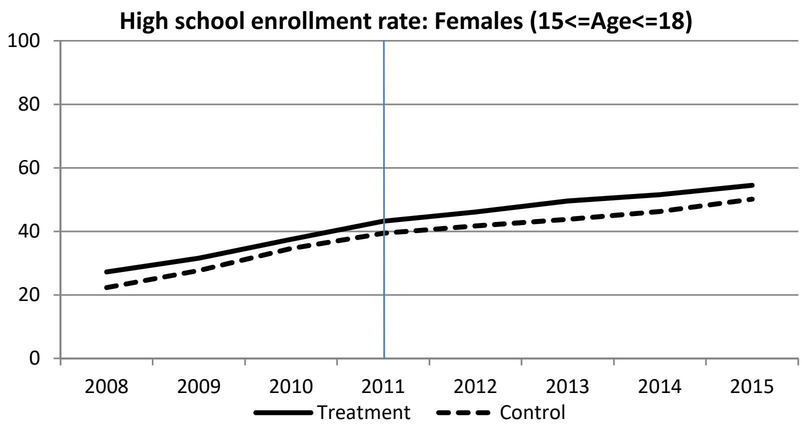


Figure 4: Raw high school enrollment rates before and after the refugee influx (calculated from HLFS). To highlight the trends, the raw enrollment rates are expressed as three-year moving averages—for example, the enrollment rate for 2008 is the average of 2006, 2007, and 2008. The blue vertical line indicates the date that the refugee influx started. The data source is the Turkish Household Labor Force Survey.

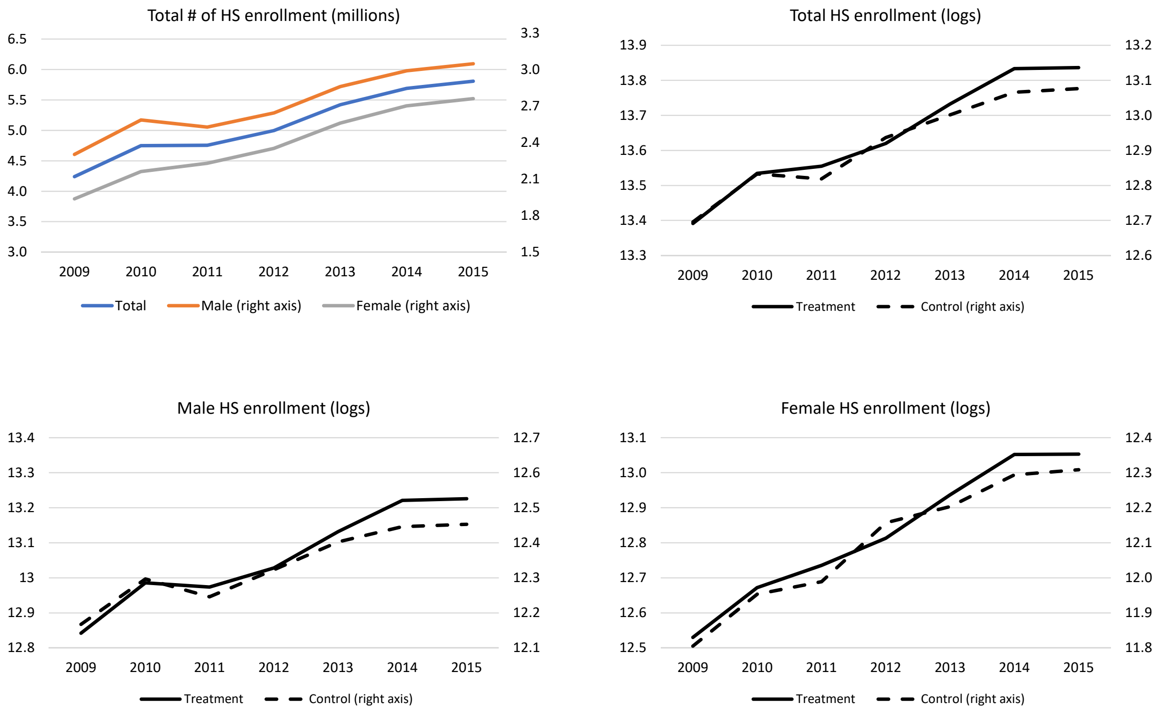


Figure 5: Trends in high school enrollment rates (calculated from National Education Statistics). The figures display the trends in high school enrollment rates from 2009 to 2015. Data source is the National Education Statistics Yearbooks of the Ministry of National Education. The upper left figure plots the trends for the entire country, while the rest of the figures compare the trends in treatment and control regions. The range of left and right axes are set in such a way that the trends are comparable across series.

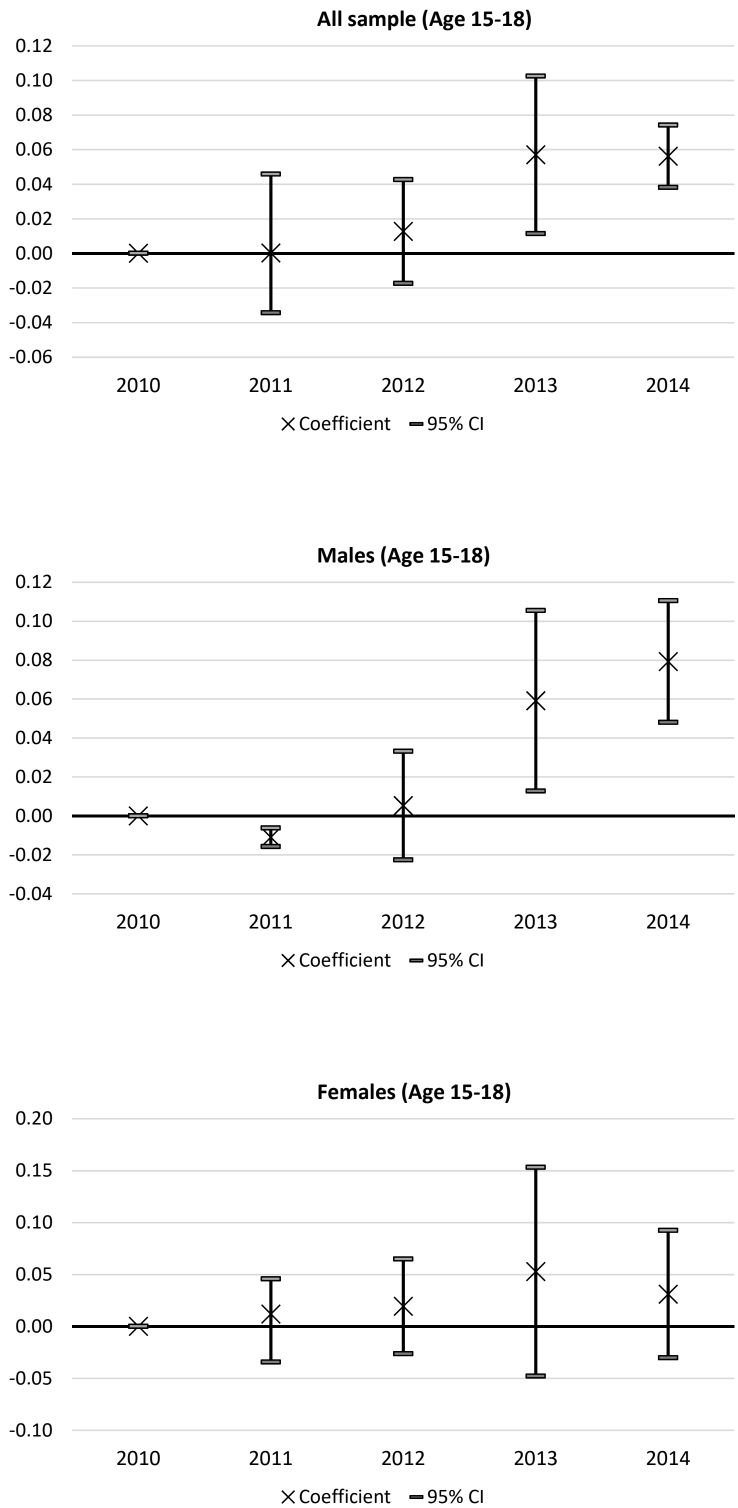


Figure 6: Testing the common trends assumption. Estimated coefficients of the interaction between treatment and year dummies are plotted together with the 95% confidence intervals—following Autor (2003). Standard errors are clustered at region-year level. The estimation procedure is described in Section 4.1.

Refugee vs native characteristics (2016)

	Natives	Refugees
Age		
0-14	23.7	41.7
15-64	68.0	56.5
65+	8.3	1.8
Gender		
Male	50.2	53.1
Female	49.8	46.9
Education		
Unknown	0.7	27.8
No degree	14.9	44.4
Less than high school	51.6	22.3
High school	19.2	3.5
College & above	13.7	2.0

Table 1: Comparison between refugee vs native characteristics as of 2016. The numbers indicate percentages. The summary data on refugee characteristics comes from the Ministry of Interior, Directorate General of Migration Management. The data for natives come from the Turkish Statistical Institute. The age and gender statistics are compiled from the Address Based Population Registration System Results. The education statistics are taken from the National Education Statistics Database.

Summary statistics (diff-in-diff specification)

	2010-2011		2012-2014	
	Treatment	Control	Treatment	Control
Male	0.528	0.512	0.516	0.526
Age 15	0.264	0.279	0.269	0.267
Age 16	0.263	0.264	0.264	0.274
Age 17	0.255	0.252	0.252	0.252
Age 18	0.218	0.205	0.216	0.207
Enrolled in high school	0.529	0.494	0.592	0.549
Parent less than high school	0.925	0.916	0.917	0.911
Parent high school & above	0.075	0.084	0.083	0.089
Intact parents	0.913	0.924	0.920	0.942
Household size	6.502	7.053	5.728	5.882
Parent formally employed	0.305	0.290	0.353	0.333
Parent informally employed	0.327	0.422	0.283	0.431
Parent unemployed	0.076	0.059	0.075	0.047
Parent not in labor force	0.292	0.229	0.289	0.189
# of observations	15,688	9,269	23,660	14,910

Table 2: Summary statistics for the diff-in-diff specification (2010-14). Individuals of high-school age ($15 \leq \text{age} \leq 18$) are included into the calculations. Each cell in the table indicates the mean of the corresponding variable. "Parent" refers to the head of the household. The data source is the Turkish Household Labor Force Survey.

Summary statistics (IV specification)

	2010	2011	2014	2015	2016
Male	0.519	0.517	0.523	0.520	0.525
Age 15	0.272	0.265	0.268	0.273	0.271
Age 16	0.266	0.268	0.275	0.260	0.272
Age 17	0.257	0.259	0.259	0.267	0.256
Age 18	0.205	0.208	0.198	0.200	0.201
Enrolled in high school	0.606	0.622	0.696	0.722	0.746
Parent less than high school	0.870	0.865	0.860	0.854	0.839
Parent high school & above	0.130	0.135	0.140	0.146	0.161
Intact parents	0.913	0.914	0.920	0.922	0.925
Household size	5.529	5.469	5.122	5.361	5.173
Parent formally employed	0.411	0.444	0.480	0.485	0.502
Parent informally employed	0.277	0.273	0.261	0.256	0.248
Parent unemployed	0.073	0.057	0.054	0.057	0.059
Parent not in labor force	0.240	0.226	0.205	0.202	0.191
# of observations	35,542	35,324	36,140	35,772	33,619

Table 3: Summary statistics for the IV specification (2010-11, 2014-16). Individuals of high-school age ($15 \leq \text{age} \leq 18$) are included into the calculations. Each cell in the table indicates the mean of the corresponding variable. "Parent" refers to the head of the household. The data source is the Turkish Household Labor Force Survey.

Diff-in-diff estimation: Baseline specification

	2010-2014			2009-2015		
	All	Male	Female	All	Male	Female
Refugee effect	0.0109	0.0329*	-0.0134	0.0186	0.0380**	-0.0031
	(0.0131)	(0.0189)	(0.0151)	(0.0159)	(0.0185)	(0.0202)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
# of obs.	63,527	33,083	30,444	90,405	46,976	43,429
R^2	0.0878	0.0794	0.0988	0.0939	0.0813	0.1075

Table 4: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 ($15 \leq \text{age} \leq 18$), who are not enrolled in college. The treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions. In the first three columns, the pre-influx and post-influx years are 2010-2011 and 2012-2013-2014, respectively. In the last three columns, the pre-influx and post-influx years are defined as 2009-2010-2011 and 2012-2013-2014-2015, respectively. The "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Diff-in-diff estimation: Parental, regional, and capacity controls are included

	Parental controls			Regional controls added			Capacity controls added		
	All	Male	Female	All	Male	Female	All	Male	Female
Refugee effect	0.0269*	0.0460**	0.0049	0.0278*	0.0429**	0.0097	0.0357*	0.0438**	0.0264
	(0.0159)	(0.0205)	(0.0153)	(0.0160)	(0.0191)	(0.0188)	(0.0206)	(0.0221)	(0.0209)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Per capita real GDP (log)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Real trade volume (log)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
# of high schools (log)	No	No	No	No	No	No	Yes	Yes	Yes
# of high school teachers (log)	No	No	No	No	No	No	Yes	Yes	Yes
# of obs.	63,527	33,083	30,444	63,527	33,083	30,444	63,527	33,083	30,444
R^2	0.1607	0.1441	0.1810	0.1607	0.1443	0.1814	0.1610	0.1446	0.1819

Table 5: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 (15≤age≤18), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. The treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions. The pre-influx and post-influx years are defined as 2010-2011 and 2012-2013-2014, respectively. The "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Diff-in-diff estimation: The role of parental education

	Parental education < HS			Parental education ≥ HS		
	All	Male	Female	All	Male	Female
Refugee effect	0.0376**	0.0444*	0.0292	0.0040	0.0018	0.0053
	(0.0202)	(0.0229)	(0.0212)	(0.0243)	(0.0318)	(0.0252)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes	Yes	Yes
Per capita real GDP (log)	Yes	Yes	Yes	Yes	Yes	Yes
Real trade volume (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high schools (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high school teachers (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of obs.	58,264	30,392	27,872	5,263	2,691	2,572
R^2	0.1358	0.1200	0.1567	0.2481	0.2351	0.2604

Table 6: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 ($15 \leq \text{age} \leq 18$), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. The treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions. The pre-influx and post-influx years are defined as 2010-2011 and 2012-2013-2014, respectively. The "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Diff-in-diff estimation: The role of parental employment (males only)

	Formal emp.	Informal emp.	Unemp.	Not in LF
Refugee effect	0.0143	0.0532**	0.0935**	0.0685***
	(0.0398)	(0.0225)	(0.0447)	(0.0202)
Year dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes
Per capita real GDP (log)	Yes	Yes	Yes	Yes
Real trade volume (log)	Yes	Yes	Yes	Yes
# of high schools (log)	Yes	Yes	Yes	Yes
# of high school teachers (log)	Yes	Yes	Yes	Yes
# of obs.	10,853	11,621	2,149	8,640
R^2	0.1546	0.1127	0.1303	0.0992

Table 7: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes males between age 15 and 18 ($15 \leq age \leq 18$), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. The treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions. The pre-influx and post-influx years are defined as 2010-2011 and 2012-2013-2014, respectively. The "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Diff-in-diff estimation: Placebo before/after dates

After influx Before influx	2007/08 2005/06	2009/10 2005/06	2010/11 2005/06	2008/09 2006/07	2010/11 2006/07	2010/11 2008/09
Refugee effect	-0.0027	-0.0084	-0.0043	-0.0050	-0.0008	-0.0021
	(0.0144)	(0.0087)	(0.0052)	(0.0114)	(0.0045)	(0.0166)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes	Yes	Yes
Per capita real GDP (log)	Yes	Yes	Yes	Yes	Yes	Yes
Real trade volume (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high schools (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high school teachers (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of obs.	35,705	37,089	37,171	49,979	49,000	48,893
R^2	0.1958	0.1871	0.1905	0.1904	0.1973	0.1733

Table 8: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 (15≤age≤18), who are not enrolled in college. The treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions. The "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

IV estimation (2SLS)

	Baseline estimates			Controls added		
	All	Male	Female	All	Male	Female
Refugee share (OLS)	-0.1714	0.1364	-0.5065**	-0.0367	0.2595*	-0.3547*
	(0.1758)	(0.1464)	(0.2506)	(0.1513)	(0.1551)	(0.1873)
Instrument (Red. form)	0.0029*	0.0059***	-0.0003	0.0037**	0.0067***	0.0005
	(0.0016)	(0.0014)	(0.0018)	(0.0015)	(0.0005)	(0.0028)
1st stage	0.0071***	0.0071***	0.0070***	0.0086***	0.0086***	0.0086***
	(0.0027)	(0.0028)	(0.0027)	(0.0022)	(0.0022)	(0.0022)
Refugee share (2SLS)	0.4103**	0.8297***	-0.0374	0.4262*	0.7780***	0.0612
	(0.1707)	(0.1962)	(0.2429)	(0.2285)	(0.1724)	(0.3346)
F-statistic	6.79	6.73	6.85	15.20	15.07	15.36
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Log distance	Yes	Yes	Yes	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household size	No	No	No	Yes	Yes	Yes
Parental education	No	No	No	Yes	Yes	Yes
Intact family	No	No	No	Yes	Yes	Yes
Real trade volume (log)	No	No	No	Yes	Yes	Yes
# of high schools (log)	No	No	No	Yes	Yes	Yes
# of high school teachers (log)	No	No	No	Yes	Yes	Yes
# of obs.	176,397	91,880	84,517	176,397	91,880	84,517

Table 9: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 ($15 \leq \text{age} \leq 18$), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. 26 NUTS2 regions are included in all regressions. Years of observation include 2010, 2011, 2014, 2015, and 2016. The regressions include a year-specific distance (in natural logs) to the nearest border crossing. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

IV estimation (2SLS): The role of parental education

	Parental education < HS			Parental education ≥ HS		
	All	Male	Female	All	Male	Female
Refugee share (OLS)	-0.0685	0.2427	-0.4013**	0.1465	0.1664	0.0852
	(0.1500)	(0.1541)	(0.1937)	(0.2599)	(0.1751)	(0.4133)
Instrument (Red. form)	0.0022	0.0058***	0.0028	0.0111***	0.0078**	0.0143***
	(0.0017)	(0.0007)	(0.0035)	(0.0026)	(0.0039)	(0.0037)
1st stage	0.0085***	0.0086***	0.0085***	0.0097***	0.0096***	0.0098***
	(0.0023)	(0.0023)	(0.0022)	(0.0018)	(0.0018)	(0.0018)
Refugee share (2SLS)	0.2521	0.6823***	-0.1867	1.1411***	0.8119*	1.4333***
	(0.2337)	(0.1488)	(0.3503)	(0.3943)	(0.4941)	(0.4681)
F-statistic	14.20	14.09	14.32	28.97	28.80	29.10
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Log distance	Yes	Yes	Yes	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes	Yes	Yes
Log trade volume	Yes	Yes	Yes	Yes	Yes	Yes
# of high schools (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high school teachers (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of obs.	151,315	78,957	54,176	25,082	12,923	12,159

Table 10: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 (15≤age≤18), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. 26 NUTS2 regions are included in all regressions. Years of observation include 2010, 2011, 2014, 2015, and 2016. The regressions include a year-specific distance (in natural logs) to the nearest border crossing. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

The common-trends assumption is relaxed

	Diff-in-Diff			IV		
	All	Male	Female	All	Male	Female
Refugee effect	0.0512***	0.0643***	0.0393**	0.3221	0.6826***	-0.0759
	(0.0136)	(0.0186)	(0.0184)	(0.2228)	(0.2205)	(0.2797)
F-statistic	-	-	-	85.57	93.90	77.40
R^2	0.1616	0.1452	0.1834	-	-	-
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region (NUTS1) $\times$ Year	Yes	Yes	Yes	Yes	Yes	Yes
Log distance	No	No	No	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes	Yes	Yes
Per capita real GDP (log)	Yes	Yes	Yes	No	No	No
Real trade volume (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high schools (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high school teachers (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of obs.	63,527	33,083	30,444	176,397	91,880	84,517

Table 11: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 18 ($15 \leq \text{age} \leq 18$), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. *For the diff-in-diff analysis*, the treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions; the pre-influx and post-influx years are defined as 2010-2011 and 2012-2013-2014, respectively; the "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. *For the IV analysis*, 26 NUTS2 regions are included in all regressions; years of observation include 2010, 2011, 2014, 2015, and 2016; the IV regressions include a year-specific distance (in natural logs) to the nearest border crossing. Region-specific (NUTS1 level) year effects are included to relax the common-trends assumption. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Testing the effect of military service enrollment: Age 18 is excluded

	Diff-in-Diff			**IV**		
	All	Male	Female	All	Male	Female
Refugee effect	0.0295	0.0399*	0.0166	0.3937	0.9034***	-0.1660
	(0.0199)	(0.0212)	(0.0210)	(0.3548)	(0.2505)	(0.5300)
F-statistic	-	-	-	15.21	15.37	15.08
R^2	0.1555	0.1357	0.1810	-	-	-
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Log distance	No	No	No	Yes	Yes	Yes
Gender dummies	Yes	No	No	Yes	No	No
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household size	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Intact family	Yes	Yes	Yes	Yes	Yes	Yes
Per capita real GDP (log)	Yes	Yes	Yes	Yes	Yes	Yes
Real trade volume (log)	Yes	Yes	Yes	No	No	No
# of high schools (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of high school teachers (log)	Yes	Yes	Yes	Yes	Yes	Yes
# of obs.	50,019	25,811	24,208	140,627	72,324	68,303

Table 12: The dependent variable is a dummy taking 1 if the individual is enrolled in high school and 0 otherwise. The sample includes individuals between age 15 and 17 ($15 \leq \text{age} \leq 17$), who are not enrolled in college. Household size is the number of family members living in the same house. Parental education is the final educational degree—defined by six categories ranging from no degree to college & above—obtained by the household head. Intact family is a dummy variable taking 1 if the parents live together and 0 otherwise. *For the diff-in-diff analysis*, the treated regions are Adana, Gaziantep, Hatay, Mardin, and Sanliurfa NUTS2 regions, while the control regions include Malatya, Erzurum, Agri, and Van regions; the pre-influx and post-influx years are defined as 2010-2011 and 2012-2013-2014, respectively; the "refugee effect" variable is obtained by interacting the treated-untreated and before-after dummies. *For the IV analysis*, 26 NUTS2 regions are included in all regressions; years of observation include 2010, 2011, 2014, 2015, and 2016; the IV regressions include a year-specific distance (in natural logs) to the nearest border crossing. Standard errors are clustered at region-year level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.