Is It Merely A Labor Supply Shock?

Impacts of Syrian Migrants on Local Economies in Turkey

Doruk Cengiz,[*] Hasan Tekguc[†]

Abstract

We use a large and geographically varying inflow of over 2.5 million Syrian migrants in Turkey between 2012 and 2015 to study the effect of migration on local economies. Using recently available province-level residence data of Syrian population in Turkey, we do not find adverse employment or wage effects for native-born Turkish workers overall, or those without a high school degree. These results are robust to a range of strategies to construct reliable control groups. On the other hand, we find evidence for a number of channels indicating demand side effects of migration that helped offset the impact of a labor supply shock. Turkish workers' participation in the formal sector rose in response to the migration, consistent with complementarity of migrants and native born workers. In addition, migration led to an increase in residential building construction, with the number of new dwelling units increasing by more than 33%. Finally, Syrian migration brought in capital and entrepreneurs to the host regions, spurring new business creation: the migration led to a more than 24% increase in new companies, reflecting an increase in both Syrian-founded and non-Syrian founded companies. Our findings suggest that migration-induced increases in regional demand and capital supply enable local labor markets to absorb inflow of migrant labor, and prevent sizable wage decline or job loss for native workers.

Keywords: Syrian Immigration, Labor Market, Construction Industry, Firm Formation, Generalized Synthetic Control

Jel Codes: F20, J61, R21

[*]University of Massachusetts, Amherst

[†]Kadir Has University

We thank Arindrajit Dube, Ina Ganguli, Erin Conlon, İpek İlkkaracan, participants at 2nd TIPES Interdisciplinary Workshop, 2017 EEA Conference, 2017 The New School-UMass Economics Graduate Student Workshop for very helpful comments.

1 Introduction

The Syrian Civil War started in 2011, has led more than 5 million Syrians to leave their country. Such a huge displacement affects countries worldwide. As a result, the topic of "immigration" has taken the center stage in political debates. The primary question for large segments of societies is: "What will happen to my job if they come here?" The canonical economic model paints a relatively pessimistic picture for the affected groups; predicts a wage, and potentially employment, decline (Borjas, 2013; Dustmann et al., 2016). The empirical evidence, however, is mixed, and the debate on the labor market effects of the migration still continues (Akgündüz et al., 2015b; Borjas and Monras, 2016; Card, 2009; Clemens and Hunt, 2017; Peri, 2016).

In this paper, we contribute to this debate by exploiting an unusually large and sudden migration flow of more than 2.5 million Syrians into Turkey between 2012 and 2015, displayed in figure 1; and assess whether effects of the migration can be reduced to a labor-supply shock. We examine the effects of the Syrian migrants on Turkish workforce and on the native wage distribution. We provide evidence for the lack of adverse employment and wage effects on affected native workers. Regarding the migration solely as a labor-supply shock does not coincide with the empirical findings. The descriptive supply-demand framework requires a migrant-induced rise in labor-demand that offsets the supply increase (Constant, 2014; Peri, 2014). To document the labor-demand rise, we also provide evidence on two of the mechanisms, rise in housing demand and new firm formation, that may enable local labor markets to absorb the labor-supply shock.

One main empirical challenge in estimating effects of migration is that migrants potentially prefer to go to regions that are experiencing an economic boom, so the underlying pre-existing trends might bias the estimates. This is probably less relevant in the case, since the primary reason for the migration is the war in the home country, and Syrian migrants in Turkey reside largely in the border regions (see figure 2). However, there is no certainty that these regions and the rest of the country have been following a similar path. To address the potential endogeneity, we employ the generalized synthetic control method (GSC) that purges underlying pre-existing trends using a data-driven procedure (Xu, 2017), in addition to standard difference-in-differences and two-stage least squares models.

As initial evidence of the impacts, we present estimates for employment and wage effects of the forced migration from Syria on natives with informal jobs, with less than high school degree (LTHS), and with high school and above education (HSG). More than 90% of the migrants do not have a high school degree. They cannot formally work, thus they have entered Turkish labor market through informal employment. Nevertheless, the entry has not adversely affected native workers. We show that the average wage of the informal jobs has considerably declined; yet LTHS natives did not experience an employment or wage loss. On the contrary, they have been able to move to formal employment, suggesting imperfect substitution between migrant and native labors. The migration has benefited HSG natives, significantly increased their wages. Examining the wage distribution corroborates these results. We find that the migration has increased wages of some native workers to the national minimum wage who would otherwise earn below the minimum wage. In addition, it enlarged the share of upper-middle income workers, and had virtually no effect on top wage-earners.

The non-finding of an adverse wage or employment effect even after such a massive shock indicates a counteracting labor-demand rise. In exploring the mechanisms that may offset the labor supply shock, we first consider developments in the residential construction sector. Based on the fact that the industry is relatively large and low-skilled labor intensive, the migrant-induced demand increase for the residential buildings can absorb a portion of the labor supply shock. This is particularly relevant in the case, since approximately 90% of the migrants reside outside Temporary Protection Camps (TPC). We estimate that the migration has a major positive impact on the residential construction sector. The construction of new dwelling units has risen by more than 33% in the host provinces.

An alternative channel that enables local labor markets to absorb the labor-supply shock is the rise in new firm formation. Karahan et al. (2016) notes that migration in the U.S. attracts capital to host regions, and leads to an increased new firm formation. This effect is valid also in Turkey. We estimate that the total number of new firms in the host provinces has risen by 24%. The increase is partly due to Syrian entrepreneurs. New firms with at least one Syrian co-founder has increased from 81 in 2011 to 1,599 in 2015. On the other hand, the new firm formation with no Syrian co-founders has also increased, by more than 20%, reflecting the Syrian entrepreneurs have not crowded-out non-Syrians.

The canonical accounting framework in the immigration literature employs a constant elasticity of substitution production function. In section 6, we examine whether the effects of the migration can be reduced to a labor-supply shock. We show that the theoretical predictions overshoot the empirical wage decline estimates even when capital is assumed to adjust perfectly and a moderate level of imperfect substitution between native and migrant laborers is allowed. This indicates that other parameters of the framework are also affected by migration. Our findings on the effects on residential building and new firm formation suggest that the distribution parameter between low-skilled and high-skilled workers, and the technology parameter also change with migration.

Currently, the debate on the impact of the Syrian migration to Turkey revolves around its effects on the employment. The findings from empirical studies on the subject are mixed. Three of the studies that are closely related to ours are Akgündüz et al. (2015b), Del Carpio and Wagner (2015), and Tumen (2016).[1] The former study argues that there is no significant employment effect on native workers; while the latter two claim a significant decline in informal employment. Our re-examination of the latter two studies reveals that the discrepancies are largely due to the difference in selected control regions, and the methods of statistical inference. In Appendix B, we replicate the baseline estimate of Tumen (2016). Our analysis shows that (i) the control regions of their baseline regression model have followed a path that is dissimilar to the rest of the country, and the findings qualitatively change for alternative control regions; and (ii) allowing for within region correlation of the errors prevents us from rejecting the no-effect hypothesis. Del Carpio and Wagner (2015) do not address the serial correlation either, hence their estimates are potentially too precise as well. In addition, based on the reported results in the study, the implied counterfactuals of the treated regions do not follow a similar path as the actual ones before the migration shock, weakening the causal interpretation.

On the other hand, the literature on quasi-experimental migration shocks generally focuses on changes in wages of natives whose education levels are similar to the migrants. Our findings mostly coincide with those that estimate no or small adverse effects on the native wages (Clemens and Hunt, 2017; Peri and Yasenov, 2017). Our baseline 95% confidence interval indicates that the average wage of native LTHS workers has not declined by more than 4.8%, and has not risen by more than 5.6% in the regions where the migrant population is at least 10% of natives. Given the size of the shock, this rules out some of the estimates in the literature obtained from different refugee waves, including -1.3 wage elasticity in table 3 of Borjas and Monras (2016), and the implied wage elasticities ranging

[1] An expanded version of Tumen (2016) is published as Ceritoglu et al. (2017).

between -0.5 and -1.5 in Borjas (2015) for *Marielitos* in the US in the early 1980s.

The rest of the paper is organized as follows. Section 2 lays out the canonical model and its underlying assumptions. The dataset is presented in Section 3. Section 4 develops the empirical methodology, briefly explains the generalized synthetic control method, and the inference methods. Section 5 presents the empirical findings about the effects of the migration on native workforce and wage distribution, on the residential building construction sector, and on the new firm establishments. Section 6 discusses the empirical findings using the canonical framework. We conclude in section 7.

2 Migrant-induced labor demand shocks integrated into the canonical production accounting framework

This section presents the canonical production function and examines the underlying assumptions that limit the analysis of migration to the labor supply shock.[2] The model is presented in Borjas (2013) and Borjas (2014). It is the standard two-level nested constant elasticity of substitution (CES) production function where aggregate production, Q, depends on the capital stock, K, and the number of laborers in efficiency unit, L. In calculating L, another CES function is employed to homogenize different types of labor (H subscript for high-skilled and L for low-skilled workers). Then, the production function is;

$$Q = A((1-\alpha)K^{\rho} + \alpha L^{\rho})^{\frac{1}{\rho}} \tag{1a}$$

$$L = (\theta L_L^{\gamma} + (1-\theta)L_H^{\gamma})^{\frac{1}{\gamma}} \tag{1b}$$

where $\rho < 1$, $\gamma < 1$, α (κ) corresponds to the distribution parameter between K and L (L_L and L_H), and A is the residual (factor neutral technology coefficient). Assuming that each factor is paid according to its marginal contribution, the real wage of a low-skilled laborer is $w_{L_L} = \alpha L^{\rho-1} Q^{1-\rho} A^{\rho} \theta L_L^{\gamma-1} L^{1-\gamma}$.

Low-skilled migration causes L_L (low-skilled workers) to increase. Its effect on the wages of native low-skilled workers can be found by the following:

[2]We do not present a survey of the immigration literature. For a comprehensive literature review see Kerr and Kerr (2011) as well as Borjas (2014).

$$
\begin{aligned}
dlnw_{L_N} = dln(\alpha) + (\rho - 1)dln(L) + dln(A) + (1-\rho)(s_L dln(L) + s_K dln(K)) + \\
dln(\theta) + (\gamma - 1)dln(L_L) + (1-\gamma)dln(L).
\end{aligned} \tag{2}
$$

Three of the assumptions made by the canonical model in the short-run are the following: (1) low-skilled immigrants and natives are perfectly substitutable; (2) the change in L_L affects only L and Q in the short-run, with the latter effect is through the former; and (3) capital adjustment is perfect ($dlnL = dlnK$) in the long-run.[3] Then, in the short-run, we have

$$
dln(A) = dln(\alpha) = dln(\theta) = dln(K) = 0,\ dln(Q) = dln(L) * s_L, \tag{3}
$$

where s_L is the share of (homogenized) labor. Assuming diminishing marginal returns, since percentage increase in homogenized labor is higher than that of aggregate production ($s_L < 1$), as well as the percentage increase in low-skilled labor ($dln(L_L)$) is larger than that of homogenized labor ($dln(L)$), the change in the average wage is always negative. For the given assumptions, any low-skilled migration is predicted to lead a decline in wages of low-skilled native workers.

Peri and Sparber (2009) and Ottaviano and Peri (2012) show empirically that native and immigrant workers are not perfectly substitutable. The model can incorporate less than perfect substitution between natives and immigrants by adding one more level of CES aggregator;

$$
L_L = (\omega_L L_{L,N}^{\delta_L} + (1-\omega_L) L_{L,I}^{\delta_L})^{\frac{1}{\delta_L}}, \tag{4}
$$

where $L_{L,N}$ indicates supply of low-skilled native workers, $L_{L,I}$ supply of low-skilled migrant workers, and the elasticity of substitution between low-skilled native and migrant workers is $\sigma_{L_{L,I},L_{L,N}} = \frac{1}{1-\delta_L}$.[4] Then, native low-skilled wages are $w_{L_{L,N}} = \alpha L^{\rho-1} Q^{1-\rho} \theta L_L^{\gamma-1} L^{1-\gamma} \omega L_{L,N}^{\delta_L - 1} L_L^{1-\delta_L}$. Assuming the supply of native workers does not change, the effect of migration on the marginal product of low-skilled native workers is;

$$
dlnw_{L_{L,N}} = dlnw_{L_L} + dln(\omega_L) + (1-\delta_L)dln(L_L), \tag{5}
$$

[3] Certainly, this list of assumptions is by no means exhaustive.

[4] Simply adding up the number of low-skilled natives and migrants to obtain L_L implicitly assumes $\delta_L = 1$, where natives and immigrants are perfectly substitutable.

where the last term, strictly positive, diminishes the size of the predicted negative effect.

The main argument of this paper complements the growing literature on native-immigrant complementarity. We claim that migrants bring their purchasing power, wealth to the host country. On the one hand, in the short-run, the increased purchasing power attracts capital, and migrant and native entrepreneurs establish new firms. So, the capital stock might increase in the short-run as well ($dlnK \neq 0$). Additionally, the boost in the new firm formation can lead to productivity increases ($dlnA > 0$) (Decker et al., 2014). On the other hand, the migration also increases the relative demand for the basic needs goods, such as residential buildings, potentially affect the relative sizes of industries, and change the skill intensities of the aggregate production function in the host regions ($dln\theta \neq 0$).

3 Legal status of the migrants, and data

The first Syrian migrants who seek refuge in Turkey have arrived in 2012, the official number has reached over a million in 2014 and continued to increase in 2015. Syrians running away from the civil war are not eligible to apply for refugee status and considered as "guests" in Turkey (Özden, 2013). Guests cannot seek asylum in another country and cannot work formally. In 2014, Turkish government started to distribute identity cards that enabled them to access to certain services including aid, healthcare, and education. The cards do not allow them to work formally, though. Before 2016, employers could apply for work permits only if the potential Syrian employee had entered Turkey with a valid visa. This policy prevented almost all Syrian guests from working formally. By January 2016, only 7,351 of them had work permits. The number, on the other hand, greatly underestimates the number of Syrians employed. Üstun (2016) reports that approximately 400,000 Syrians are employed informally in 2015.

We get the data on the number of Syrian guests in Turkey from the Ministry of Interior Directory General of Migration Management (MoI) database.[5] The available data on the total number of Syrian guests in Turkey starts from 2011, yet it is potentially underestimated before 2014. We pick 2012 as the first year of the wave. Since 2015, MoI reports the number of Syrian migrants at

[5]We provide detailed variable descriptions and data sources in appendix table A.1.

province level; their age and educational distribution at national-level.[6,7] We acquire the historical distribution of Arabic speaking population from the 1965 Turkish Census.

For labor market effects of the migration, we rely on statistics published by TURKSTAT, the official statistical institute of Turkey. Using 3,921,420 individuals aged between 15-64 from 2004-2015 Household Labor Force Surveys (LFS), we generate annual employment counts at the NUTS-2 regional level.[8] The survey does not include guests or refugees, so the analysis focuses exclusively on citizen workers.[9]

The wage variable from the same data set reports individual's monthly wage income from the primary job.[10] In TURKSTAT LFS, the question is asked only if the individual declares herself as a wage worker, hence excludes self-employed individuals. As a result, 1,010,230 observations, approximately a fourth of the sample, report positive wages. The number of observations drop to less than 1,000 for couple of NUTS-2 regions in some years. This affects the accuracy and the precision of the estimates, and is potentially one of the reasons why the point of contention is the employment effect, instead of the wage effect.[11]

In Turkey, the compulsory education is currently 12 years, yet it was 5 years until 1997, and 8 years until 2012. This implies that, except for very young workers, having a high school diplome is a signal for high skill. Based on this, similar to Akgündüz et al. (2015b) and Tumen (2016), we define natives with no high school degree as low-skilled workers ($L_{L,N}$). According to the data, 37.5% of all employed natives in the sample belong to the group.

New residential building and occupancy permits data as well as province-level GDP information are also from TURKSTAT.[12] The former two are administrative data covering 2004-2015; thus they do not include squatter housing units (*gecekondu*), which is relatively common in Turkey, particularly among low-income households. If the migration cause an increase in newly built *gecekondu*s, the total

[6]The educational attainment distribution of the migrants data is from May 2016.

[7]In 2014, the Ministry of Interior made a public statement on the number of Syrian guests in each province. Although the relative Syrian densities in the statement is highly similiar to the recent data, the figures are too round to be exact.

[8]The survey does not report province of residence or work information. NUTS-2 regions are composed of combination of multiple provinces with populations less than 3 million habitants

[9]We do not include 2016 in the analysis due to the political and economic turmoil in the aftermath of the attempted coup.

[10]Share of workers with two jobs is less than 3%.

[11]Del Carpio and Wagner (2015) is the only study that finds a statistically significant wage effect, though positive. However, the authors admit that they may be over-estimating, since the findings are affected greatly from the choice of the reference year, as explained below.

[12]Regional and province level GDP data ends in 2014.

effect on the new residential buildings is partly invisible to the data we use. This potentially pulls the estimates towards zero. Therefore, we interpret our estimates as they constitute lower-bounds.

We use the administrative data on new firm establishments published by The Union of Chambers and Commodity Exchanges of Turkey (TOBB). Since 2010, TOBB collects and reports province level information on new company establishments and their start-up capital on behalf of TURKSTAT.[13] The data also provides total amount of new Syrian co-founded firms and the capital invested in Turkey.

Descriptive Statistics

Table 1 presents a descriptive summary of the demographic characteristics of (15+) Syrian migrants at national-level. For comparative purposes, we also present comparable statistics for all (15-64) natives, and (15-64) natives that reside in the regions where the ratio of migrant population to the natives is greater than 10%.[14] The table shows that the guests have less education than both native samples. While 92.4% of them have no high school degree, this number is 66.1% for all natives, and 76.6% for the latter sample. They are also younger and less likely to be woman than natives.

Table 2 summarizes the data on native employment and wages, residential permits, and new firm establishments. We divide the sample into 6, according to the Syrian guest density (less than 2%, between 2% and 10%, and more than 10%) and the period (2004-2011 and 2012-2015).

In terms of the overall native population, the regions with high and low Syrian density are similar. Employment rate, on the other hand, is remarkably lower for the former regions than others before 2012. Decomposing it into formal and informal employment rates reveals that the discrepancy is primarily due to the share of formally employed workers. The share of individuals that are formally employed is considerably lower in the latter regions. More than two thirds of workers in high Syrian density regions are informally employed, whereas this number is below 50% for other region groups. This partly explains the pre-2012 difference across regions in the average wages of all workers, and of those without a high school degree. The employment rates and wages have considerably increased; and the informal employment rate has declined after 2012 in all region groups.[15]

[13]Data on new company establishments starts from 2009 and data on start-up capital investment from 2010.
[14]We exclude natives older than 65, because they are mostly retired, and not in the labor force.
[15]The increases and the decrease do not occur due to a specific recession or a boom year, but a general trend

Building, and new firm statistics reveal that the size of the economic activity is similar in high and low Syrian density regions, and remarkably larger in the medium density regions. The latter is primarily due to Istanbul and Izmir, two provinces whose combined gross provincial products amount to more than 35% of Turkish GDP. After 2012, we observe substantial increases in the residential building statistics for all region groups, and in the new firm statistics for high and medium Syrian density regions.

4 Econometric Framework

Difference-in-Differences Specification

Following the previous studies (Akgündüz et al., 2015b; Tumen, 2016), we first employ a difference-in-differences (DiD) model to assess the effects of Syrians in Turkey. The DiD econometric model is;

$$Y_{i,t} = \beta * T_{i,t} + \lambda_i + \mu_t + \epsilon_{i,t}, \tag{6}$$

where i indicates the region or province, t time period, and $Y_{i,t}$ is the outcome variable. The main outcome variables are informal, less than high school (LTHS), high school and above graduates (HSG), and overall employment rates and wages; the number of new residential buildings; and log of the number of new firm establishments. λ_i and μ_t are region or province, and year specific effects, respectively. $T_{i,t}$ is a binary variable that takes on the value of 1 for the years after 2012 in the regions with relatively dense Syrian population, and 0 otherwise. β, the coefficient of interest, yields the mean difference between control and treated regions after purging national macroeconomic shocks and time-invariant region or province effects. This model implicitly assumes that the dependent variables would follow similar paths in the control and treated regions in the absence of the migration shock.

There are Syrians in all regions of Turkey, but it is possible to determine where they constitute a substantial share of regional population. As shown in figure 3, as of 2015, the total number of Syrian guests always constitute less than 2% of the native population in 16 NUTS-2 regions (62 provinces). We consider these regions (provinces) as controls. Between 2% and 10%, there are 7

affecting throughout Turkey.

NUTS-2 regions (14 provinces). They are excluded entirely from the sample in the DiD. Though artificial, this allows us to create a sizable difference between treated and control regions in terms of Syrian guest population as a share of native population. In 3 NUTS-2 regions (5 provinces), the Syrian population is at least as large as 10% of the native population. Then, the data employed in this specification includes 19 NUTS-2 regions (67 provinces) where 3 (5) of them are treated.

We normalize the total counts of workers by working-age native population of the corresponding demographic group. Similarly, the number of building permits are normalized by the province's GDP. Using current population or GDP may yield misleading results if the migration causes native out-migration or affects economic activity in the province (Borjas, 2003; Card and Peri, 2016). To prevent this from biasing the findings, we use the working-age population and province's GDP in 2011 to normalize. For the sake of interpretability, we divide the estimates from the latter regressions by the mean of the dependent variable to get percentage change (%ΔY);[16]

$$\%\Delta Y \approx \frac{\Delta Y}{\overline{Y}} = \frac{\beta}{\overline{Y}}. \tag{7}$$

Generalized Synthetic Control

We employ the generalized synthetic control (GSC) method to prevent underlying regional trends from affecting the estimates. In the presence of unobserved time varying confounders, such as regional trends, the identifying assumption of difference-in-differences estimator, namely the parallel trends assumption, might be violated. The GSC model overcomes this problem by purging the patterns in the error term that can be formulated as interactions of region-specific intercepts (factor loadings) and time varying coefficients (latent factors).

Specifically, if the error term of the equation 6 is $\epsilon_{i,t} = \Gamma_i f_t + u_{i,t}$ and $\Gamma_i f_t$ are unobservable time-varying confounders that are correlated with the treatment T and $u_{i,t}$ is idiosyncratic error, then the estimates of the DiD model are biased and inconsistent. Bai (2009) has proposed a way to surmount this problem. His interactive fixed effects (IFE) model solves it by finding and purging the patterns in the composite error term $\epsilon_{i,t}$. More clearly, given $\Gamma_i f_t$,

[16]We wish to emphasize that the approximations have no qualitative effects on the findings.

$$\hat{\beta}(\mathbf{\Gamma}, \boldsymbol{f}) = \sum_{i=1}^{N}(X_i'X_i)^{-1}\sum_{i-1}^{N}X_i(Y_i - f\Gamma_i), \tag{8}$$

where X are all observable controls, fixed effects and the variable of interest.[17] Subtracting the time-varying confounders from the outcome eliminates the part of the error term causing the endogeneity and renders the estimates unbiased and consistent.

Given $\hat{\beta}$, the unobservable confounders can be found by the principal component analysis (PCA): By turning the vector $W_{i,t} = Y_{i,t} - X_{i,t}\beta$ of length $T * N$ into a TxN matrix, we estimate its principal components. The PCA is an unsupervised machine learning method that unravels patterns in the feature space that can be decomposed as $C_{i,t} = D_i * E_t$.[18] For instance, NUTS-2 region-specific trends of any polynomial degree can be detected by PCA. So, given β and $W_i = f\Gamma_i + e_i$, we can calculate $f\Gamma_i$. In addition, given $f\Gamma_i$, we can find $\hat{\beta}$ from equation 8. As long as the number of principal components are specified, this scheme can be solved by iteration.

Xu (2017) proposes that leave-one-out cross-validation procedure can be used to determine the number of unobserved factors to be purged. For any number of principal components, the procedure leaves out one pre-treatment observation from each of the treated units at a time and builds a model using the control units to predict the left-out observations. The number of unobserved factors to be purged is determined from the model with the smallest mean squared prediction error (MSPE). Then, the GSC builds the baseline model and creates the counterfactual; namely the hypothetical treated regions in the post-treatment periods that have not received the treatment. The counterfactual and the actual treated regions are compared to estimate the impact of the treatment.

Before proceeding, Xu (2017) cautions the users when the number of pre-treatment periods is fewer than 10 or the number of control units is fewer than 40. The use of data-driven selection methods requires more data than the basic fixed-effects estimators. This prevents us from employing GSC in estimating the effect on the number of new firm establishments. Additionally, we note that there are 9 pre-treatment years and 16 control regions in our employment and wage regressions.

[17]Note that in the absence of $\Gamma_i f_t$ term, this is the standard fixed-effects estimator.

[18]Mathematically, it provides "low-dimensional linear surfaces that are closest to the observations." (James et al., 2013) The number of dimensions are selected by the user .

Alternative Specification: OLS

Alternatively, instead of pooling all the treated regions and dropping some regions from the sample, we estimate the same model using regional treatment intensity variable calculated by interacting the binary treatment variable with the 2015 ratio of Syrian guest population to natives. All NUTS-2 regions or provinces are used in this specification. Then, the alternative estimation equation is as follows;

$$Y_{i,t} = \omega * Syr_{i,t} + \lambda_i + \mu_t + \epsilon_{i,t}, \tag{9}$$

where $Syr_{i,t}$is the treatment intensity or Syrian density variable. It takes non-zero values starting from 2012. Although findings are not affected qualitatively, we add the quadratic term of the treatment intensity variable in the building permits and firm establishment regressions, since the relationships exhibit a concave pattern. The quadratic model is $Y_{i,t} = \omega_1 * Syr_{i,t} + \omega_1 * Syr^2_{i,t} + \lambda_i + \mu_t + \epsilon_{i,t}$. The estimated coefficients of the alternative specification are not directly comparable to the previous ones. Therefore, we approximate ΔY by dividing the parameter of interest, ω, to the average difference in treatment intensity in the post-treatment period between treated Syr_{tr} and control regions Syr_{co};

$$\beta = \Delta Y * \frac{\overline{\Delta Syr}}{\overline{\Delta Syr}} \approx \omega * \overline{\Delta Syr}, \tag{10}$$

where $\overline{\Delta Syr} = Syr_{tr} - Syr_{co}$. In the regressions with the quadratic term, we approximate β by $\omega_1 * \overline{\Delta Syr} + \omega_2 * (Syr^2_{tr} - Syr^2_{co})$.

Alternative Specification: 2SLS

A general concern in the literature is that underlying trends that improve the economic outcomes in a region might also pull immigrants. The rise might be incorrectly attributed to immigrants due to the way the regression model is constructed. Although the GSC arguably overcomes the problem, we also employ instruments that rely on the distance (Llull, 2017; Peri, 2012) and historical ethnic enclaves (Borjas and Monras, 2016; Card and Lewis, 2007; Card, 2009).

We use an indicator variable for border regions (or provinces) as an instrument. It is natural to expect that border regions host disproportionately more migrants than others, since it is the least costly way to be away from the war in terms of transportation expenses. The cost is borne by the guests and it is unrelated to the macroeconomic environment. Furthermore, occasionally, a part of the family stays in Syria and, depending on the conditions of the war, visited during religious holidays by the leavers.

Secondly, we redistribute the total number of Syrian guests according to 1965 province-level Arabic speaking shares.[19] The distribution of Arabic speaking population in 1965 is independent of current regional trends. Networks established by older generations attract newcomers. One issue arises here is that there were 67 provinces in Turkey in 1965; however there are 81 provinces since 1999. Hence, we build a crosswalk between 1965 provinces and 1999 provinces according to the pre-treatment native population. If a province in 1965 is split into two or more in later years, we distribute the Arabic speakers according to the 2011 native population. If a new province is composed of parts of multiple provinces of 1965, then we group them and distribute the speakers similarly. [20]

We use the past distribution of Arabic speakers for the guest distribution, and normalize it by pre-treatment population to obtain conjectured Syrian density. Formally, we calculate the conjectured number of Syrian migrants in province i as follows;

$$(Total\,Guests_{2015} * Arabic\,speaking_{1965,i} * C_i). \tag{11}$$

$Total\,Guests_{2015}$ is the total number of Syrian guests in 2015, C_i is the conversion coefficient between 1965 and 1999 provinces, $Arabic\,speaking_{1965,i}$ is the ratio of Arabic speaking population in province i to all Arabic speaking population in Turkey in 1965. Note that $\sum_i Arabic\,speaking\,share_{1965,i} * C_i = 1$. Then, our second instrument is the conjectured number of Syrian migrants divided by the pre-treatment population in province i.

In regressions with the quadratic endogenous term, we add a third instrument. To capture the non-linear term, we follow the suggestion in Wooldridge (2010). First, we predict the Syrian density using our exogenous variables. Then, we create a variable by taking the square of the predicted values and include it in the instruments. By construction, this instrument is able to

[19] This instrument is highly similar to the one used in Altindag and Kaushal (2017).

[20] This method implicitly assumes that the population ratios between provinces in 2011 reflect Arabic speaking ratios in 1965. We acknowledge that this assumption does not necessarily reflect the reality. However, using the population information from other years has virtually no effect on findings.

retrieve the non-linear exogenous component, and it satisfies the exogeneity assumption as long as other instruments do. We employ equation 10 to approximate β.

Given the assumptions, the DiD estimate can be directly interpreted as the causal effect. However, as shown below, the parallel trends assumption are violated in some cases. Therefore, when feasible, we present the estimate from the GSC method as our baseline estimate and others as supplementary evidences. In the firm formation regressions, we cannot use the GSC due to very short time span of the data, and consider estimates from the other models as equally likely to capture the truth.

Inference

In DiD models with fixed effects, the inference is based on the assumption that the number of treated and control groups grow as the sample size gets larger. An intuitive exposition of this problem is presented in Conley and Taber (2011). Briefly, group-time random effects in the error term of the treated group do not vanish as the sample gets bigger if the number of treated units is fixed and small. The confidence interval, although unbiased, is inconsistent; because it captures part of the error term as well as the true effect.

Conley and Taber (2011)'s randomization inference procedure overcomes this problem. The main idea is that the variability of the coefficient of interest can be estimated from the distribution of the coefficients estimated from the placebo treatments. The latter coefficients are obtained by assigning placebo treatments to the control sample. This requires the "exchangeability" assumption, namely the treated units are chosen at random, and the observations from different NUTS-2 regions have the same joint distribution(Doudchenko and Imbens, 2016).

MacKinnon and Webb (2016) shows that randomization inference based on coefficients (RI-β) does not perform well in the existence of heterogeneity; but RI-t recovers very quickly as the number of treated groups increase. For instance, when the number of treated units is 3 and the standard error of the treated sample is twice (half) the control sample, RI-t rejects the true null hypothesis less than 10% (more than 2%) of the time when the test size is set to 5%. The same procedure shows that the true null is rejected in more than quarter of the trials with RI-β procedure.

In the regressions where the panel variable is at NUTS-2 level, to obtain the empirical distribution of placebo t-statistics for the null zero, we assign placebo treatments to all groups of 3 regions other than the original group. We get 968 placebo t-statistics.[21] This number is large enough to produce a cumulative distribution that resembles a continuous function. We produce the p-values of the estimates using the distribution of the placebo t-statistics.

The GSC standard errors are produced using a parametric bootstrap technique (Xu, 2017). Firstly, we create a set composed of the residuals of the control sample. Secondly, we randomly declare one control region as the "treated region", re-sample the control regions with replacement, apply the GSC and obtain a vector of prediction residuals B_1 times. Then, a second bootstrap with B_2 iterations commences. The former (latter) set of residuals are sampled with replacement to impose randomness to the control (true treated) regions. Using the B_2 "randomness imposed" data, the GSC estimates are computed at each iteration. The estimates are used to construct the confidence intervals and standard errors.

In the regressions with continuous variable of interest, the strict difference between treated and control units is blurred. For employment and wage regressions, the small number of clusters issue persists, though. As noted by Angrist and Pischke (2008) and Cameron and Miller (2015), with few clusters, the standard errors tend to be underestimated. Similar to Akgündüz et al. (2015b), we assess the reliability of the test using the cluster-robust standard errors in these regressions by reporting the p-values derived using the wild cluster bootstrap methods (CGM), which are shown by Cameron et al. (2008) to perform well.[22] We report the p-values from the "wild restricted residual bootstrap" (WRR) for the 2SLS regressions when the number of clusters is fewer than 30 (Davidson and MacKinnon, 2010).

5 Empirical Results

5.1 Effects on native workforce

5.1.1 Employment effect

Dustmann et al. (2016) remarks that the labor supply of native workers is not necessarily inelastic. Natives might not work below some wage levels and a portion of the impact of the migration might be absorbed as a fall in native employment. Additionally, the debate on the effects of the

[21] $\binom{19}{3} - 1 = 968$.

[22] MacKinnon (2006) presents evidence that the bootstrap methods work well even for univariate regressions.

Syrian migration mostly revolves around its impact on native employment. Therefore, we start by examining the effect on Turkish workers.

We do not directly observe Syrians in the LFS dataset. In determining the most impacted native groups, we exploit the fact that Syrians cannot formally work. We are able to identify workers working informally from the survey since TURKSTAT asks respondents whether they have social security coverage. According to the Turkish law, every employee should respond affirmatively. This implies that no social security coverage indicates informal employment. 55.55% of workers with no high school degree are employed informally in 2011; whereas this number is 23.02% (7.38%) for workers with high school degree (some college education or above). Additionally, in terms of educational credentials, MoI reports that more than 92.5% of the migrants do not have a high school diploma. We consider individuals with no high school degree as the most impacted demographic group, and jobs that do not provide social security coverage are the most impacted jobs. We also present estimates for high school and above graduates as well as on the average native worker.[23]

Table 3 shows our findings on the employment effects of the Syrian migration. Panel A uses the DiD, panel B the GSC, panel C the alternative model that incorporates intensity of the migration shock by replacing the binary treatment variable with the continuous regional Syrian density variable, and panel D the 2SLS model. The first two rows of all panels indicate the estimated coefficient of interest, and the cluster robust or the GSC standard errors. The third rows calculate the p-values using the randomization inference (RI-t), bootstrap results from the GSC, the wild cluster bootstrap (CGM), or the wild restricted residual bootstrap (WRR). Each panel reports the number of clusters. Panels A and B report the number of regions that are defined as treated as well. In addition, panel B reports the number of purged unobserved confounders. Lastly, we present the first-stage F-statistics in panel D to assess the weak identification of the endogenous variable.

The first column of panel A shows that the informal employment rate has not changed after the migration. The estimated effect is 0.008, statistically indistinguishable from zero. When we focus on the highly impacted demographic group, we find that the LTHS employment-rate has risen by 2%, though the corresponding p-value and standard error are large ($p = 0.539$ and $se = 0.025$), preventing us from rejecting the null hypothesis. Similarly, the employment-rate of high school

[23]We provide additional findings for sub-groups of the highly impacted demographic group, and jobs in the appendix figure C.1.

and above graduates (HSG) increase by 6.3% after the migration . However, the estimate is quite imprecise ($se = 0.047$) due to relatively small number of HSG population in the treated regions. Column 4 reports that the estimated change in overall employment rate is also very close to zero ($\hat{\beta} = 0.013$).

Panel B uses the GSC model. For informal employment, the cross-validation procedure finds that one unobserved factor needs to be purged. Doing this produces small negative estimate for the effect of the migration ($-0.003(0.03)$). Similarly, with one unobserved factor, the employment effect is essentially zero for LTHS individuals ($0.000(0.047)$). The cross-validation procedure suggests that the DiD performs the best, and we do not need to purge any unobserved factors. Therefore, we leave the last two columns in panel B blank, since they are the same as those of panel A.

Panel C employs the alternative specification. In this specification, we include all the NUTS-2 regions in our analysis, and assess the reliability of the cluster-robust standard errors by using the wild cluster bootstrap procedure. As noted, for comparability, we divide estimated coefficients by the difference in the Syrian densities between treated and control regions. All of the estimates are qualitatively and quantitatively similar to the findings from previous specifications, denoting no adverse employment effect of the migration. Findings from the two-stage least squares model in the panel D are also pointing to the non-rejection of the null hypothesis. The first-stage F-statistic of the model indicates that the parameter of interest is strongly identified. [24]

Pooling all the post-treatment years and comparing them with the entire pre-treatment period might be misleading in the existence of dynamic effects. In addition, the time path of the effect can be used to further assess the validity of the parallel trends assumption of our baseline model, as in Allegretto et al. (2011) and Dube (2017). We present the evolution of the LTHS and informal employment rate using the GSC in Figure 4. The solid line in the figure shows the estimated difference in the employment rate between the treated units and the counterfactuals; whereas the horizontal dash line at 0 indicates the perfect coincidence of the two. The vertical line is for the year 2012, the first year of the shock. The estimated line is very close to zero during both the pre- and post-treatment periods for both LTHS individuals and informal employment, displaying no pre-existing trends or lagged displacement effects.

[24] We have one endogenous and two instrumental variables. The Hansen's J statistics never reject the null hypothesis that the model is over-identified, with p-values always greater than 0.10.

The outcome variable of the regressions estimating the effect on native LTHS employment pools formal and informal workers. This potentially conceals the effect the share of LTHS workers employed formally. The migrants might be employed informally, hence push natives toward formal employment. We confirm the argument in figure 5. The formally employed share of LTHS workers has increased considerably after 2012. The GSC estimate for the increase after the shock is 0.033(0.021) with two unobservable factors.[25] The rise is persistent and the point estimates increase over time.

While migrant workers are informally employed, natives have found formal job opportunities. The evidence is also compatible with Peri and Sparber (2009)'s argument that natives switch to less routine-manual-task-intensive jobs with migration. The informal employment is prevalent in elementary occupations (9 according to ISCO-08) and skilled agricultural, forestry, and fishery workers (6 according to ISCO-08). 62% and 83% of all workers in the occupations are informally employed, and slightly more than half of all LTHS workers are employed in these jobs. They are highly routine and physically demanding. The rise in the share of formally employed native LTHS workers suggests that Turkish workers have partly protected themselves by specializing in relatively less manual routine intensive jobs.

Our findings on employment effect of Syrian migration differ from both Tumen (2016) and Del Carpio and Wagner (2015) who both find negative and statistically significant effects of Syrian migration on the natives' informal employment. We delineate our differences in Appendix B in detail. Briefly, we show that their findings are driven primarily by the choice of control regions, and the implied counterfactuals that exhibit underlying economic trends different from the treated regions. Additionally, the significance of their findings are potentially a result of underestimated standard errors. Within-region correlation of the error terms is not addressed in either study.

5.1.2 Wage effect

The other channel through which the migration wave might impact the labor market is wages. Natives might accept lower wages to protect their employment.

In studies that analyze the wage effects of the migration, it is common to adjust the data

[25] The DiD estimate for the change is 0.057 (0.026). The findings of no change in informal rate and rising formal share of LTHS employment might appear conflicting at first glance. However, we note that this is primarily driven by the level differences of the formal employment rate of LTHS individuals between the treated and control regions. Although the formal share of LTHS workers rise considerably and informal employment rate is unaffected, the effect of the former on the total LTHS employment is limited due to the small number of formally employed LTHS individuals.

for demographic composition of workers (Borjas, 2015; Card, 2009; Monras, 2015). To decrease potential contaminating effects of incidental changes of the composition due to the relatively small size of the sample and to eliminate effects of differential levels and trends of the returns to personal characteristics at national-level, we time-demean the log-transformed wage variable at individual-level for each of the age-by-education-by-gender groups, and obtain residual wages. Specifically, we calculate the residual wages as follows;

$$\tilde{Y}_{n,i,t} = Y_{n,g,i,t} - \overline{Y_{g,t}}. \tag{12}$$

n indicates individual, g the demographic group. $Y_{n,g,i,t}$ is individual's log wage, and $\overline{Y^m_{g,t}}$ is the average wage of workers in group g.[26]

After having obtained the residual wages, we collapse the data at region-by-year level. In the appendix, figure C.2 shows the benefits of the demeaning for average wages. Except for using the GSC model with residual wages as the outcome, all other cases display a U-shaped pre-existing trend, invalidating the causal interpretation.

Armed with this information, we only report the GSC estimates for the (residual) wage effects of the migration wave in table 4.[27] Firstly, as opposed to the employment estimates, we find that wages of the workers informally employed have dropped, on average, considerably after the inflow of migrants ($\%\Delta Y = -0.11\,(0.13)$). However, the estimated wage effect we obtain for native workers with no high school education is $-0.02\,(0.02)$. In addition, as shown in figure 6, the average wage of informally employed workers decreases every year after 2012, yet that of native LTHS workers recovers quickly, and the estimate is -0.015 and -0.014 in 2014 and 2015.

This is unexpected at first glance, since most of the LTHS workers are informally employed. The rise in formal share of native LTHS workers explains the conundrum. It is likely that the informally employed LTHS workers that were already earning higher have moved to formal employment with the migration. Their wages have increased; while wages of those remained informally employed have slightly fallen. Thus, the average wage of LTHS workers have not changed.

[26] In other words, we regress log of wage on time dummies for each age-by-education-by-gender group. Further region-demeaning slightly improves the precision, and has no effect on the point estimates.

[27] In Appendix Table C.2, we present the estimated wage effects using other model specifications for completeness. The violation of parallel trends nullifies the causality, yet the estimates still convey descriptive information of how the wages of selected demographic groups changed after 2012 in the treated regions compared to the control regions.

Given the imperfect substitution between LTHS and HSG workers, we expect a positive effect on the average wages of native HSG workers. Column 3 of table 4 confirms the expectation and reports that the average wage of native HSG rises statistically significantly by 5.7% with the migration. Lastly, column 4 reports that the average wage of all native workers has slightly increased ($\%\Delta Y = 0.02\,(0.02)$).

By focusing on the estimates for 2015, the year with the official province-level Syrian migrant counts are available, we can compare our findings to the larger literature on the effects of quasi-experimental migration shocks. The data reports that Syrian density in the treated regions is 0.151. Even when we conservatively assume that the ratio of migrant LTHS workers to native LTHS workers is 10% in the treated regions, our 95% confidence interval rejects the supply shock has led LTHS native wages to decline by more than 4.8% and to rise by more than 5.6%.[28] This coincides with studies that do not find a significant wage decline after a migration shock (Clemens and Hunt, 2017; Peri and Yasenov, 2017). However, it rules out some of the estimates in the literature obtained from different refugee waves, including -1.3 wage elasticity in table 3 of Borjas and Monras (2016), and the implied wage elasticities ranging from -0.5 to -1.5 in Borjas (2015) for *Marielitos* in the U.S. in the early 1980s .

Effects on native wage distribution

To further delineate the impact, we provide estimates for the effects of the Syrian migration on the shares above various multiples of the national minimum wage. We normalize wages by the minimum wage due to the fact that informal jobs are directly affected by the migration and the minimum wage is not binding for them. Put differently, the changes that occur around the minimum wage might reveal both the shift towards the formal employment and the effects on the informally employed workers with low wages.

Similar to the wage regressions, we do not employ region-by-year shares as is, but the shares adjusted for demographic characteristics. To do this, we calculate the annual nation-wide share of workers earning above given threshold for each of the age-by-education-by-gender group. Then,

[28] A more reasonable estimate for the number of migrant LTHS workers is 200,000, as explained in section 6. This corresponds to 13.7% (13.3%) of the native LTHS worker count in 2011 (2015). Using the number, our 95% confidence interval rejects that 10% increase in the low-skilled workforce due to the migration decreases wages of low-skilled natives by 3.5% (3.6%).

we subtract the average shares from observed outcomes that take the value of 1 if the individual earns above the threshold, and 0 otherwise. Finally, we calculate the mean of the residuals at region-by-year level and get (residual) shares.[29]

Therefore, the regression equations we use to estimate the change in share of workers earning more than given threshold is as follows;

$$\tilde{Y}^{m}_{i,t} = \beta * T_{i,t} + \lambda_i + \mu_t + \epsilon_{i,t}, \tag{13}$$

where $\tilde{Y}^{m}_{i,t}$ is the (residual) share of workers earning at least m% of the minimum wage, and m is multiples of 25 between 50 and 500. There are 19 regression estimates in total. Unless the GSC recommends the DiD, we employ the GSC. Hence, $\tilde{Y}^{m}_{i,t}$ is also free from unobserved factors that are deemed important by the cross-validation procedure.

Figure 7 presents the findings. We make four observations: First, the share of workers earning at or above the minimum wage has significantly increased by 2.50 percentage points in the treated regions compared to the counterfactuals. This is compatible with our finding that migrants have pushed natives to formal jobs that are relatively higher wage jobs. Some low-wage informal workers who would earn slightly less than the minimum wage are minimum wage workers after the migration. Second, the migration has increased share of workers earning upper-middle income. The shares of workers earning more than 175%, 200%, 225%, and 250% of the minimum wage have increased by 2.69, 2.11, 2.10, 2.19 percentage points. Third, the share of workers earning more than 50% of the minimum wage has not changed, yet those with wages less than 75% of the minimum wage has risen by 1.40 percentage point. A portion of the workers who would earn slightly less than the minimum wage has experienced a wage decline in the treated regions with the migration shock. Fourth, the migration had almost no effect on very-high-wage workers in the treated regions.

To sum up, our findings confirm that Syrian migrants enter the labor force through informal employment. This causes a reduction in the average wage of informal jobs. However, after the shock, an important portion of affected native workers are more likely to find formal jobs that pay relatively higher wages; hence, on average, are protected from the potential adverse wage effects. Therefore, we conclude that there is scant evidence for adverse effects of the migrants on native

[29] In terms of equation 12, $Y^{m}_{n,g,i,t}$ is 1 if the individual's wage is greater than the threshold m, and 0 if not. $\overline{Y^{m}_{g,t}}$ is the average share for workers in group g earning above m% of the minimum wage in time t.

wages. The non-finding of employment and wage effects even after such a huge migration shock implies that the effects of the migration are not limited to the labor supply.

5.2 Effects on residential building construction sector

Migrants improve the economy in the regions they reside. This is not only due to the additional labor-power; but they also bring their wealth and purchasing power to the host country. One of the industries that potentially benefit the most by migrants is the residential construction. This is especially relevant in Turkey, since around 90% of Syrian population live outside the camps.[30]

Table 5 shows the impact of the migration shock on residential building permits. The first column measures the number of new residential building permits in m^2; whereas the second in dwelling units and the third in new buildings. The dependent variables are all normalized by the pre-treatment province level GDP. For interpretation, we further divide the estimated coefficients by the mean value of the dependent variable to get the percentage change due to the migration shock. Thanks to the availability of province-level data, the cluster-robust standard errors can be used for p-values; yet we keep the structure of the table similar to the previous ones for consistency.[31] Following Solon et al. (2015)'s suggestion, we weight the regressions by province-level GDP for better precision.[32]

Panel A reports the estimates using the DiD specification. On average, the new residential building permits in m^2 has increased by 44% in areas with dense Syrian population than the counterfactual case. In column 2, the estimate is virtually the same and more precise. Column 3 shows that the number of permits in new buildings has also risen by 33%. We reject the null hypothesis of no effect in all columns at conventional levels.

Employing the GSC in panel B produces smaller, though qualitatively similar estimates. When m^2 is used for measurement unit, the cross-validation suggests the validity of the DiD model. Two unobservable factors are deemed important and purged in columns 2 and 3. The GSC estimates indicate that the number of residential dwelling unit permits has increased by 34%, and the number of permits in new buildings has increased by 25% after the shock. All estimates are statistically

[30] Howard (2017) observes a similar migration-induced residential demand boost in the United States.

[31] Though not as necessary as previous regressions, it is possible to construct confidence intervals and estimate p-values using RI-t. However, it is very computationally demanding. For each regression, it requires to calculate $\binom{67}{5} - 1 = 9,657,648$ placebo t-statistics.

[32] The weighting has no qualitative impact on the results.

significant.

Using continuous variable of interest and its quadratic term in panel C, we estimate that the arrival of migrants has increased the number of new residential building permits by 35% (34%) when m^2 (the number of dwelling unit) is used for measurement; whereas the increase is 24% in the last column. Panel D employs the two-stage least squares regression, and points towards the same direction of change.[33] Although the alternative model estimates are smaller compared to the ones in panels A and B; all models show that the migration boosted the residential construction industry.

Plotting the time pattern of the treatment effect presents a more transparent picture and enables us to assess the parallel trends assumption. The figure 8 shows the GSC estimates for the evolution of the percentage change in the new residential dwelling unit permits. New dwelling unit permits in the treated regions have increased and remained higher than the counterfactual during the entire post-migration-wave period. The sustained increase is reassuring since the migrant flow from Syria is not a one-shot event, as shown in figure 1.

One possible concern is that the rise in the new building permits only means that the government has allowed entrepreneurs to build on the designated lot, it does not indicate whether buildings are erected on the site. In the appendix, figure C.3 confirms that the buildings are constructed and ready to be occupied. The number of occupancy permits in dwelling units rises by 33% in 2014, and 42% in 2015.

Summarizing the impact of Syrian guests on the residential construction industry, we find that migrants boost the construction of new residential buildings. According to the 2012-2015 LFS, this is a large and relatively low-skilled intensive industry. More than 10% of all LTHS workers are construction workers, and 80% of the native construction workers do not have a high school diploma. Hence, the rise in the industry potentially leads to an increase in the demand for low-skilled workers, and change the distribution parameter between low- and high-skilled workers ($dln\theta > 0$).[34]

[33] There are two first stage regressions, since we include the quadratic term of the Syrian guest intensity variable. We report Sanderson-Windmeijer F-statistics for each of them (Sanderson and Windmeijer, 2016).For the first two columns, the Hansen's J statistic does not reject the null hypothesis that the model is over-identified, with p-values well above 0.2. It is 0.07 for the third column. We also estimate the model with the LIML estimator that performs better than 2SLS in the existence of weak instruments. The estimates are 0.559, 0.429 and 0.178.

[34] According to our estimates, the direct effect appears to partly disappear in the long run. Similar pattern is also observed in Howard (2017) (see figure 8 in his paper), though he concludes that migration-induced housing demand boom is an accelerator, hence has long-term effects.

5.3 Effects on new firm establishments

Another channel through which we observe the labor-demand effect of the migration is the new company establishments. High regional demand attracts capital and leads to increased new company formation, both of which create new employment opportunities (Baptista et al., 2008; Karahasan, 2015; Van Stel and Suddle, 2008). Moreover, migrants themselves might start new businesses.

Table 6 documents the increase in new company establishments as a result of the Syrian migration.[35] The first two columns report the percentage change in the number of new company establishments; while the last two columns the total start-up capital invested. If migrant capital crowded out other capitals, then the increase in the number of new companies would be at the expense of non-Syrian entrepreneurs. To assess this, in columns 2 and 4, we exclude all the companies with at least one Syrian founders, and all capital invested by Syrians. The short time span of the data before 2012 prevents us from employing the GSC model.

According to the DiD model, the number of new company establishments has increased by 36.2% due to the migration shock. The second column shows that it is not only Syrian entrepreneurs that enjoyed the better economic environment. The number of new company establishments with no Syrian founders has also risen by 26.5%. The corresponding estimates are larger when we focus on the total initial capital invested in columns 3 and 4 (52.1% when start-up capital invested by entrepreneurs of all nationalities are included, and 45% when we exclude the Syrian share). The alternative models, OLS and 2SLS, in panels B and C indicate similarly sized and statistically significant positive effects of the migration.[36]

The difference between columns 1 and 2 (or 3 and 4) implies that Syrian migrants played a direct role in boosting the firm formation. Figure 9 presents the evolution of the total number of companies with Syrian founders (top graphs) and the real initial capital invested by Syrian nationals (bottom graphs). The left-hand side graphs report the absolute numbers and the right-hand side ones the shares out of all non-native start-ups.

First, all graphs point to a massive increase in the Syrian business creation. In 2010, the total

[35]The data on new company establishments is at province level; hence the few clusters issue does not arise in this case. In addition, there are no zero values in the new company establishments data, so we log-transform the dependent variables and do not approximate using equation 7.

[36]The 2SLS regressions include the linear and quadratic terms of the Syrian guest intensity variable. The p-values corresponding to the Hansen's J statistics of each of the regressions are greater than 0.1, leading us to accept the null hypothesis of no over-identification. LIML estimates of columns 1-4 are 0.23, 0.21, 0.56 and 0.54.

number of new companies with Syrian founders was virtually zero. It is 1,599 in 2015. While the Syrian share in total number of new companies with foreign founders was 2.3% in 2010; it is 32% in 2015. Examining the start-up capital invested by Syrian nationals displays a similar pattern. In 2010, the total start-up capital invested by Syrians in Turkey was 7.93 million TL. This corresponds to 1.3% of all initial capital investments by foreign entrepreneurs. In 2015, the Syrian share in total foreign start-up capital is 22.9% and its outstanding amount reaches to 159.9 million TL (in 2010 TL). Lastly, we also note that Syrian-non Syrian partnerships have increased with the migration. From 2010 to 2015, the total initial capital investment by non-Syrian partners in Syrian co-founded firms rose from 2,067,780 TL to 14,023,571 TL (in 2010 TL). We also note the similarity of the graphs with figure 1. Increase in total number of Syrians is highly correlated to the total number of firms founded by Syrians. More Syrian migrants in Turkey has translated into more Syrian companies in Turkey (Akgündüz et al., 2015a).

Syrian wealth and purchasing power have created a fruitful environment for entrepreneurs in Turkey. Even in the short-run, capital is attracted to regions with high Syrian density. The migrants boost the local demand and the rise urges entrepreneurs to invest more ($dlnK > 0$). Syrians themselves are also among the entrepreneurs that benefited from it.

6 Discussion

In this section, we discuss whether the predictions of the canonical model coincide with the empirical estimates. We show that the model that assumes perfect substitution between migrant and native workers overshoots the native LTHS wage decline even if capital is assumed to adjust instantaneously. Allowing for the native-migrant complementarity decreases the discrepancy between empirical estimates and theoretical predictions, yet it is not sufficient either. Based upon our findings on the effects on residential building and new firm formation, we suggest that the technology parameter (A), and the distribution parameter between low-skilled and high-skilled workers (θ) of the canonical accounting framework are also affected by the migration.

Imperfect substitution between native and migrant workers

To calculate the theoretical predictions, we need elasticities of substitution between LTHS and HSG workers, and between native and migrant workers.

We cannot directly calculate the elasticity of substitution between low-skilled migrants and natives ($\sigma_{L_{L,N}L_{L,I}}$) because we only have data on migrants' residence and no data on their wage and place-of-work-by-education. Instead, we devise an algorithm that searches $\sigma_{L_{L,N}L_{L,I}}$ by employing the canonical accounting framework, the wage and employment estimates for native workers, and elasticity of substitution between LTHS and HSG workers (σ_{L_H,L_L}) for given number of LTHS and HSG-equivalent migrant workers in the treated regions.[37] Specifically, the search employs our baseline wage and (composite) employment changes in the treated regions, and search for $\sigma_{L_{L,N}L_{L,I}}$ that allow us to obtain the elasticity of substitution between LTHS and HSG workers.

We can estimate the elasticity of substitution between LTHS and HSG workers at national-level by employing a standard Katz and Murphy (1992) type of regression equation using the 2004-2011 TURKSTAT LFS.[38] The regression equation is as follows;

$$ln(\frac{w_{L_{H,N}}}{w_{L_{L,N}}})_t = ln(\frac{1-\theta}{\theta})_t - \frac{1}{\sigma_{L_H,L_L}} * ln(\frac{L_H}{L_L})_t. \tag{14}$$

The coefficient of interest is $\frac{1}{\sigma_{L_H,L_L}}$. The first term in the right hand side of the equation 14 is a linear time trend to capture the relative demand shifts.

The estimation yields $ln(\frac{w_{L_{H,N}}}{w_{L_{L,N}}})_t = 0.031\,(0.004) * year - 0.679\,(0.155) * ln(\frac{L_H}{L_L})_t + constant$. Then, the implied elasticity of substitution between LTHS and HSG workers is $\hat{\sigma}_{L_H,L_L} = 1.473$. It is similar to the ones commonly estimated in the literature (Card, 2009), despite the fact that we alter the definition of high-skilled workers to include high school graduates in order to account for the education distribution of Turkish workforce .

Having obtained $\hat{\sigma}_{L_L,L_H} = 1.473$, we can search for the elasticity of substitution between LTHS native and migrant workers using the estimated employment and wage changes in the treated regions compared to the controls. In other words, we aim to obtain $\sigma_{L_{L,I}L_{L,N}}$(and δ_L, since

[37] We differentiate "HSG migrant worker" and "HSG-equivalent migrant worker". The latter considers skill downgrading of the migrant workers.

[38] We leave out the migration period from the analysis. Running the same regression for 2004-2015 period with the assumptions and estimates below yields essentially the same estimates and has no qualitative effect on the results.

$\sigma_{L_{L,I},L_{L,N}} = \frac{1}{1-\delta_L}$) that satisfy the reduced form wage estimates and the $\hat{\sigma}_{L_L,L_H}$ at the same time. Mathematically, the formula we employ is;

$$\%\Delta w_{L_{H,N}} - \%\Delta w_{L_{L,N}} = (\gamma - \delta_H) \times \%\Delta L_H - (\gamma - \delta_L) \times \%\Delta L_L, \tag{15}$$

where $\%\Delta L_L$ and $\%\Delta L_H$ stand for percentage change in L_L and L_H, calculated according to equation (1b).

The search requires assumptions on i) the place-of-work-by-education distribution of the migrants to obtain changes in L_L and L_H in the treated regions, and ii) relative shares of native and migrant workers (ω). Moreover, its validity requires similarity of σ_{L_L,L_H} at the national-level and underlying σ_{L_L,L_H} for the treated regions. We acknowledge that these are relatively strong assumptions, hence the results should be considered as a set of possible scenarios.

We compare 2011 and 2015 estimates for the calculation, since there is essentially no guests in 2011 and we have reliable data for the 2015 province level distribution of Syrian forced migrants.[39] Guided by the previous studies (Ottaviano and Peri, 2012), we assume the elasticity of substitution between HSG-equivalent native and migrant workers,$\sigma_{L_{H,I},L_{H,N}}$ to be high. We set $\sigma_{L_{H,I},L_{H,N}} = 33$. Assuming that $\omega = 0.5$ (according to the framework in section 2), table 7 produces implied elasticities of substitution between LTHS and HSG labor for alternative values of total employed LTHS and HSG-equivalent migrants in the treated regions.

We pick bounds for the number of LTHS migrant workers (200,000 and 225,000), HSG-equivalent migrant workers (0 and 2,000). 200,000 LTHS migrant workers correspond to 13.7% (13.3%) of the native LTHS workers in 2011 (2015) in the treated regions. Although the employment rate of migrants is lower than natives, the ratio of LTHS to HSG individuals is larger for migrants. Hence, we expect the ratio of the LTHS migrant workers to LTHS native workers to be slightly less than the population ratio (15.1%). We can also reach the similar number for LTHS migrant workers in the treated regions if we use the residency information as a proxy for province-level employment distribution of Syrian migrants. Given that the total number of employed migrants is roughly 400,000, and slightly more than 50% of the migrants reside in the treated regions, we again obtain

[39] Total LTHS (HSG) employment is 1,463,343 (571,664) in the treated regions in 2011. Between 2011 and 2015, the estimated percentage changes in the average wages (employment) of native HSG, and LTHS workers are 9% and -0.9% (5.7% and 3.5%).

200,000 LTHS migrant workers. Making conjectures on the number of HSG-equivalent employed migrants in the treated regions is relatively more challenging. Although their employment-rate might be larger than LTHS migrants, they suffer from skill downgrading. They cannot work formally, and almost all of them are non-Turkophone (Akgündüz et al., 2015b). In addition, they are likely to migrate to large cities, such as Istanbul and Izmir, that provide more diverse employment opportunities. Hence, we expect the migrant HSG-equivalent workers in the treated regions to be very close to zero. We also provide implied σ_{L_L,L_H} for 225,000 LTHS and 2,000 HSG migrant workers.

According to the table 7, we obtain $\sigma_{L_H,L_L} = 1.473$ when there is a modest elasticity of substitution between migrant and native LTHS workers (the exact range for implied σ_{L_L,L_H} is between 5.75 (row 2) and 16.13 (row 3)). This is largely compatible with the ones reported in Manacorda et al. (2012) for the U.K. (5 for recent immigrants), and in Ottaviano and Peri (2012) for the U.S. (11 for low-skilled workers).[40]

By construction, implied $\sigma_{L_{L,N}L_{L,I}}$ values for given numbers of LTHS and HSG-equivalent migrant workers satisfy the relative wage estimates. Another potentially revealing question is whether solely relaxing the perfect substitution between migrant and native workers is sufficient to obtain the reduced form wage estimates. Specifically, do we obtain LTHS wage decline of 0.9% for the year 2015 compared to 2011, when we plug in $\hat{\sigma}_{L_{L,N}L_{L,I}}$, $\hat{\sigma}_{L_H,L_L}$ to the equation 2, and assume that capital-labor ratio returns back to its original level ($dln(K) = dln(L)$) ?

The canonical model that assumes perfect substitution between natives and migrants predicts the average native LTHS wage to decline by 3.6% for the mid-point of the cases (212,500 LTHS and 1,000 HSG-equivalent migrant workers). Relaxing the perfect substitution assumption, the decrease becomes 3.2% for $\sigma_{L_{L,N},L_{L,I}} = 16.13$, and 2% for $\sigma_{L_{L,N},L_{L,I}} = 5.75$. When we assume 200,000 LTHS and 2,000 HSG-equivalent migrant workers to predict the lower bound for the wage decline, we obtain $\%\Delta w_{L_{L,N}} = -3.2\%$ for $\sigma_{L_{L,N}L_{L,I}} = \infty$; and $\%\Delta w_{L_{L,N}} = -1.7$ for $\sigma_{L_{L,N}L_{L,I}} = 5.75$. Although we have assumed perfect capital adjustment, all the predicted changes overestimate the decline. This implies that, for the canonical nested CES model, migration alters some of the parameters

[40]This implies that we could search for the number of HSG-equivalent and LTHS migrant workers using the wage and employment estimates for native workers, $\hat{\sigma}_{L_L,L_H} = 1.473$, and the estimated elasticities of substitution between migrant and native workers estimated in the literature. Unless perfect substitution between native and migrant workers is assumed, the search yields 200,000-225,000 LTHS, and 0-2,000 HSG migrant workers. To obtain $\sigma_{L_{L,N}L_{L,I}} = 50$, we would need to assume 240,000 LTHS and 0 HSG-equivalent migrant workers in the treated regions.

that are assumed to be unrelated with it.

Productivity and LTHS share effects of the migration

Our empirical findings of the effects on company formation and residential construction suggest that the technology parameter (A), and the parameter that determines distribution between low-skilled and high-skilled workers (θ) have also changed after Syrian migrants entered Turkey.

In the context of the U.S., Alon et al. (2017) empirically examines links between startup rates and the productivity. Based on the fact that young firms are more likely to exit the market than to contract, the study argues that younger firms are the locomotive of the productivity growth for two main reasons: Low productivity young firms lose their market share very quickly to the high productivity ones (re-allocation), and they exit the market (selection). They find that, in the long-run, 1 percentage point increase in startup rate leads state-level productivity to rise by 1.7% in the U.S (see Table 2 column 4 in their paper).

The numbers are estimated for the U.S., and the effects possibly vary across countries. However, if the true estimate for the treated regions were similar, we could directly quantify the long-run effect of the migration on the factor neutral technology parameter (A). The Ministry of Customs and Trade (2014) reports that the startup rate in Turkey is approximately 4.2%. Then, as long as the start-up rate in the treated regions is similar to the rest of the country, the most conservative estimate for the rise in company formation in table 6 (0.238) would yield that the productivity would rise by 1.7% in the treated regions in the long-run.[41]

The boost in the startup rate is only one of the channels that migration might affect the productivity. Peri (2012) claims that efficient specialization of native and migrant workers in different task intensity occupations is another important channel that increases the productivity. Given the routine-manual-task-intensity of the informal jobs, our finding of the relative rise in formally employed LTHS workers in the treated regions is compatible with the argument.

Second, the rise in residential construction demand might alter the regional production function, and affects sectoral composition in the treated regions. The relative demand for the LTHS workers

[41] 1 percentage-point increase from 4.2% to 5.2% corresponds to 23.8% increase, virtually the same as the smallest estimate in column 1 of table 6 (0.238).

increases and pushes their wages upwards ($dln\theta > 0$). Additionally, Peri (2012) suggests that, in the long-run, adoption of low-skilled-efficient technologies is another channel that increases θ after the migration.

7 Conclusion

In this paper, we study the impact of more than 2.5 million Syrian migrants on host local economies in Turkey. We begin with its employment and wage effects on the native workers. We find no robust evidence of an employment decline for the demographic group that is expected to be adversely affected the most (individuals with no high school diploma), nor for the informal employment. In terms of the wage effects, we find some evidence indicating that the wages of informally employed natives have declined by more than 10%. However, a similarly sized effect does not occur for the native less than high school (LTHS) workers thanks to a rise in formal employment share of the group. The minimum wage is binding for formally employed workers, hence some workers who would otherwise earn less than it have become minimum wage workers. On the other hand, workers with at least high school degree (HSG) have benefited from the shock. Their wages have, on average, increased by 5.7%.

The non-finding of an adverse effect in the short-run even after such a large inflow of migrants requires an explanation. According to the descriptive demand-supply framework, there must occur a labor-demand shift, along with the labor-supply shift, so that migration has no effect on native employment and wages. We use two indicators for the positive demand shock specifically germane to the case of Turkey; the number of new residential buildings and the number of new company establishments. Compared to the counterfactual case, the former has increased by 34%, and the latter by more than 24% after the migration in the host regions. Given that the construction industry is relatively low-skilled labor intensive, it can partly absorb the shock. Additionally, the rise in new firms suggests a general increase in job creation. Furthermore, Syrian migrants also co-found new firms and directly employ people. New companies with at least one Syrian co-founder was fewer than 100 in 2011, whereas it is 1,599 in 2015.

It is common in the literature to consider the immigration as merely a labor-supply shock. According to the canonical model, in the short-run, the only affected parameter is total labor supply

(L in equation 1a, and L_L and L_H in equation 1b); whereas capital (K in equation 1a) increases in the long-run. Our findings suggest that even when we assume capital to be fully adjusted, the canonical model severely overshoots the empirically estimated wage decline. The predictions approach to the empirical findings when we introduce imperfect substitutability between migrant and native workers, yet it is not sufficient either. Other parameters of the framework need to change as well. We argue that the boost in the residential construction industry might alter the regional production function in favor of low-skilled workers. Additionally, migrant-induced boost in the new firm formation potentially increases the productivity and pulls all native wages upwards.

One of the main policy questions related to the immigration is whether native workers are adversely affected. We show that the affirmative answer to the question is not empirically warranted in the case of Turkey. Nonetheless, two main characteristics of the current case raise potential questions about the generalizability of these findings. The first one is related to the informality of the migrants: What if the migrants had work permits? It might be argued that the migrants with work permits would substantially increase the intensity of the shock up to a point that native workers are adversely affected. Nevertheless, the government enforcement for formal employment is relatively weak in general, and in the host regions in particular. In Turkey (the latter regions), 41% (57%) of all workers aged 15-64 were informally employed in 2011. The ratio is even higher, 55% (70%), when we limit the sample to LTHS workers. The relatively high informality rate suggests that not having a work permit might not have substantially deterred Syrian guests from competing for jobs. In other words, we conjecture that granting work permits to Syrians would likely not considerably change their effect on native workforce. Second, Turkey is a semi-industrialized developing country. Would the increase in supply of low-skilled workers alter technologies in industries, and potentially the development path of the country? As we mention in the discussion section, our results suggest some alterations in the production functions at least during the first couple of years of intense migration from Syria. Providing definite answers to the questions is beyond the scope of the paper, and presents a program for the future research.

References

Akgündüz, Yusuf, Marcel Van den Berg, and Wolter HJ Hassink. 2015a. "The Impact of Refugee Crises on Firm Dynamics and Internal Migration: Evidence from the Syrian Refugee Crisis in Turkey."

——— 2015b. "The impact of refugee crises on host labor markets: the case of the Syrian refugee crisis in Turkey."

Allegretto, Sylvia A, Arindrajit Dube, and Michael Reich. 2011. "Do minimum wages really reduce teen employment? Accounting for heterogeneity and selectivity in state panel data," *Industrial Relations: A Journal of Economy and Society*, 50(2): 205–240.

Alon, Titan, David Berger, Robert Dent, and Benjamin Pugsley. 2017. "Older and slower: The startup deficit's lasting effects on aggregate productivity growth," *Journal of Monetary Economics*.

Altindag, Onur and Neeraj Kaushal. 2017. "Do Refugees Impact Voting Behavior in the Host Country? Evidence from Syrian Refugee Inflows in Turkey," IZA DP No. 10849.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*: Princeton university press.

Bai, Jushan. 2009. "Panel data models with interactive fixed effects," *Econometrica*, 77(4): 1229–1279.

Balcılar, Mehmet and Jeffrey B Nugent. 2016. "The Migration of Fear: An Analysis of Migration Choices of Syrian Refugees."

Baptista, Rui, Vítor Escária, and Paulo Madruga. 2008. "Entrepreneurship, regional development and job creation: the case of Portugal," *Small Business Economics*, 30(1): 49–58.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Borjas, George J. 2003. "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market," *The quarterly journal of economics*, 118(4): 1335–1374.

——— 2013. “The analytics of the wage effect of immigration,” *IZA Journal of Migration*, 2(1), p. 22.

——— 2014. *Immigration economics*: Harvard University Press.

——— 2015. “The wage impact of the Marielitos: A reappraisal,” *ILR Review*, p. 0019793917692945.

Borjas, George J and Joan Monras. 2016. “The Labor Market Consequences of Refugee Supply Shocks,” NBER Working Paper No. 22656.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. “Bootstrap-based improvements for inference with clustered errors,” *The Review of Economics and Statistics*, 90(3): 414–427.

Cameron, A Colin and Douglas L Miller. 2015. “A practitioner’s guide to cluster-robust inference,” *Journal of Human Resources*, 50(2): 317–372.

Card, David. 2009. “Immigration and Inequality,” *The American Economic Review*, 99(2), p. 1.

Card, David and Ethan G Lewis. 2007. “The diffusion of Mexican immigrants during the 1990s: Explanations and impacts,” in *Mexican immigration to the United States*: 193–228, University of Chicago Press.

Card, David and Giovanni Peri. 2016. “Immigration Economics: A Review,” *Unpublished paper, University of California.*

Ceritoglu, Evren, H Burcu Gurcihan Yunculer, Huzeyfe Torun, and Semih 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), p. 5.

Clemens, Michael A and Jennifer Hunt. 2017. “The Labor Market Effects of Refugee Waves: Reconciling Conflicting Results,”Technical report, National Bureau of Economic Research.

Conley, Timothy G and Christopher R Taber. 2011. “Inference with "difference in differences" with a small number of policy changes,” *The Review of Economics and Statistics*, 93(1): 113–125.

Constant, Amelie F. 2014. “Do migrants take the jobs of native workers?” *IZA World of Labor*.

Davidson, Russell and James G MacKinnon. 2010. “Wild bootstrap tests for IV regression,” *Journal of Business & Economic Statistics*, 28(1): 128–144.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. “The role of entrepreneurship in US job creation and economic dynamism,” *The Journal of Economic Perspectives*, 28(3): 3–24.

Del Carpio, Ximena V and Mathis C Wagner. 2015. “The impact of Syrian refugees on the Turkish labor market,” Policy Research Working Paper Series 7402, The World Bank.

Doudchenko, Nikolay and Guido W Imbens. 2016. “Balancing, regression, difference-in-differences and synthetic control methods: A synthesis,”Technical report, National Bureau of Economic Research.

Dube, Arindrajit. 2017. “Minimum wages and the distribution of family incomes.”

Dustmann, Christian, Uta Schönberg, and Jan Stuhler. 2016. “The Impact of Immigration: Why Do Studies Reach Such Different Results?” *Journal of Economic Perspectives*, 30(4): 31–56.

Gaston, Noel and Douglas R Nelson. 2013. “Bridging trade theory and labour econometrics: the effects of international migration,” *Journal of economic surveys*, 27(1): 98–139.

Howard, Greg. 2017. “The Migration Accelerator: Labor Mobility, Housing, and Aggregate Demand,” Unpublished manuscript.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An introduction to statistical learning*, 6: Springer.

Karahan, Fatih, Benjamin Pugsley, and Aysegül Sahin. 2016. “Demographic Origins of the Startup Deficit,” Mimeo.

Karahasan, Burhan Can. 2015. “Dynamics of regional new firm formation in Turkey,” *Review of urban & regional development studies*, 27(1): 18–39.

Katz, Lawrence F and Kevin M Murphy. 1992. “Changes in relative wages, 1963–1987: supply and demand factors,” *The quarterly journal of economics*, 107(1): 35–78.

Kerr, Sari Pekkala and William R Kerr. 2011. "Economic Impacts of Immigration: A Survey.," *Finnish Economic Papers*, 24(1).

Llull, Joan. 2017. "The effect of immigration on wages: exploiting exogenous variation at the national level," *Journal of Human Resources*: 0315–7032R2.

MacKinnon, James G. 2006. "Bootstrap methods in econometrics," *Economic Record*, 82(s1).

MacKinnon, James G and Matthew D Webb. 2016. "Difference-in-Differences Inference with Few Treated Clusters."

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth. 2012. "The impact of immigration on the structure of wages: theory and evidence from Britain," *Journal of the European Economic Association*, 10(1): 120–151.

Monras, Joan. 2015. "Minimum Wages and Spatial Equilibrium: Theory and Evidence," IZA DP No. 9460.

Ottaviano, Gianmarco IP and Giovanni Peri. 2012. "Rethinking the effect of immigration on wages," *Journal of the European economic association*, 10(1): 152–197.

Özden, Senay. 2013. "Syrian refugees in Turkey."

Özpınar, Esra, Seda Başıhoş, and Aycan Kulaksız. 2015. "Göçün Ardından Suriye ile Ticari İlişkiler."

Peri, Giovanni. 2012. "The effect of immigration on productivity: Evidence from US states," *Review of Economics and Statistics*, 94(1): 348–358.

——— 2014. "Do immigrant workers depress the wages of native workers?" *IZA world of Labor*.

——— 2016. "Immigrants, Productivity, and Labor Markets," *The Journal of Economic Perspectives*, 30(4): 3–29.

Peri, Giovanni and Chad Sparber. 2009. "Task specialization, immigration, and wages," *American Economic Journal: Applied Economics*, 1(3): 135–169.

Peri, Giovanni and Vasil Yasenov. 2017. "The Labor Market Effects of a Refugee Wave: Synthetic Control Method Meets the Mariel Boatlift," IZA DP No. 10605.

Sanderson, Eleanor and Frank Windmeijer. 2016. "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 190(2): 212–221.

Solon, Gary, Steven J Haider, and Jeffrey M Wooldridge. 2015. "What are we weighting for?" *Journal of Human resources*, 50(2): 301–316.

The Ministry of Customs and Trade. 2014. "Şirket İstatistikleri Bülteni, Haziran."

Tumen, Semih. 2016. "The Economic Impact of Syrian Refugees on Host Countries: Quasi-Experimental Evidence from Turkey," *The American Economic Review*, 106(5): 456–460.

Üstun, Nazlı. 2016. "Suriyelilerin Türk İşgücü Piyasasına Entegrasyonu."

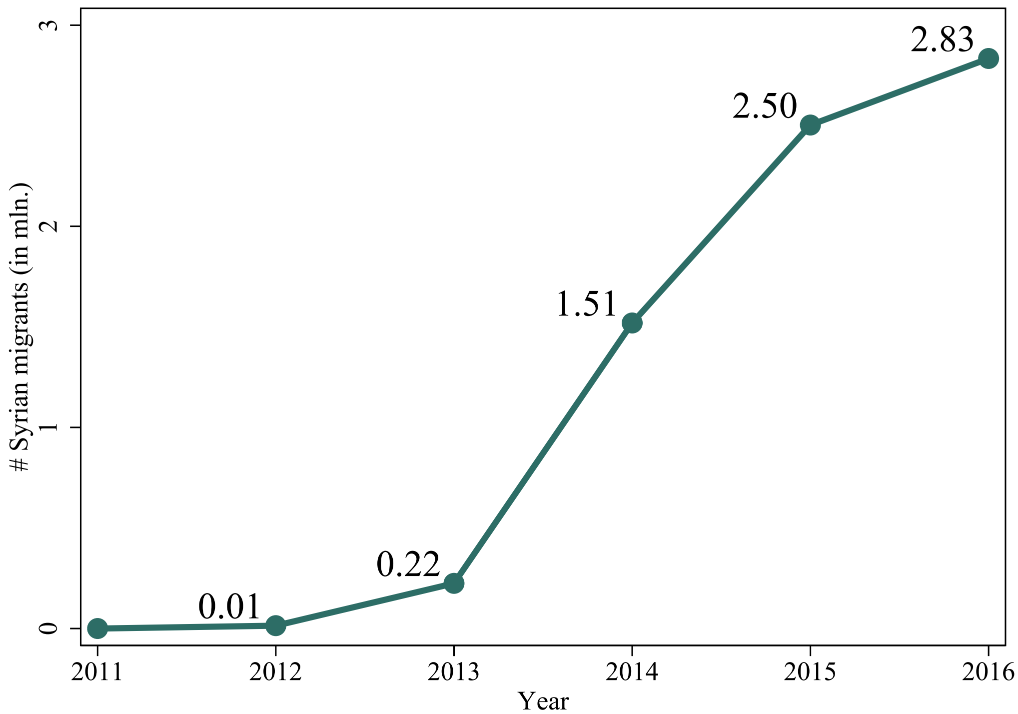
Van Stel, André and Kashifa Suddle. 2008. "The impact of new firm formation on regional development in the Netherlands," *Small Business Economics*, 30(1): 31–47.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*: MIT press.

Xu, Yiqing. 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models," *Political Analysis*, 25(1): 57–76.

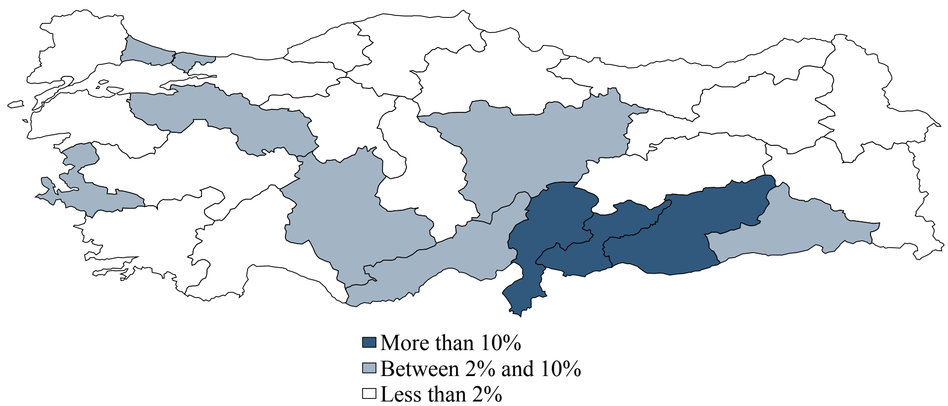
Zipperer, Ben. 2016. "Did the minimum wage or the Great Recession reduce low-wage employment? Comments on Clemens and Wither (2016)."

Figure 1: Total number of Syrian forced migrants in Turkey

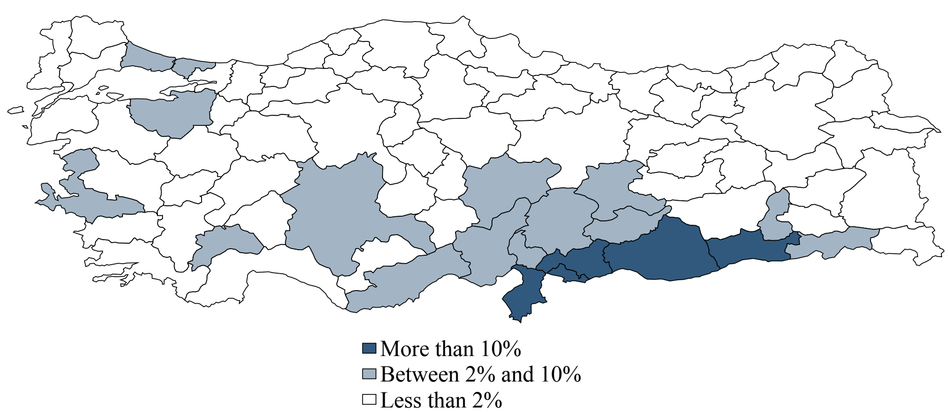


Notes: The figure shows the total number of Syrian forced migrants in Turkey between 2011 and 2015. Only Syrian nationals fled their country due to the war are considered. The data provided by Ministry of Interior Directory General of Migration Management is used.

Figure 2: Regional distribution of Syrian migrants



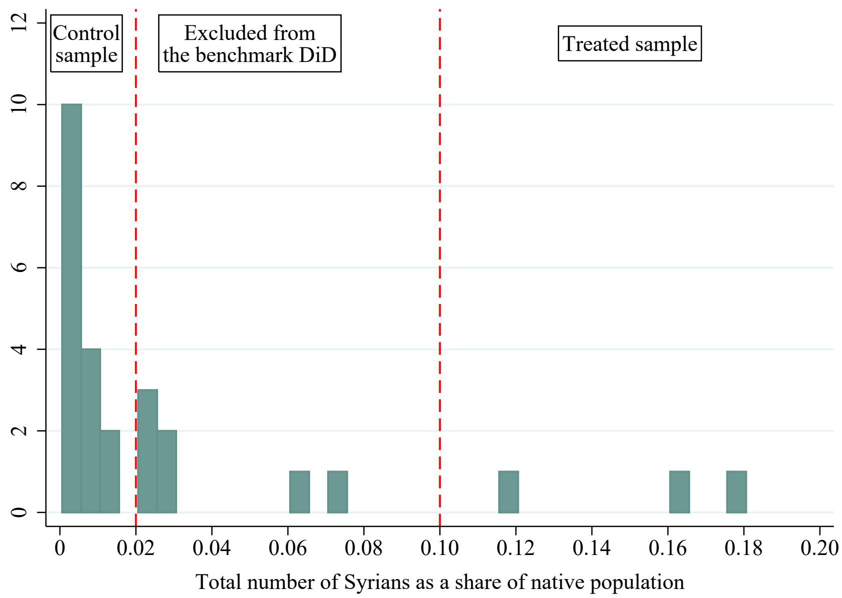
(a) NUTS-2 level



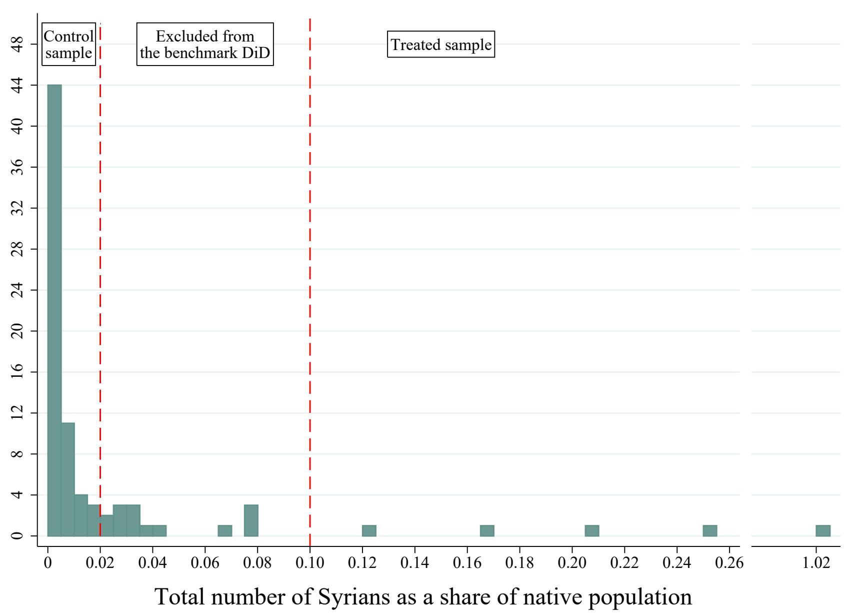
(b) Province level

Notes: The graphs plot the NUTS-2, and provincial distributions of the number of Syrian guests as a share of native population in 2015. Dark blue areas indicate that the Syrian guest population is at least 10% of the native population, white areas at most 2%, and light blue areas between 2% and 10%. The data provided by Ministry of Interior Directory General of Migration Management is used.

Figure 3: Frequency distribution of Syrian migrants as a share of native population



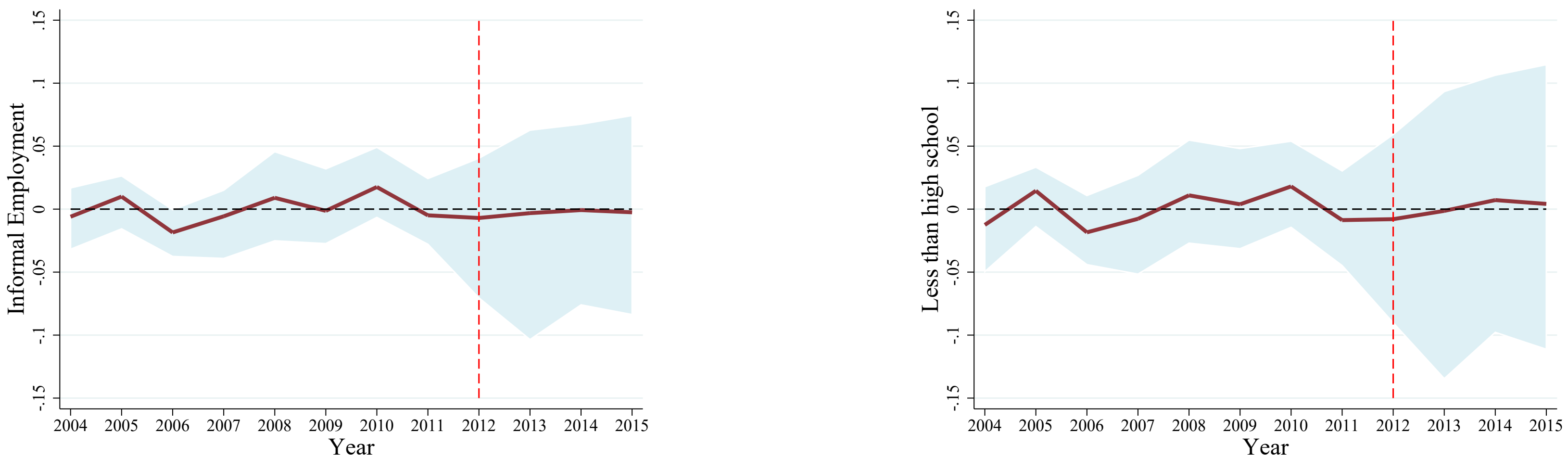
(a) NUTS-2 level



(b) Province level

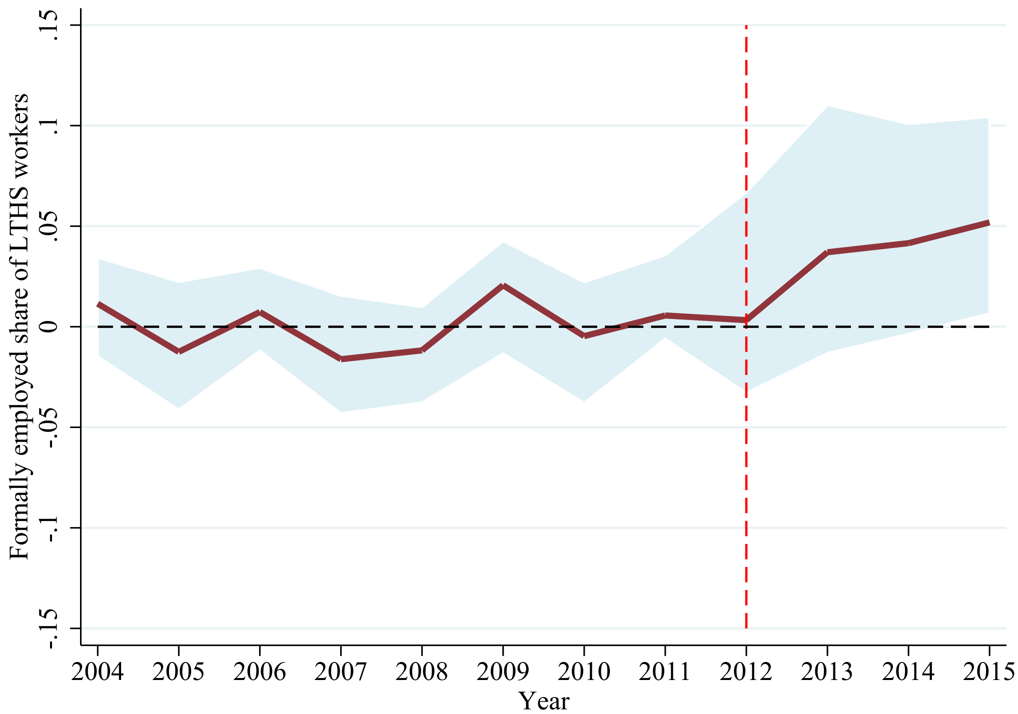
Notes: The graphs plot the frequency distributions of the guest density in 2015 at NUTS-2, and at province level. The vertical dash lines separate control regions, the regions excluded in the DiD and the GSC specifications, and the treated regions. The x-axis in panel (b) is broken due to unusually large Syrian density in one province (Kilis). The data provided by Ministry of Interior Directory General of Migration Management is used.

Figure 4: Impact of Syrian migrants on native informal, and LTHS employment rates over time; the GSC



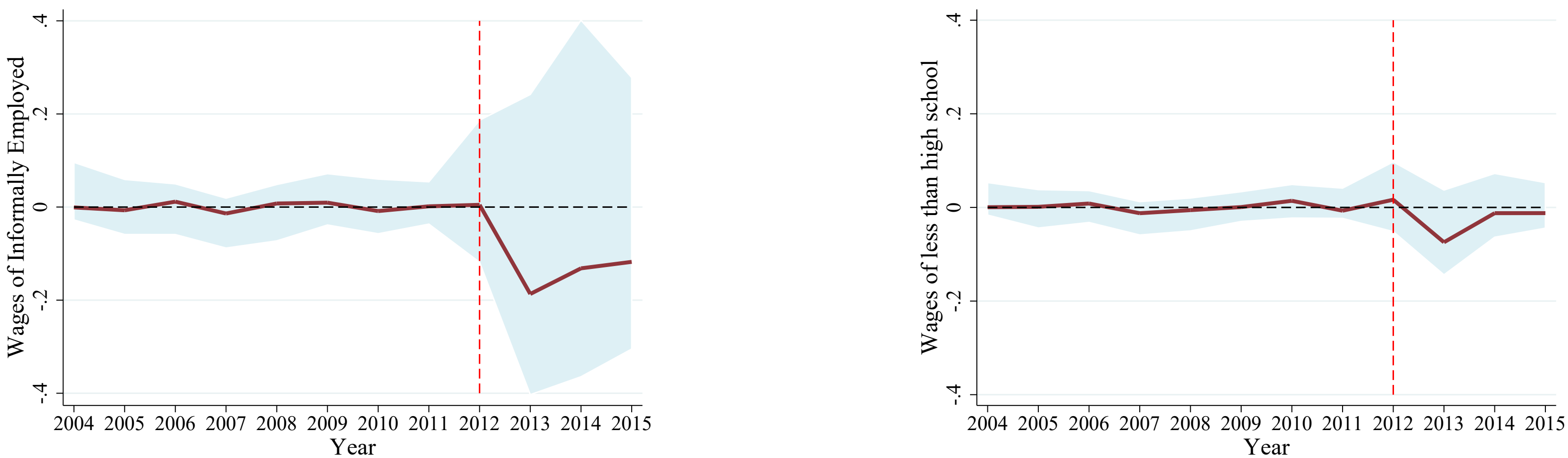
Notes: The graphs plot the changes in the native informal, and less than high school (LTHS) employment rates in the treated regions compared to the counterfactual, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. The vertical dash lines indicate the first year of the migration shock. The generalized synthetic control method (GSC) is employed to estimate the effects. The shaded areas show 95% confidence intervals, calculated using the parametric bootstrap of the GSC.

Figure 5: Impact of Syrian migrants on formally employed share of native LTHS workers over time; the GSC



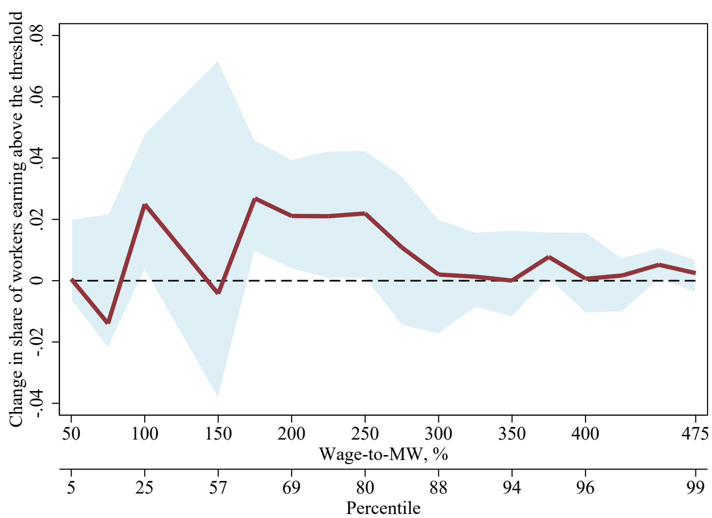
Notes: The figure plots the change in the share of formally employed native LTHS workers out of all native LTHS workers in the treated regions compared to the counterfactual, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. The vertical dash line indicates the first year of the migration shock. The generalized synthetic control method (GSC) is employed. The shaded areas show 95% confidence intervals, calculated using the parametric bootstrap of the GSC.

Figure 6: Impact of Syrian migrants on average wages of native informally employed, and LTHS workers over time; the GSC



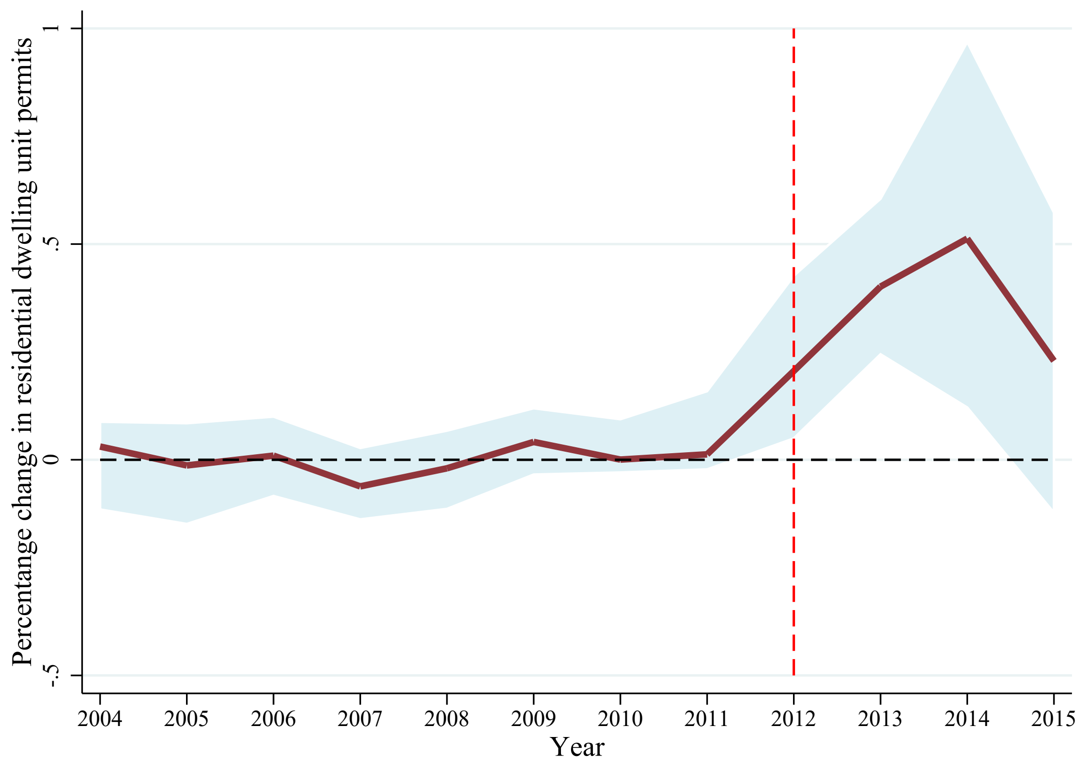
Notes: The graphs plot the percentage change of the average (residual) wages of native informally employed, and LTHS workers in the treated regions compared to the counterfactual, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. The vertical dash lines indicate the first year of the migration shock. The generalized synthetic control method is employed to estimate the effects on the residual wages. The shaded areas show 95% confidence intervals, calculated using the parametric bootstrap of the GSC.

Figure 7: Impact of Syrian migrants on the wage distribution of native workers



Notes: The figure plots the percentage point change in share of workers earning above multiples of the national minimum wage in the treated regions compared to the counterfactual, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. Unless there is no unobserved confounders deemed important, the generalized synthetic control method is employed to estimate the effects. The shaded areas show 90% confidence intervals, calculated using the parametric bootstrap of the GSC. The second x-axis is the average wage percentile value corresponding to the multiples of the national minimum wage.

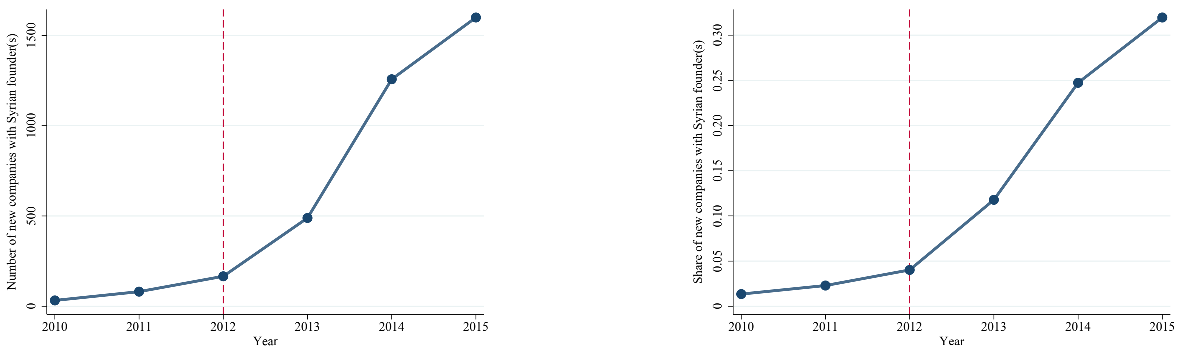
Figure 8: Impact of Syrian migrants on the number of new dwelling unit permits over time; the GSC



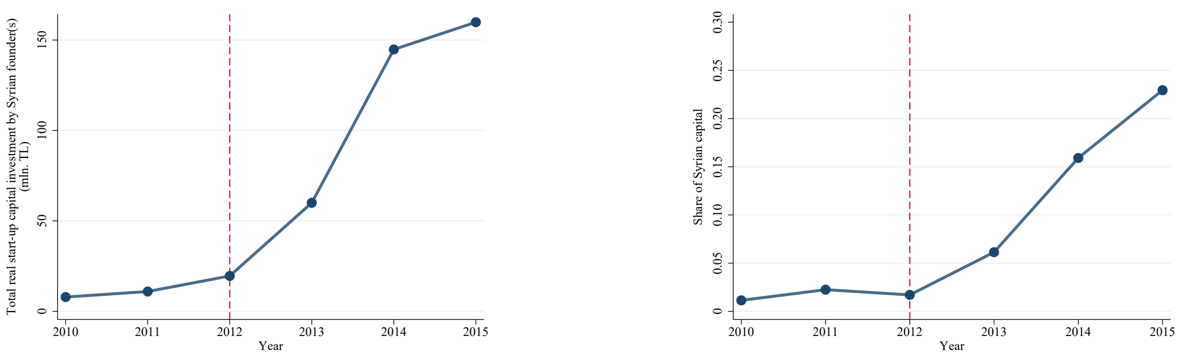
Notes: The figure plots the percentage change in the new dwelling unit building permits in the treated regions compared to the counterfactual, using 2004-2015 province-by-year aggregated TURKSTAT Building Statistics. The vertical dash lines indicate the first year of the migration shock. The generalized synthetic control method is employed to estimate the effects. The shaded area shows the 95% confidence intervals, calculated using the parametric bootstrap of the GSC. For better precision, the regressions are weighted by province's GDP.

Figure 9: Progression of Syrian capital in Turkey

Panel A: Syrian (co-)founded firms



Panel B: Initial capital invested by Syrians



Notes: The graphs plot the evolution of the number of companies founded, and the real initial capital invested by Syrian nationals, using new firm statistics from TOBB. The vertical dash lines indicate the first year of the migration shock. The figures on the left-hand side display the absolute number of Syrian firms or the total start-up capital investment (in 2010 TL) by Syrians; whereas those on the right-hand side show the Syrian shares out of total number of new companies (co-)founded by foreign investors, or out of total foreign start-up capital investments.

Table 1: Characteristics of Syrian migrants, and natives

	Syrian migrant (Age: 15+)	Native (Age: 15-64)	Native (Age: 15-64)
Educational Attainment			
No degree	0.623	0.116	0.234
Primary School	0.215	0.321	0.280
Secondary School	0.086	0.224	0.252
High school	0.047	0.191	0.143
Some college and above	0.027	0.148	0.092
Age groups			
15-18	0.182	0.123	0.176
19-24	0.189	0.106	0.127
25-29	0.154	0.119	0.120
30-34	0.129	0.124	0.116
35-39	0.095	0.117	0.104
40-44	0.069	0.107	0.099
45-49	0.055	0.089	0.078
50-54	0.044	0.088	0.075
55-59	0.030	0.070	0.056
60-64	0.021	0.058	0.050
65+	0.032	-	-
Gender			
Man	0.531	0.501	0.490
Woman	0.469	0.499	0.510
Sample:	Turkey	Turkey	Syrian Density > 10%

Notes. The first column reports the demographic characteristics of the (15+) Syrian migrants in 2015. For comparison, we also provide comparable numbers for all (15-64) natives, and (15-64) natives in the regions where the ratio of Syrian migrant population to natives is at least 10% calculated from 2015 TURKSTAT Household Labor Force Survey. Data on Syrian migrants is from Ministry of Interior Directory General of Migration Management.

Table 2: Descriptive statistics

Variables	Pre-2012 Averages: Density > 10%	Pre-2012 Averages: 10% ≥ Density ≥ 2%	Pre-2012 Averages: 2% > Density	2012-2015 Averages: Density > 10%	2012-2015 Averages: 10% ≥ Density ≥ 2%	2012-2015 Averages: 2% > Density
Labor Force Statistics						
Working age population	1,559,380	2,833,012	1,323,757	1,832,563	3,135,181	1,444,050
Employment rate	0.366	0.435	0.488	0.393	0.487	0.528
Informal employment rate	0.228	0.155	0.235	0.195	0.130	0.207
Employment rate of LTHS	0.335	0.379	0.448	0.354	0.418	0.479
Average wage (in 2010 TL)	793.182	1,013.708	992.321	944.653	1,146.054	1,125.238
Average wage in informal employment (in 2010 TL)	486.547	632.988	541.072	558.988	630.005	593.086
Average wage of LTHS (in 2010 TL)	596.169	744.461	711.090	684.889	797.797	778.216
Building Statistics						
Resid. building permits (m^2)	816,598	2,639,413	768,649	2,219,706	4,086,398	1,127,139
Resid. building permits (# dwelling units)	4,798	17,219	5,043	12,872	26,465	7,086
Resid. building permits (# buildings)	689	2,509	841	1,148	2,999	910
Resid. occupancy permits (# dwelling units)	2,383	8,414	3,635	7,005	20,114	6,051
New Firm Statistics						
# New firm establishments	470.600	1,988.667	303.699	563.250	2,246.107	303.544
Start-up capital investment (in 2010 mln. TL)	173.547	667.043	79.246	133.015	435.109	53.020

Notes. The table reports the mean values for the outcomes. The sample is divided into 6, according to the Syrian guest density, and time dimensions. Pre-2012 corresponds to 2004-2011, 2009-2011, and 2010-2011 for labor and building statistics, the number of new firm establishments, and the total start-up capital investment, respectively. Labor statistics are from TURKSTAT Household Labor Force Survey, building statistics from TURKSTAT, and new firm statistics from TOBB.

Table 3: Impact of Syrian migrants on employment

	(1)	(2)	(3)	(4)
Panel A: DiD				
$\hat{\beta}$	0.008	0.020	0.063	0.013
Clustered se	(0.023)	(0.025)	(0.047)	(0.021)
RI-t p-value	0.779	0.539	0.286	0.607
# Clusters	19	19	19	19
# Treated clusters	3	3	3	3
Observations	228	228	228	228
Panel B: GSC				
$\hat{\beta}$	-0.003	0.000	-	-
GSC SE	0.034	0.047	-	-
GSC p-value	0.898	0.991	-	-
# Unobserved factors	1	1	0	0
# Clusters	19	19	19	19
# Treated clusters	3	3	3	3
Observations	228	228	228	228
Panel C: Alternative model; OLS				
$\hat{\beta}$	0.005	0.017	0.076	0.015
Clustered SE	(0.019)	(0.023)	(0.038)	(0.021)
CGM p-value	0.775	0.457	0.072	0.491
# Clusters	26	26	26	26
Observations	312	312	312	312
Panel D: Alternative model; 2SLS				
$\hat{\beta}$	0.009	0.015	0.069	0.005
Clustered se	(0.022)	(0.023)	(0.043)	(0.020)
WRR p-value	0.762	0.577	0.281	0.806
First-stage F-test	35.925	35.925	35.925	35.925
# Clusters	26	26	26	26
Observations	312	312	312	312
Groups:	Informal	LTHS	HSG	Overall

Notes. The table reports the change in the native informal employment rate, employment rate of the natives with no high school diplome (LTHS), employment rate of natives with at least high school degree (HSG), and overall employment rate in the treated regions after the migration shock, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. The dependent variables are the native informal, LTHS, HSG and overall employment counts normalized by 2011 working-age population or the population of the demographic group. Panels A, B, C, and D employ the DiD, the GSC, the alternative model specification, and the 2SLS to estimate the effects, respectively. In addition to the standard errors clustered at NUTS-2 level and the GSC standard errors, the p-values produced by the randomization inference by t-statistic (RI-t), the parametric bootstrap technique of the GSC, the wild cluster bootstrap (CGM), and the wild restricted residual bootstrap (WRR) are reported for inference. Panel B reports the number of unobserved factors purged by the GSC, and panel D reports the first-stage F-statistics.

Table 4: Impact of Syrian migrants on wages

	(1)	(2)	(3)	(4)
$\hat{\beta}$	-0.113	-0.021	0.057	0.017
GSC SE	0.128	0.025	0.023	0.023
GSC p-value	0.391	0.396	0.011	0.466
# Unobserved factors	3	3	2	3
# Clusters	19	19	19	19
# Treated clusters	3	3	3	3
Observations	228	228	228	228
Groups:	Informal	LTHS	HSG	Overall

Notes. The table reports the percentage change in the (residual) wages of native informal, less than high school (LTHS), high school and above graduates (HSG) and overall workers in the treated regions after the migration shock, using NUTS-2-by-year aggregated 2004-2015 TURKSTAT Household Labor Force Survey. Only the GSC estimates are reported; because the other specifications are likely to suffer from the violation of parallel trends assumption. The GSC standard errors and p-values are calculated using the parametric bootstrap technique. The number of unobserved factors purged by the GSC is reported.

Table 5: Impact of Syrian migrants on residential buildings

	(1)	(2)	(3)
Panel A: DiD			
$\hat{\beta}/\overline{Y}$	0.444	0.443	0.330
Clustered SE	(0.093)	(0.058)	(0.077)
C.R. p-value	0.000	0.000	0.000
# Clusters	67	67	67
# Treated clusters	5	5	5
Observations	804	804	804
Panel B: GSC			
$\hat{\beta}/\overline{Y}$	-	0.337	0.246
GSC SE	-	(0.121)	(0.147)
GSC p-value	-	0.006	0.008
# Unobserved factors	0	2	2
# Clusters	67	67	67
# Treated clusters	5	5	5
Observations	804	804	804
Panel C: Alternative model; OLS			
$\hat{\beta}/\overline{Y}$	0.347	0.341	0.243
Clustered SE	(0.064)	(0.066)	(0.083)
C.R. p-value	0.000	0.000	0.005
# Clusters	81	81	81
Observations	972	972	972
Panel D: Alternative model; 2SLS			
$\hat{\beta}/\overline{Y}$	0.542	0.430	0.196
Clustered se	(0.210)	(0.157)	(0.074)
C.R. p-value	0.012	0.007	0.010
SW F-test; lin. term	7.793	7.793	7.793
SW F-test; quad. term	4.225	4.225	4.225
# Clusters	81	81	81
Observations	972	972	972
Measurement unit:	m^2	dwelling unit	building

Notes. The table reports the percentage change in the residential building permits in the treated regions after the migration shock, using province-by-year aggregated 2004-2015 TURKSTAT building statistics. The dependent variable is the total new building permits in m^2, in the number of dwelling units, and in the number of buildings divided by 2011 GDP of the province. The percentage change is approximated using the equation 8. Panels A, B, C, and D employ the DiD, the GSC, the alternative model specification, and the 2SLS to estimate the effects, respectively. Standard errors are clustered at province level or the GSC standard errors are reported. The corresponding p-values are reported for inference. Panel B reports the number of unobserved factors purged by the GSC, and panel D reports Sanderson-Windmeijer first-stage F-statistics. For better precision, the regressions are weighted by province's GDP.

Table 6: Impact of Syrian migrants on new company establishments

	Number of Firms		Start-up Capital	
	(1)	(2)	(3)	(4)
Panel A: DiD				
$\hat{\beta}$	0.362	0.265	0.521	0.449
Clustered se	(0.093)	(0.063)	(0.102)	(0.078)
C.R. p-value	0.000	0.000	0.000	0.000
# Clusters	67	67	67	67
# Treated clusters	5	5	5	5
Observations	469	469	402	402
Panel B: Alternative model, OLS				
$\hat{\beta}$	0.337	0.278	0.500	0.464
Clustered se	(0.099)	(0.094)	(0.130)	(0.128)
C.R. p-value	0.001	0.004	0.000	0.000
# Clusters	81	81	81	81
Observations	567	567	486	486
Panel C: Alternative model, 2SLS				
$\hat{\beta}$	0.240	0.207	0.561	0.540
Clustered SE	(0.101)	(0.111)	(0.203)	(0.210)
C.R. p-value	0.020	0.066	0.007	0.012
SW F-test; lin. term	8.772	8.772	8.770	8.770
SW F-test; quad. term	7.458	7.458	7.456	7.456
# Clusters	81	81	81	81
# Treated clusters				
Observations	567	567	486	486
Syrian share excluded :		Y		Y

Notes. The table reports the percentage change in the new company establishments, and real start-up capital invested in the treated regions after the migration shock, using province-by-year aggregated 2009-2015 and 2010-2015 TOBB firm statistics. The dependent variables are log-transformed number of new company establishments, and log-transformed real start-up capital invested. The second and fourth columns exclude companies with at least one Syrian co-founder, and the Syrian capital. Panels A, B, and C employ the DiD, the alternative model specification, and the 2SLS to estimate the effects, respectively. Standard errors clustered at province level, and the corresponding p-values are reported for the precision and the inference. Panel C reports the Sanderson-Windmeijer first-stage F-statistics.

Table 7: Search for the elasticity of substitution between HSG and LTHS workers

LTHS migrants	HSG-equivalent migrants	$\sigma_{L_{L,N},L_{L,I}} = 6$	$\sigma_{L_{L,N},L_{L,I}} = 9$	$\sigma_{L_{L,N},L_{L,I}} = 12$	$\sigma_{L_{L,N},L_{L,I}} = 15$	$\sigma_{L_{L,N},L_{L,I}} = 18$	$\sigma_{L_{L,N},L_{L,I}} = 21$
200,000	0	1.491	1.404	1.354	1.321	1.299	1.282
200,000	2,000	1.462	1.371	1.318	1.284	1.261	1.244
225,000	0	1.617	1.551	1.509	1.481	1.462	1.448
225,000	2,000	1.589	1.518	1.474	1.445	1.424	1.409

Notes. The table calculates the implied elasticity of substitution between less than high school (LTHS) and high school and above graduate (HSG) workers using the accounting framework in section 2, and the GSC estimates for native employment and wage changes between 2011 and 2015 for alternative values of LTHS and HSG-equivalent migrant workers in the treated regions, and the elasticity of substitution between migrants and natives, $\sigma_{L_{L,N},L_{L,I}}$. The empirical estimate for the elasticity of substitution between HSG and LTHS workers is 1.473.

Appendix A: Data Appendix

Table A.1: Data Appendix

Variables	Description	Panel Structure/Source
Total Number of Syrian Guests in Turkey	Total number of Syrian migrants temporary protection in Turkey	Annual, National-level / Ministry of Interior, Directorate General of Migration Management
Province-level residence data of Syrian guests in 2015	Province level distribution of Syrians under temporary protection in 2015	Province level / Ministry of Interior, Directorate General of Migration Management
Employment rate of Syrian guests	Employment rate of Syrian migrants at national level	National-level / Balcılar and Nugent (2016)
Treatment Regions (Provinces)	Regions (Provinces) that the number of Syrian migrants in 2015 is more than 10% of the native population are considered as treated regions. The first treatment year is 2012. Used in the DiD and the GSC.	Annual, NUTS-2 or province-level / Constructed variable
Control Regions (Provinces)	Regions (Provinces) that the number of Syrian migrants in 2015 is less than 2% of the native population are considered as control regions. Used in the DiD and the GSC.	Annual, NUTS-2 or province-level / Constructed variable
Native Population	The total number of native population.	Annual, province level / TURKSTAT
Native Working Age Population	The number of native population of ages 15-64.	Annual, NUTS-2 level / TURKSTAT Household Labor Force Survey
Employment	The number of native working population between ages 15-64.	Individual level / TURKSTAT Household Labor Force Survey
Informal Employment	The number of native working population between ages 15-64 with no social security coverage	Individual level / TURKSTAT Household Labor Force Survey
Education	The educational level of the native population between ages 15-64. Categories are; less than primary school, primary school, middle school, high school, vocational high school, some college or college, graduate school	Individual level / TURKSTAT Household Labor Force Survey

Table A.1 : Data Appendix (continued)

Variables	Description	Panel Structure/Source
Age	The categorical age variable. Categories are [15, 20), [20, 25) … [60, 65).	Individual level / TURKSTAT Household Labor Force Survey
Wage	Monthly after tax wage data of the native working population between ages 15-64. Includes bonuses, performance pays.	Individual level / TURKSTAT Household Labor Force Survey
New Residential Building Permits	The number of new building permits given for dwelling purposes. Administrative data.	Annual, Province level / TURKSTAT
New Residential Occupancy Permits	The number of new occupancy permits given for completed buildings for dwelling purposes. Administrative data.	Annual, Province level / TURKSTAT
Total number of new company establishments	The number of new company establishments in each province. Administrative data.	Annual, Province level / TOBB
Total number of firm establishments by Syrian founders, province-level	Similar to above, only by Syrian nationals. Administrative data.	Annual, Province level / TOBB, Özpınar et al. (2015)
Total amount of start-up capital invested	Total amount of capital invested initially in new firms. Administrative data.	Annual, Province level / TOBB
Total amount of start-up capital invested by Syrian founders	Similar to above, only by Syrian nationals. Administrative data.	Annual, Province level / TOBB
Gross Provincial Product	The value which is equal to the sum of the values of taxes minus subsidies and gross value added by province.	Annual, Province level / TURKSTAT
Arabic Speaking Population in 1965	Total number of people with Arabic as the first language	Province level / TURKSTAT

Notes: The table reports all the variables, the descriptions, and the data sources used.

Appendix B: Re-examination of Tumen (2016) and Del Carpio and Wagner (2015)

This section presents our observations on the main findings of Tumen (2016), and Del Carpio and Wagner (2015).

We begin with Tumen (2016). The study estimates that the informal employment rate significantly declines by 2.3% in the regions that host a relatively large share of Syrian migrants. The four main differences between the baseline regression model of Tumen (2016) and the DiD specification used in panel A of table 3 are 1) treated and control regions included in the regression, 2) time period, 3) the method used in calculating standard errors, and 4) demographic controls. [42]

There are five treated and four control regions in Tumen (2016). Their definition for the treated region is more inclusive, so it includes all of our treated regions and two of the regions we drop in the primary sample. The four control regions are geographically proximate to the treated ones. Our primary control sample includes all of them. Secondly, the sample Tumen (2016) employs is from 2010 to 2013. Thirdly, Tumen (2016) accounts for neither the fact that the value of the variable of interest does not change across individuals for a given region-year nor the serial correlation. Lastly, Tumen (2016) includes gender, marital status, age, education, age-by-education, and urban area controls in the regressions.

In table B.1, we search for the factors causing the discrepancy in conclusions. In column 1, we replicate Tumen (2016)'s main finding: Informal employment rate falls by 2.3% in the regions with high Syrian density. Our replication matches perfectly with the reported estimate (first row) and the heteroskedasticity-consistent standard error (second row) in Tumen (2016). In the third row, we cluster the standard errors at NUTS-2 regional level. This more than quintuples the standard errors and renders the estimate insignificant. As shown in Bertrand et al. (2004), when the data is at individual-level, the use of heteroskedasticity-consistent standard errors severely over-rejects the

[42] They drop observations that report birth places other than Turkey and defines two NUTS-2 regions (Adana and Mersin) that are dropped from our primary sample as treated regions. The difference due to dropping people born abroad is negligible, constitutes less than %1 of Tumen (2016)'s sample, and the variable is absent in earlier surveys. Hence, we keep these observations.

true null hypothesis if the treatment affects every individual in the region. Then, the first reason behind the discrepancy in findings is due to the use of different inference methods. Column 1 of table B.1 shows that allowing for within-region correlation of the errors renders Tumen (2016)'s main estimate insignificant.

In column 2, we do not control for demographic factors. The finding of this column is qualitatively same as the first one: The informal employment rate declines by 2.4%, and the 95% confidence interval estimated using the cluster-robust standard error is too large to reject the null hypothesis. In column 3, we expand the time span of the sample to explore whether it is the longer panel causing the conflict. This increases the absolute magnitude of the negative estimate ($\hat{\beta} = -0.04$); hence, we conclude that neither the time window nor the use of demographic controls is causing the discrepancy.

In columns 4 and 5 of table B.1, we estimate a geographic placebo test along the lines of Zipperer (2016). Essentially, we drop the treated regions as defined by Tumen (2016) from the sample and assign a placebo shock to their 4 control regions. In column 4, we use all the other 17 NUTS-2 regions in Turkey for control regions; while in column 5, we only use the control regions in our primary sample (4 proximate ones as placebo treated regions and remaining 12 as controls). Both regressions produce essentially the same estimate. Compared to other regions in Turkey, the control regions of Tumen (2016) have experienced 4% increase in the informal employment rate. In other words, the negative estimated effect in Tumen (2016) is essentially due to the increase in informal employment rate in the control regions compared to the rest of Turkey. When other non-Western regions with small Syrian population are used as controls for their treated regions in column 6, the negative estimate disappears. The estimated effect is positive and very close to zero ($\hat{\beta} = 0.01$).

In Ceritoglu et al. (2017), they check the robustness of baseline estimates in Tumen (2016) by trying 12 different alternative control samples that contain 4 regions. In addition to the aforementioned covariates, this test controls for the log of regional foreign trade volume. They claim to have confirmed Tumen (2016)'s conclusion. We have two criticisms for the test. Firstly, migration affects foreign trade volume, therefore including the latter as a control removes one of the channels that counteracts the potential adverse effects (Gaston and Nelson, 2013). Secondly, there are $\binom{26-9}{4} = 2380$ different possible control samples. It is unclear why only 12 of these regression results are reported.

To summarize, table B.1 shows that the two factors behind the discrepancy are that the control regions of Tumen (2016) followed a path dissimilar to all other regions, and their confidence intervals are too narrow.[43] Although the confidence intervals in table B.1 include the baseline estimate of Tumen (2016), they cannot reject the null hypothesis.

Another study that reports significant negative effect of Syrian migration on informal employment in Turkey is Del Carpio and Wagner (2015). They compare 2011 and 2014 informal employment rates and employ an instrumental variable strategy. They control for the distance to the border, primarily to prevent changing trade patterns with Syria due to the war and different underlying economic trends in distant regions from affecting the estimates. They instrument Syrian guest share with the travel distance from origin governorates in Syria to NUTS-2 regions in Turkey using the Google Maps. Thus, the identification depends on the multiple border-crossings between Turkey and Syria. To assess pre-existing trends, they perform placebo tests by changing the timing of the shock. In the test, 2009 is declared as the reference year and 2011 is the post-treatment period.

We make three observations on the study: First, based on the reasoning behind the use of distance-to-border, it is not entirely clear if an accurate counterfactual is produced. Changing trade patterns with Syria might as well be due to the migration, hence should not be controlled but included in the analysis of the effects of the migration. Additionally, by implicitly increasing the importance of the neighboring regions in the control sample, the issues of Tumen (2016) control sample might be present here as well. Second, based on the reported estimates in placebo table, it appears that in the absence of the treatment, the actuals and counterfactuals do not follow a parallel trend. Most of the placebo estimates are, in absolute terms, at least half the size of their benchmark estimates. This makes the estimates sensitive to the reference year. Concretely, if the year of 2009 or the average of 2009 and 2011 were declared as the reference, most of the negative employment estimates would be substantially smaller in size and statistically insignificant. Third, the standard errors are clustered at region-by-year level, so they do not account for serial correlation. Doing so might reveal that standard errors are underestimated.[44]

[43]Unreported RI-t or CGM confidence intervals also contain the null hypothesis.

[44]The authors note that there are 26 regions in Turkey; hence clustering at regional level has caused the over-rejection of the null, and resulted in smaller standard errors. As we point out in the text, this is expected when the number of clusters is few. However, as shown in Cameron and Miller (2015), there are methods to address the issue that do not require assuming away the serial correlation.

Table B.1: Re-examination of Tumen (2016)

Time Span:	Yrs 2010 to 2013 (1)	Yrs 2010 to 2013 (2)	Yrs 2004 to 2015 (3)	Yrs 2004 to 2015 (4)	Yrs 2004 to 2015 (5)	Yrs 2004 to 2015 (6)
$\hat{\beta}$	-0.0226	-0.0237	-0.0425	0.0392*	0.0402	0.0082
H.C. se	(0.0028)	(0.0030)	(0.0018)	(0.0015)	(0.0016)	(0.0014)
C.R. se	(0.0183)	(0.0211)	(0.0230)	(0.0218)	(0.0236)	(0.0207)
C.R. p-value	0.252	0.294	0.101	0.087	0.109	0.699
# Clusters	9	9	9	21	16	12
# Treated clusters	5	5	5	4	4	5
Observations	354,513	354,513	1,074,587	3,250,172	2,095,268	1,580,976
Specification						
Demographic controls	Y					
Tumen (2016) regions	Y	Y	Y			
Placebo				Y	Y	
Baseline sample control regions					Y	Y
Western & Tumen (2016) control regions excluded						Y

Notes. The table reports the change in the informal employment rate in the treated regions after the migration shock, using individual-level 2004-2015 TURKSTAT Household Labor Force Survey. The dependent variable is the indicator for informal employment. The first column replicates the column 1 of table 1 in Tumen (2016). First line reports the point estimate, the second line the heteroskedasticity-consistent standard errors, and the third line the robust standard errors clustered at NUTS-2 level. Demographic controls include gender, marital status, age, education, age-by-education, and urban area controls. Placebo specification indicates the regressions where we define Tumen (2016) control regions as the treated regions, and the treated regions in Tumen (2016) are excluded from the sample. Tumen (2016) control regions are Erzurum, Agri, Malatya and Van. Baseline sample control regions are the ones shown in figure 3, and the Western regions are Istanbul (TR1), West Marmara (TR2), Aegean (TR3), and West Anatolia (TR5) regions.

Appendix C: Additional Figures and Tables

This section presents additional figures and tables.

Figure C.1 shows the empirical cumulative distribution function of the t-statistics obtained from placebo employment regressions. The acceptance region for the null hypothesis when the test size is 5% is relatively large and slightly skewed, yet the curves are relatively smoothly S-shaped with one visibly clear inflection point. This indicates that the outliers have not affected RI-t p-values substantially.

Figure C.2 shows importance of employing the GSC model and the residual wages as the dependent variable. The DiD specification as well as using the log wages for the outcome always produce a U-shaped pattern before the shock. This indicates violation of the parallel trends assumption.

Figure C.3 plots the evolution of occupancy permits. The rise in the new building permits only means that the government has allowed entrepreneurs to build on the designated lot. It does not indicate whether buildings are erected on the site. The figure confirms that the buildings are in fact constructed and ready to be occupied. With approximately two years of lag, starting from the year of 2014, the increase in the number of residential occupancy permits in dwelling units is considerable and its size is comparable to that of the building permits. [45]

Table C.1 reports the impact of Syrian migrants on sub-groups of the groups in table 3. The first column reports the change in teen employment-rate. 77% of teens in 2011 are informally employed, hence they are expected to be highly affected by the shock.[46] The second and third columns examine LTHS man and woman separately to assess whether the effects are similar for men and women. The fourth column reports the change in informally employed LTHS workers, the intersection of highly impacted group and highly impacted jobs. The last column excludes middle school graduates from the highly impacted group, and only considers individuals with less than middle school degree. We find a sizable or statistically significant negative effect for none of the sub-groups. The absolute

[45] The GSC indicates that the DiD performs the best. Estimating the model with one or two unobserved factors produces virtually the same results.

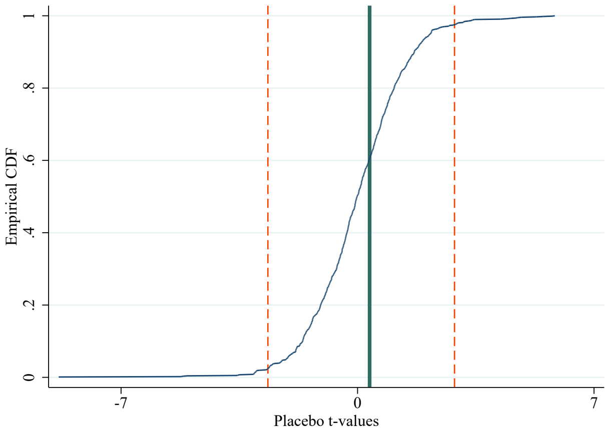
[46] Note that the compulsory education has increased from 8 years to 12 years in 2012. Thus, the estimates for teen employment might be partly affected by differential enforcement of the policy.

magnitudes of the estimates are smaller than 0.01, corroborating our conclusion that the migrants have not led to employment loss for natives.

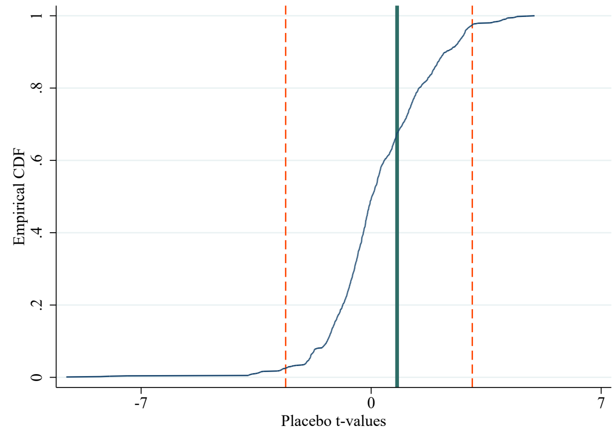
For completeness, we present estimated (residual) wage effects using the DiD, the alternative models (OLS and 2SLS) in table C.2. We wish to caution the reader that these models are potentially suffering from the pre-existing trends.

Due to the existence of one observation with zero residential building permits, we do not employ log transformation to estimate the percentage change in the number of building permits in the main text. However, this model does not require the approximation in equation 7, hence can be considered as a more direct approach for estimating the percentage change. In table C.3, we present estimated percentage increases in the residential building permits using the log-transformed dependent variable. We replace missing observation with 0. The findings are qualitatively same as those in 5. The GSC recommends the DiD in all columns.

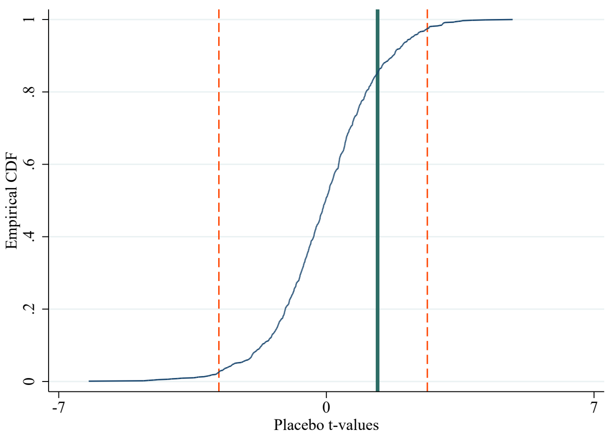
Figure C.1: Empirical CDF of employment estimates



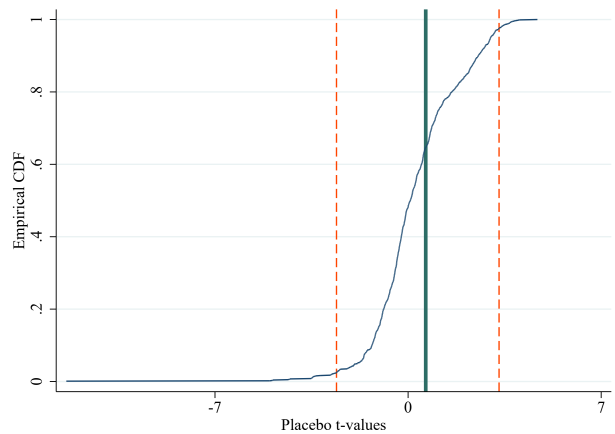
(a) Informal employment



(b) LTHS



(c) HSG

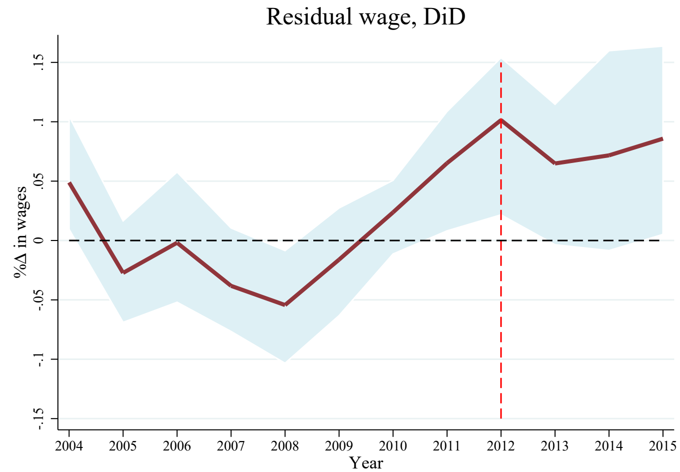


(d) Overall

Notes: The empirical cumulative distribution functions of placebo t-values obtained from placebo regressions of native informal, LTHS (less than high school), HSG (high school graduate and above), and overall employment rates are plotted. The vertical dash lines indicate 95% acceptance region of the null hypothesis of no effect, and the vertical straight line shows the true point estimate.

Figure C.2: Average wages of all workers over time

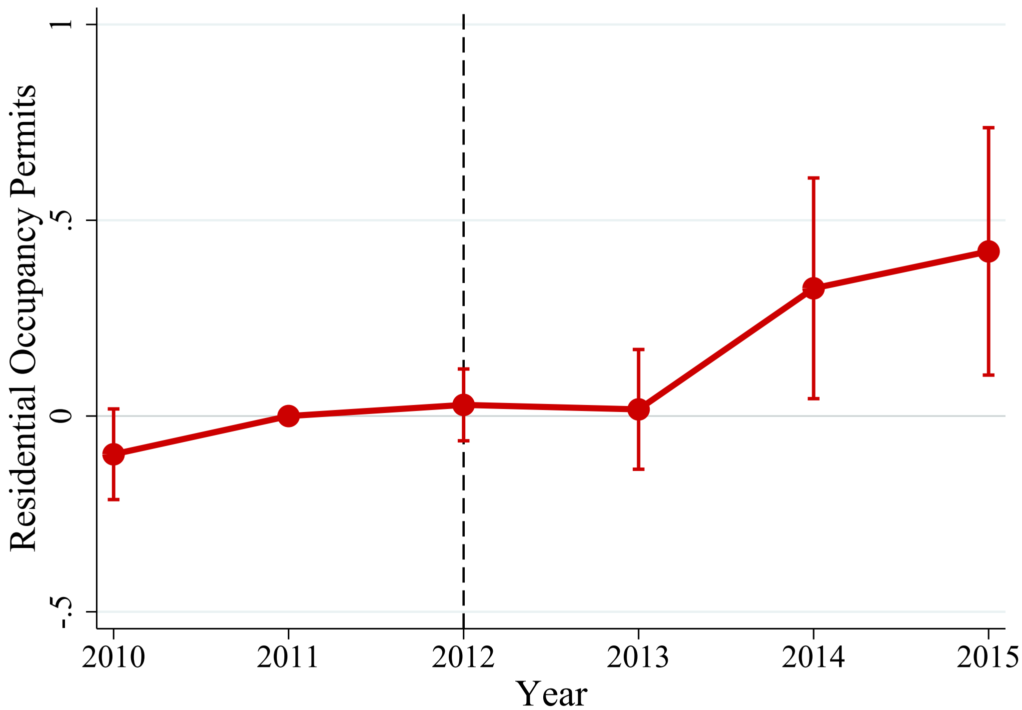
Panel A: Difference-in-differences (DiD)



Panel B: Generalized synthetic control (GSC)

Notes: The graphs plot the evolution of the average wages in the treated regions, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. The vertical dash lines indicate the first year of the migration shock. The top two figures employ the DiD, and the bottom two the generalized synthetic control method. The dependent variables in the graphs on the left-hand side are the average wage; whereas it is the average residual wage for the ones on the right-hand side. The shaded areas show 95% confidence intervals.

Figure C.3: Impact of Syrian migrants on new residential occupancy permits over time



Notes: The figure plots the evolution of the new dwelling unit occupancy permits in the treated regions, using 2004-2015 province-by-year aggregated TURKSTAT Building Statistics. The vertical dash line indicates the first year of the migration shock. The GSC suggests the validity of the difference-in-differences model, hence it is employed. The capped spikes show 95% confidence intervals, calculated using the standard errors clustered at NUTS-2 level. The regressions are weighted by province's GDP.

Table C.1: Impact of Syrian migrants on employment; additional results from sub-groups

	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	0.007	0.000	-0.001	-0.002	0.001
SE	(0.053)	(0.044)	(0.025)	(0.043)	(0.052)
P-value	0.717	0.949	0.965	0.920	0.999
# Unobserved factors	1	2	0	1	1
# Clusters	19	19	19	19	19
# Treated clusters	3	3	3	3	3
Observations	228	228	228	228	228
Groups:	Teen	LTHS, man	LTHS, woman	LTHS, informal	Less than middle school

Notes. The table reports the change in the native teen, LTHS male, LTHS woman, LTHS informal employment rates, and employment rate of native individuals with no middle school degree (LTMS) in the treated regions after the migration shock, using 2004-2015 NUTS-2-by-year aggregated TURKSTAT Household Labor Force Survey. The dependent variables are the native teen, LTHS male, LTHS woman, LTHS informal, and LTMS employment counts normalized by 2011 population of the demographic group. In columns (1), (2), (4), and (5), the GSC is employed. In column (3), the GSC recommends the DiD. Reported standard errors and the p-values are the GSC standard errors and corresponding p-values, except in column (3). In column (3), we report the standard error clustered at NUTS-2 level, and the p-value produced by the randomization inference by t-statistic (RI-t). The number of unobserved factors purged by the GSC is reported.

Table C.2: Impact of Syrian migrants on wages; DiD and alternative specifications

	(1)	(2)	(3)	(4)
Panel A: DiD				
$\hat{\beta}$	0.024	0.031	0.053	0.041
Clustered SE	(0.053)	(0.027)	(0.015)	(0.017)
RI-t p-value	0.706	0.369	0.042	0.122
# Clusters	19	19	19	19
# Treated clusters	3	3	3	3
Observations	228	228	228	228
Panel B: Alternative model; OLS				
$\hat{\beta}$	0.022	0.042	0.051	0.050
Clustered SE	(0.054)	(0.026)	(0.014)	(0.014)
CGM p-value	0.695	0.141	0.053	0.058
# Clusters	26	26	26	26
Observations	312	312	312	312
Panel C: Alternative model; 2SLS				
$\hat{\beta}$	0.048	0.040	0.055	0.049
Clustered se	(0.059)	(0.025)	(0.013)	(0.014)
WRR p-value	0.368	0.137	0.071	0.075
First-stage F-test	35.925	35.925	35.925	35.925
# Clusters	26	26	26	26
Observations	312	312	312	312
Groups:	Informal	LTHS	HSG	Overall

Notes. The table reports the percentage change in the (residual) wages of LTHS, informal and overall workers in the treated regions after the migration shock, using NUTS-2-by-year aggregated 2004-2015 TURKSTAT Household Labor Force Survey. Panels A, B, and C employ the DiD, the alternative model specification, and the 2SLS to estimate the effects, respectively. Standard errors clustered at NUTS-2 level, and the corresponding p-values are reported for the precision and the inference. Panel C reports the first-stage F-statistics.

Table C.3: Impact of Syrian migrants on residential buildings; Log-transformed dependent variable

	(1)	(2)	(3)
Panel A: DiD			
$\hat{\beta}$	0.630	0.673	0.472
Clustered se	(0.074)	(0.069)	(0.073)
C.R. p-value	0.000	0.000	0.000
# Clusters	67	67	67
# Treated clusters	5	5	5
Observations	804	804	804
Panel B: GSC			
$\hat{\beta}$	-	-	-
GSC SE	-	-	-
GSC p-value	-	-	-
# Unobserved factors	0	0	0
# Clusters	67	67	67
# Treated clusters	5	5	5
Observations	804	804	804
Panel C: Alternative model; OLS			
$\hat{\beta}$	0.519	0.554	0.365
Clustered se	(0.074)	(0.093)	(0.092)
C.R. p-value	0.000	0.000	0.000
# Clusters	81	81	81
Observations	972	972	972
Panel D: Alternative model; 2SLS			
$\hat{\beta}$	0.748	0.683	0.302
Clustered se	(0.176)	(0.179)	(0.120)
C.R. p-value	0.000	0.000	0.014
SW F-test; lin. term	7.793	7.793	7.793
SW F-test; quad. term	4.225	4.225	4.225
# Clusters	81	81	81
Observations	972	972	972
Measurement unit	m^2	dwelling units	buildings

Notes. The table reports the percentage change in the residential building permits in the treated regions after the migration shock, using log-transformed province-by-year aggregated 2004-2015 TURKSTAT building statistics. The dependent variable is the log-transformed building permits in m^2, in the number of dwelling units, and in the number of buildings. Panels A, B, C, and D employ the DiD, the GSC, the alternative model specification, and the 2SLS to estimate the effects, respectively. Standard errors are clustered at province level or the GSC standard errors are reported. The corresponding p-values are reported for inference. Panel B reports the number of unobserved factors purged by the GSC, and panel D reports Sanderson-Windmeijer first-stage F-statistics. For better precision, the regressions are weighted by province's GDP.